

RESEARCH HANDBOOK ON **Inventory Management**

Edited by
Jing-Sheng Jeannette Song



RESEARCH HANDBOOK ON INVENTORY MANAGEMENT

To Shucheng, Peter, and Daphne

Research Handbook on Inventory Management

Edited by

Jing-Sheng Jeannette Song

R. David Thomas Professor of Business Administration and Professor of Operations Management, The Fuqua School of Business, Duke University, USA



Cheltenham, UK • Northampton, MA, USA

© Jing-Sheng Jeannette Song 2023

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2023939803

This book is available electronically in the **Elgaronline**
Business subject collection
<http://dx.doi.org/10.4337/9781800377103>

ISBN 978 1 80037 709 7 (cased)
ISBN 978 1 80037 710 3 (eBook)

Contents

<i>List of contributors</i>	vii
<i>Preface</i>	xiv
<i>Acknowledgments</i>	xxi

PART I FUNDAMENTALS – THEORY AND METHODOLOGIES

1	Lost-sales inventory systems <i>Marco Bijvank, Woonghee Tim Huh, and Ganesh Janakiraman</i>	2
2	Perishable inventory systems <i>Qing Li and Peiwen Yu</i>	27
3	Capacitated inventory systems <i>Roman Kapuściński and Rodney P. Parker</i>	46
4	Generalizations of the Clark–Scarf model and analysis <i>Alexandar Angelus</i>	73
5	Single-stage approximations of multi-echelon inventory models <i>Kevin H. Shang, Jing-Sheng Jeannette Song, and Sean X. Zhou</i>	99
6	Single-unit analysis <i>Alp Muharremoglu, Xin Geng, and Nan Yang</i>	125
7	Robust inventory management <i>Michael R. Wagner</i>	147
8	Dual-sourcing, dual-mode dynamic stochastic inventory models <i>Linwei Xin and Jan A. Van Mieghem</i>	165
9	Assemble-to-order systems <i>Levi DeValve, Jing-Sheng Jeannette Song, and Yehua Wei</i>	191
10	Inventory models with returns and remanufacturing <i>Xiting Gong and Sean X. Zhou</i>	213
11	Approximation algorithms for stochastic inventory systems <i>Cong Shi</i>	234

PART II INTERFACES

12	Information and incentives in inventory management <i>Bharadwaj Kadiyala, Hau Lee, and Özalp Özer</i>	261

13	Joint pricing and inventory decisions <i>Xin Chen, Peng Hu, and Zhenyu Hu</i>	288
14	Statistical learning in inventory management <i>Wang Chi Cheung and David Simchi-Levi</i>	307
15	Online learning in inventory and pricing optimization <i>Xiuli Chao, Boxiao Chen, and Huanan Zhang</i>	333
16	Inventory models with financial flows <i>Kevin H. Shang and Jing-Sheng Jeannette Song</i>	379
17	Behavioral inventory management <i>Andrew M. Davis and Jordan D. Tong</i>	400
 PART III CONTEXT SPECIFIC MODELS AND METHODS		
18	Healthcare inventory management <i>Turgay Ayer, Chelsea C. White III, and Can Zhang</i>	431
19	Spare parts inventory planning <i>Rob Basten and Geert-Jan van Houtum</i>	455
20	Retail inventory systems <i>Stefan Minner and Anna-Lena Sachs</i>	476
21	Online retailing inventory management <i>Mengxin Wang and Zuo-Jun Max Shen</i>	500
	<i>Index</i>	524

Contributors

Alexandar Angelus is Assistant Professor in the Department of Information and Operations Management at Mays Business School of Texas A&M University. He received his Ph.D. in Operations, Information and Technology from the Graduate School of Business at Stanford University, and B.Sc. in Mathematics from MIT. Prior to joining Texas A&M, he was a faculty member at the University of Texas, Dallas, and Singapore Management University. His main research interests are in supply chain management and renewable energy. In particular, he works on capacity and inventory planning for multi-stage supply chains, and mathematical frameworks for investment in different renewable energy systems. His research has been published in *Management Science*, *Manufacturing Service & Operations Management*, *Operations Research*, and *Production and Operations Management*.

Turgay Ayer is the Virginia C. and Joseph C. Mello Chair and the Research Director for Healthcare Analytics and Business Intelligence in the Center for Health & Humanitarian Systems at Georgia Tech. In addition, Dr. Ayer holds a courtesy appointment at Emory Medical School and serves as a senior scientist at the Centers for Disease Control and Prevention (CDC). Ayer's research focus is analytics for healthcare, social good, and socially responsible operations. Ayer's research findings have been published in top-tier business, engineering, medical, and health policy journals and are widely covered in popular media outlets. In support of his research, Ayer has received over \$3 million in grant funding and several awards for his work, including an NSF CAREER Award. Ayer is actively involved in INFORMS, the largest international association of operations research and analytics professionals, and is a past president of the INFORMS Health Applications Society.

Rob Basten is Associate Professor at the Department of Industrial Engineering & Innovation Sciences of the Eindhoven University of Technology. He studies after-sales services for high-tech equipment. He is especially interested in exploiting new technologies, such as 3D printing of spare parts and new sensing and communication technology that enables predictive maintenance. He is also active in behavioral operations management, with a particular interest in the interaction between humans and decision support systems. Much of his research is interdisciplinary and practice-based.

Marco Bijvank is Associate Professor at the Operations and Supply Chain Management area of the Haskayne School of Business, University of Calgary. His main research interests include stochastic modeling and data analytics with applications in supply chain management (in particular inventory management and retail operations) and healthcare operations (in particular at the emergency department).

Xiuli Chao is the Ralph L. Disney Chair Professor of Industrial and Operations Engineering at the University of Michigan, Ann Arbor. His research interests includes queueing, scheduling, financial engineering, inventory control, supply chain management, and data-driven optimization. He is a co-developer of the Legin Scheduling System, and is the co-author of two books, *Operations Scheduling with Applications in Manufacturing and Services*

(Irwin/McGraw Hill, 1998), and *Queueing Networks: Customers, Signals, and Product Form Solutions* (Irwin/McGraw Hill, 1998). He is a fellow of both IISE and INFORMS.

Boxiao Chen is Associate Professor at the College of Business Administration at the University of Illinois, Chicago. In her research, she applies techniques from stochastic analyses, machine learning, and optimization to develop data-driven algorithms for decision-making. Some of her works include dynamic pricing, inventory control and supply chain management, assortment planning, retailing, energy, and capacity expansion.

Xin Chen is James C. Edenfield Chair and Professor at the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. Prior to this appointment, he was Professor of Industrial Engineering at the University of Illinois at Urbana-Champaign. His research interest lies in optimization, data analytics, revenue management and supply chain management. He received the INFORMS revenue management and pricing section prize in 2009. He is the co-author of the book *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management* (Second Edition, 2005, & Third Edition, 2014), and serving as the department editor of logistics and supply chain management of *Naval Research Logistics* and an associate editor of several leading journals including *Operations Research*, *Management Science*, and *Production and Operations Management*.

Wang Chi Cheung is Assistant Professor at the Department of Industrial Systems Engineering and Management at the National University of Singapore. He completed his Ph.D. at the Operations Research Center at the Massachusetts Institute of Technology, supervised by David Simchi-Levi. Wang Chi is interested in data-driven optimization, with applications to revenue management and inventory control models. He is a recipient of the Singapore Agency for Science, Technology and Research scholarships from 2007 to 2010, and from 2011 to 2016. Wang Chi was also a finalist in the George Nicholson Student Paper Competition in 2015.

Andrew M. Davis is Associate Professor of Operations, Technology, and Information Management at the Johnson Graduate School of Management at Cornell University in Ithaca, NY. His research examines problems within supply chain management and service operations, including procurement, inventory, contracting, bargaining, and the sharing economy. To investigate these problems, he employs a combination of behavioral experiments and analytical models, with the aim of identifying practical insights for managers.

Levi DeValve is Assistant Professor of Operations Management at the University of Chicago Booth School of Business. Professor DeValve's research applies optimization tools to a variety of supply chain problems, including assemble-to-order systems, inventory management, e-commerce fulfillment, and network design, and has been recognized with awards in the MSOM Practice-Based Research Competition and Service Science Paper Competition. He received his Ph.D. in Decision Sciences at Duke University's Fuqua School of Business in 2019.

Xin Geng is Assistant Professor of Management at Miami Herbert Business School, University of Miami. His research interests include supply chain strategies, issues in OM/Marketing interface, and optimization models for pricing and revenue management. His work has been published in *Management Science*, *Manufacturing & Service Operations Management*, and *Production and Operations Management*. Xin received his Ph.D. in Management Science from the University of British Columbia in 2015, and has worked at the University of Miami since then.

Xiting Gong is Associate Professor at the Department of Decision Sciences and Managerial Economics, at the Chinese University of Hong Kong (CUHK) Business School. His research interests include operations management, stochastic inventory theory and applications, revenue management and pricing, and approximation and data-driven algorithms.

Peng Hu is Professor at the School of Management at Huazhong University of Science and Technology. He received his Ph.D. degree in Operations Research from the University of Illinois at Urbana-Champaign, his master's degree in Control & Operations Research from the Chinese Academy of Sciences, and his bachelor's degree in Statistics from Peking University. His research interests include optimal pricing and inventory management, and customer behavior in operations management.

Zhenyu Hu is Associate Professor at the Department of Analytics & Operations of the NUS Business School, National University of Singapore. He received his Ph.D. in Industrial Engineering from the University of Illinois at Urbana-Champaign, and B.Sc. in Applied Mathematics from Sun Yat-sen University. His research focuses on dynamic pricing, revenue management, and inventory and supply chain management.

Woonghee Tim Huh is Professor at the Sauder School of Business at the University of British Columbia. His current research interests include inventory control, assortment optimization, and pricing. He holds Canada Research Chair in Operations Excellence and Business Analytics.

Ganesh Janakiraman is Ashbel Smith Professor in the Operations Management area of the Naveen Jindal School of Management, at the University of Texas at Dallas. The primary research methodologies he employs are stochastic, dynamic optimization, and mechanism design. He applies these to inventory theory, sourcing, and real-time bidding in online advertising, etc.

Bharadwaj Kadiyala is an Assistant Professor of Operations Management at the David Eccles School of Business, University of Utah. Using both theory and data, his research investigates operational issues that arise in dynamic incomplete information settings in pricing, revenue and inventory management. His research is published in *Management Science* and *Production and Operations Management*. He received his Ph.D. in Management Science from the University of Texas at Dallas in September 2017.

Roman Kapuściński is the John Psarouthakis Research Professor of Manufacturing Management, Technology, and Operations Management at the University of Michigan Ross School of Business. His primary research interests include stochastic modeling with applications to supply chain management, revenue management, with recent emphasis on energy markets.

Hau Lee is the Thomas Professor of Operations, Information and Technology at the Stanford Graduate School of Business. His areas of specialization include global value chain innovations, supply chain management, global logistics, inventory modeling, and environmental and social responsibility. Professor Lee was elected to the National Academy of Engineering in 2010. He received the Harold Lardner Prize for International Distinction in Operations Research, Canadian Operations Research Society, 2003. He was elected a Fellow of Manufacturing and Service Operations Management, INFORMS, 2001; Production and Operations Management Society, 2005, and INFORMS, 2005.

Qing Li is Professor of Operations Management at the HKUST Business School and the Academic Director for the M.Sc. in Global Operations (MSGO) program. Besides perishable inventory systems, his current research projects include high-volume recruitment, land-based salmon production and harvesting, self-control in managing projects, and new media. He received his degrees from the University of British Columbia, Fudan University, and Tsinghua University. He is an avid runner and hiker.

Stefan Minner is Full Professor of Logistics and Supply Chain Management at TUM and a core member of the Munich Data Science Institute. His research interests are in global supply chain design, transportation optimization, and inventory management. His work has been published in *Management Science*, *Manufacturing & Service Operations Management* and *Operations Research*, *Production and Operations Management*. Currently, Stefan Minner is the editor-in-chief of the *International Journal of Production Economics* and a fellow of the International Society for Inventory Research.

Alp Muharremoglu is Senior Principal Scientist in the Supply Chain Optimization Technologies team at Amazon. His research interests are in sequential decision-making under uncertainty, focusing on applications in inventory and revenue management. He received his Ph.D. in Operations Research at MIT, and before joining Amazon he was on the faculty of Columbia Business School and the University of Texas at Dallas.

Özalp Özer is an executive leader at Amazon and the George and Fonsa Brody Professor of Management Science at The University of Texas at Dallas. At Amazon, he is helping build the science required to manage and optimize the supply chain of the world's biggest marketplace. His areas of specialty include end-to-end management and coordination of global value chains, entrepreneurship and innovation, strategic investment decisions, capacity and inventory planning, market timing, distribution channel management, procurement contract design, and retail and pricing management. Professor Özer was voted the Favorite Professor by Poets & Quants for Executives in 2015, endowed with the Ashbel Smith Professorship from the UT system, the Wickham Skinner Early-Career Research Accomplishment Award from the POM Society, the Hellman faculty fellowship, and the Terman faculty fellowship from Stanford.

Rodney P. Parker is Associate Professor of Operations Management and Fetting/Whirlpool Faculty Fellow at the Indiana University Kelley School of Business. He received his Ph.D. from the University of Michigan and has held prior faculty positions at Yale University, Cornell University, and the University of Chicago. His research interests include supply chain management, inventory theory particularly with capacity limits, healthcare operations management, finance-operations interface, and Markov games.

Anna-Lena Sachs is Senior Lecturer in Predictive Analytics at Lancaster University. She completed her Ph.D. at the Technical University of Munich and then joined the University of Cologne as Assistant Professor in Supply Chain Management. Her main research interests are inventory management, behavioral operations management, and forecasting with a focus on retailing.

Kevin H. Shang is Joseph J. Ruvane, Jr. Distinguished Professor of Business Administration and a Professor of Operations Management at the Fuqua School of Business, Duke University. His research mainly focuses on developing simple and effective inventory policies for supply chain systems. He also conducts research in the interface of operations and finance and renewable energy systems.

Zuo-Jun Max Shen is Chair Professor of Logistics and Supply Chain Management at the University of Hong Kong, jointly holding an appointment with the Faculty of Business and Economics and the Faculty of Engineering. He is on leave from UC Berkeley, where he is Professor in the Department of Industrial Engineering and Operations Research and the Department of Civil and Environmental Engineering. He received his Ph.D. from Northwestern University. He has been active in the following research areas: integrated supply chain design and management, operations management, data-driven optimization algorithms and applications, energy systems, and transportation system planning and optimization. Max has extensive research collaborations with government agencies as well as private companies (both US and international). He serves as the Chief Supply Chain Scientist for JD.com. Max is the president-elect for the Production and Operations Management Society, a Department Editor for Production and Operations Management, and Associate Editors for several leading journals. Max received the CAREER award from National Science Foundation, and the Franz Edelman Laureate Award from INFORMS, won several best paper awards, and was elected Fellow of INFORMS in 2018.

Cong Shi is Associate Professor of operations research at the University of Michigan at Ann Arbor. His research is focused on the design and analysis of efficient algorithms for stochastic optimization models in operations management. Main areas of applications include inventory control, supply chain management, revenue management, service operations, and human-robot interaction. He received his Ph.D. in Operations Research at MIT in 2012, and his B.Sc. in Mathematics from the National University of Singapore in 2007.

David Simchi-Levi is Professor of Engineering Systems at MIT and serves as the head of the MIT Data Science Lab. He is considered one of the premier thought leaders in supply chain management and business analytics. Professor Simchi-Levi is the current Editor-in-Chief of *Management Science*, one of the two flagship journals of INFORMS. He served as the editor-in-chief for *Operations Research* (2006–2012), the other flagship journal of INFORMS and for *Naval Research Logistics* (2003–2005). In 2020, he was awarded the prestigious INFORMS Impact Prize for playing a leading role in developing and disseminating a new highly impactful paradigm for the identification and mitigation of risks in global supply chains.

Jing-Sheng Jeannette Song is the R. David Thomas Professor of Business Administration and a Professor of Operations Management at the Fuqua School of Business of Duke University. Her research expertise is in supply chain management and operations strategy, with particular interests in inventory management under uncertainty, assemble-to-order systems, supply chain digitization, and global supply chain risk mitigation. She is an INFORMS fellow and a MSOM fellow.

Jordan D. Tong is the Wisconsin Naming Partners Professor and Associate Professor in the Department of Operations and Information Management at the Wisconsin School of Business. His research primarily examines how human cognitive limitations interact with broader system dynamics to inform operations design. He received his Ph.D. in Operations Management from Duke University and his B.A. in Mathematics from Pomona College.

Geert-Jan van Houtum is Professor of Maintenance and Reliability and Vice-Dean at the Department Industrial Engineering & Innovation Sciences of the Eindhoven University of Technology. He studies predictive maintenance, maintenance optimization, availability management of capital goods, spare parts management, and the effect of design decisions on the

total cost of ownership of capital goods. Much of this research is practice-based. He is the department editor at *Service Science*, associate editor at *Operations Research Letters*, and editorial board member at the *International Journal of Production Economics*.

Jan A. Van Mieghem is A.C. Buehler Distinguished Professor of Operations Management at the Kellogg School of Management of Northwestern University. His research studies service and supply chain operations from a strategic and tactical perspective and addresses both theory and practice. Research methodologies include mathematical modeling, stochastic analysis, game theory, machine learning, optimization and control theory, as well as empirical estimation, prediction and causal inference. Current research focuses on collaboration in people-intensive processes, including healthcare; supply chain flexibility, dual-sourcing, and near-shoring; machine learning in digital operations; and social enterprise.

Michael R. Wagner is Associate Professor of Operations Management and a Neal and Jan Dempsey Endowed Faculty Fellow in the Foster School of Business, at the University of Washington. He has a Ph.D. degree in Operations Research, an MEng degree in Electrical Engineering and Computer Science, and dual B.S. degrees in Electrical Engineering and Computer Science, and Mathematics, all from MIT. His research interests include inventory management, supply chain management, logistics, crowdsourcing, and (stochastic and robust) optimization.

Mengxin Wang is Ph.D. Candidate at the Department of Industrial Engineering and Operations Research at the University of California, Berkeley. Prior to coming to UC Berkeley, Mengxin earned her B.E. in Industrial Engineering at Tsinghua University. Mengxin's research focuses on algorithm and strategy design for operations management applications of online platforms.

Yehua Wei is Associate Professor of Business Administration in Decision Sciences at Fuqua School of Business. He received his Ph.D. in Operations Research from MIT in 2013. His research interest lies in the broad area of decisions under uncertainty. Specifically, he studies problems in flexibility design, dynamic resource allocation, vehicle routing, queueing networks, strategic routing, and e-commerce fulfillment. His research has been recognized by several awards, including the George Nicholson Paper Competition, Daniel H. Wagner Prize for Excellence in Operations Research Practice, MSOM Service Management SIG Best Paper Prize, and MSOM Practice-Based Research Competition.

Chelsea C. White III holds the Schneider National Chair of Transportation and Logistics at the H. Milton Stewart School of Industrial and Systems Engineering of Georgia Institute of Technology. He is the former Director of the A.P. Sloan Foundation Trucking Industry Program. While at the University of Michigan, he was the founding engineering co-director of the Tauber Institute for Global Operations. He is a member of the board of directors of the Industry Studies Association, a Fellow of the IEEE and of INFORMS, and an INFORMS Edelman Laureate. He is a former member of the WEF Trade Facilitation Council.

Linwei Xin is Assistant Professor of Operations Management at the Booth School of Business, University of Chicago. His research is on inventory and supply chain management: designing models and algorithms for organizations to effectively “match supply to demand” in various contexts with uncertainty. His research on stochastic inventory theory by using asymptotic analysis has been recognized with several INFORMS paper competition awards, including the Applied Probability Society Best Publication Award (2019) and First Place in the George

E. Nicholson Student Paper Competition (2015). His work with JD.com on dispatching algorithms for robots in intelligent warehouses was recognized as a finalist for the INFORMS 2021 Franz Edelman Award, with an estimate of billions of dollars in savings. His research has been published in journals such as *Operations Research* and *Management Science*.

Nan Yang is Leslie O. Barnes Professor of Management at Miami Herbert Business School, University of Miami. Her research interests include inventory management, integrated supply and demand management, supply risk management, competition in supply chains, and health-care operations. Her work has been published in *Management Science*, *Operations Research*, *Manufacturing and Service Operations Management*, and *Production and Operations Management*. She received her Ph.D. in Decision, Risk and Operations from Columbia University in 2007, and worked at Cornell University and Washington University in St. Louis before joining the University of Miami in 2016.

Peiwen Yu is Professor at the School of Economics and Business Administration, at Chongqing University. He received his Ph.D. degree from HKUST Business School. His research interest is in perishable inventory management, behavioral operations management, and data-driven operations management.

Can Zhang is Assistant Professor in the Operations Management area at the Fuqua School of Business at Duke University. His research focuses on socially responsible and sustainable operations with an emphasis on underserved populations. More specifically, his research studies nonprofit and public sector operations, health and humanitarian supply chains, and smallholder agricultural supply chains. His work has received several recognitions, including the winner of the MSOM Best Paper Award, MSOM Award for Responsible Research in OM, MSOM Practice-Based Research Competition, MSOM Student Paper Competition, honorable mention for George B. Dantzig Dissertation Award, and a finalist for the Public Sector Operations Research Best Paper Award, Franz Edelman Award, and George Nicholson Student Paper Competition.

Huanan Zhang is Assistant Professor at the Leeds School of Business, University of Colorado Boulder. He received his Ph.D. in Industrial and Operation Engineering at the University of Michigan in 2017. Before joining Leeds, he was an Assistant Professor in the Department of Industrial and Manufacturing Engineering at Penn State University. Huanan's research interests include the design of data-driven algorithms, approximation algorithms, and their applications in inventory and supply chain management, revenue management, and service operations.

Sean X. Zhou is Professor at the Department of Decision Sciences and Managerial Economics, The Chinese University of Hong Kong (CUHK) Business School. His main research area is supply chain management, with particular interests in inventory control, dynamic pricing, sustainable operations, and data-driven supply chain optimization.

Preface

The advent and spread of the COVID-19 pandemic crippled the supply chains for hundreds and thousands of products. This disruption was eminently felt by nearly all countries in the early stages of the pandemic when medical supplies such as masks, personal protection equipment, and ventilators were in severe shortage for healthcare providers. At the beginning of the lockdowns, consumers often found empty shelves at local stores for groceries and daily consumables, and automakers had to halt production because of a shortage of semiconductor chips. Furthermore, during this past summer, many leading retailers had to slash their massive inventory because they failed to predict the changes in consumer taste and spending behavior that resulted from the pandemic. Never before have the words “supply chains” and “inventory” appeared so frequently in the news and media, nor has the impact of poor inventory management been so intimately felt in everyday life. More than ever, organizations and companies now realize how ineffective inventory management can threaten the bottom line, and managers are keen on searching for sound solutions. The impacts of geopolitical conflicts, global economic uncertainty, and climate change can only intensify the need for prudent inventory management.

Thanks to the separation of production and consumption in both space and time, economies-of-scale and uncertainties inventory are necessary for all stages of supply chains, from raw materials to components and parts, and from semi-finished products to final products. Managed well, inventory is considered the lubricant of the economy. As such, there has been a rich academic literature on inventory planning and optimization starting from the 1950s, when the classical quantitative inventory models and theory were first developed. We refer the reader to Arrow et al. (1958), Axsäter (2006), de Kok and Graves (2003), Graves et al. (1993), Hadley and Whitin (1963), Porteus (2002), and Zipkin (2000) for the developments up to the early 2000s.

Despite over half a century’s progress, inventory management remains a challenge, as evidenced by the recent pandemic. One major reason is the sheer complexity of a supply chain, which may involve numerous suppliers, factories, logistics providers, and retailers in different locations and organizations, resulting in high uncertainty in supply and demand at every stage. Moreover, supply chains are highly dynamic, evolving rapidly in response to technological advances and globalization, such as e-commerce and global sourcing; hence, new issues keep arising. Consequently, researchers continue to build models to capture new features and apply advanced analytical tools to develop tractable solutions. This handbook summarizes the key developments in inventory research in the past two decades, with references to the earlier literature.

More specifically, this handbook offers a comprehensive overview of the state-of-the-art research in quantitative models for inventory management. It is an essential original reference tool for researchers and students in this area. The book consists of 21 chapters in three parts – Fundamentals, Interfaces, and Context Specific Models. The book’s contributors are all active inventory researchers from major research universities worldwide, many of whom are world-renowned scholars.

The first part of the handbook, “Fundamentals – theory and methodologies,” contains 11 chapters, where inventory decisions are the primary concern, assuming the conditions for inventory control are given, such as demand distributions, lead times, and cost parameters. These chapters present the mathematical modeling of various inventory systems, including systems with lost sales, perishable inventory, finite production capacity, remanufacturing and product returns, expediting, dual-sourcing, and assemble-to order. Both single-stage, single-item systems and multi-stage, single/multi-item systems are considered. The chapters cover structural properties of the optimal and asymptotically optimal inventory policies, performance evaluation and optimization tools for a given type of inventory policy, and efficient approximation algorithms to compute effective policies and error bounds. In particular:

- Chapter 1 – “Lost-sales inventory systems”, by Marco Bijvank, Woonghee Tim Huh, and Ganesh Janakiraman, reviews the literature on lost-sales inventory systems, i.e., unsatisfied demands due to inventory shortage are lost. The chapter focuses on periodic-review models, the structure of the optimal policy, near-optimal or asymptotically optimal policies, and heuristic policies. The authors also provide references for continuous-review models and models with partially observed (i.e., censored) demand.
- Chapter 2 – “Perishable inventory systems”, by Qing Li and Peiwen Yu, summarizes recent literature on inventory systems for perishable products, such as chemical, food, and pharmaceutical products. It covers models with one class or multiple classes of demand in one location and one class of demand in multiple locations. The focus is on the structural properties of optimal policies and intuitive heuristic policies. The chapter also discusses empirical studies and several ideas for future research.
- Chapter 3 – “Capacitated inventory systems”, by Roman Kapuściński and Rodney P. Parker, describes the evolution of inventory models with production capacity limits. The authors first discuss how a core single-installation model is changed with capacity limits, and how this core model may be augmented with features such as pricing, outsourcing, and uncertain capacity, without losing the fundamental base-stock optimality result. They then present results on shortfall analysis, capacity investment, multi-item production, and fixed costs. Finally, they present results on serial and assembly systems, with variants such as information, vendor-managed inventory, and competition.
- Chapter 4 – “Generalizations of the Clark–Scarf model and analysis”, by Alexandar Angelus, reviews recent progress in multi-echelon inventory theory. It discusses the research in both series-system and assembly-system generalizations of the original Clark and Scarf (1960) model. The focus is on optimal and heuristic policies that render the more general models analytically and numerically tractable. The author also identifies several open research problems of theoretical interest and practical relevance.
- Chapter 5 – “Single-stage approximations of multiechelon inventory models”, by Kevin H. Shang, Jing-Sheng Jeannette Song, and Sean X. Zhou, summarizes recent developments on simple heuristics for optimal inventory policies of multi-echelon inventory systems with or without fixed order costs. These simple heuristics are based on solving a sequence of single-stage inventory problems with primitive model parameters. The chapter focuses on series systems under continuous-review and periodic-review schemes and briefly discusses the key results for the assembly and distribution systems.
- Chapter 6 – “Single unit analysis”, by Alp Muharremoglu, Xin Geng, and Nan Yang, introduces the single unit analysis and its applications in solving some types of multi-echelon

inventory systems such as serial systems and assembly systems. By focusing on the movement of a particular unit-customer pair, the single unit analysis can provide simple proof techniques, generate ideas for developing efficient algorithms, and offer novel insights for management.

- Chapter 7 – “Robust inventory management”, by Michael R. Wagner, presents robust inventory optimization models that minimize cumulative shortage, holding, and ordering costs over T periods. Stochastic demand is possibly serially correlated, possibly non-identically distributed. The means and covariance matrix of stochastic demand are known; the distributions are not needed. The models consider uncertainty sets motivated by the limit theorems of probability, including the Central Limit Theorem, the Strong Law of Large Numbers, and the Law of the Iterated Logarithm. The chapter presents closed-form ordering quantities for these sets for both static and dynamic rolling horizon models.
- Chapter 8 – “Dual-sourcing, dual-mode dynamic stochastic inventory models”, by Linwei Xin and Jan A. Van Mieghem, reviews dynamic inventory models that study the replenishment of inventory from two sources (or two transportation modes) in the presence of demand uncertainty. The two sources (modes) differ by replenishment lead times and costs. The chapter covers both discrete- and continuous-time models and provides self-contained proofs of several fundamental results in the literature. It also highlights recently advanced theory using asymptotic analysis.
- Chapter 9 – “Assemble-to-order systems”, by Levi DeValve, Jing-Sheng Jeannette Song, and Yehua Wei, reviews the recent literature on assemble-to-order (ATO) systems, which stock components that are then assembled into final products upon receipt of customer demand. The key feature of such systems, component commonality (i.e., products may share common components), provides benefits from inventory pooling, but also complicates replenishment and allocation decisions. Hence, much academic research has been devoted to optimizing these decisions. The chapter reviews both one-period and dynamic models. It provides a brief overview of the traditional literature and emphasizes the details of recent advances not covered in the previous reviews. It concludes with several promising directions for future research, including applications to online resource allocation, e-retailing, system design, and pricing.
- Chapter 10 – “Inventory models with returns and remanufacturing”, by Xiting Gong and Sean X. Zhou, summarizes recent studies on periodic-review inventory models with product returns and remanufacturing. It covers single-stage single-return models, multi-echelon or multi-return models, and models with differentiated remanufactured and new products. At the end, it identifies some directions for future research.
- Chapter 11 – “Approximation algorithms for stochastic inventory systems”, by Cong Shi, provides an in-depth introduction to approximation algorithms for stochastic inventory control problems. It presents the classical multi-period inventory control problem and its variants, including fixed costs and perishable products. The predominant approach in the inventory literature has been dynamic programming. Unfortunately, due to the well-known curse of dimensionality, these stochastic models are very hard to solve for optimality both in theory and practice. There has been a stream of recent and growing studies on the design and analysis of approximation algorithms for these models. These approximation algorithms are computationally efficient, conceptually sound, and admit worst-case performance guarantees. The chapter surveys state-of-the-art methods and techniques and discusses several promising and important future research avenues.

The second part of the book, “Interfaces,” includes six chapters discussing interdisciplinary topics in response to technological advancement and globalization. Here, some conditions for inventory control are not given, for example, demand distributions may be endogenously dependent on the pricing decisions, or demand distributions are unknown and need to be learned from historical data, or inventory records may be inaccurate, or demand information may not be truthfully shared across the supply chain unless proper incentives are in place, etc. These chapters review models in which the inventory decisions are made jointly with other considerations, such as pricing, information and incentives, financial flows, parametric and nonparametric data-driven learning, and behavioral issues. Specifically:

- Chapter 12 – “Information and incentives in inventory management”, by Bharadwaj Kadiyala, Hau Lee, and Özalp Özer, discusses problems arising at the confluence of demand- or supply-side information and incentives of firms over the past two decades. The impact of technology, firm-level collaborative practices, and, more recently, the explosion of data has added several dimensions to explore inventory-management problems. Alongside, advances in management science have provided scholars with enhanced toolkits to study and analyze complexities inherent in these problems. The chapter highlights models that capture the dynamic nature of information and incentives and how they interact in inventory management.
- Chapter 13 – “Joint pricing and inventory decisions”, by Xin Chen, Peng Hu, and Zhenyu Hu, reviews recent developments in analyzing single product dynamic models of joint inventory and pricing decisions. Recent literature has significantly generalized the classic model on both demand and supply dimensions. The authors emphasize models that incorporate inter-temporal effects on demand via modeling consumer behaviors and models that incorporate inter-temporal effects on supply by considering replenishment lead times and perishable products.
- Chapter 14 – “Statistical learning in inventory management”, by Wang Chi Cheung and David Simchi-Levi, surveys the recent development of inventory control problems through the lens of statistical learning. The chapter consists of two parts. The first part focuses on the single period setting, also known as the newsvendor model. The authors first showcase the performance guarantee of the sample average approximation (SAA) method. They then discuss other non-SAA approaches, followed by a generalization based on the supervised learning model. The second part focuses on multi-period models. It reviews works that explore the intersection of statistical learning and multi-period inventory control in multiple settings, such as capacity constraints on ordering quantity, pricing decisions, serial systems, fixed costs, and censored demand data.
- Chapter 15 – “Online learning in inventory and pricing optimization”, by Xiuli Chao, Boxiao Chen, and Huanan Zhang, discusses solving inventory control problems using online learning when little or limited prior demand information is available. The chapter starts by defining the objective function, called regret, which is the cost difference between the designed algorithm and that of the clairvoyant solution when complete demand information is available, and then reviews some common approaches in online learning. Due to the complexity of system dynamics in different inventory problems, such as inventory carryover, lead times, and censored demand observation, these methods may not be directly applicable, and tailored solutions need to be developed. The authors first focus on models with inventory decisions only, including periodic-review inventory

systems with censored demand, perishable inventory systems, lost-sales inventory systems with positive lead time, multiple-product systems with substitution or warehouse capacity constraints, systems with fixed ordering cost, and dual-sourcing inventory systems. The authors then discuss joint inventory control and pricing optimization problems, including periodic-review inventory problems with backlogging, lost-sales problems with censored demand, and the case that the number of price changes is constrained. The chapter concludes with some discussions on future research directions.

- Chapter 16 – “Inventory models with financial flows”, by Kevin H. Shang and Jing-Sheng Jeannette Song, reviews the recent developments in dynamic inventory models with financial flow considerations. The focus is on the literature that introduces cash flow dynamics into the classic inventory models that do not explicitly consider the interactions between physical (or material) and financial flows. These augmented models serve two important purposes. First, they help understand the impact of financial flows on inventory dynamics and decisions. Second, with the connection to the classic inventory models, one can leverage the extant results to derive the optimal control policy or to evaluate/optimize any given policy and reveal insights. The authors summarize models for both single-stage and multi-stage inventory systems and discuss the implications and applications of decentralized systems within a broader topic of supply chain finance.
- Chapter 17 – “Behavioral inventory management”, by Andrew M. Davis and Jordan D. Tong, introduces behavioral inventory-management research. This research area leverages ideas and methodologies commonly used in behavioral economics and psychology to advance our understanding of the role of human behavior in inventory management. The chapter first reviews some of the early papers in this area organized by problem settings. It describes the history and notable papers in two most significant streams: the newsvendor and serial supply chain settings. It then discusses how behavioral science can advance inventory research by describing three pathways for improving inventory performance: correcting sub-optimal decision behavior, responding to others’ behavioral patterns, and designing processes and systems for behavioral regularities. For each pathway, the authors provide recent examples from the literature. The chapter concludes with a discussion on current research opportunities in this sub-field involving other types of human decisions beyond the inventory decisions typically considered.

Finally, the third part of the book, “Context-specific models,” contains four chapters discussing tailored modeling and analytical solutions of inventory systems in the fast-changing industries of healthcare, spare parts logistics, retailing, and online retailing. Each of these industries has unique features and challenges, such as multi-stakeholders involvement, high system availability requirement, endogenous and correlated demand, and fast order fulfillment combined with vast stock-keeping units (SKUs). To be more specific:

- Chapter 18 – “Healthcare inventory management”, by Turgay Ayer, Chelsea C. White III, and Can Zhang, provides a brief discussion of how inventory-management theory and concepts can be applied to the healthcare sector. The chapter particularly focuses on the inventory management of two essential healthcare products: blood products and pharmaceuticals, by considering the perishability of these products and the multi-stakeholder nature of the healthcare system. The chapter concludes with several future research opportunities in healthcare inventory management.

- Chapter 19 – “Spare parts inventory planning”, by Rob Basten and Geert-Jan van Houtum, is devoted to spare parts inventory models for advanced capital goods that are essential for the primary processes of their users. The users generally require high system availability, implying low fractions of time that a system is down because of a lack of spare parts. This asks for multi-item inventory models for networks of central and local warehouses and a minimization of the total relevant costs (i.e., inventory holding costs and costs for emergency and lateral shipments) subject to constraints in terms of system availability or another system-oriented service measure. The chapter gives insights into system-oriented spare parts models and recent research on how to benefit from new enabling technologies such as 3D printing, sensor technology, and the Internet of Things.
- Chapter 20 – “Retail inventory systems”, by Stefan Minner and Anna-Lena Sachs, reviews different problems in retail inventory management. First, the authors provide a summary of the required data and performance indicators. Then, they give an overview of strategic, tactical, and operational decision support. The authors also emphasize the multi-product, multi-stage, and multi-location aspects of retail inventory management under dynamic and stochastic demands.
- Chapter 21 – “Online retailing inventory management”, by Mengxin Wang and Zuo-Jun Max Shen, focuses on online retail inventory management. Inventory management is a key driver of success in online retailing. The ability to quickly fulfill a vast number of diverse orders determines customer satisfaction and allows retailers to carve out a niche in this competitive field. However, managing the inventory of a diverse range of products in a large fulfillment network is a challenge. The chapter opens with an overview of the literature and retail industry practices, focusing on two major e-retailing players JD.com and Amazon. It then presents the two problems that recently gained attention in online retailing: inventory placement and order picking. Inventory placement is a tactical-level decision on allocating different items in a multi-warehouse network. Order picking is an operational-level decision regarding order fulfillment inside a warehouse. The chapter concludes by discussing the future directions for these two problems and the online retailing literature.

Instructors can use selected chapters in their first- or second-year Ph.D. courses in operations and supply chain management to introduce research in this area. Researchers can use the book to learn topics outside their research focus or use it as a guide to begin research on a new topic. Research-based practitioners can also use this book as a reference to learn about the state-of-art of any particular topic and the relevant literature.

November 2022, Jing-Sheng Jeannette Song

REFERENCES

- Arrow, K. J., Karlin, S., Scarf, H. E., et al. (1958). *Studies in the mathematical theory of inventory and production*. Stanford University Press.
- Axsäter, S. (2006). *Inventory control*. Springer.
- de Kok, A. D., & Graves, S. C. (2003). *Supply chain management: Design, coordination and operation*. Elsevier.
- Graves, S. C., Kan, A. R., & Zipkin, P. H. (1993). *Logistics of production and inventory* (Vol. 4). Elsevier.

- Hadley, G., & Within, T. (1963). *Analysis of inventory systems*. Prentice-Hall.
- Porteus, E. L. (2002). *Foundations of stochastic inventory theory*. Stanford University Press.
- Silver, E. A., Pyke, D. F., & Petterson, R. (1988). *Inventory management and production planning and scheduling*. John Wiley & Sons.
- Simchi-Levi, D., Chen, X., & Bramel, J. (2014). *The logic of logistics: Theory, algorithms, and applications for logistics and supply chain management*. Springer.
- Snyder, L. V., & Shen, Z.-J. M. (2019). *Fundamentals of supply chain theory*. John Wiley & Sons.
- Zipkin, P. (2000). *Foundations of inventory management*. McGraw Hill.

Acknowledgments

I am grateful to all the contributing authors for their excellent contributions. I am also deeply indebted to Xin Chen, Steve Graves, Ganesh Janakiraman, Cong Shi, and Geert-Jan van Houtum for their helpful discussions and encouragement during the inception of this handbook. Last, but not least, I would like to thank Shuyu Chen for his tremendous technical help in compiling the book, and the editorial staff at Edward Elgar Publishing for their assistance, guidance, and patience.

PART I

FUNDAMENTALS – THEORY AND METHODOLOGIES

1. Lost-sales inventory systems

*Marco Bijvank, Woonghee Tim Huh, and
Ganesh Janakiraman*

1.1 INTRODUCTION

The subject of this chapter is *lost-sales inventory systems*, both single-stage and multi-stage (serial). The defining feature of these systems is that when available inventory is insufficient to meet demand the excess demand is assumed to be *lost*. This should be contrasted with the assumption that excess demand is *backordered*. The dominant assumption in the literature on dynamic (multi-period) inventory models is that of backordering. This dominance is primarily driven by the analytical tractability that this assumption confers. For instance, order-up-to (OUT) policies are optimal in the standard single-stage, multi-period inventory model with backordering and lead times. We will see that for the same model with lost sales, OUT policies are *not* optimal – in fact, the simultaneous presence of lead times and lost sales complicates the structure of optimal policies. As a result, lost-sales inventory systems have been much harder to deliver neat analytical results for. Fortunately, the last two decades have seen important progress in the study of these systems.

Before we review the literature, we remark on the practical relevance of the backordering and lost-sales assumptions. In settings where a firm is selling a product to another firm (i.e., B2B), it is likely that in stock-out situations the buying firm can be persuaded to wait for some time before its demand can be fulfilled; this is backordering. In settings where a retail firm is selling to consumers (i.e., B2C), it is more common that in stock-out situations the consumer will take their demand elsewhere; this leads to lost sales. Also, in some settings including examples in service-part supply chains, it is necessary to use a secondary source of supply (more expensive than the primary source but “instantly” available) when regular supply is insufficient to meet demand in a period. Such settings can also be viewed as lost-sales inventory systems where the lost-sales cost is taken to be the premium incurred for using the secondary source of supply. In fact, one can view the lost-sales system as a special case of the dual-sourcing system, and a more general and precise connection between lost-sales inventory systems and dual-sourcing systems is studied by Sheopuri et al. (2010). See also the chapter in this edited volume entitled, “Dual-sourcing, dual-mode dynamic stochastic inventory models: A review.”

1.1.1 Preliminary Remarks

Our focus in this chapter is on analytical results. Most results of this type in the literature on lost-sales systems are based on discrete time models. Consequently, we will also focus on such models. Some of these results in the literature have been derived based on a finite-horizon cost, while others use the infinite-horizon cost (discounted or long-run average). A partial resolution of this difference is a set of techniques and results for extending finite-horizon results for stochastic inventory problems to the infinite horizon (see Huh et al. (2011b)). Our presentation will be faithful to the results in the literature in their original forms.

1.2 SINGLE STAGE

Consider a periodically reviewed, single-item inventory system that replenishes its inventory from an external supplier with a lead time of $L \geq 1$ periods. (The case in which $L=0$ is nothing but the multi-period newsvendor problem for which an OUT policy is easily shown to be optimal. The analysis is almost identical to that under the backordering assumption.) Excess demand that cannot be satisfied immediately is lost. In each period, the following sequence of events occurs: (1) receipt of delivery of the replenishment order placed L periods earlier, (2) order placement, and (3) demand realization. The demand in each period is satisfied to the extent possible, and any demand that cannot be satisfied immediately is lost. Let $p \geq 0$ denote the per-unit lost-sales penalty cost, and let $h \geq 0$ denote the unit holding cost for inventory on hand. These costs are charged at the end of a period. Let x_0 denote the number of units of on-hand inventory at the beginning of a period (after receiving the replenishment due that period). For any $l \in \{1, 2, \dots, L-1\}$, let $x_l \geq 0$ denote the quantity to be delivered l periods later (equivalently, this is the quantity ordered $L-l$ periods earlier). Let q denote the quantity of replenishment ordered in this period. Thus, the vector $(x_0, x_1, \dots, x_{L-1})$ represents the state at the beginning of a period after receiving delivery, and q represents the action in a period. We use $\mathbf{x}$ to denote the vector $(x_0, x_1, \dots, x_{L-1})$. Let D_t denote the demand in period t . We assume that demands are independently distributed across periods and that the distributions of these demands are known. Whenever we discuss the infinite-horizon model, we automatically assume that demands are also identically distributed across periods.

1.2.1 Dynamic Programming Formulation

Next, we define the dynamic program for this system for a planning horizon of T periods indexed by $t = 1, 2, \dots, T$. Periods are indexed forwards, i.e., period $t+1$ follows period t . Given a state-action pair of $\mathbf{x}$ and q in period t , the inventory on hand in period $t+1$ is $(x_0 - D_t)^+ + x_1$. The amounts of inventory in transit move by one component in the state, while q takes its position in the last component of the state. Mathematically, the state in the next period (period $t+1$) is

$$\left((x_0 - D_t)^+ + x_1, x_2, x_3, \dots, x_{L-1}, q \right).$$

Let $f_{T+1}(\mathbf{x}) = 0$ for all $\mathbf{x}$. Consider any $t \in \{1, 2, \dots, T\}$. Let $\alpha \in (0, 1]$ denote the discount factor for capturing the time value of money. Define

$$f_t(\mathbf{x}) = \min_{q \geq 0} \tilde{g}_t(\mathbf{x}, q),$$

where

$$\begin{aligned} \tilde{g}_t(\mathbf{x}, q) := & E \left[h \cdot (x_0 - D_t)^+ + p \cdot (D_t - x_0)^+ \right] \\ & + \alpha \cdot E \left[f_{t+1} \left((x_0 - D_t)^+ + x_1, x_2, x_3, \dots, x_{L-1}, q \right) \right]. \end{aligned}$$

Let $q_t^*(\mathbf{x})$ denote a minimizer in the optimization problem above. (When the optimization problem above does not have a unique solution, any statements we make on properties of

q_t^* should be taken to mean the existence of an optimal selector $q_t^*(\mathbf{x})$ (precisely defined in Section 3.4 of Huh and Janakiraman (2010b) with those properties.)

1.2.2 Optimal Policy Structure

The above dynamic program was first studied by Karlin (1958) for the special case where $L = 1$. In this case, the state is single dimensional, consisting of the amount of on-hand inventory x_0 only. Their major results were the following: (a) An OUT policy (defined in Section 1.2.3) is not optimal. This policy is not optimal since the sum of the on-hand inventory and the order quantity is insufficient to characterize the inventory availability in future periods. While the on-hand inventory in the current period may be sold and become unavailable in the future, the entire quantity ordered in the current period becomes available in the next period due to the lost-sales assumption. (b) Under an optimal policy, the quantity ordered in a period is a decreasing function of the amount of inventory on hand, and the rate of this decrease is at most 1. Their analysis is based on studying the single-variable function $f_t(x_0)$ and, for any x_0 , the single-variable function $g_t(x_0, q)$. Next, Morton (1969) generalizes this analysis to the case where the lead time L is a general positive integer. He also derives upper and lower bounds on the optimal order quantity in a period as functions of the state vector. Morton's analysis is based on first principles. More recently, Zipkin (2008a) uses the concept of L -natural-convexity ($L^\#$ -convexity) from the field of discrete convexity (see Murota (2003)) to present an alternate, elegant derivation of Morton's results on the structure of the optimal policy for arbitrary lead times. These results are summarized below.

Theorem 1.1 Assume that q_t^* is a differentiable function of $\mathbf{x}$. Then, the following inequalities hold:

$$-1 \leq \frac{\partial q_t^*}{\partial x_{L-1}} \leq \frac{\partial q_t^*}{\partial x_{L-2}} \leq \dots \leq \frac{\partial q_t^*}{\partial x_0} \leq 0.$$

If q_t^* is not differentiable, then the inequalities above hold after replacing every quantity of the form $\frac{\partial q_t^*}{\partial x_j}$ with $\frac{q_t^*(\mathbf{x} + \varepsilon_j) - q_t^*(\mathbf{x})}{\varepsilon}$, where ε_j is a vector with ε in its j^{th} component and zero in all the other components and ε is any strictly positive number.

In words, Theorem 1.1 states that the optimal order quantity is a decreasing function of the quantity scheduled to be delivered l periods later, for any $1 \leq l \leq L-1$, and that the rate of this decrease is smaller than 1. Furthermore, the sensitivity of the optimal order quantity to x_l is greater than that to $x_{l'}$ if $1 \leq l' \leq l \leq L-1$.

When L is large, finding the optimal solution becomes challenging due to the curse of dimensionality. For example, Zipkin (2008b) discusses a problem instance in which the size of the state space grows from 31 to 228,581 as L changes from 1 to 4.

1.2.3 Order-Up-To Policies

An *order-up-to policy*, or OUT policy for short, with parameter (i.e., order-up-to level) S is defined as follows: For any period t , given any state $\mathbf{x}$, the quantity ordered is

$$q(\mathbf{x}) = \left(S - \sum_{l=0}^{L-1} x_l \right)^+.$$

OUT policies are often referred to as base-stock policies. While Karlin (1958) has shown that such a policy cannot be optimal (when $L=1$), this policy has practical appeal and is, therefore, a commonly used policy. Two important questions arise: (a) How can we optimize within the class of OUT policies? (b) What is the performance loss (i.e., optimality gap) suffered by restricting attention to this class? In particular, under what conditions do we know that this performance loss is small? We address these questions below.

1.2.3.1 Optimizing S

Consider an order-up-to- S policy. Assume that the starting state is such that $x_0=S$ and $x_l=0$ for all $l \in \{1, 2, \dots, L-1\}$. Let $f_{\alpha, T}^{OUTP}(S)$ and $f_{\alpha, \infty}^{OUTP}(S)$ denote the expected discounted sum of costs under this policy over a finite horizon of T periods and over an infinite horizon, respectively. Let $f_{Avg}^{OUTP}(S)$ denote the infinite-horizon long-run average cost. The following results are based on Downs et al. (2001) and Janakiraman and Roundy (2004).

Theorem 1.2 *The functions $f_{\alpha, T}^{OUTP}(S)$ and $f_{\alpha, \infty}^{OUTP}(S)$ are convex. Further, when demands are integer-valued and the probability of zero demand in a period is strictly positive, then $f_{Avg}^{OUTP}(S)$ is convex in S .*

The proof for these results is based on a stronger observation that, for every realization of demands, the sum of the costs over the first t periods is convex in S for any $t \geq 1$. In other words, the proof establishes the sample path properties of the cumulative cost incurred over $[1, t]$. Interestingly, the cost incurred in any given period t need not be convex in S but the sum of these costs over $[1, t]$ does. Thus, the analysis involves a detailed examination of the dynamics of the inventory process as opposed to appealing directly to existing results such as the convexity of the newsvendor cost function. While the result stated above assumes L is deterministic, Janakiraman and Roundy (2004) allow a certain class of stochastic lead time processes under which orders do not cross or overtake each other.

Theorem 1.2 implies that the parameter S can be optimized using efficient techniques such as bisection search in conjunction with simulation to evaluate the objective function at any given value of S .

1.2.3.2 Asymptotic optimality

Here, we discuss a paper by Huh et al. (2009). In several important settings, the per-unit lost-sales penalty cost p is much larger than the per-unit per-period holding cost h . This motivates the authors to study the asymptotic behavior of $\min_{S \geq 0} f_{Avg}^{OUTP}(S; p)$ relative to the optimal long-run average cost (over all non-anticipatory policies), which we denote with $f_{Avg}^*(p)$ (here, we include p as an argument in $f_{Avg}^{OUTP}(S; p)$ and $f_{Avg}^*(p)$ for clarity), as p becomes infinitely large. Let $\mathbf{D} \sim_d \sum_{i=1}^{L+1} D_i$ denote the total demand over $L+1$ periods, representing the total demand over the lead time including the period when we place the order. Assume that $\mathbf{D}$ is not deterministic, since the deterministic problem is trivial to solve. Let G denote the cumulative distribution function of $\mathbf{D}$. Let $\bar{M} \in \mathbb{R}_+ \cup \{+\infty\}$ denote the upper end of the support of G ; that is,

$$\bar{M} = \sup\{x : G(x) < 1\}.$$

We allow $\overline{M}$ to be infinity when the demand $\mathbf{D}$ is unbounded. For any $t \geq 1$, we define the *mean residual life* (MRL) $m_{\mathbf{D}}(t)$ as follows:

$$m_{\mathbf{D}}(t) = \begin{cases} E[\mathbf{D} - t \mid \mathbf{D} > t], & \text{if } t < \overline{M}, \\ 0, & \text{otherwise.} \end{cases}$$

The following assumption is made by Huh et al. (2009).

1.2.3.3 Assumption (MRL assumption)

$$\lim_{t \rightarrow \overline{M}} m_{\mathbf{D}}(t) / t = 0.$$

Sufficient conditions for this assumption are the following.

Theorem 1.3 *The total demand over the lead time $\mathbf{D}$ satisfies the MRL assumption if any of the following conditions hold:*

- (a) *The demand D in each period (either discrete or continuous) is bounded; that is, $\overline{M} < \infty$.*
- (b) *The demand D in each period (either discrete or continuous) has an increasing failure rate (IFR) distribution.*
- (c) *$\mathbf{D}$ has a finite variance. Furthermore, the distribution G of $\mathbf{D}$ has a density function g and a failure rate function $r(t)$ of G that does not decrease to zero faster than $1/t$; that is,*

$$\lim_{t \rightarrow \infty} t \cdot r(t) = \infty,$$

where, for any $t \geq 0$, we define $r(t) = g(t) / (1 - G(t))$.

Thus, the MRL assumption encompasses a large class of discrete and continuous demand distributions used in many supply chain models, including any bounded demand random variables. For unbounded demand, part (b) of Theorem 1.3 shows that many commonly used distributions also satisfy the MRL assumption. Examples include geometric distributions, Poisson distributions, negative binomial distributions with parameter $r > 0$ and $0 < p < 1$, exponential distributions, and Gaussian distributions. When the demand distribution does not exhibit the IFR property, part (c) of the above theorem shows that the MRL assumption remains satisfied as long as the failure rate does not decrease to zero too quickly.

It is well known that the OUT policy with parameter $S^{\mathcal{B}^*}(b) := G^{-1}\left(\frac{b}{b+h}\right)$ minimizes the long-run average cost per period in a backorder system, which is identical to the lost-sales system under analysis with the exception that the excess demand at the end of every period is backordered at a penalty cost of b per unit. We denote that backorder system with $\mathcal{B}(b)$.

Theorem 1.4 *For any $h \geq 0$ and $p \geq 0$, let $S_{p+Lh} = S^{\mathcal{B}^*}(p + Lh)$ denote the optimal order-up-to level in the backorder system $\mathcal{B}(p + Lh)$. Then, under the MRL assumption, the order-up-to S_{p+Lh} policy is asymptotically optimal; that is,*

$$\lim_{p \rightarrow \infty} \frac{\min_{S \geq 0} f_{Avg}^{OUTP}(S, p)}{f_{Avg}^*(p)} = \lim_{p \rightarrow \infty} \frac{f_{Avg}^{OUTP}(S_{p+Lh}, p)}{f_{Avg}^*(p)} = 1.$$

Bijvank et al. (2014) demonstrate that this result is robust in the following sense. There is a large family of functions $b(p)$ such that an order-up-to $S^{B*}(b(p))$ policy is asymptotically optimal as p approaches ∞ .

1.2.3.4 Main ideas in the proof of Theorem 1.4

Since $f_{Avg}^*(p)$ is the cost associated with the optimal policy over all policies, and $f_{Avg}^{OUTP}(S, p)$ is the cost associated with the OUT policy with the order-up-to level S , we have

$$f_{Avg}^*(p) \leq \min_{S \geq 0} f_{Avg}^{OUTP}(S, p) \leq f_{Avg}^{OUTP}(S_{p+Lh}, p).$$

It is challenging to compare these costs since the structure of the lost-sales system is not simple. Instead, we work with lower and upper bounds derived using backorder systems with suitably modified parameters. Then, we show that the ratio of these backorder system-based bounds becomes 1 as $p \rightarrow \infty$.

We first consider a lower bound on the optimal cost $f_{Avg}^*(p)$. Janakiraman et al. (2007) show that the optimal cost of a lost-sales system is bounded below by the optimal cost of a backorder system that has the same parameters as the lost-sales system except that the backorder parameter is only $1/(L+1)$ of the lost-sales parameter. (This result stems from an intuition that a customer in the backorder system waits at most $L+1$ periods for a backordered unit.)

Next, we can show that the optimal cost in the backorder system with the backorder penalty $p + L \cdot h$ is an upper bound on $f_{Avg}^{OUTP}(S_{p+Lh}, p)$. To prove this, we compare a lost-sales system with the lost-sales penalty p and managed using an OUT policy to a backorder system with the backorder penalty $p + L \cdot h$, which is also managed using the same policy. Both systems have the same holding cost h . One can show that the above-mentioned lost-sales system does not incur more cost than this backorder system. (This follows because of the following two observations: First, the lost-sales system incurs fewer shortages than the backorder system. Second, as far as the holding cost is concerned, the definition of the backorder penalty here as $p + L \cdot h$ accounts for the fact that a unit of unmet demand can lead to an extra unit in inventory in the lost-sales system, compared to the backorder system, for at most L periods.) This result is valid for any order-up-to level, and it holds in particular when the order-up-to level is S_{p+Lh} . Note that S_{p+Lh} is the optimal order-up-to level for a backorder system with the backorder penalty $p + L \cdot h$, which is exactly the same backorder system discussed above. Thus, the cost in the lost-sales system under the order-up-to S_{p+Lh} policy is bounded above by the optimal cost in $\mathcal{B}(p + Lh)$.

Thus, we have lower and upper bounds, both based on backorder systems. The lower bound uses the backorder penalty of $p/(L+1)$, and the upper bound has the penalty $p + L \cdot h$. Consider the limit as p becomes arbitrarily large. We compare the cost of these two backorder systems. As the backorder penalty grows large, the optimal stocking amount becomes large, incurring a high inventory holding cost. Meanwhile, the backorder amount becomes small. While it is tempting to conclude that the backorder cost approaches 0, it requires caution as the

backorder penalty parameter increases arbitrarily. We can show that, under the MRL assumption, the holding cost is the primary determinant of the backorder system's cost when $p \rightarrow \infty$. It can then be shown that the ratio of the holding costs (and, in turn, the total costs) of the two backordering systems that appear in our lower bound on $f_{\text{Avg}}^*(p)$ and the upper bound on $f_{\text{Avg}}^{\text{OUTP}}(S_{p+Lh}, p)$ approaches 1 as p approaches ∞ . This then implies that the ratio between the order-up-to cost ($f_{\text{Avg}}^{\text{OUTP}}(S_{p+Lh}, p)$) and the optimal cost ($f_{\text{Avg}}^*(p)$) also approaches 1.

We mention that a recent survey on asymptotic analysis in inventory systems by Goldberg et al. (2021) provides a more detailed proof sketch of this theorem as well as the asymptotic optimality of constant-order policies (Theorem 1.5 below) in a more unified manner. It observes that both analyses share the commonality of upper and lower bounding a complex system with simpler systems, and recognizing that the difference between the two bounds is negligible in the asymptotic regime.

1.2.4 Constant-Order Policies

A *constant-order policy* with parameter r (i.e., the constant-order quantity) is defined as follows: For any period t , given any state $\mathbf{x}$, the quantity ordered is

$$q(\mathbf{x}) = r.$$

This is the simplest ordering policy one can think of, in fact, the most naive policy. Before we discuss the results pertaining to these policies, we explain the natural reason for considering these policies.

When $L = 0$, we know that an OUT policy is optimal (see Karlin (1958)). For $L > 0$, we have seen that OUT policies work well when p is sufficiently large. Moreover, the numerical investigations in Zipkin (2008b) and Huh et al. (2009) show that, for a given p , the optimality gap suffered by OUT policies increases with L . Notice that the OUT policy is completely sensitive to demand (more precisely, sales) in the sense that every unit of sales experienced in a period triggers the order of a corresponding unit in the next period. The opposite of this is the constant-order policy, which is completely insensitive to demand. Recall from Theorem 1.1 that the optimal order quantity is a function of the inventory level with a slope between -1 and 0 , or conversely, a function of the sales with a slope between 0 and 1 . Note that the OUT policy specifies a function for the order quantity that depends on the sales with a slope of exactly 1 , and the constant-order policy with a slope of exactly 0 . Thus, it is worth exploring whether constant-order policies work well when L is large.

Another intuitive reason for such a conjecture is this: An order placed in period t arrives in period $t + L$. If L is large, there is a significant amount of uncertainty regarding the inventory level in period $t + L$, as seen from period t . Thus, the ability to affect future states by dynamically choosing the order quantity in a period in response to the state in that period is limited. The constant-order policy takes this to the extreme by being unresponsive to the state. This decoupling strategy of the order quantity from the state of the inventory system should therefore work well for large values of L .

The papers by Goldberg et al. (2016) and Xin and Goldberg (2016) prove that a (specific) constant-order policy is asymptotically optimal as the lead time L approaches ∞ . The former proves this result under the finite-horizon framework by studying a sequence of problems parametrized by their horizon lengths and their lead times. The latter proves this result under

the infinite horizon average cost framework. We present the latter result here. A precursor to these results is Reiman (2004) who proved, in a continuous-review model with a Poisson demand arrival process, that constant-order policies perform better than OUT policies for sufficiently large lead times.

Let $f_{\text{Avg}}^{\text{COP}}(r; L)$ denote the long-run average cost per period under the constant-order policy with parameter r and lead time L . It is easy to see that if $r \geq E[D]$ and D is not deterministic, then, the expected steady-state inventory on hand equals ∞ . Consequently, for $r \geq E[D]$, we define $f_{\text{Avg}}^{\text{COP}}(r) := \infty$. Furthermore, for any $r \in [0, E[D]]$, $f_{\text{Avg}}^{\text{COP}}(r; L)$ does not depend on L . Thus, we can eliminate the lead time argument from $f_{\text{Avg}}^{\text{COP}}(r; L)$ and denote this cost with $f_{\text{Avg}}^{\text{COP}}(r)$.

1.2.4.1 Optimizing the constant-order quantity, r

Define

$$r^* = \underset{r \geq 0}{\operatorname{argmin}} \{ f_{\text{Avg}}^{\text{COP}}(r) \}.$$

When the minimizer above is not unique, r^* is defined as the infimum of the set of all minimizers. Xin and Goldberg (2016) prove that $f_{\text{Avg}}^{\text{COP}}(r)$ is a convex function. This observation suggests that computationally efficient methods can be used to determine r^* .

1.2.4.2 Asymptotic optimality

The main result of Xin and Goldberg (2016) is given below.

Theorem 1.5 *Assume $E[D] < \infty$ and that D is not deterministic. Then, there exist constants $A > 0$ and $\gamma \in [0, 1)$, both independent of L , such that*

$$\frac{f_{\text{Avg}}^{\text{COP}}(r^*) - f_{\text{Avg}}^*(L)}{f_{\text{Avg}}^*(L)} \leq A\gamma^{L+1},$$

where $f_{\text{Avg}}^*(L)$ represents the optimal long-run average cost of the system with the lead time L . The above theorem shows that constant-order policies are asymptotically optimal as the lead time L becomes large. If $\varepsilon > 0$ denotes a desired upper bound on the optimality gap, this theorem implies that, as ε is reduced, the minimum lead time required is $O(\log(1/\varepsilon))$.

For integer-valued demand and order quantities (i.e., non-divisible products), Bai et al. (2023) show that constant-order policies are typically not asymptotically optimal. Instead, the authors study a *bracket policy*, which defines a fixed sequence of order quantities $\lfloor r \rfloor$ and $\lceil r \rceil$ such that the average order quantity is exactly r . The authors show that this bracket policy is asymptotically optimal for non-divisible products.

1.2.4.3 Main ideas in the proof of Theorem 1.5

A key observation is that, under the constant-order policy with parameter r , the inventory system receives the same amount of r units in each period, and the demands across time periods are independent and identically distributed. Consequently, the on-hand inventory level evolves in the same manner as the waiting-time process of a single-server queuing system ($GI/G/1$ queue), which is initially empty and faces the inter-arrival distribution $\{D_i\}$ and the constant

service time r . The proof takes advantage of the connection between the inventory level and waiting times.

Let I_∞^r denote the inventory on hand at the end of a period, in a steady state, under the constant-order policy with parameter r . It is easy to see that I_∞^r has the same distribution as the waiting time in the queue described above. Then, it is well known that

$$I_\infty^r \sim \sup_{j \geq 0} \left(jr - \sum_{i=1}^j D_i \right).$$

Also, by definition,

$$f_{\text{Avg}}^{\text{COP}}(r) = h \cdot E[I_\infty^r] + p \cdot (E[D] - r).$$

Similarly, we define

$$f_L(r) := h \cdot E[I_L^r] + p \cdot (E[D] - r),$$

where

$$I_L^r = \max_{j=0, \dots, L} \left\{ jr - \sum_{i=1}^j D_i \right\} = \left(\left((r - D_1)^+ + r - D_2 \right)^+ + \dots \right)^+ + r - D_L.$$

Also define

$$r_L = \underset{r \geq 0}{\operatorname{argmin}} \{ f_L(r) \}.$$

To gain intuition, it is useful to note that I_L^r is the inventory on hand at the end of period $L + 1$ in a lost-sales system if the inventory at the beginning of period 1 is zero (before delivery) and if a delivery of r is received in each of the periods $1, 2, \dots, L + 1$. Similarly, I_L^r can also be interpreted as the waiting time of the L^{th} customer in the $GI/G/1$ queue referred to earlier.

By comparing the expressions for I_L^r and I_∞^r , we can see that $I_L^r \leq_d I_\infty^r$. Thus, the inventory holding cost that appears in $f_L(r)$ is lower than that in $f_{\text{Avg}}^{\text{COP}}(r)$; it is intuitive then to conjecture that the minimizers of these two functions satisfy

$$r_L \geq r^*,$$

which Xin and Goldberg (2016) prove.

The next key observation is about the optimal policy. Let $(I, q_1, q_2, \dots, q_L)$ denote the stationary distribution of the state of the system under the optimal policy; here, I is the inventory on hand at the beginning of a period after receiving delivery and q_k ($1 \leq k \leq L$) is the quantity scheduled to be delivered k periods later. Then, by stationarity, we can write the following relations:

$$E[q_k] = E[q_{k'}], \quad 1 \leq k, k' \leq L,$$

$$f_{\text{Avg}}^*(L) = h \cdot E[I] + p \cdot E[D - q_1], \quad \text{and}$$

$$E[I] = E \left[\left(\left(\left((I - D_1)^+ + q_1 - D_2 \right)^+ + \dots + q_{L-1} - D_L \right)^+ + q_L \right) \right].$$

The expression within the expectation on the right-hand side of the equation above is a convex function of $(I, q_1, q_2, \dots, q_L)$. Thus, Jensen's inequality implies that

$$E[I] \geq E \left[\left(\left(\left(E[I] - D_1 \right)^+ + E[q_1] - D_2 \right)^+ + \dots + E[q_{L-1}] - D_L \right)^+ + E[q_L] \right].$$

Since I is non-negative and $E[q_1] = E[q_2] = \dots = E[q_L]$, we can use the relations derived above along with the expressions for $f_{\text{Avg}}^*(L)$ and $f_L(r)$ to conclude that

$$f_{\text{Avg}}^*(L) \geq f_L(E[q_1]).$$

Since r_L minimizes $f_L(r)$, we have

$$f_{\text{Avg}}^*(L) \geq f_L(r_L).$$

This implies that

$$f_{\text{Avg}}^{\text{COP}}(r^*) - f_{\text{Avg}}^*(L) \leq h \cdot (E[I_\infty^r] - E[I_L^r]) + p \cdot (r_L - r^*).$$

The right-hand side above can be rewritten as

$$h \cdot (E[I_\infty^r] - E[I_L^r]) + h \cdot (E[I_L^r] - E[I_L^L]) + p \cdot (r_L - r^*).$$

In the expression above, the first term is bounded from above by a term of order γ^{L+1} for some $\gamma \in [0, 1)$ based on a result by Kingman (1962) on the convergence of the waiting time in a $GI/G/1$ queue to its steady state. The bounding of the sum of the second and third terms involves several steps and many of the facts established above. In particular, Kingman's results are used to derive analytical expressions for $E[I_L^r]$ and $E[I_\infty^r]$ for any $r < E[D]$; also, since r^* minimizes $f_{\text{Avg}}^{\text{COP}}(r)$, the first order conditions can be used to show that

$$\sum_{n=1}^{\infty} P \left(n \cdot r^* \geq \sum_{i=1}^n D_i \right) \geq p / h.$$

This analysis and the fact that $r_L \geq r^*$ are then used together with Chernoff's inequality to show that $h \cdot (E[I_L^r] - E[I_L^L]) + p \cdot (r_L - r^*)$ is bounded from above by a quantity which is of the order of γ^{L+1} .

1.2.5 Combining Order-Up-To and Constant-Order Policies

The collective wisdom of Sections 1.2.3 and 1.2.4 is that OUT policies work well when L is small, whereas constant-order policies work well when L is large. This leads to the intuitive conjecture that a policy that intelligently combines the order-up-to and constant-order structures could be more effective across all values of L . Xin (2021) explores this idea by proposing the *capped base-stock policy*, which has two parameters $S \geq 0$ and $r \in [0, S]$ and is defined as follows: For any period t , given any state $\mathbf{x}$, the quantity ordered is

$$q(\mathbf{x}) = \min \left\{ \left(S - \sum_0^{L-1} x_l \right)^+, r \right\}.$$

Thus, this policy is identical to the OUT policy with parameter S except that the order quantity in any period is capped at a level r . Notice that this capped base-stock policy coincides with the order-up-to S policy when $r = S$ and coincides with the constant-order policy with parameter r when $S = \infty$. Thus, the class of capped base-stock policies subsumes both the classes of OUT policies and constant-order policies. This same policy has also been studied by Bijvank and Johansen (2012).

Let $f_{\text{Avg}}^{\text{CBP}}(S, r)$ denote the long-run average cost of a capped base-stock policy with parameters S and r . Therefore, the class of capped base-stock policies inherits the asymptotic optimality properties of OUT policies and constant-order policies proved in Theorems 1.4 and 1.5. Moreover, Xin (2021) derives a finer relationship between capped base-stock policies and constant-order policies in the following statement.

Theorem 1.6 *For each $r \in [0, E[D]]$ and $S > (L+1)r$, there exist two constants $\theta > 0$ and $B \geq 0$ which depend only on r and the distribution of D such that $f_{\text{Avg}}^{\text{CBP}}(S, r) - f_{\text{Avg}}^{\text{COP}}(r) \leq rpe^{-\theta(S-(L+1)r)}B$.*

In words, the cost increase (if any), by using a capped base-stock policy with parameters (S, r) relative to a constant-order policy with parameter r , converges to zero exponentially as $S \rightarrow \infty$. Xin (2021) also presents a numerical investigation in which he demonstrates the superior performance of capped base-stock policies to several other policies proposed in the literature.

A recent working paper by Janakiraman and Wu (2020) proposes another policy that combines an order-up-to S policy and a constant-order policy with parameter r . Their policy is additionally parametrized by $\beta \in [0, 1]$ and is defined as follows: For a given realization of demands $(D_1, D_2, \dots, D_t)$ in the first t periods, let q_t^1 and q_t^2 denote the quantities ordered by the OUT policy and the constant-order policy, respectively, in period t ; the proposed policy orders the quantity $\beta q_t^1 + (1-\beta)q_t^2$ in period t . This policy is unconventional in the sense that the order quantity in a period cannot be specified by knowing only the state of the system in that period; the knowledge of the demand history is also required. This is because the state in period t under the OUT policy and that under the constant-order policy could be different from each other and from the state in the focal system. Note that when $\beta = 0$, the policy coincides with the constant-order policy, and that when $\beta = 1$ it coincides with the OUT policy; thus, by optimizing over β , the resulting policy inherits the asymptotic optimality properties

of OUT policies and constant-order policies proved in Theorems 1.4 and 1.5. The paper shows that the long-run average cost is convex in β which facilitates an efficient optimization over β .

The performance of constant-order policies, base-stock policies, and capped base-stock policies in comparison to optimal replenishment policies is numerically illustrated for a small set of test problems. For all instances, we used a Poisson demand distribution with a mean of 10 and $h = 1$. The lost-sales penalty cost p equals the values 2, 4, 9, 19, and 39, and the lead time ranges from one to four review periods. The results are presented in Table 1.1.

The constant-order policy performs well when the lead time L increases and when the penalty cost p is small. When the penalty cost is high, the lead time needs to be significantly larger in order for this replenishment policy to perform well. The base-stock policy performs better when the constant-order policy does not and vice versa. The capped base-stock policy takes the best of both policies and performs well in all instances (within 1% of the cost for the optimal policy). This is not surprising as capped base-stock policies are a generalization of base-stock and constant-order policies. However, finding the best values for the policy parameters can be a computationally challenging task. A number of other heuristic approaches are discussed in our next section.

1.2.6 Heuristic Procedures and Policies

A number of heuristic procedures have been proposed in the literature to set the parameter values for lost-sales inventory systems controlled by (modified) base-stock policies, as well as several other heuristic replenishment policies. Before we discuss any of them, we first introduce new notation regarding the demand distribution. Let $G_\tau(y)$ denote the cumulative distribution function of the demand over τ periods with expectation $E[D_\tau]$, and $\mathcal{G}_\tau(y) = E[y - D_\tau]^+ = \sum_{d=0}^{y-1} G_\tau(d)$.

1.2.6.1 Base-stock policies

The simplest base-stock policy is to use the best base-stock value in the backlog system, or

$$S^B = S^{B*}(p) = \min \{y : G_{L+1}(y) \geq p / (p + h)\}.$$

One of the differences between the backorder system and the lost-sales system is that each unit of backordered demand is charged a penalty cost in each period that it is on backlog. This means that a backlogged unit can be charged a penalty cost more than once. Rosling (2002) studies a backorder system where a penalty cost is incurred only one time, i.e., in the period where that demand exceeds the on-hand inventory. This results in a base-stock level

$$S^R = \operatorname{argmin}_{y \geq 0} \left\{ h \cdot \mathcal{G}_{L+1}(y) + p \cdot (\mathcal{G}_{L+1}(y) - \mathcal{G}_L(y) + E[D_1]) \right\}.$$

Archibald (1981) studies a continuous-review lost-sales system and proposes to subtract the expected demand short during the lead time from the base-stock level. This motivated Bijvank et al. (2014) to propose the following base-stock level,

$$S^A = S^R - E[D_{L+1} - S^R]^+ = 2 \cdot S^R - \mathcal{G}_{L+1}(S^R) - E[D_{L+1}].$$

A simple heuristic base-stock policy is proposed by Huh et al. (2009), namely

$$S^H = \frac{p}{p+h} \cdot S^B + \frac{h}{p+h} \cdot S_0,$$

where S_0 is the optimal base-stock level in a backorder system with no lead time (i.e., $S_0 = \min\{y : G_1(y) \geq p/(p+h)\}$).

Finally, Bijvank and Johansen (2012) approximate the average expected total cost for a given base-stock level S by introducing a correction factor $\tilde{c}(S)$ to correct the demand distribution for any lost demand, and set the base-stock level such that it minimizes this cost expression:

$$\tilde{S} = \operatorname{argmin}_{y \geq 0} \left\{ h \cdot \tilde{c}(y) \cdot \mathcal{G}_{L+1}(y) + p \cdot \left(E[D_1] + \hat{c}(y) \cdot [\mathcal{G}_L(y) - \mathcal{G}_{L+1}(y)] \right) \right\},$$

where the correction factor is provided by

$$\tilde{c}(S) = S \cdot \left[L \cdot (\mathcal{G}_L(S) - \mathcal{G}_{L+1}(S)) + \mathcal{G}_L(S) \right]^{-1}.$$

For the capped base-stock policy (referred to as modified base-stock policy by the authors), Bijvank and Johansen (2012) propose to set the maximum order quantity to $\tilde{r} = \tilde{S}/(L+1)$ rounded to the nearest integer.

1.2.6.2 Other heuristic replenishment policies

A number of alternative (heuristic) replenishment policies have been proposed in the literature. Morton (1969) derives bounds on the optimal policy and introduces a heuristic policy in which the order quantity q is set to the largest value satisfying these bounds. Zipkin (2008b) calls this policy the *standard vector base-stock policy* as it is a vector equivalent of a base-stock policy: when all elements in the vector are set to the same value, it corresponds to a base-stock policy.

In addition, Morton (1971) proposes a *myopic policy* in which the order quantity q is set such that the probability of not stocking out until the next order delivery equals the news vendor ratio. Define,

$$P_{NS}^{(t)}(y_0, y_1, \dots, y_t) = \begin{cases} G_1(y_0), & \text{if } t = 0, \\ \sum_{d=0}^{y_0} g_1(d) \cdot P_{NS}^{(t-1)}(y_0 - d + y_1, y_2, \dots, y_t) \\ + [1 - G_1(y_0)] \cdot P_{NS}^{(t-1)}(y_1, y_2, \dots, y_t), & \text{if } t \geq 1. \end{cases}$$

The myopic policy of Morton (1971) prescribes to order q units, where q is the smallest value such that $P_{NS}^{(L)}(\mathbf{x}, q) \geq p/(p+h)$. The order quantities in this myopic policy are an upper bound on the optimal order quantities (see Morton (1969)). They can also be found by minimizing the expected one-period cost in the period of order delivery. Therefore, the myopic policy is considered a one-period optimization problem. Alternatively, the problem can be extended to a horizon of T periods. This is called the myopic- T policy, where $T = 1$ corresponds to Morton's myopic policy.

Furthermore, Morton (1971) proposes a *restricted base-stock policy* with order quantities given by

$$q = \min \left\{ S_0, S^B - \sum_{t=0}^{L-1} x_t \right\}.$$

Note that this policy is equivalent to a capped base-stock policy, where the base-stock level is S^B and the maximum order quantity is S_0 .

In all policies mentioned so far, the sum of the expected holding and penalty costs is minimized. Lost-sales penalty costs are incurred due to the risk of ordering too few units, whereas holding costs are incurred due to the risk of ordering too many units. Levi et al. (2008) propose a *dual-balancing policy*, in which the two risks are balanced (see Section 1.2.7 for more details). The authors prove that the expected cost of this policy is at most twice the expected cost of the optimal policy.

1.2.6.3 Performance of heuristic policies: computational analysis

The performances of all heuristic policies mentioned in this section are presented in Table 1.2 for the same instances as Table 1.1. Since the base-stock policy with S^A is proposed to improve the policies with S^B and S^R , we omit the results of the latter two policies in Table 1.2. A comparison of their performances is discussed by Bijvank et al. (2014), where the authors conclude that the approach inspired by Archibald (1981) (i.e., S^A) outperforms the two other heuristic approaches to set the base-stock level.

From Table 1.2 we observe that the base-stock policies with S^A and S^H perform best when the penalty cost is high, whereas the base-stock policy with $\tilde{S}$ performs better for moderate to low values of p . Furthermore, it seems that $\tilde{S} \leq S^*$ (also observed by Bijvank and Johansen (2012)), whereas S^A is too high for most instances and S^H is too low when the penalty cost is low and too high when the penalty cost is high. A more extensive numerical analysis performed by Bijvank et al. (2014) reveals that the base-stock policy with S^A performs best of the latter two base-stock levels. The base-stock policy with $\tilde{S}$ performs best over all heuristic base-stock policies, and it can be improved when the order quantities are capped by $\tilde{r}$. In particular, this capped base-stock policy outperforms the best base-stock policy for low to moderate penalty costs.

Of the non-parametric heuristic policies, the myopic policy outperforms the capped base-stock policy $(\tilde{S}, \tilde{r})$ for moderate to high values of p , whereas the myopic-2 policy is consistently the best heuristic policy (also observed by Zipkin (2008b)). The standard vector base-stock policy only performs well when the penalty cost is high and the lead time is low, whereas the dual-balancing policy performs well when the penalty cost is low to moderate and the policy improves when the lead time grows. Based on this small set of instances, it is premature to draw firm conclusions but capped base-stock policies seem to provide some insights into what order quantities are close to optimal. An interesting heuristic to explore for setting the parameter values is to take the (weighted) average of $\tilde{S}$ and S^A as base-stock level and use the rule proposed by Bijvank and Johansen (2012) to set the maximum order quantity (i.e., base-stock level divided by $L + 1$).

1.2.7 More General Demand Models and Balancing Policies

Levi et al. (2008) develop what they refer to as dual-balancing policies for the lost-sales problem – this work is inspired by Levi et al. (2007), which had earlier developed such policies for

Table 1.2 Average total cost and relative cost increase compared to the optimal policy for the heuristic policies

Policy	$p = 2$	$p = 4$			$p = 9$			$p = 19$			$p = 39$				
$L = 1$															
BS - S^A	21	4.38	5.5%	23	5.88	3.3%	25	7.71	1.4%	27	9.36	0.8%	28	10.99	1.6%
BS - S^H	18	4.64	11.8%	22	5.87	3.1%	25	7.71	1.4%	27	9.36	0.8%	29	10.88	0.6%
BS - $\tilde{S}$	20	4.32	4.1%	22	5.87	3.1%	24	7.83	3.1%	26	9.55	2.7%	28	10.99	1.6%
CBS - S^B, S_0	22,11	4.40	5.9%	24,13	5.99	5.2%	26,14	7.78	2.4%	28,15	9.48	2.0%	29,17	10.87	0.4%
CBS - $\tilde{S}, \tilde{r}$	20,10	4.19	0.8%	22,11	5.76	1.1%	24,12	7.78	2.3%	26,13	9.52	2.5%	28,14	10.98	1.5%
Std. vector BS		4.40	5.9%		5.99	5.2%		7.78	2.4%		9.48	2.1%		10.87	0.4%
Myopic		4.20	1.2%		5.76	1.2%		7.67	1.0%		9.35	0.6%		10.82	0.0%
Myopic 2		4.16	0.1%		5.70	0.1%		7.60	0.0%		9.31	0.2%		10.85	0.3%
Dual balancing		4.17	0.4%		5.78	1.6%		8.03	5.7%		10.35	11.4%		12.72	17.6%
$L = 2$															
BS - S^A	30	4.83	9.7%	34	6.89	10.9%	36	8.92	4.2%	38	10.91	1.9%	40	12.78	1.0%
BS - S^H	25	5.26	19.4%	31	6.60	6.2%	35	8.84	3.3%	38	10.91	1.9%	40	12.78	1.0%
BS - $\tilde{S}$	29	4.76	8.0%	31	6.60	6.2%	34	8.96	4.7%	37	10.99	2.7%	39	12.97	2.6%
CBS - S^B, S_0	32,11	4.94	12.2%	35,13	7.08	13.8%	37,14	9.07	6.0%	39,15	11.02	2.9%	41,17	12.89	1.9%
CBS - $\tilde{S}, \tilde{r}$	29,10	4.49	1.8%	31,10	6.37	2.5%	34,11	8.82	3.1%	37,12	10.89	1.8%	39,13	12.93	2.2%
Std. vector BS		4.94	12.2%		7.02	12.8%		9.04	5.6%		11.00	2.8%		12.87	1.7%
Myopic		4.56	3.5%		6.40	2.9%		8.74	2.1%		10.81	1.0%		12.72	0.5%
Myopic 2		4.41	0.1%		6.22	0.0%		8.57	0.1%		10.71	0.0%		12.65	0.0%
Dual balancing		4.43	0.6%		6.26	0.7%		8.88	3.8%		11.66	8.9%		14.55	15.0%

(Continued)

Table 1.2 (Continued)

Policy	$p = 2$		$p = 4$		$p = 9$		$p = 19$		$p = 39$						
$L = 3$															
BS - $\mathcal{S}^A$	40	5.31	17.5%	44	7.52	15.3%	47	9.95	8.2%	50	12.35	5.5%	52	14.38	2.6%
BS - $\mathcal{S}^H$	32	5.62	24.4%	39	7.19	10.3%	45	9.65	5.0%	49	12.11	3.5%	52	14.38	2.6%
BS - $\tilde{\mathcal{S}}$	37	5.02	11.0%	40	7.07	8.5%	44	9.71	5.7%	47	12.22	4.4%	50	14.39	2.7%
CBS - $\mathcal{S}^B, \mathcal{S}_0$	43,11	5.70	26.1%	45,13	7.68	17.8%	48,14	10.17	10.6%	51,15	12.66	8.2%	53,17	14.68	4.8%
CBS - $\tilde{\mathcal{S}}, \tilde{r}$	37,9	4.61	2.0%	40,10	6.71	3.0%	44,11	9.42	2.5%	47,12	12.02	2.7%	50,13	14.25	1.7%
Std. vector BS		5.65	25.2%		7.56	16.0%		10.09	9.7%		12.61	7.7%		14.64	4.5%
Myopic		4.75	5.1%		6.85	5.0%		9.53	3.7%		11.98	2.3%		14.22	1.5%
Myopic 2		4.53	0.2%		6.57	0.8%		9.23	0.4%		11.73	0.2%		14.02	0.1%
Dual balancing		4.57	1.2%		6.54	0.3%		9.43	2.6%		12.55	7.2%		15.85	13.1%
$L = 4$															
BS - $\mathcal{S}^A$	51	6.01	31.0%	55	8.39	25.0%	58	10.91	13.1%	61	13.46	7.9%	63	15.63	3.6%
BS - $\mathcal{S}^H$	39	5.85	27.5%	47	7.64	13.9%	55	10.27	6.6%	60	13.15	5.5%	63	15.63	3.6%
BS - $\tilde{\mathcal{S}}$	45	5.20	13.3%	49	7.41	10.4%	54	10.29	6.7%	57	13.15	5.4%	60	15.78	4.6%
CBS - $\mathcal{S}^B, \mathcal{S}_0$	53,11	6.12	33.3%	56,13	8.59	28.1%	59,14	11.19	16.0%	62,15	13.81	10.7%	64,17	15.95	5.7%
CBS - $\tilde{\mathcal{S}}, \tilde{r}$	45,9	4.74	3.2%	49,10	6.94	3.4%	54,11	9.86	2.3%	57,11	12.93	3.7%	60,12	15.64	3.6%
Std vector BS		6.06	32.0%		8.38	24.8%		11.05	14.6%		13.73	10.1%		15.9	5.3%
Myopic		4.92	7.2%		7.16	6.6%		10.18	5.6%		12.94	3.8%		15.41	2.1%
Myopic 2		4.62	0.6%		6.83	1.7%		9.73	0.9%		12.54	0.5%		15.13	0.3%
Dual balancing		4.65	1.4%		6.73	0.3%		9.81	1.8%		13.21	5.9%		16.84	11.6%

Note: For the base-stock (BS) policies and capped base-stock (CBS) policies, we also report the corresponding base-stock level (and maximum order size).

inventory problems with backordering. These policies are guaranteed to achieve a cost which is no more than twice the optimal cost. An important feature of this work is that it allows a very general demand model – in particular, their model includes as special cases several important models of correlated demand, including several auto-regressive demand models.

The dual-balancing policy considers, as with other lost-sales models, two types of cost incurred in the planning horizon: the cost of ordering too little (underage cost) and the cost of ordering too much (overage cost). It attributes these costs to a decision made by the system in each period. The underage cost it assigns to a given period is the expected lost-sales penalty cost incurred in the period when the order arrives. The underage cost associated with period t occurs in period $t+L$. In comparison, the overage cost associated with period t is the expected marginal holding cost incurred by the units ordered in that period (period t) over the entire horizon $\{t+L, t+L+1, \dots, T\}$. The ordering decision in period t considers both of these costs. Instead of minimizing the sum of the underage and overage costs, the dual-balancing policy sets the ordering quantity such that these two costs are equal.

The analysis of dual-balancing is more complicated than its statement. It uses the notion of truncated inventory position, which is defined as the on-hand inventory and the sum of outstanding units ordered by a certain period in the past. The truncated inventory is parameterized by when this cut-off point in the past is located, and generalizes the classical notion of the inventory position, which includes all outstanding orders. The analysis depends on comparing how the dual-balancing policy behaves compared to the optimal policy on each sample path. It partitions the time periods into two sets H and Π , depending on whether the optimal policy has more inventory than the corresponding quantity in the dual-balancing policy. In the first set of periods (set H), the dual-balancing policy has less inventory than the optimal policy, and the dual-balancing policy incurs less holding cost than the optimal policy. Similarly, in the latter set of periods (set Π), the dual-balancing policy incurs less penalty cost than the optimal policy. The above comparisons are between the dual-balancing policy and the optimal policy. One can also compare the holding costs within the dual-balancing policy. Since set H refers to the periods in which the dual-balancing policy does not hold much inventory, one can show that the dual-balancing policy's expected single-period holding cost in a period conditioned that it belongs to H is less than the corresponding quantity without this condition. We can also obtain a similar result about the penalty cost. Also, recall how the dual-balancing policy chooses the order quantity to balance these two costs in each period. Putting these results together, the dual-balancing policy is shown to achieve, in the worst case, twice the optimal cost.

We note that the extent of notation required to discuss Levi et al. (2008) rigorously is formidable; hence our choice of the written description above.

1.2.8 The Projected Inventory Level Policy

Recently, van Jaarsveld and Arts (2021) proposed a single-parameter policy in which the order quantity is determined such that the expected inventory level at the time of order arrival matches the given parameter. They show that this policy, called the Projected Inventory Level policy, is asymptotically optimal as the penalty cost p approaches ∞ . Moreover, when the demand distribution is exponential, they show that this policy is asymptotically optimal as the lead time L approaches ∞ .

1.3 MULTI-ECHELON SERIAL SYSTEMS

Consider a serial inventory system with lost sales under periodic review. The system consists of $L \geq 1$ stages, indexed by $l = 0, 1, \dots, L-1$. (Recall that we used L to denote the lead time in Section 1.2. We find it convenient to use L to denote the number of stages in this section for reasons that will become clear once we present the results.) The lowest stage facing exogenous demand is represented by stage 0, stage l orders from stage $l+1$ where $l \in \{0, \dots, L-2\}$, and stage $L-1$ orders from an outside supplier with infinite supply. The replenishment lead time from one stage to another is deterministic, and we assume, for expositional convenience, that each lead time is 1 period. In each period, the following sequence of events occurs: (1) receipt of delivery at every stage, (2) order placement at every stage, and (3) demand realization. The demand in each period is satisfied to the extent possible, and any demand that cannot be satisfied immediately is lost. Let $p \geq 0$ represent the per-unit lost-sales penalty cost, and let $H_l \geq 0$ denote the holding cost at stage l ; that is, the cost for holding one unit at stage l for a period. These costs are charged at the end of a period. For any $l \in \{0, 1, \dots, L-1\}$, let $x_l \geq 0$ denote the stage- l inventory level after receiving deliveries at the beginning of a period, and let $q_l \in [0, x_{l+1}]$, where x_L is taken to be ∞ , denote the quantity ordered by stage l in a period. Thus, the vector $(x_0, x_1, \dots, x_{L-1})$ represents the state at the beginning of a period after receiving deliveries, and the vector $(q_0, q_1, \dots, q_{L-1})$ represents the action in a period. We use $\mathbf{x}$ and $\mathbf{q}$ to denote the vectors $(x_0, \dots, x_{L-1})$ and $(q_0, \dots, q_{L-1})$, respectively. Let D_t denote the demand in period t . We assume that demands are independently distributed across periods.

1.3.1 Dynamic Programming Formulation

Huh and Janakiraman (2010b) study the dynamic program for this system and present important results on the structure of the optimal policy. The dynamic program for this system for a planning horizon of T periods indexed by $t = 1, 2, \dots, T$ is presented below. Periods are indexed forwards, i.e., period $t+1$ follows period t . Given a state-action pair of $\mathbf{x}$ and $\mathbf{q}$ in period t , the stage- l inventory level in period $t+1$ is given by

$$\begin{cases} (x_0 - D_t)^+ + q_0, & \text{if } l = 0, \\ x_l - q_{l-1} + q_l, & \text{if } l \in \{1, \dots, L-1\}. \end{cases}$$

Let $f_{T+1}(\mathbf{x}) = 0$ for all $\mathbf{x}$. Consider any $t \in \{1, \dots, T\}$. Let $\alpha \in (0, 1]$ denote the discount factor for capturing the time value of money. Define

$$\begin{aligned} f_t(\mathbf{x}) = & E \left[H_0 \cdot (x_0 - D_t)^+ + \sum_{l=1}^{L-1} H_l \cdot x_l + p \cdot (D_t - x_0)^+ \right] \\ & + \alpha \cdot \min_{\mathbf{q}} E \left[f_{t+1} \left((x_0 - D_t)^+ + q_0, x_1 - q_0 + q_1, \dots, x_{L-1} - q_{L-2} + q_{L-1} \right) \right] \\ \text{s. t.} \quad & 0 \leq q_l \leq x_{l+1}, l = 0, \dots, L-1. \end{aligned}$$

Let $\mathbf{q}_t^*(\mathbf{x})$ denote a minimizing vector $\mathbf{q}$ in the optimization problem above. Thus, the l^{th} component of $\mathbf{q}_t^*$ denotes the order quantity for stage l , according to this vector. (When the

optimization problem above does not have a unique solution, any statements we make on properties of $\mathbf{q}_i^*$ should be taken to mean the existence of an *optimal selector* $\mathbf{q}_i^*(\mathbf{x})$ (precisely defined in Section 3.4 of Huh and Janakiraman (2010b) with those properties.)

Huh and Janakiraman (2010b) derive the following structural results that generalize the result of Theorem 1.1 using a sophisticated application of $L^\natural$ -convexity.

Theorem 1.7 *Assume that $\mathbf{q}_i^*$ is a differentiable function of $\mathbf{x}$. Then, the following inequalities hold:*

$$\begin{aligned} -1 &\leq \frac{\partial q_k^*}{\partial x_k} \leq \dots \leq \frac{\partial q_k^*}{\partial x_0} \leq 0 \text{ for every } k \in \{0, \dots, L-1\}, \\ 0 &\leq \frac{\partial q_k^*}{\partial x_{L-1}} \leq \dots \leq \frac{\partial q_k^*}{\partial x_{k+1}} \leq 1 \text{ for every } k \in \{0, \dots, L-2\}, \text{ and} \\ &\frac{\partial q_k^*}{\partial x_{k+1}} - \frac{\partial q_k^*}{\partial x_k} \leq 1 \text{ for every } k \in \{0, \dots, L-2\}. \end{aligned}$$

If $\mathbf{q}_i^*$ is not differentiable, then the inequalities above hold after replacing every quantity of the form $\frac{\partial q_k^*}{\partial x_l}$ with $\frac{q_k^*(\mathbf{x} + \boldsymbol{\varepsilon}^l) - q_k^*(\mathbf{x})}{\varepsilon}$, where $\boldsymbol{\varepsilon}^l$ is a vector with ε in its l^{th} component and zero in all the other components and ε is any strictly positive number.

In words, Theorem 1.7 states that the optimal order quantity at stage k is a decreasing function of the inventory at any downstream stage j (i.e., $j \leq k$), and that the rate of this decrease is smaller than 1. On the other hand, the optimal order quantity at stage k is an increasing function of the inventory at any upstream stage j (i.e., $j > k$), and the rate of this increase is also smaller than 1. Furthermore, we consider the impact of the location on these sensitivity results (i.e., the dependence of the optimal quantity at stage k on the inventory amount at another stage j). We see that the closer two stages are, the greater the sensitivity's magnitude. Finally, the rate of decrease of the optimal order quantity at a stage with respect to any downstream stage's inventory plus the rate of its increase with respect to any upstream stage's inventory is itself smaller than 1.

1.3.2 Echelon OUT Policies

An *echelon OUT policy* with order-up-to vector $\mathbf{S}$ is defined as follows: For any period t , given any state $\mathbf{x}$, the quantity shipped from stage $l+1$ to stage l , $0 \leq l \leq L-1$ (stage L is the external supplier), is

$$q_l(\mathbf{x}) = \min\{x_{l+1}, (S_l - (x_0 + x_1 + \dots + x_l))^+\},$$

where $x_L := \infty$. In other words, it brings each echelon- l inventory position to S_l .

1.3.2.1 Optimizing $\mathbf{S}$

Consider an echelon order-up-to $\mathbf{S}$ policy. Assume that the starting state is such that $x_l = S_l$ for all $l \in \{0, 1, \dots, L-1\}$. Let $f_{\alpha, T}^{OUTP}(\mathbf{S})$ and $f_{\alpha, \infty}^{OUTP}(\mathbf{S})$ denote the expected discounted sum of

costs under this policy over a finite horizon of T periods and over an infinite horizon, respectively. Let $f_{\text{Avg}}^{\text{OUTP}}(\mathbf{S})$ denote the infinite-horizon, long-run average cost. We use $\text{Lost}_{\alpha,T}^{\text{OUTP}}(\mathbf{S})$, $\text{Lost}_{\alpha,\infty}^{\text{OUTP}}(\mathbf{S})$ and $\text{Lost}_{\text{Avg}}^{\text{OUTP}}(\mathbf{S})$ to denote the lost-sales costs in these three settings. The following results are based on Huh and Janakiraman (2010a), which are the counterparts of Theorem 1.2 for serial inventory systems.

Theorem 1.8 *The functions $\text{Lost}_{\alpha,T}^{\text{OUTP}}(\mathbf{S})$, $\text{Lost}_{\alpha,\infty}^{\text{OUTP}}(\mathbf{S})$, and $\text{Lost}_{\text{Avg}}^{\text{OUTP}}(\mathbf{S})$ are convex in $\mathbf{S}$. Moreover, each of the functions $f_{\alpha,T}^{\text{OUTP}}(\mathbf{S})$, $f_{\alpha,\infty}^{\text{OUTP}}(\mathbf{S})$, and $f_{\text{Avg}}^{\text{OUTP}}(\mathbf{S})$ can be written as a sum of a convex function of $\mathbf{S}$ and a concave function of $\mathbf{S}$ (or, equivalently, as the difference between two convex functions of $\mathbf{S}$).*

In contrast to what might be expected from Theorem 1.2 for the single-stage inventory system, the expected discounted total cost and long-run average cost for the serial inventory system are non-convex functions in $\mathbf{S}$. However, special global optimization techniques are available for minimizing the difference between two convex functions subject to polyhedral constraints (e.g., see Chapter 4 of Horst et al. (2000)). Such techniques can be used to optimize the desired cost metric ($f_{\alpha,T}^{\text{OUTP}}(\mathbf{S})$, $f_{\alpha,\infty}^{\text{OUTP}}(\mathbf{S})$, or $f_{\text{Avg}}^{\text{OUTP}}(\mathbf{S})$) over the set $\{\mathbf{S} : S_L \geq S_{L-1} \geq \dots \geq S_0\}$.

1.3.2.2 Asymptotic optimality

A recent working paper by Bijvank et al. (2020) generalizes the results of asymptotic optimality of OUT policies in single-stage systems as p approaches ∞ (see Theorem 1.4) by demonstrating the asymptotic optimality of a family of echelon OUT policies in serial systems. It is well known that an echelon OUT policy minimizes the long-run average cost per period in a backorder system identical to the lost-sales system under analysis with the exception that excess demand in every period is backordered at a penalty cost of b per unit (see Clark and Scarf (1960) or Federgruen and Zipkin (1984)). We denote that backorder system with $\mathcal{B}(b)$ and the corresponding echelon order-up-to vector in the optimal policy as $\mathbf{S}_b^{\mathcal{B}*}$. Note that $\mathbf{S}_b^{\mathcal{B}*}$ can be computed using the algorithm described in Federgruen and Zipkin (1984), which is based on the algorithm of Clark and Scarf (1960) for the finite-horizon problem.

Let $f_{\text{Avg}}^{\text{OUTP}}(\mathbf{S}; p)$ denote the long-run average cost of the echelon OUT policy as a function of the echelon order-up-to vector $\mathbf{S}$ and the lost-sales penalty cost p . Let $f_{\text{Avg}}^*(p)$ denote the optimal cost (over all non-anticipatory policies). Bijvank et al. (2020) show the following.

Theorem 1.9 *Assume random variable D (single-period demand) satisfies one of the following conditions:*

- (i) *D is unbounded, i.e., $\sup\{x : P(D \leq x) < 1\} = \infty$, and $0 < E[D] < \infty$. For any positive integer k ,*

$$\lim_{d \rightarrow \infty} E[D[1, k] - d \mid D[1, k] > d] / d = 0.$$

- (ii) *D is bounded.*
Then,

$$\lim_{p \rightarrow \infty} \left(\frac{f_{\text{Avg}}^{\text{OUTP}}(\mathbf{S}; p)}{f_{\text{Avg}}^*(p)} \right) = 1,$$

where $\mathbf{S} = \mathbf{S}^{B*}$ and $H_L := 0$.

$$p + \sum_{i=0}^{L-1} (H_i - H_{i+1})(L-i)$$

1.4 POINTERS TO RELATED LOST-SALES MODELS

1.4.1 Continuous-Review Models

When replenishment orders can be placed in continuous time instead of after prespecified time intervals, similar characteristics are known about the structure of an optimal replenishment policy as for the periodic review counterpart (as discussed in Section 1.2.2). The constant-order policy discussed in Section 1.2.4 that Reiman (2004) proposed was initially in the continuous-review setting. When a fixed cost is imposed on each order, Johansen and Thorstenson (1996) show that optimal order quantities are decreasing in the inventory position with a rate of decrease between zero and one, and there is a threshold value for the inventory position above which no order is placed. As for most models that study continuous-review inventory systems with lost sales, the authors make the assumption that the demand process is Poisson, the lead time is either constant or exponentially distributed and at most one order can be outstanding. Consequently, the inventory system can be modeled as a one-dimensional semi-Markov process where the decision epochs to order are either when a demand occurs and no order is outstanding, or at order delivery. In the literature, heuristic replenishment policies are mostly explored instead of an optimal policy. In this section, we only focus on such studies with negligible fixed replenishment order cost, and we refer to Bijvank and Vis (2011) otherwise.

In base-stock policies, every customer's demand is immediately reordered. For inventory systems with Poisson demand and (non-crossing) Erlangian lead times that are controlled by such a policy, Johansen (2005) shows that the stationary distribution of the number of outstanding replenishment orders is a modified truncated version of the negative binomial distribution. The author also derives an expression for the fraction of the demand lost, which is convex at the base-stock level. Consequently, computing the long-run average cost per unit time and finding the optimal order-up-to level is straightforward. More general lead times are studied by Johansen (2021). Chen et al. (2011) derive the exact stationary distribution for the number of units on order when demand follows a stuttering Poisson process (i.e., a Poisson arrival process with geometrically distributed demand sizes) and the lead times have a general distribution with a finite mean. These results are extended to the general compound Poisson demand case by Koukia et al. (2019), who also study the stock-out frequency and base-stock levels that minimize the average expected cost under complete and partial *rejection policies*.

Base-stock policies with a delay are proposed by Hill (1999), Hill (2007), and Johansen (2013), where new replenishment orders are only placed when the time between successive orders is at least κ time units. The value of κ can be either state-dependent or a constant. Note that base-stock policies with an order cap in periodic review inventory systems (see Section 1.2.5) also delay the ordering of units that exceed the threshold on the order size. The reason that both modified base-stock policies work well is the same: The order placement is delayed when many units are demanded in a short amount of time, since it is likely that the inventory system already has sufficient units in stock if the ordered units would have been delivered earlier. Furthermore, note that orders are more evenly spread over time when a minimum delay is imposed. In the extreme case, orders can be placed at every κ time unit (similar to constant-order policies for the periodic review inventory system discussed in Section 1.2.4).

In the numerical instances investigated by Johansen (2013), the proposed delayed base-stock policy reduces the average total cost compared to the (simple) base-stock policy by 2 to 8% in most instances. The relative cost reductions become smaller for shorter lead times or larger unit penalty costs (compared to the unit holding cost).

1.4.2 Demand Learning with Lost Sales and Censored Demand

When the demand distribution is not known *a priori*, the manager makes decisions to balance learning and immediate cost. In each period, the manager decides the inventory amount and then observes the sales quantity, which is the smaller of the inventory level and realized demand. With lost sales, it is natural to assume that the manager does not know realized demand and only observes the sales quantity, the censored demand. Extending the well-known sample-average-approximation (SAA) approach to the newsvendor problem with censored demand, Huh et al. (2011a) propose a policy motivated by the Kaplan-Meier estimator and show convergence to optimality. Besbes and Muharremoglu (2013) develop a cycle-based method with $O(\log T)$ regret bound and a matching lower bound. In a parametric setting with the Weibull demand distribution, Besbes et al. (2022) show that the myopic policy achieves a regret of $O(\log T)$. Huh and Rusmevichientong (2009) propose a stochastic gradient descent (SGD) method that is asymptotically optimal with a regret of $O(\sqrt{T})$, which can be improved to $O(\log T)$ under a technical condition. Zhang et al. (2020) apply an SGD method to the lost-sales system with positive lead times and show that their algorithms achieve the performance gap of at most $O(\sqrt{T})$ against the best base-stock policy. While this gap is sublinear in T , it is exponential in the lead time L , and Agrawal and Jia (2019) develop a learning algorithm with a regret bound that is linear in L .

REFERENCES

- Agrawal, S., & Jia, R. (2019). Learning in structured mdps with convex cost functions: Improved regret bounds for inventory management. *Proceedings of the 2019 ACM Conference on Economics and Computation*, 743–744.
- Archibald, B. (1981). Continuous review (s, S) policies with lost sales. *Management Science*, 27(10), 1171–1177.
- Bai, X., Chen, X., Li, M., & Stolyar, A. (2023). Asymptotic Optimality of Open-Loop Policies in Lost-Sales Inventory Models with Stochastic Lead Times. *Working Paper*.
- Besbes, O., & Muharremoglu, A. (2013). On implications of demand censoring in the newsvendor problem. *Management Science*, 59(6), 1407–1424.
- Besbes, O., Chaneton, J. M., & Moallemi, C. C. (2022). The exploration-exploitation trade-off in the newsvendor problem. *Stochastic Systems*.
- Bijvank, M., Huh, W., & Janakiraman, G. (2020). Asymptotic optimality of echelon ($s-1, s$) and (r, nq) replenishment policies for serial inventory systems with lost sale. *Working Paper*.
- Bijvank, M., & Johansen, S. (2012). Periodic review lost-sales inventory models with compound Poisson demand and constant lead times of any length. *European Journal of Operational Research*, 220(1), 106–114.
- Bijvank, M., Huh, W. T., Janakiraman, G., & Kang, W. (2014). Robustness of order-up-to policies in lost-sales inventory systems. *Operations Research*, 62(5), 1040–1047.
- Bijvank, M., & Vis, I. F. (2011). Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215(1), 1–13.
- Chen, J., Jackson, P. L., & Muckstadt, J. A. (2011). Exact analysis of a lost sales model under stuttering Poisson demand. *Operations Research*, 59(1), 249–253.

- Clark, A., & Scarf, H. (1960). Optimal policies for a multiechelon inventory problem. *Management Science*, 6, 475–490.
- Downs, B., Metters, R., & Semple, J. (2001). Managing inventory with multiple products, lags in delivery, resource constraints, and lost sales: A mathematical programming approach. *Management Science*, 47(3), 464–479.
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836.
- Goldberg, D. A., Katz-Rogozhnikov, D. A., Lu, Y., Sharma, M., & Squillante, M. S. (2016). Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. *Mathematics of Operations Research*, 41(3), 898–913.
- Goldberg, D. A., Reiman, M. I., & Wang, Q. (2021). A survey of recent progress in the asymptotic analysis of inventory systems. *Production and Operations Management*, 30(6), 1718–1750.
- Hill, R. (1999). On the suboptimality of $(s-1, s)$ lost sales inventory policies. *International Journal of Production Economics*, 59, 387–393.
- Hill, R. (2007). Continuous-review, lost-sales inventory models with Poisson demand, a fixed lead time and no fixed order cost. *European Journal of Operational Research*, 176(2), 956–963.
- Horst, R., Pardalos, P. M., & Van Thoai, N. (2000). *Introduction to global optimization*. Springer Science & Business Media.
- Huh, W. T., Levi, R., Rusmevichientong, P., & Orlin, J. B. (2011a). Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research*, 59(4), 929–941.
- Huh, W. T., & Rusmevichientong, P. (2009). A nonparametric asymptotic analysis of inventory planning with censored demand. *Mathematics of Operations Research*, 34(1), 103–123.
- Huh, W. T., & Janakiraman, G. (2010a). Base-stock policies in capacitated assembly systems: Convexity properties. *Naval Research Logistics*, 57(2), 109–118.
- Huh, W. T., & Janakiraman, G. (2010b). On the optimal policy structure in serial inventory systems with lost sales. *Operations Research*, 58(2), 486–491.
- Huh, W. T., Janakiraman, G., Muckstadt, J. A., & Rusmevichientong, P. (2009). Asymptotic optimality of order-up-to policies in lost sales inventory systems. *Management Science*, 55(3), 404–420.
- Huh, W. T., Janakiraman, G., & Nagarajan, M. (2011b). Average cost single-stage inventory models: An analysis using a vanishing discount approach. *Operations Research*, 59(1), 143–155.
- Janakiraman, G., & Wu, Q. (2020). Convex combinations of asymptotically optimal policies in stochastic inventory systems. *Working Paper*.
- Janakiraman, G., & Roundy, R. O. (2004). Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Operations Research*, 52(5), 795–803.
- Janakiraman, G., Seshadri, S., & Shanthikumar, J. G. (2007). A comparison of the optimal costs of two canonical inventory systems. *Operations Research*, 55(5), 866–875.
- Johansen, S. (2005). Base-stock policies for the lost sales inventory system with Poisson demand and Erlangian lead times. *International Journal of Production Economics*, 93–94, 429–437.
- Johansen, S. (2013). Modified base-stock policies for continuous-review, lost-sales inventory models with Poisson demand and a fixed lead time. *International Journal of Production Economics*, 143(2), 379–384.
- Johansen, S. (2021). The Markov model for base-stock control of an inventory system with Poisson demand, non-crossing lead times and lost sales. *International Journal of Production Economics*, 231.
- Johansen, S., & Thorstenson, A. (1996). Optimal (r, q) inventory policies with Poisson demands and lost sales: discounted and undiscounted cases. *International Journal of Production Economics*, 46–47, 359–371.
- Karlin, S. (1958). Inventory models of the Arrow-Harris-Marschak type with time lag. *Studies in the Mathematical Theory of Inventory and Production*. Stanford University Press.
- Kingman, J. (1962). Some inequalities for the queue GI/G/1. *Biometrika*, 49(3/4), 315–324.
- Koukia, C., Babaib, M. Z., Jemaic, Z., & Minner, S. (2019). Solution procedures for lost sales base-stock inventory systems with compound Poisson demand. *International Journal of Production Economics*, 209, 172–182.
- Levi, R., Janakiraman, G., & Nagarajan, M. (2008). A 2-approximation algorithm for stochastic inventory control models with lost sales. *Mathematics of Operations Research*, 33(2), 351–374.

- Levi, R., Pál, M., Roundy, R. O., & Shmoys, D. B. (2007). Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, 32(2), 284–302.
- Morton, T. (1969). Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM Review*, 11(4), 572–596.
- Morton, T. (1971). The near-myopic nature of the lagged-proportional-cost inventory problem with lost sales. *Operations Research*, 19(7), 1708–1716.
- Murota, K. (2003). *Discrete Convex Analysis*, SIAM Monographs on Discrete Mathematics and Applications, vol. 10. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Reiman, M. I. (2004). A new and simple policy for the continuous review lost sales inventory model. *Unpublished manuscript*.
- Rosling, K. (2002). Inventory cost rate functions with non-linear shortage costs. *Operations Research*, 50(6), 1007–1017.
- Sheopuri, A., Janakiraman, G., & Seshadri, S. (2010). New policies for the stochastic inventory control problem with two supply sources. *Operations Research*, 58(3), 734–745.
- van Jaarsveld, W., & Arts, J. (2021). Projected inventory level policies for lost sales inventory systems: Asymptotic optimality in two regimes. *arXiv preprint arXiv:2101.07519*.
- Xin, L. (2021). Understanding the performance of capped base-stock policies in lost-sales inventory models. *Operations Research*, 69(1), 61–70.
- Xin, L., & Goldberg, D. A. (2016). Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Operations Research*, 64(6), 1556–1565.
- Zhang, H., Chao, X., & Shi, C. (2020). Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Science*, 66(5), 1962–1980.
- Zipkin, P. (2008a). On the structure of lost-sales inventory models. *Operations Research*, 56(4), 937–944.
- Zipkin, P. (2008b). Old and new methods for lost-sales inventory systems. *Operations Research*, 56(5), 1256–1263.

2. Perishable inventory systems

Qing Li and Peiwen Yu

2.1 INTRODUCTION

Perishable inventory management is one of the most researched areas in Operations Management/Operations Research. There have been at least four reviews of the literature: Prastacos (1982), Nahmias (1982), Karaesmen et al. (2011), and Nahmias (2011). The review by Prastacos (1982) focuses on blood products, and the other three are more general. Since the last two reviews in 2011, the area has been experiencing a resurgence of interest. In this review, we focus on recent results that are not discussed in the earlier reviews.

In our view, there are three reasons for the renewed interest in this research area. First, driven by market development, there has been much interest in e-commerce, health care operations, and sustainable operations (in particular waste reduction), and many problems in these areas are related to perishable inventory systems. Second, although perishable inventory systems are known to be difficult to analyze, this research area has received a boost from the successful application of tools such as multimodularity and L^1 -convexity in large-scale dynamic programs. Third, perishable inventory systems, which are typically computationally challenging, have proved to be great test beds for new ideas in algorithm design, another area that has been experiencing increased research recently.

In this chapter, we focus on research on the structure of optimal inventory decisions. Research on perishable inventory management with pricing decisions is reviewed in Chapter 14 of this book (e.g., Chen et al., 2014; Hu et al., 2016). Research on approximation and learning algorithms for perishable inventory systems is reviewed in Chapters 12, 15, and 16 of this book (e.g., Chao et al., 2015, 2018). The remainder of this review is organized as follows. Section 2.2 reviews the results for joint replenishment and clearance sales when there is only one class of demand and inventories are depleted either on a first-in-first-out (FIFO) basis or last-in-first-out (LIFO) basis. Section 2.3 discusses how the results under the FIFO rule can be generalized to allow multiple classes of demand. Section 2.4 reviews the results of perishable inventory systems that involve multiple locations. Section 2.5 reviews models with endogenous product lifetimes. Section 2.6 discusses the results of empirical research. The review concludes with a discussion of potential future research directions in Section 2.7.

2.2 MODELS WITH ONE CLASS OF DEMAND AND LOCATION

Consider a firm that sells perishable products with an n -period lifetime. The firm purchases products at unit cost c . The products can be sold either at a regular price, r , or a clearance sale price, s . With a regular price, the demand in a period is random. Let D represent the regular demand. Unmet demand is lost. Demand under the clearance sales price scenario is abundant, and thus a firm can control how many items it sells for a given price. Without loss of generality, it is assumed that items have zero value once they expire. Items that expire incur outdating

cost θ per unit and are removed from the shelf and disposed of. Items that are carried over to the next period incur holding cost h per unit. Profits received in future periods are discounted by discount factor α . Clearance sales are a common strategy to reduce a mismatch between supply and demand for perishable goods. To effectively use clearance sales to reduce a mismatch, it is important that retailers choose the right timing and sales depth and coordinate such sales with replenishment decisions.

Depending on the sequence by which items in different age groups are used to fulfill the regular demand, the analysis of these models requires different analytical tools and leads to different optimal replenishment and clearance sale policies.

Li and Yu (2014) study a problem in which a firm controls inventory issuance. This problem is particularly relevant to blood banks and e-commerce platforms that sell grocery products. The problem's sequence of events is as follows. At the beginning of each period, the firm's initial state is $\mathbf{y} = (y_1, y_2, \dots, y_{n-1})$, which represents the inventory level after regular demand is fulfilled but before any clearance sales. Here, y_i represents inventory with a remaining lifetime of i periods. The firm must decide on an order quantity q of new items and the amount of inventory that will be carried over to the next period, denoted by $\mathbf{z} = (z_1, z_2, \dots, z_{n-1})$. Then, after regular demand is realized, the firm decides on an issuing policy to meet the demand. At the end of each period, the items that expired in that period are removed and disposed of. Let d_i denote the amount of regular demand that is met by the inventories with a remaining lifetime of i periods. Let $\mathbf{d} = (d_1, d_2, \dots, d_n)$ and $O(D) = \{\mathbf{d} : \sum_{i=1}^n d_i \leq D, 0 \leq d_n \leq q, 0 \leq d_i \leq z_i \text{ for } 1 \leq i \leq n-1\}$.

The dynamic programming formulation is as follows:

$$\pi_t(\mathbf{y}) = s \sum_{i=1}^{n-1} y_i + \max_{0 \leq z \leq y, q \geq 0} u_t(\mathbf{z}, q),$$

where

$$\begin{aligned} u_t(\mathbf{z}, q) = & -(s + h) \sum_{i=1}^{n-1} z_i - \alpha c q + \alpha \mathbb{E} \max_{\mathbf{d} \in O(D)} \left\{ r \sum_{i=1}^n d_i - \theta(z_1 - d_1) \right. \\ & \left. + \pi_{t+1}(z_2 - d_2, \dots, z_{n-1} - d_{n-1}, q - d_n) \right\}, \end{aligned}$$

and $\pi_{t+1}(\mathbf{y}) = s \sum_{i=1}^{n-1} y_i$. In the formulation above, $s \sum_{i=1}^{n-1} (y_i - z_i)$ represents the revenue from clearance sales. Since $\sum_{i=1}^{n-1} z_i$ units of inventory are carried over to the next period, they incur a total holding cost of $h \sum_{i=1}^{n-1} z_i$. Given the demand fulfillment vector $\mathbf{d}$, the revenue from regular sales is $r \sum_{i=1}^n d_i$, the amount of inventory that will be outdated is equal to $z_1 - d_1$, and the initial state in the next period is $(z_2 - d_2, \dots, z_{n-1} - d_{n-1}, q - d_n)$. Let $(\mathbf{z}_t^*, q_t^*)$, which denotes the optimal solution.

This problem is unique in two ways. First, it has a multidimensional state space. Second, the state and decision variables are economic substitutes. The standard approach used to prove the preservation of structural properties in dynamic programs can not be applied directly to substitutes. Li and Yu (2014) use the concept of multimodularity to establish the structural properties for this problem, as follows.

An n -dimensional set $X \subseteq \mathbb{R}^n$ is called a *multimodular set* if there exist $\mathbf{a}_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$ such that $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i \cdot \mathbf{x} \geq b_i, i = 1, 2, \dots, m\}$ and $\mathbf{a}_i$ has the form $(0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$; that is, the nonzero components of $\mathbf{a}_i$ are either consecutive 1s or consecutive -1s. Let $\mathbf{x} = (x_1, \dots, x_n)$. An n -dimensional function $f(\mathbf{x})$ defined on a multimodular set $X \subseteq \mathbb{R}^n$ is *multimodular* (*anti-multimodular*) if $f(x_1 - z, x_2 - x_1, \dots, x_n - x_{n-1})$ is submodular (supermodular) in $(\mathbf{x}, z)$.

Anti-multimodularity implies decreasing difference, and it thus can be used to analyze models with substitutable variables. Anti-multimodular functions have some useful properties. A continuous anti-multimodular function is jointly concave. If $g(x)$ is a one-dimensional concave function, then $f(\mathbf{x}) = g(x_1 + x_2 + \dots + x_n)$ is anti-multimodular in $\mathbf{x}$. The sum of anti-multimodular functions is still anti-multimodular; that is, if $f(\mathbf{x})$ and $g(\mathbf{x})$ are anti-multimodular, then $f(\mathbf{x}) + g(\mathbf{x})$ is anti-multimodular, and if $f(\mathbf{x}, d)$ is anti-multimodular in $\mathbf{x}$ for any given d and D is a random variable, then $\mathbb{E}_D f(\mathbf{x}, D)$ is anti-multimodular in $\mathbf{x}$.

Anti-multimodularity is preserved under maximization and the maximizer of an anti-multimodular function has monotonicity properties with bounded sensitivity. This property makes anti-multimodularity a useful tool for identifying structural properties in dynamic programs. To be more specific, if $f(\mathbf{x}, y)$ is an $n + 1$ dimensional anti-multimodular function and $\{(\mathbf{x}, y) \mid \mathbf{x} \in X, y \in Y(\mathbf{x})\}$ is a multimodular set, then

$$g(\mathbf{x}) = \max_{y \in Y(\mathbf{x})} f(\mathbf{x}, y)$$

is anti-multimodular in $\mathbf{x}$. The optimal solution, denoted by $\bar{y}(\mathbf{x})$, satisfies the following inequalities:

$$-1 \leq \Delta_{x_n} \bar{y} \leq \Delta_{x_{n-1}} \bar{y} \leq \dots \leq \Delta_{x_1} \bar{y} \leq 0.$$

Throughout the remainder of the chapter, the notation $\Delta_{x_i} J(\mathbf{x})$ is used to represent $(J(\mathbf{x} + \delta \mathbf{e}_i) - J(\mathbf{x})) / \delta$, where $\mathbf{e}_i$ is a vector with 1 in its i^{th} component and 0 in all the other components and δ is a small positive number. When $J(\mathbf{x})$ is differentiable, then $\Delta_{x_i} J(\mathbf{x})$ means that $\partial J(\mathbf{x}) / \partial x_i$.

Multimodularity is closely related to $L^\natural$ -convexity, a stronger notion of complementarity than submodularity (Lu & Song, 2005; Zipkin, 2008). Multimodular functions and $L^\natural$ -convex functions are related through unimodular coordinate transformations. For the models in which the state and decision variables are economic substitutes, to use $L^\natural$ -convexity to show structural properties, one must first transform the original variables into complementary variables, then show structural properties with respect to the new variables by showing $L^\natural$ -convexity, and finally transform the properties back to those with respect to the original variables. In spite of their mathematical equivalence, they represent two conceptually different paths to the same destination. Whereas one tackles the problems directly, the other takes a detour by transforming them into problems of complementarity.

Using multimodularity, Li and Yu (2014) show that the optimal inventory issuing rule is FIFO, and that both the maximal profit function, π_i , and the objective function, u_i , are anti-multimodular. Note that without the option of clearance sales (e.g., Fries, 1975; Nandakumar & Morton, 1993), FIFO may not be the optimal issuing rule and multimodularity may not hold. For a given inventory of a certain age, the optimal policy on clearance sales has a clear-down-to structure; that is, there is a clear-down-to level such that a clearance sale will take

place if and only if the inventory level is above the clear-down-to level and the clearance sale always reduces the inventory to that level. The details are summarized in Theorem 2.1 below.

Theorem 2.1

- (i) *The functions $u_t(\mathbf{z}, q)$ and $\pi_t(\mathbf{y})$ are anti-multimodular.*
- (ii) *The optimal policy for clearance sales is characterized by $\bar{z}_{t,i}$, where $\bar{z}_{t,i}$ is a decreasing function of $y_{i+1}, y_{i+2}, \dots, y_{n-1}$ and is independent of $y_1, y_2, \dots, y_i$. The optimal policy is:*

$$z_{t,i}^*(\mathbf{y}) = \begin{cases} y_i & \text{if } y_i \leq \bar{z}_{t,i}; \\ \bar{z}_{t,i} & \text{if } y_i > \bar{z}_{t,i}. \end{cases}$$

In addition, the following inequalities hold:

$$-1 \leq \Delta_{y_{i+1}} \bar{z}_{t,i} \leq \Delta_{y_{i+2}} \bar{z}_{t,i} \leq \dots \leq \Delta_{y_{n-1}} \bar{z}_{t,i} \leq 0.$$

- (iii) *The optimal replenishment quantity $q_t^*(\mathbf{y})$ is a decreasing function of $y_1, y_2, \dots, y_{n-1}$, and the following inequalities hold:*

$$-1 \leq \Delta_{y_{n-1}} q_t^* \leq \Delta_{y_{n-2}} q_t^* \leq \dots \leq \Delta_{y_1} q_t^* \leq 0.$$

The quantities $\bar{z}_{t,i}$ are state-dependent thresholds, and they depend only on inventories that are newer than i . Specifically, the more inventories with a remaining lifetime longer than i , the less inventory with an i -period remaining lifetime should be carried to the next period. In addition, the thresholds $\bar{z}_{t,i}$ are more sensitive to the inventories with a remaining lifetime closer to i . Similarly, the inequalities about the optimal order quantity confirm that the order quantity is more sensitive to newer inventory than to older inventory.

This model is extended by Liu et al. (2019) to allow for two perishable products with a dependent supply. Their research is based on the case of a blood center that periodically collects whole blood and processes it into multiple blood products such as red blood cells and platelets. They similarly characterize the optimal clearance sale and replenishment policies by showing the multimodularity of the objective function.

While both Li and Yu (2014) and Liu et al. (2019) assume that firms control how inventory is depleted and hence that the FIFO issuing rule is optimal, Li et al. (2016) study a problem in which consumers control inventory depletion and hence the LIFO rule is appropriate. In offline retailing, consumers observe expiration dates and decide which items to pick. Therefore, fresher items typically sell first on a LIFO basis. The problem is challenging not only because the state space is large but also because inventory systems under LIFO are known to lack the common technical properties such as concavity needed for analysis, let alone stronger properties such as multimodularity.

The sequence of the events is the same as that in the previous models. As the regular demand is fulfilled using the LIFO rule, given an initial state $\mathbf{y} = (y_1, y_2, \dots, y_{n-1})$, the state transition becomes $\mathbf{Y}(q, \mathbf{z}, D) = (Y_1, Y_2, \dots, Y_{n-1})$, where for $1 \leq i \leq n-2$

$$Y_i(q, \mathbf{z}, D) = (z_{i+1} - (D - q - \sum_{k=i+2}^{n-1} z_k)^+)^+$$

and

$$Y_{n-1}(q, \mathbf{z}, D) = (q - D)^+.$$

Here, the notation $x^+ = \max\{x, 0\}$. The quantity of outdated inventory is

$$S(q, \mathbf{z}, D) = (z_1 - (D - q - \sum_{i=2}^{n-1} z_i)^+)^+.$$

The dynamic programming formulation is as follows:

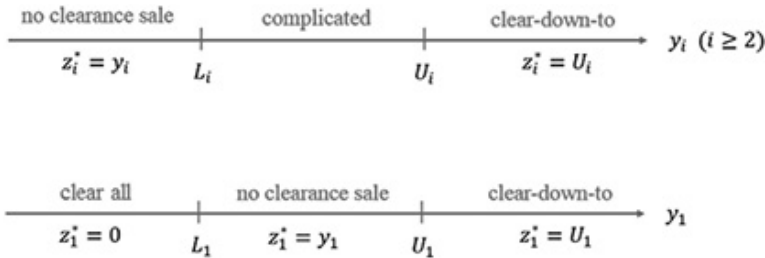
$$\pi_t(\mathbf{y}) = s \sum_{i=1}^{n-1} y_i + \max_{0 \leq \mathbf{z} \leq \mathbf{y}, q \geq 0} u_t(\mathbf{z}, q),$$

where

$$u_t(\mathbf{z}, q) = -(s + h) \sum_{i=1}^{n-1} z_i - \alpha c q + \alpha \mathbb{E} \left\{ r \min(q + \sum_{i=1}^{n-1} z_i, D) - \theta S(q, \mathbf{z}, D) + \pi_{t+1}(\mathbf{Y}(q, \mathbf{z}, D)) \right\}$$

The optimal clearance sale policies are shown in Figure 2.1. Under the LIFO rule, there are two thresholds for each age group of inventory: a lower and an upper threshold. For an age group with a remaining lifetime of two periods or more, if its inventory level is below the lower threshold, then there is no clearance sale; if it is above the upper threshold, then it will be cleared down to the upper threshold.

The optimal policy for the age group with a remaining lifetime of one period is very different, however. Clearance sales may take place if its inventory level is above the upper threshold or below the lower threshold. The lower the initial inventory level, the more new supply is needed to meet demand in the current period. However, the more new supply there is, the less likely it is that the oldest items will be used to meet demand because customers always select



Source: Li et al. (2016).

Figure 2.1 Optimal clearance sale policies for inventory with different remaining lifetimes under LIFO

the newest items first. The retailer is therefore better off clearing the small number of oldest items to recoup some revenue and avoid outdated. The practice of avoiding having the newest and the oldest items in the system at the same time through clearance sales is unique to the inventory systems under LIFO.

Implementing the model in practice requires solving a dynamic program with a multidimensional state space and a non-concave objective function, which is challenging. Motivated by the structural properties of the optimal policy, Li et al. (2016) consider two myopic heuristics. For both heuristics, the value-to-go function is approximated by $\pi_{t+1}(\mathbf{y}) = \frac{s + \alpha c - h}{2} \sum_{i=1}^{n-1} y_i$ because the marginal values of inventories are bounded between s and $\alpha c - h$. In the first heuristic, in computing the order quantity and clearance sale quantity, all inventories on hand are treated as if they would expire in one period. Let $y = \sum_{i=1}^{n-1} y_i$. The heuristic policies are then derived from the following one-period problem:

$$\begin{aligned} \max_{0 \leq z \leq y, q \geq 0} & -(s+h)z - \alpha c q + \alpha \mathbb{E}[r \min(q+z, D) - \theta(z - (D-q)^+)^+ \\ & + \frac{s + \alpha c - h}{2} (q - D)^+]. \end{aligned}$$

In the second heuristic, in addition to the total inventory level y , information about y_1 , which is the inventory level of items with a one-period lifetime remaining, is also needed. The heuristic policies are obtained by solving the following:

$$\max_{z_1, z, q} -(s+h)(z_1 + z) - \alpha c q + \alpha \mathbb{E}[r \min(q + z_1 + z, D) - \theta(z_1 - (D - q - z)^+)^+]$$

subject to the constraints: $0 \leq z_1 \leq y_1, 0 \leq z \leq y - y_1, q \geq 0$.

In numerical experiments, the first heuristic may generate as much as 7% less expected profit than the optimal policy. The second heuristic significantly outperforms the first heuristic, and its profit is consistently very close to the optimal profit.

The analysis by Li et al. (2016) demonstrates that inventory with a one-period remaining lifetime (i.e., the oldest inventory) plays a qualitatively different role than inventories of other age groups. First, the optimal order quantity is monotonic in the oldest inventory, but it is not necessarily so in other inventories. Second, the optimal policy on clearance sales with respect to the oldest inventory is to clear all, to not clear, and then to clear down to a certain level when the inventory level increases. The optimal policies with respect to other inventories, however, are different. In particular, clearance sales won't happen when inventories are low enough.

Finally, it is critically important to keep a record of the oldest inventory, and the performance of myopic heuristics that take advantage of that record are consistently close to that of the optimal policy. The value of keeping a record of other inventories, however, is insignificant. There is no age information in the bar codes currently used in retailing. In practice, retailers typically check and remove the expired items manually. Putting items on clearance sales is also done manually. The information about the oldest inventory can be obtained during these manual processes, and the additional effort may not be significant. If this is done, then the second heuristic above can be implemented without the need to include the full age information in bar codes.

2.3 MULTIPLE CLASSES OF DEMAND

Firms managing perishable inventory systems usually face multiple classes of customers that differ in their requirements for product freshness. For example, in health care, patients with different illnesses may require platelets of different ages. In grocery retailing, some customers may be more sensitive to product freshness than others. In these settings, in addition to the coordination of clearance sales and replenishment decisions discussed earlier, firms also need to determine the optimal allocation of perishable inventory to different classes of customers.

Abouee-Mehrzi et al. (2019) study the problem assuming that unmet demand is back-ordered and that there is a positive lead time for replenishment. Let L denote the lead time. A firm sells a perishable product with a lifetime of $n + L$ periods. Thus, the product has a remaining lifetime of n periods when the firm receives it. Without loss of generality, assume that there are n classes of customers indexed by $k = 1, \dots, n$, where class k customers only accept products with a remaining lifetime longer than or equal to k . Let D^k denote the class k demand. Let $\mathbf{D} = (D^1, \dots, D^n)$. Let b_k denote the per unit back-order penalty cost for class k demand. Assume that $b_n \geq b_{n-1} \geq \dots \geq b_1 > 0$; that is, class i has a higher priority than class j if $i > j$.

The sequence of events is the same as that in the models reviewed in Section 2.2. Here, the initial state is $\mathbf{y} = (\mathbf{y}_o, \mathbf{y}_p)$, where $\mathbf{y}_o = (y_1, y_2, \dots, y_n)$ and $\mathbf{y}_p = (y_{n+1}, y_{n+2}, \dots, y_{n+L-1})$. The variable y_i represents the inventory position for on-hand items with a remaining lifetime of i periods for $1 \leq i \leq n$; and y_{n+i} represents the pipeline inventory that will arrive in i periods for $1 \leq i \leq L-1$. The variable q is still used to represent the order quantity and $\mathbf{z} = (z_1, z_2, \dots, z_n)$ to denote the amount of on-hand inventory that is carried over to the next period. If $y_i \geq 0$, then $y_i - z_i$ denotes items with a remaining lifetime of i periods that are sold at clearance sales. If $y_i < 0$, that is, if there is unmet back-ordered demand, z_i must be equal to y_i .

Let d_j^k denote the amount of the class k demand that is met by using inventories with a remaining lifetime of j periods. Let $\mathbf{d}^k = (d_1^k, d_2^k, \dots, d_n^k)$ and $\mathbf{d} = (\mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^n)$. Denote $O(\mathbf{D}, \mathbf{z}) = \{\mathbf{d} : \sum_{j=k}^n d_j^k \leq D^k - (z_k)^-, \sum_{k=j}^n d_j^k \leq (z_j)^+, d_j^k \geq 0 \text{ for } 1 \leq j, k \leq n\}$. Here, $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$. Let r , s , c , h , and θ denote the regular price, clearance price, unit purchasing cost, unit holding cost, and outdating cost, respectively. The dynamic programming formulation is as follows:

$$\pi_t(\mathbf{y}) = s \sum_{i=1}^n y_i + \max_{\substack{(\mathbf{y}_o)^- \leq \mathbf{z} \leq \mathbf{y}_o \\ q \geq 0}} u_t(\mathbf{z}, \mathbf{y}_p, q),$$

where

$$\begin{aligned} u_t(\mathbf{z}, \mathbf{y}_p, q) = & -s \sum_{i=1}^n z_i - h \left(\sum_{i=1}^n z_i \right)^+ - \alpha c q \\ & + \alpha \mathbb{E} \max_{\mathbf{d} \in O(\mathbf{D}, \mathbf{z})} \left\{ r \left(\sum_{k=1}^n \sum_{j=1}^n d_j^k \right) - \sum_{k=1}^n b_k (D^k - (z_k)^- - \sum_{j=k}^n d_j^k)^+ \right. \\ & \left. - \theta (z_1 - d_1^1)^+ + \pi_{t+1}(z_2 - \sum_{k=1}^2 d_2^k, \dots, z_n - \sum_{k=1}^n d_n^k, \mathbf{y}_p, q) \right\}, \end{aligned}$$

and the terminal condition is $\pi_{T+L+1}(\mathbf{y}) = 0$.

It can be proved that the value function $\pi_t(\mathbf{y})$ and the objective function $u_t(\mathbf{z}, \mathbf{y}_p, q)$ are both anti-multimodular. The optimal clearance sales strategy follows a clear-down-to structure and the structural properties in Theorem 2.1 continue to hold in the presence of multiple classes of customers.

The optimal allocation policy is a sequential rationing policy. The demand with the highest priority is first satisfied, then the demand with the second-highest priority, and so on. In fulfilling each demand class, it is optimal to reserve fresher inventories at certain thresholds to meet future demand with a higher priority and use the remainder to fulfill the demand of that class as much as possible following the FIFO rule. The rationing threshold of inventory with a remaining lifetime i when fulfilling class k demand depends on the inventory levels both of products with a remaining lifetime longer than i and products with a remaining lifetime shorter than k . Anti-multimodularity also implies that these thresholds have bounded sensitivities with respect to the state variables.

Finally, through numerical studies, Abouee-Mehrzi et al. (2019) compare three strategies a firm can use to improve the management of perishables: decreasing the lead time, increasing the lifetime of products, and increasing customers' willingness to accept older products. They show that, among these three strategies, decreasing lead time is the most efficient, with a potential cost benefit of 20% in the numerical examples. Increasing the lifetime is more efficient than increasing customers' willingness to accept older products.

Chen et al. (2019a) study a similar problem but under the assumption that the unmet demand is lost and the replenishment lead time is zero. D^k is again used to represent the demand from class k customers who only accept products with a remaining lifetime of at least k periods. Let p_i denote the per unit lost-sales penalty cost for class i demand. Assume that $p_n \geq p_{n-1} \geq \dots \geq p_1 > 0$; that is, class i has a higher priority than class j if $i > j$.

The sequence of events and notation for the state and decision variables are the same as those in Li and Yu (2014). In what follows, q and z_n are used interchangeably. Let d_j^k denote the amount of the class k demand that is met by using inventories with a remaining lifetime of j periods. Let $\mathbf{d}^k = (d_1^k, d_2^k, \dots, d_n^k)$ and $\mathbf{d} = (\mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^n)$. Denote $O(\mathbf{D}) = \{\mathbf{d} : \sum_{j=k}^n d_j^k \leq D^k, \sum_{k=j}^n d_j^k \leq z_j, d_j^k \geq 0 \text{ for } 1 \leq j, k \leq n\}$. Let r, s, c, h , and θ denote the regular price, clearance price, unit purchasing cost, unit holding cost and outdating cost, respectively. The dynamic programming formulation is as follows:

$$\pi_t(\mathbf{y}) = s \sum_{i=1}^{n-1} y_i + \max_{\substack{0 \leq \mathbf{z} \leq \mathbf{y} \\ q \geq 0}} u_t(\mathbf{z}, q),$$

where

$$\begin{aligned} u_t(\mathbf{z}, q) = & -(s+h) \sum_{i=1}^{n-1} z_i - \alpha c q + \alpha \mathbb{E} \max_{\mathbf{d} \in O(\mathbf{D})} \left\{ r \left(\sum_{k=1}^n \sum_{j=1}^n d_j^k \right) - \theta (z_1 - d_1^1)^+ \right. \\ & \left. - \sum_{k=1}^n p_k (D^k - \sum_{j=k}^n d_j^k)^+ + \pi_{t+1}(z_2 - \sum_{k=1}^2 d_2^k, \dots, z_n - \sum_{k=1}^n d_n^k) \right\}, \end{aligned}$$

and the terminal condition is $\pi_{T+1}(\mathbf{y}) = 0$.

Chen et al. (2019a) prove that Theorem 2.1 still holds in this model. That is, the value function $\pi_t(\mathbf{y})$ and the objective function $u_t(\mathbf{z}, \mathbf{y}_p, q)$ are both anti-multimodular; the optimal strategy on clearance sales follows a clear-down-to structure and the optimal order quantity is decreasing with inventory level and is more sensitive to changes in newer inventory. The optimal allocation policy is simpler than that in Abouee-Mehrizi et al. (2019) due to the assumption of zero lead time. In particular, demand with a higher priority should be satisfied as much as possible and on a FIFO basis before demand with a lower priority.

Based on the structural properties of the optimal value function, Chen et al. (2019a) then develop an adaptive approximation approach to overcome the curse of dimensionality in solving the dynamic program. The essential idea is to approximate the value function by a linear combination of a one-dimensional function $B_j(x)$, i.e., letting the approximate value function be $\hat{\pi}_{t+1}(\mathbf{y}) = \sum_{j=1}^{n-1} \eta_{t+1,j} B_{t+1,n-1} \left(\sum_{l=j}^{n-1} y_l \right)$. Here, for each $j \in \{1, 2, \dots, n-1\}$, $B_{t,j}(y)$ can be recursively solved by the following one-dimensional dynamic program:

$$B_{t,j}(y) = sy + \max_{\substack{\mathbf{0} \leq \mathbf{z} \leq \mathbf{y} \mathbf{e}_j \\ q \geq 0}} u_t(\mathbf{z}, q),$$

where

$$u_t(\mathbf{z}, q) = -(s+h) \sum_{i=1}^{n-1} z_i - \alpha c q + \alpha \mathbb{E} \max_{\mathbf{d} \in O(\mathbf{D})} \left\{ r \left(\sum_{k=1}^n \sum_{j=1}^n d_j^k \right) - \theta (z_1 - d_1^1)^+ - \sum_{k=1}^n p_k \left(D^k - \sum_{j=k}^n d_j^k \right)^+ + \hat{\pi}_{t+1} \left(z_2 - \sum_{k=1}^2 d_2^k, \dots, z_n - \sum_{k=1}^n d_n^k \right) \right\}.$$

The weights $\eta_{t+1,j}$ for $j \in \{1, 2, \dots, n-1\}$ are calculated based on the relationship between the marginal values of $B_{t,j}$ and $B_{t,n-1}$. Constructed in this way, $\hat{\pi}_t(\mathbf{y})$ retains the anti-multimodularity of the optimal value function. Fresher inventory has a higher marginal value under this approximate value function.

Under this approximation scheme, the heuristic policy generated by the approximate value function retains the same structural properties as the optimal policy. Numerical studies demonstrate that the proposed approximation approach is nearly optimal, with a tan average optimality gap of 0.30%, and significantly outperforms the other heuristics in the literature.

Chen et al. (2019b) apply some of the ideas in Chen et al. (2019a) to a setting in which a blood center faces two classes of age-differentiated demand for platelets from downstream hospitals. Using the tool of multimodularity, they characterize the structure of the optimal policy for whole blood collection, platelet production, and inventory issuing, rationing, and disposal. Fu et al. (2019) also study perishable inventory systems with multiple classes of demand. However, as they consider product returns and remanufacturing, their model and results are reviewed in Chapter 11 of this book.

2.4 MULTIPLE LOCATIONS

Research on non-perishable inventory systems shows that transshipment can balance inventories in different locations and hence simultaneously reduce overages at some locations and shortages at others. This section introduces two papers that provide new insights into the roles and value of transshipment in perishable inventory systems.

Li et al. (2021) explore the idea of transshipment in an offline retailer with a LIFO inventory issuing rule. The retailer owns two outlets, indexed by superscript $i = 1, 2$. The products they sell have an n -period lifetime. The products can be sold at either a regular price p , or a clearance sale price s . Under a regular price, the demand at each outlet is random and is modeled by random variable D^i . The demand under a clearance sale is sufficiently high (or s is sufficiently low) that the products on sale will never go unsold. Assume that D^1 and D^2 are identically distributed but not necessarily independent.

The sequence of events is as follows. 1) At the beginning of a period, the retailer determines how much to order, how much should be sold in a clearance sale and how much and what should be transshipped from one outlet to the other. 2) The random demand for regular sales is realized and satisfied. 3) At the end of the period, the unsold inventory with a remaining lifetime of one period expires. Assume that there is no transshipment cost in the model.

For outlet i , the initial inventory is represented by a vector $\mathbf{x}^i = (x_1^i, x_2^i, \dots, x_{n-1}^i)$, where x_j^i represents the inventory with a remaining lifetime of j periods at outlet i . Let $x_j = x_j^1 + x_j^2$. The system state is captured by $\mathbf{x} = (x_1, x_2, \dots, x_{n-1})$. Let q^i be the order quantity of new items at outlet i . Let $\mathbf{z}^i = (z_1^i, \dots, z_{n-1}^i)$, where z_j^i is the inventory with a remaining lifetime of j periods that retail outlet i has after transshipment and clearance sales. As such, the total amount of inventory with a remaining lifetime of j periods available for regular sale is $z_j^1 + z_j^2$ and the amount sold in clearance sales is $x_j - z_j^1 - z_j^2$. Customers will always choose the freshest products first; that is, inventory leaves the retail shelf on a LIFO basis. Suppose that the system state becomes $\mathbf{Y}^i(q^i, \mathbf{z}^i, D^i) = (Y_1^i, Y_2^i, \dots, Y_{n-1}^i)$ in the next period. Then, for $1 \leq j \leq n-2$

$$Y_j^i(q^i, \mathbf{z}^i, D^i) = (z_{j+1}^i - (D^i - q^i - \sum_{k=j+2}^{n-1} z_k^i)^+)^+$$

and

$$Y_{n-1}^i(q^i, \mathbf{z}^i, D^i) = (q^i - D^i)^+.$$

The amount of outdated inventory is

$$S(q^i, \mathbf{z}^i, D^i) = (z_1^i - (D^i - q^i - \sum_{j=2}^{n-1} z_j^i)^+)^+.$$

Let c , θ , and α be the ordering cost, outdated cost, and the discounting factor, respectively. Without loss of generality, assume that there is no holding cost. The dynamic programming formulation is then as follows:

$$J_t(\mathbf{z}^i, q^i) = -s \sum_{j=1}^{n-1} z_j^i - cq^i + p \mathbb{E} \min(q^i + \sum_{j=1}^{n-1} z_j^i, D^i) - \theta \mathbb{E} S(q^i, \mathbf{z}^i, D^i) \quad (2.1)$$

and

$$v_t(\mathbf{x}) = s \sum_{j=1}^{n-1} x_j + \max\{J_t(\mathbf{z}^1, q^1) + J_t(\mathbf{z}^2, q^2) + \alpha \mathbb{E} v_{t+1}(\sum_{i=1}^2 \mathbf{Y}^i(q^i, \mathbf{z}^i, D^i))\}, \quad (2.2)$$

subject to $z_j^1 + z_j^2 \leq x_j$, $z_j^i \geq 0$, $q^i \geq 0$ for all $i = 1, 2$ and $j = 1, 2, \dots, n-1$. On the right-hand side of Equation (2.1), the second term is the purchasing cost, the third term is the revenue from regular sales, and the last term is the outdating cost. The sum of the first terms on the right-hand sides of Equations (2.1) and (2.2) represents the revenue from clearance sales. Hence $J_t(\mathbf{z}^i, q^i)$ is the one-period profit generated at outlet i . The planning horizon is T and the terminal condition is $v_{T+1}(\mathbf{x}) = s \sum_{i=1}^{n-1} x_i$. The optimal solution to Equation (2.2) is denoted by $(\bar{z}_j^i, \bar{q}^i)$, $j = 1, 2, \dots, n-1$ and $i = 1, 2$.

The following theorem shows that transshipment plays two roles for perishable inventory systems under the LIFO rule. One is inventory balancing, which is well known in the literature. The other is inventory separation, which is new to the literature. In Theorem 2.2, we assume that the demand distribution is PF_2 , which is a common assumption in the inventory literature (e.g., Porteus, 2002; Huggins & Olsen, 2010; Li & Yu, 2012). The class of PF_2 distributions includes many commonly used distributions such as the exponential, the uniform, the Erlang, the normal, and convolutions of such distributions.

Theorem 2.2 Suppose that D^i has a PF_2 distribution. If $\bar{z}_i^1 = \bar{z}_i^2$ for $2 \leq i \leq n-1$, then there is an optimal policy such that at least one of $\bar{z}_1^1$, $\bar{z}_1^2$, $\bar{q}^1$ and $\bar{q}^2$ is zero.

Theorem 2.2 includes two special cases. The first case is when $x_i = 0$ for all $i = 2, \dots, n-1$, and the second is when the lifetime is $n = 2$. In both cases, the condition $\bar{z}_i^1 = \bar{z}_i^2$ for $2 \leq i \leq n-1$ is obviously satisfied. Transshipment allows the retailer to send the oldest inventory to one outlet and the newest inventory to the other (i.e., separation of inventories), and to send inventory from the outlet with excess inventory to the outlet with a shortage (i.e., balance of inventories). Inventory separation occurs when one of $\bar{z}_1^1$, $\bar{z}_1^2$, $\bar{q}^1$ or $\bar{q}^2$ is zero, or when two are zero and one outlet holds only the oldest inventory and the other holds only the newest inventory.

To understand how inventories should be separated and how much benefit transshipment can generate, Li et al. (2021) consider an approximation. Under the approximation, the computation of the optimal policy relies on only two pieces of information, namely, the number of items expiring in one period (old inventory) x_1 and the number of remaining items (new inventory), denoted by $x_{[2]} = \sum_{j=2}^{n-1} x_j$. The profit-to-go is approximated by a linear function. That is, in period t , let $v_{t+1}(\mathbf{x}) = v \sum_{j=1}^{n-1} x_j$, where v is a number bounded by c and s because the marginal value of inventory is bounded by c and s .

Let z_1^i and $z_{[2]}^i$ represent the amount of old inventory and new inventory, respectively, allocated to outlet i for regular sale. Let y^i be the amount of new inventory after ordering at outlet i , that is, the order-up-to level for new inventory at outlet i . Let

$$J(z_1, y) = -sz_1 - cy + p\mathbb{E}\min(D, z_1 + y) \\ - \theta\mathbb{E}(z_1 - (D - y)^+)^+ + \alpha v\mathbb{E}(y - D)^+,$$

and, to find a heuristic policy, solve the following one-period optimization problem:

$$\max\{(c - s)(z_{[2]}^1 + z_{[2]}^2) + J(z_1^1, y^1) + J(z_1^2, y^2)\}$$

subject to $z_1^1 + z_1^2 \leq x_1$, $z_{[2]}^1 + z_{[2]}^2 \leq x_{[2]}$, $z_1^i \geq 0$, $z_{[2]}^i \geq 0$, $y^i \geq z_{[2]}^i$ for $i = 1, 2$.

Li et al. (2021) then provide a theoretical bound for the gap between the performance of the approximation and the optimal profit. When the demand is compound Poisson, the bound approaches zero as the arrival rate approaches infinity. The optimal policy under the approximation is characterized by two increasing switching curves that divide the entire state space into three regions. In the first region, only one outlet holds old items but both hold new items. In the second, one outlet holds only old items and the other holds only new inventory. In the third, only one outlet holds new items while both hold old items.

Numerical studies show that transshipment and clearance sales are substitutes in terms of both increasing profit and reducing waste. Transshipment can increase profit by as much as several percentage points. It is most valuable in increasing profit when the variable cost of products is high, the outdated cost is high, the clearance sale price is low or the demand variability is high.

Zhang et al. (2022) study the transshipment of perishable inventory under a FIFO rule and exogenous base-stock levels (the model is extended to allow general ordering policies in Zhang et al., 2022 and this is reviewed in Chapter 19 of this book). Their research is motivated by a platelet (a blood product with a shelf-life of three days) inventory management problem in two hospitals that belong to the same integrated healthcare system.

In the model, there are two locations indexed by superscripts $i = 1, 2$. The product has a lifetime of n periods. The sequence of events is as follows. 1) At the beginning of a generic period, the initial inventory at location i is $\mathbf{x}^i = (x_1^i, \dots, x_{n-1}^i)$, where x_j^i is the inventory level of products of age j . The base-stock level at location i is S^i , so $x_0^i = (S^i - \sum_{j=1}^{n-1} x_j^i)^+$ items of fresh products are ordered. 2) The random demand D^i at each location is realized and satisfied. 3) Products are transshipped from one location to the other in a FIFO manner; that is, older items are shipped first. Let u denote the total items transshipped from location 1 to location 2. A negative u implies transshipment from location 2 to location 1. Then, $-D^1 \leq u \leq D^2$. 4) After transshipment, the products at each location are issued to satisfy demand in a FIFO manner, and unmet demand is lost. 5) At the end of each period, products reaching age n are disposed of.

Let $\mathbf{X}^i = (X_1^i, \dots, X_{n-1}^i)$ denote the initial inventory at location i at the beginning of the next period. Then,

$$X_j^i = (x_{j-1}^i - (D^i + u(-1)^{i+1} - \sum_{k=j}^{n-1} x_k^i)^+)^+.$$

Let p^i , h^i , and θ^i denote the unit shortage, holding and outdated cost, respectively, at location i . Denote r^i as the unit transshipment cost from location i to the other location. Without loss

of generality, assume that the unit ordering cost at each location is zero. Also, assume that the system starts with zero inventory. The one-period cost function $L(\mathbf{x}^1, \mathbf{x}^2, u)$ is then given by:

$$\sum_{i=1}^2 [p^i (D^i + u(-1)^{i+1} - S^i)^+ + h^i (S^i - D^i + u(-1)^i)^+ + \theta^i (x_{n-1}^i - D^i + u(-1)^i)^+ + r^i (u(-1)^{i+1})^+].$$

Let $C_t(\mathbf{x}^1, \mathbf{x}^2)$ denote the optimal expected cost-to-go function at period t . The optimality equation is then defined as:

$$C_t(\mathbf{x}^1, \mathbf{x}^2) = \mathbb{E} \left[\min_{-D^1 \leq u \leq D^2} L(\mathbf{x}^1, \mathbf{x}^2, u) + \alpha C_{t+1}(\mathbf{X}^1, \mathbf{X}^2) \right],$$

where α is the discount factor. The terminal condition is $C_{T+1}(\mathbf{x}^1, \mathbf{x}^2) = 0$.

Zhang et al. (2022) first provide a partial characterization of the direction of optimal transshipment. In the case of non-perishable inventory, the direction is determined by whether a location experiences a surplus or a shortage under mild cost conditions. However, for the case of perishable inventory, they show that an important additional factor is the quantity of the oldest inventory x_{n-1}^i , because of the outdating cost.

When the outdating cost is sufficiently small compared with the unit transshipment cost, the optimal transshipment policy for the perishable case is the same as that for the non-perishable case. The details are given in Theorem 2.3, where $-i$ denotes the location other than i and u^* and u^{N*} denote the optimal transshipment quantity for the perishable and non-perishable cases, respectively.

Theorem 2.3

$$(i) \ u^{N*} = \begin{cases} \min\{(S^1 - D^1)^+, (D^2 - S^2)^+\} & \text{if } S^1 \geq D^1, D^2 \geq S^2 \\ -\min\{(D^1 - S^1)^+, (S^2 - D^2)^+\} & \text{if } D^1 \geq S^1, S^2 \geq D^2 \\ 0 & \text{otherwise.} \end{cases}$$

$$(ii) \ |u^*| \geq |u^{N*}|.$$

$$(iii) \ \text{If } \theta^i \leq r^i - h^i + h^{-i} \text{ for } i = 1, 2, \text{ then } u^* = u^{N*}.$$

Theorem 2.3 shows that, in general, the optimal transshipment quantity for the non-perishable case provides a lower bound than that for the perishable case. Zhang et al. (2022) further present an example to show that this lower bound is non-tight. The implication of these results is that when managing perishable inventory, one should expect transshipments to occur more often or in larger quantities than for non-perishable inventory, because in the perishable case, transshipments are valuable not only for reducing shortages but also for balancing the age of products at different locations, thus reducing outdating.

They then investigate how the optimal transshipment quantity changes with the inventory level at each location. For a special case with a two-period product lifetime, they prove that the optimal cost function is $L^{\natural}$ -convex, which implies that the optimal transshipment quantity is monotonic in the inventory level at each location. They further show via a counterexample

that the property of $L^\natural$ -convexity, however, does not hold in the general case of longer product lifetimes.

These findings motivate Zhang et al. (2022) to develop a simple transshipment policy that satisfies the monotonicity property. Under this policy, transshipment is triggered when there is either a shortage or immediate outdating. In this case, only the oldest products at each location are transshipped unless there is a shortage at the other location. They then derive approximations of the expected cost functions, which they then use to compute the base-stock levels for both locations. Using real-life data from platelet inventory management in hospitals, they show that the proposed policy performs well and significantly reduces the total cost compared with benchmark policies.

Finally, through numerical studies, Zhang et al. (2022) show that the value of inventory sharing for perishable products is typically higher than for non-perishable products. Interestingly, unlike for non-perishable products, the value of inventory sharing for perishable products can be strictly positive and substantial, even when demand at one location is deterministic, because old perishable products in a location with random demand can be transshipped to a location with deterministic demand to reduce outdates. The implication of this result is that when products are perishable, transshipment should be considered even though the results in the non-perishable inventory literature suggest that it has little or no value.

2.5 CONTROL OF LIFETIMES

Firms can control the lifetimes of inventories when they enter the inventory system. For example, retailers of perishable goods are often faced with a choice between more expensive packaging that can extend the shelf-life of their products and less expensive packaging that cannot. Other examples include situations in which firms can buy from multiple sources with different product lifetimes and costs.

Li et al. (2017) model a retailer who must determine in each period the optimal order quantities and types of packaging. There are two types of packaging, which they call “regular” and “active”. Items in a regular package will perish in one period and the variable cost, which includes the purchasing and packaging costs, is c_1 ; items in an active package have a two-period lifetime and the variable cost is c_2 , which is higher than c_1 . Items that perish carry an outdating cost m per unit.

The total demand in each period is independently and identically distributed. Let D be the random demand in a period. It is assumed that customers always prefer an item with a longer remaining lifetime to an item with a shorter remaining lifetime. That is, inventory is depleted on a LIFO basis. When items with a remaining lifetime of two periods are out of stock, a random percentage β of customers are willing to purchase items with a remaining lifetime of one period and the remainder will walk away. Any unmet demand is lost and the penalty cost of not meeting demand is p per unit. The objective is to determine the quantity of items in the two types of packaging in each period that minimizes the total expected cost.

Let q be the quantity of items in active packaging with a two-period remaining lifetime, y be the total number of items with a one-period remaining lifetime, which includes the initial inventory at the beginning of each period and items just ordered but in regular packaging, and $V_t(x)$ be the minimum total expected cost from period t to the end of horizon T , when the

initial inventory is x . The costs incurred in future periods are discounted by a discount rate α . Then,

$$V_t(x) = \min_{y \geq x, q \geq 0} J_t(y, q) - c_1 x,$$

where

$$\begin{aligned} J_t(y, q) = & c_1 y + c_2 q + p \mathbb{E}g(y, q, \beta, D) \\ & + m \mathbb{E}S(y, q, \beta, D) + \alpha \mathbb{E}V_{t+1}(Y(y, q, \beta, D)). \end{aligned}$$

The amount of unmet demand, $g(y, q, \beta, D)$, the initial inventory level in the next period $Y(y, q, \beta, D)$, and the amount of outdating in the current period $S(y, q, \beta, D)$ are given by:

$$g(y, q, \beta, D) = (1 - \beta[D - q]^+ + [\beta D - \beta q - y]^+,$$

$$Y(y, q, \beta, D) = (q - D)^+,$$

$$S(y, q, \beta, D) = [y - \beta(D - q)^+]^+.$$

Here, $[x]^+ = \max\{x, 0\}$. In the above expressions, $[\beta D - \beta q - y]^+$ and $(1 - \beta)[D - q]^+$ represent the amount of unmet demand due to the stockout of items with a one-period and two-period remaining lifetime, respectively. The terminal condition is $V_{T+1}(x) = -c_1 x$, which means that any unused inventory at the end of the horizon can be salvaged at a cost of c_1 per unit.

Li et al. (2017) consider two cases that differ depending on the source of uncertainty. In the first case, the total demand D in each period is random but the proportion of customers willing to accept less fresh items β is not. In this case, they show that if the proportion β is high enough, as the initial inventory level increases, the optimal policy changes from using active packaging only to using regular packaging only and finally to ordering nothing.

Note that the retailer here either uses active packaging or regular packaging, but never both at the same time. In deciding on its choice of packaging, the retailer must consider two critical factors. The first is the need to fulfill the demand in the current period, and the second is the likelihood of items with a two-period lifetime being carried over to the next period. Which packaging the retailer should use then depends on the incremental cost. The incremental cost of using active packaging decreases with the quantity of items in active packaging, and is lower than the cost of using regular packaging if and only if the quantity of items in active packaging is sufficiently large that the items in active packaging are highly likely to be carried over to the next period. When the initial inventory x is small, a large amount of extra supply is needed to fulfill the demand in the current period. The retailer should then use active packaging only because then the chance of items with a two-period lifetime being carried over to the next period for such a large order is high. As x increases, the extra supply needed to fulfill the current demand decreases. When x is sufficiently high, the retailer will switch to using regular packaging only. Using active packaging but ordering only a small number of items is suboptimal because the likelihood of a small number of items with a two-period lifetime being carried over to the next period is low.

In the second case, the total demand D in each period is known with certainty but the proportion of customers willing to accept less fresh items β is random. In this case, the optimal policy is to use either active packaging only or regular packaging only, depending on the cost parameters but independent of the initial inventory.

The analysis shows that regardless of the source of demand uncertainty, the optimal policy structure exhibits the same pattern of “separation”; that is, never use both types of packaging in the same period. This phenomenon is specific to the LIFO issuing rule and cannot happen under the FIFO issuing rule. The trade-off under FIFO is different. It is worth using active packaging only when there are enough items with a one-period remaining lifetime for there to be a high probability that items with a two-period remaining lifetime will be carried over to the next period. Items with a one-period remaining lifetime can either come from the initial inventory or from a new order in regular packaging in the current period. In other words, the retailer may use both regular and active packaging at the same time under FIFO. The separation phenomenon is reminiscent of the optimal clearance sales policies in Li et al. (2016) and the optimal transshipment policies in Li et al. (2021).

From a practical standpoint, Li et al. (2017) highlight the significance of coordinating inventory decisions and packaging decisions in grocery retailing. In practice, retailers appear to focus on the choice between using only regular packaging and using only active packaging. Some retailers decide to stay with regular packaging because the additional packaging cost does not justify the benefit. Li et al. (2017) argue that retailers should consider the optimal policy, which in general only requires the partial adoption of active packaging and has a lower packaging cost than the policy of active packaging only. Retailers will find it easier to justify the additional cost if they implement the optimal policy. However, from the perspective of waste reduction, Li et al. (2017) show through numerical studies that the optimal policy is almost as good as the policy of using only active packaging. These findings are useful for retail practice.

While Li et al. (2017) focus on grocery retailing, the study by Zhou et al. (2011) is motivated by hospitals’ practice of placing expedited orders for platelet inventory in addition to regular replenishments to fulfill demand. They model this problem as a perishable inventory system with dual sourcing. The platelets have a lifetime of three periods. The interval of regular orders is two periods, which is called a cycle. At the beginning of each cycle, the hospital determines the regular order quantity Q and the order-up-to level s for expedited orders in the second period of the cycle. In the analytical model, it is assumed that expedited platelets have a lifetime of two periods. All replenishments have zero lead times. Therefore, all platelets in the second period of a cycle are of the same age. These assumptions effectively reduce the dimension of the state space of the dynamic program to one.

Let x denote the inventory level at the beginning of cycle t . Let D^i denote the demand in period i within cycle t , where $i \in \{1, 2\}$. Assume that unmet demand is lost and that the issuing rule is FIFO. The amount of expedited units can then be expressed as $Q_e = (s - (Q - (D^1 - x)^+))^+$. The amount of outdated inventory is given by $O = (x - D^1)^+$. The amount of inventory at the beginning of the second period after expediting is $\tilde{X} = \max\{s, (Q - (D^1 - x)^+)^+\}$. Thus, the state at the beginning of the next cycle is $X = (\tilde{X} - D^2)^+$. The total shortage within a cycle is given by $L = (D^1 - x - Q)^+ + (D^2 - \tilde{X})$.

Without loss of generality, the regular ordering cost is assumed to be zero. Let c_e , p , and θ denote the expedited unit ordering cost, unit shortage cost and outdating cost, respectively.

Let $V_t(x)$ denote the optimal cost from cycle t to the end of planning horizon T . The dynamic program is then given by

$$V_t(x) = \min_{Q,s} c_e \mathbb{E}[Q_e] + \theta \mathbb{E}[O] + p \mathbb{E}[L] + \mathbb{E}[V_{t+1}(X)].$$

The terminal condition is given by $V_{T+1}(x) = \theta \mathbb{E}(x - D^1)^+ + p \mathbb{E}(D^1 - x)^+$.

Zhou et al. (2011) then show that when solving the dynamic program backward, the optimal solutions Q^* and s^* are uniquely determined by the first-order conditions of the objective function with respect to Q and s .

Using real-life data for platelets, Zhou et al. (2011) then numerically investigate how the optimal cost and optimal decisions vary with the model parameters. In simulation studies, they also incorporate lead times and variable product lifetimes for expedited orders. The numerical results show that the optimal cost is significantly affected by demand uncertainty, lead times, seasonality and the age of expedited orders. The optimal decisions are significantly affected by a change in expected demand but not by a change in demand variance. Furthermore, the expedited order-up-to level is relatively unchanged with respect to demand uncertainty, lead times, seasonality and age. The numerical results also imply that for small hospitals with low average demand but high demand uncertainty, the (Q, s) policy is better than the Q policy where regular orders are placed every period; for large hospitals with low demand uncertainty, the Q policy would be preferred.

In a more recent paper, Chen et al. (2020) extend the model in Zhou et al. (2011) by allowing for both returns and platelet refills during the regular ordering cycle. All of the papers reviewed in this section impose strong assumptions on lifetimes. What is the form of optimal policies when there are two sources of supply with different costs and lifetimes and lifetimes are general finite numbers? This appears to be an open question.

2.6 EMPIRICAL RESEARCH

Studies in the literature on perishable inventory control focus on pricing and inventory policies under certain assumptions about consumer behavior and suppliers. Some interesting empirical studies, albeit with different foci, can inform inventory research.

The study by Tsiros and Heilman (2005) examines consumers' behavior with respect to expiration dates for perishable grocery products. In particular, they show that consumers' willingness to pay and their frequency of checking expiration dates depend on their perceived risk associated with expiration, which varies from product to product, their consumption rates, and their ability to take measures to stop or slow the aging process of perishable products. These findings confirm that to bring perishable inventory models closer to current practice, it is necessary to model multiple classes of customers who may have different minimum acceptable remaining product lifetimes and may use different issuing rules.

While Tsiros and Heilman (2005) focus on consumer behavior, Akkas et al. (2019) focus on the supply side. They find that the main sources of product expiration in retail stores are large case sizes relative to daily consumer demand, long lead times, minimum order rules, replenishment workload, and manufacturers' incentive programs for the sales force. These findings are useful for managers developing targeted product design and information and incentives

design initiatives to reduce waste. For inventory researchers, these findings show that there are opportunities to investigate ideas for managing perishable inventory that involve the whole supply chain, as opposed to only the retailer.

2.7 FUTURE RESEARCH

In the literature, either FIFO or LIFO rule is assumed. The assumption behind the LIFO rule is that consumers are infinitely rational, therefore, in traditional bricks-and-mortar stores where consumers decide which items to pick, the LIFO rule is appropriate. However, the reality is more complex. In bricks-and-mortar stores, older items are usually placed in more convenient reach of customers on the shelves and picking the newest items requires additional effort. While some customers may be willing to make that effort, others may settle for items that are less fresh. One way to capture some of this complexity is to have multiple classes of consumers. For example, one class of consumers chooses items on a LIFO basis, whereas other classes use the FIFO rule but will not select items unless their remaining lifetimes are sufficiently long.

In e-commerce, where retailers control inventory issuance, the FIFO rule is usually assumed because it minimizes outdated. However, it may be suboptimal for retailers if consumer welfare, which is usually important to retailers, is sensitive to the remaining lifetimes of products. In a fuller model that captures this additional issue, the retailer should jointly determine ordering and issuing policies such that the utility, which includes revenue, inventory-related costs, and the impact on consumer welfare, is maximized.

One important insight contained in Operations Management textbooks is that consolidating multiple retail outlets into one can reduce the mismatch between supply and demand if demand is not affected by the consolidation, and this is beneficial to retailers. Is this result still true when consumers choose products on a LIFO basis, which is typically the case in physical stores? The foregoing discussion suggests that it may not be true in general. Having all inventories in one location means that consumers will not choose older items unless newer items are sold out. However, if inventories are placed in multiple locations, older items may be sold in some locations if newer items are sold out, even if there might still be newer items in other locations. When consolidation benefits retailers is an interesting and practical question for future research.

REFERENCES

- Abouee-Mehrzi, H., Baron, O., Berman, O., & Chen, D. (2019). Managing perishable inventory systems with multiple priority classes. *Production and Operations Management*, 28(9), 2305–2322.
- Akkas, A., Gaur, V., & Simchi-Levi, D. (2019). Drivers of product expiration in consumer packaged goods retailing. *Management Science*, 65(5), 2179–2195.
- Chao, X., Gong, X., Shi, C., Yang, C., Zhang, H., & Zhou, S. X. (2018). Approximation algorithms for capacitated perishable inventory systems with positive lead times. *Management Science*, 64(11), 5038–5061.
- Chao, X., Gong, X., Shi, C., & Zhang, H. (2015). Approximation algorithms for perishable inventory systems. *Operations Research*, 63(3), 585–601.
- Chen, K., Song, J. S., Shang, J., & Xiao, T. (2020). Managing hospital platelet inventory with mid? Cycle expedited replenishments and returns. *Production and Operations Management*, 31(5), 2015–2037.

- Chen, S., Li, Y., Yang, Y., & Zhou, W. (2019a). Managing perishable inventory systems with age-differentiated demand. Working paper, City University of Hong Kong.
- Chen, S., Li, Y., & Zhou, W. (2019b). Joint decisions for blood collection and platelet inventory control. *Production and Operations Management*, 28(7), 1674–1691.
- Chen, X., Pang, Z., & Pan, L. (2014). Coordinating inventory control and pricing strategies for perishable products. *Operations Research*, 62(2), 284–300.
- Fries, B. (1975). Optimal ordering policy for a perishable commodity with fixed lifetime. *Operations Research*, 23(1), 46–61.
- Fu, K., Gong, X., & Liang, G. (2019). Managing perishable inventory systems with product returns and remanufacturing. *Production and Operations Management*, 28(6), 1366–1386.
- Hu, P., Shum, S., & Yu, M. (2016). Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Operations Research*, 64(1), 118–134.
- Huggins, E. L., & Olsen, T. L. (2010). Inventory control with generalized expediting. *Operations Research*, 58(5), 1414–1426.
- Karaesmen, I. Z., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. In K. Kempf, P. Keskinocak, & R. Uzsoy (Eds.), *Planning production and inventories in the extended enterprise*. Springer.
- Li, Q., & Yu, P. (2012). On the quasiconcavity of lost-sales inventory models with fixed costs. *Operations Research*, 60(2), 286–291.
- Li, Q., & Yu, P. (2014). Multimodularity and its applications in three stochastic dynamic inventory problems. *Manufacturing and Service Operations Management*, 16(3), 455–463.
- Li, Q., Yu, P., & Du, L. (2021). Separation of perishable inventories in offline retailing through transshipment. *Operations Research*, Forthcoming.
- Li, Q., Yu, P., & Wu, X. (2016). Managing perishable inventories in retailing: Replenishment, clearance sales, and segregation. *Operations Research*, 64(5), 1270–1284.
- Li, Q., Yu, P., & Wu, X. (2017). Shelf life extending packaging, inventory control, and grocery retailing. *Production and Operations Management*, 26(7), 1369–1382.
- Liu, Y., Feng, Y., & Lai, G. (2019). Analysis of optimal inventory management for dual blood products with dual sources. Tech. rep., Working Paper, available at SSRN 3323514.
- Lu, Y., & Song, J. (2005). Order-based cost optimization in assemble-to-order systems. *Operations Research*, 53(1), 151–169.
- Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1), 157–180.
- Murota, K. (2003). *Discrete convex analysis*. SIAM Monographs on Discrete Mathematics and Applications (Society for Industrial and Applied Mathematics, Philadelphia).
- Nahmias, S. (1975). Optimal ordering policies for perishable inventory-ii. *Operations Research*, 23(4), 735–749.
- Nahmias, S. (1982). Perishable inventory theory: A review. *Operations Research*, 30(4), 680–707.
- Nahmias, S. (2011). *Perishable inventory systems*. Springer.
- Nandakumar, P., & Morton, T. E. (1993). Near myopic heuristics for the fixed-life perishability problem. *Management Science*, 39(12), 1490–1498.
- Porteus, E. L. (2002). *Foundations of stochastic inventory theory*. Stanford University Press.
- Prastacos, G. P. (1982). Blood inventory management: An overview of theory and practice. *Management Science*, 30(7), 777–800.
- Tsiros, M., & Heilman, C. M. (2005). The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. *Journal of Marketing*, 69(2), 114–129.
- Zhang, C., Ayer, T., White, C. C., Bodeker, J. N., & Roback, J. D. (2022). Inventory sharing for perishable products: Application to platelet inventory management in hospital blood banks. *Operations Research*, Forthcoming.
- Zhou, D., Leung, L. C., & Pierskalla, W. P. (2011). Inventory management of platelets in hospitals: Optimal inventory policy for perishable products with emergency replenishments. *Manufacturing and Service Operations Management*, 13(4), 420–438.
- Zipkin, P. (2008a). On the structure of lost-sales inventory models. *Operations Research*, 56(4), 937–944.

3. Capacitated inventory systems

Roman Kapuściński and Rodney P. Parker

3.1 WHAT IS CAPACITY AND WHY IS IT IMPORTANT FOR INVENTORIES?

The phenomenon of capacity limits is universal, whether considering service or manufacturing organizations. Often, if the available capacity is ample, models will approximate the system by omitting these capacity limits. However, there is ample evidence that a system modeled without capacity limits can often be a poor proxy for the system subject to these limits. Broadly speaking, service systems address demand in real time, in a make-to-order fashion without the facility of inventory. In these settings, the effect of capacity is somewhat transparent: it merely acts as a limit on the units of demand that may be processed in a given time. However, in manufacturing contexts, the role of capacity is more subtle since the capacity limits may be augmented with the presence of inventory. Thus, the study of inventory combined with capacity limits is of profound interest.

Several empirical papers look at the role of capacity. A prominent stream involves papers that consider to what degree capacity decreases responsiveness to demand and effectively smooths the demand (e.g., Fair, 1989; Krane & Braun, 1991).

3.1.1 How Capacity Limits Affect Systems

The range of situations where capacity is crucial is very wide. Some of them may be treated by isolating individual factors, such as one stage, or one product. In other situations, the factors are more complicated or more difficult to decouple. Trying to address the need for some answers, the operations management literature has developed a substantive analysis of single-stage systems, extending the base situation into multiple directions and focusing both on the structure of an optimal solution, as well as methods to compute the policy. For more complicated multi-stage systems the progress is fairly recent and similarly includes both structural and algorithmic developments. In some cases, the same practical problems were addressed using different alternative methodologies, say dynamic programming and queueing models, in other situations, the reality was either analyzed as random capacity or approximated with a fixed capacity. The major difficulty in analyzing capacitated systems with inventory is the lack of easy renewal – after experiencing a substantial demand, the system may take several periods to recover to the ideal state.

We attempt to follow the structure of the research publications and focus first on single-stage systems, considering a base model, extensions, and variants, before describing serial and assembly multi-stage systems. When discussing these systems, when appropriate, we distinguish between stationary and non-stationary, finite- and infinite-horizon, and optimal and approximate analysis.

3.2 SINGLE-INSTALLATION CAPACITATED INVENTORY SYSTEMS

Single-installation systems (sometimes also known as *single-stage* systems) form a backbone for other systems and actually many of the basic results extend to other systems. In this section, we describe a template through which all variants/extensions of the single-installation capacitated system may be viewed.

3.2.1 The Base Model

Capacitated systems are those where the ordering quantity or equivalently production quantity is limited. The formulation is achieved by modifying the standard inventory relationships for uncapacitated systems and examining the resulting policy.

In all inventory systems, we try to achieve a simple policy. If at the beginning of period t , we start with inventory x_t , under mild assumptions there exists the best up-to level $y_t(x_t) \geq x_t$. This obvious result is just an effect of optimization. From a practical point of view, however, this is not a satisfactory outcome. A simpler and more desirable outcome is to have a policy that has some stability, e.g., we always produce or order the same amount, or always bring the inventory to the same target. In practice, however, the external environment is changing and most organizations control inventory of many items. The simplicity of the policy allows us to adjust a single parameter to the changing external environment rather than adjusting the whole policy, possibly characterized by an infinite or very large number of parameters, and do it for every item. Not surprisingly, a dominant portion of inventory theory tries to show the existence of target levels, labeled also as *up-to* policies, or sometimes *base-stock* policies, where the policy is defined as $y_t(x_t) = \max(y_t^*, x_t)$, that is, ordering up to y_t^* , if the initial inventory is below y_t^* , and ordering nothing otherwise. This desire for simplicity presents itself in both the cases without capacity constraints and with capacity constraints.

The standard time-dependent formulation of inventory systems uses dynamic programming. We assume that t is the number of remaining periods:

$$V_t(x_t) = \min_{y_t \in A(x_t)} \left[C(x_t, y_t) + L(y_t) + \alpha EV_{t-1}(x_{t-1}(y_t)) \right] \quad (3.1)$$

where $A(x_t)$ includes all constraints on the feasible ending inventories for given starting inventory x_t , $C(x_t, y_t)$ is purchasing or production cost, $L(y_t)$ is the cost of operating the policy in the current period with y_t inventory on hand, and $x_{t-1}(y_t)$ is ending inventory in period t , which becomes starting inventory in the next period $t-1$.

Over the last 50 years there has emerged a set of standard assumptions that limit the type of functions used and serve as sufficient conditions for the desirable up-to policies. Three desirable properties can be found in a significant portion of the papers: (a) purchasing cost function $C(x_t, y_t)$ is decomposable in terms of initial inventory x_t and ending inventory y_t and, most often, assumed to be linear $C(x_t, y_t) = c(y_t - x_t)$. This implies that the right-hand side of Equation (3.1) becomes $f_1(x_t) + f_2(y_t)$, where $f_2(y_t) = cy_t + L(y_t) + \alpha EV_{t-1}(x_{t-1}(y_t))$. This allows us to disregard the effect of initial inventory when choosing the best-ending inventory.

(b) Unsatisfied demand is backlogged and is represented as negative inventory at the beginning of period $t-1$. Thus, the starting inventory in the next period depends linearly on the inventory decision y_t : $x_{t-1} = y_t - D$. (c) f_2 is quasiconvex. These jointly lead to optimality of the base-stock policy. Note that without (c) we could choose different local optima, depending on the constraint set $A(x_t)$.

Quasiconvexity, however, is often difficult to verify. Therefore, a stronger assumption is routinely used for (c): (c1) f_2 is convex, and (c2) the set of feasible actions $\cup_x A(x)$ is (geometrically) convex, as explained below.

Condition (c1) allows for great mathematical convenience. Specifically, rewriting the value function we have:

$$V_t(x_t) = -cx_t + \min_{y_t \in A(x_t)} J_t(y_t)$$

where

$$J_t(y_t) = cy_t + L(y_t) + E_{D_t} V_{t-1}(y_t - D_t)$$

Convexity is preserved when adding functions, when linearly transforming them, and when taking expectations. Thus, convexity of V_{t-1} in x_{t-1} implies, convexity V_{t-1} in y_t and, thus convexity of $J_t(y_t)$ and existence of a minimizer of J_t , y_t^* , which is the optimal base stock. Furthermore, it is straightforward to see that, due to (c2), convexity of $J_t(y)$ implies convexity of $V_t(x)$. Such properties do not hold for quasiconvex functions. Convexity also well reflects the reality of many practical situations.

Condition (b), or the backlogging assumption, implies a linear transformation between the current inventory and the future one and, therefore, allows for preserving convexity. To illustrate (c2), note that in an uncapacitated system, $A(x) = \{y, x \leq y\}$. Condition (c2) requires that in the space of all $x \in R$ and $y \in R$, the set $\cup_x A(x)$ is convex. Indeed in a system with no capacity constraints, $\cup_x \{y \mid x \leq y\}$ is simply a half space in R^2 , which is trivially convex (in a geometrical sense).

For any standard problem, satisfying (a), (b), and (c1–c2), the optimal policy will be a base-stock policy:

Lemma 3.1 *If purchasing cost is linear $C(x, y) = c(y - x)$, L is convex, unsatisfied inventory is backlogged, $\cup_x A(x)$ is convex, then there exist y_t^* such that it is for any x it is optimal to choose y_t^* if feasible, or point nearest to y_t^* that is in the feasible set $A(x)$.*

3.2.2 The Effect of Capacity

Capacity can be easily incorporated into the above system by modifying the feasible set $A(x) = \{y \mid x \leq y\}$ to $A(x) = \{y \mid x \leq y \leq x + K\}$, while leaving the formulation otherwise unchanged. Intuitively, the ending inventory cannot exceed the starting inventory by more than K .

Note that the modified set $A(x)$, $\cup_x \{y \mid x \leq y \leq x + K\}$, is a band in $\{(x, y) \mid x \leq y \leq x + K\}$ in R^2 space. Thus, Lemma 3.1 applies and the optimal policy is a base-stock policy.

This formulation leads to a number of simple extensions in terms of constraints on feasible sets: Any set of linear constraints expressed in $\cup_x A(x)$ set will preserve convexity. This includes a minimum production requirement, a maximum production requirement, as well as the maximum inventory available (such as a storage limit). (A minimum inventory may lead to infeasibility.) It appears that most of the practical applications translate into linear constraints, thus from a mathematical point of view, satisfying the convexity of $\cup_x A(x)$ in Lemma 3.1 above. One of the examples that illustrates extended constraints is Chan and Muckstadt (1999) which assumes both minimum and maximum production in every period. Liu and Tu (2008) consider inventory storage constraints.

Remark

Systems with capacity choices but that carry no inventory across periods are mostly equivalent to uncapacitated systems. Particularly, papers considering capacity choice in one-period settings from a methodology point of view can be translated into one-period inventory systems. Consequently, without the complication of managing inventory, these papers deal with more complex capacity issues, such as a mix of dedicated and flexible capacities (e.g., Van Mieghem, 1998). Some of these issues are discussed in Section 3.2.7.

3.2.3 Simple Applications

3.2.3.1 Outsourcing

It is easy to extend the logic of Section 3.2.2 to a case where in addition to a capacitated resource with unit production cost c_1 and capacity K_1 , we have one or more of more expensive additional sources: if a product may be purchased from multiple sources, $S_1, S_2, \dots, S_n$, each with linear purchasing cost c_i , possibly facing capacity K_i , then it is easy to define a nested capacitated model. Let $V_t^i(\cdot)$ denote the value of the optimal cost of using sources 1 to i in period t and $A_{i,y^{i-1}}$ captures the limited capacities of sources 1 to i . $V_t^{i+1}(y^i) = \min_{y^{i+1} \in A_{i+1,y^i}} [c_{i+1}(y^{i+1} - y^i) + V_t^i(y^{i+1})]$ with $V_t^1(y^1) = \min_{y^1 \in A_{1,y^n}} [L(y^1) + E_D V_{t-1}^n(y^1 - D)]$.

Lemma 3.2 Consider n suppliers with linear costs $c_1 < \dots < c_n$ and capacities $K_1, \dots, K_n$. If the one-period costs are convex and unsatisfied demand is backlogged, then there exist optimal base-stock levels $-\infty \leq y^{*n} < y^{*(n-1)} < \dots < y^{*1}$, such that we order from supplier $i = 1, \dots, n$ up to y^{*i} .

The base-stock levels y_i^{*i} for each source i are guaranteed to satisfy $y_i^{*(i+1)} + K_i \leq y_i^{*i}$. The solution is illustrated in Figure 3.1. While V^i captures the optimal cost when using the i cheapest resources with marginal costs $c_1, \dots, c_i$, each of the corresponding resources is limited. Therefore, an additional resource with a higher marginal cost c_{i+1} is beneficial. If that resource was unbounded, the use of resource $i+1$ would imply that every time increasing inventory benefits the firm by more than c_{i+1} per unit, it should be used. Thus, V^i would be modified by a tangent (with a slope of $-c_{i+1}$), which would become V^{i+1} . Since, however, the capacity of resource $i+1$ is K_{i+1} , only a portion of this tangent line applies, while the rest of V^i is shifted left (by a distance of K_{i+1}) to form the rest of V^{i+1} .

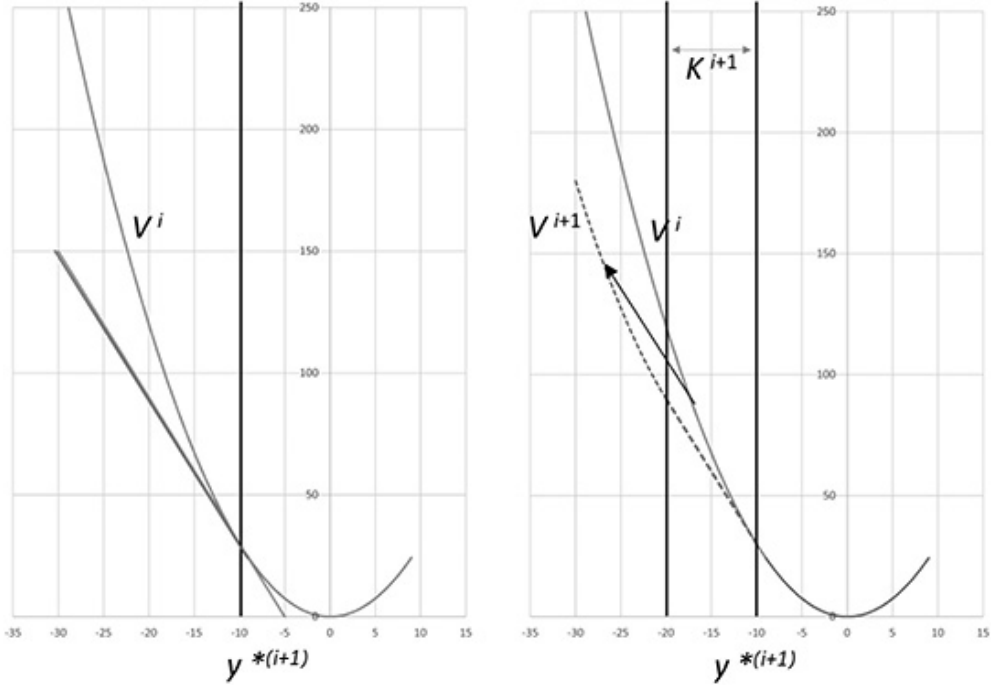


Figure 3.1 *Optimality of the outsourcing policy*

The outsourcing logic can also include fixed ordering costs if there is only one outsourcing state (Yang et al., 2005).

Other types of simple generalizations include:

1. Non-stationary demand and non-stationary cost parameters.
2. Markov-modulated formulation.

For point 1, note that we have not assumed any stationarity of demand functions D_t or of cost function $L = L_t$ or purchase costs $c = c_t$ and these are not needed for Lemma 3.1 to hold, so this generalization trivially holds.

As far as point 2, Markov-modulated formulation allows for a very broad range of generalizations. Various extensions based on Markov processes can be found in a number of publications and also not published results (e.g., Karlin & Fabens, 1959; Iglehart & Karlin, 1962; Wijngaard, 1975). Zipkin (2000) provided a description of a broad range of problems that can be classified using a Markov environment.

Assuming Markov state, m , we have:

$$V_{t,m}(x) = \min_{y \in A_{t,m}(x)} \left[c_{t,m}(y-x) + L_{t,m}(y) + E_m E_{D_{t,m}} V_{t-1}(y - D_{t,m}) \right]$$

which further leads to the existence of a base-stock policy with the target $y_{t,m}^*$ dependent on t and m .

Special cases include:

- Cyclic (periodic) inventory systems: for example, weekly demand with $M = \{0, 1, \dots, 6\}$ representing days of the week and deterministic transition $m \rightarrow m + 1$ with probability 1 (where $6 \rightarrow 0$).
- Autoregressive weather with, e.g., three temperatures $M = \{c, m, h\}$ (cold, medium, hot), or similarly state of economy, $M = \{l, m, h\}$ (low, medium, high) and corresponding cost functions and demands. Similarly, different states of technology (Krankel et al., 2006), breakdowns (Demirel et al., 2018), or learning can all be modeled in a Markov framework.

While in the basic examples, the state of the world M develops independently of demand, the next state of the Markov chain could be a function of demand in the current period, or both demand and current state, $m' = g(D_{t,m}, m)$.

3.2.3.2 Pricing

Several papers have included pricing in inventory models. Federgruen and Heching (1999) is one of the examples. By including pricing, the relationship is modified to:

$$V_t(x) = cx + \max_{y \in A(x), p \geq 0} J(y, p) \quad (3.2)$$

Joint concavity of $J(y, p) = -cy + R(y, p) + EV_{t-1}(y - d(p))$ can be easily justified if the one-period profit function $R(y, p)$ is jointly concave in inventory y and price p and if $d(p)$ is concave decreasing. This combination of assumptions allows for an inductive proof, showing the optimality of a base-stock list-price policy. The base-stock list-price policy consists of an optimal up-to level and any initial inventory below this level is paired with a list price; however, the optimal price decreases from the list price for any initial inventory above the optimal up-to level. Clearly adding any convex constraints in the feasibility set, will maintain concavity of the objective function. Hence, the results for inventory pricing models in a one-stage system can incorporate a capacity constraint (Allon & Zeevi, 2011).

3.2.3.3 Uncertain capacity

An interesting extension is a case where the current period capacity is uncertain. Capacity in a given period has a probability distribution with pdf of $f(\alpha)$. That is, if the decision maker attempts to produce $y - x$, the actual production is $\min(y - x, \alpha)$, with probability $f(\alpha)$.

This case has been studied first by Ciarallo et al. (1994) and is typically expressed using a modification of condition (c) in Lemma 3.1, where the convexity condition is replaced by showing that the objective function V is convex, while J is a quasiconvex function (actually it is decreasing down to its minimum and then increasing convex). Compared to Equation (3.1), the decoupling of initial and ending inventory is broken due to uncertain capacity.

$$V_t(x) = \min_{y \in A(x)} E_\alpha \left[J_t(\min(x + \alpha, y)) \right] \quad (3.3)$$

The right-hand side cannot be expressed solely in terms of y but instead, with different probabilities, states $x + \alpha$ are involved. However, convexity of V_{t-1} implies convexity of $M(y | x) = c(y - x) + L(y) + EV_{t-1}(y - D)$ in y . Let us define

$$\begin{aligned}
N(y | x) &:= E_p \left(M \left(\min(x + p, y) \right) \right) \\
&= (1 - F(y - x)) M(y) + \int_0^{y-x} M(p) dF(p).
\end{aligned} \tag{3.4}$$

Taking a derivative with respect to y immediately leads to $N'(y) = (1 - F(y - x)) M'(y)$. Since M is convex, we have that M' is increasing, implying that N' is positive-negative, and thus N is quasiconvex with the same minimizer as M . Note that the same logic can be applied in any non-stationary and Markov-modulated setting.

This technique has been applied in several papers extending the case of uncertain capacity. For example, Hu et al. (2008) show the structure of the optimal policy (inventory available to be transshipped will be rationed and the firms produce up-to state-dependent targets) when two locations are willing to transship goods but have uncertain capacities, or the existence of coordinating prices when two locations are decentralized (Hu et al., 2007).

3.2.4 Calculating the Optimal Base-Stock Target

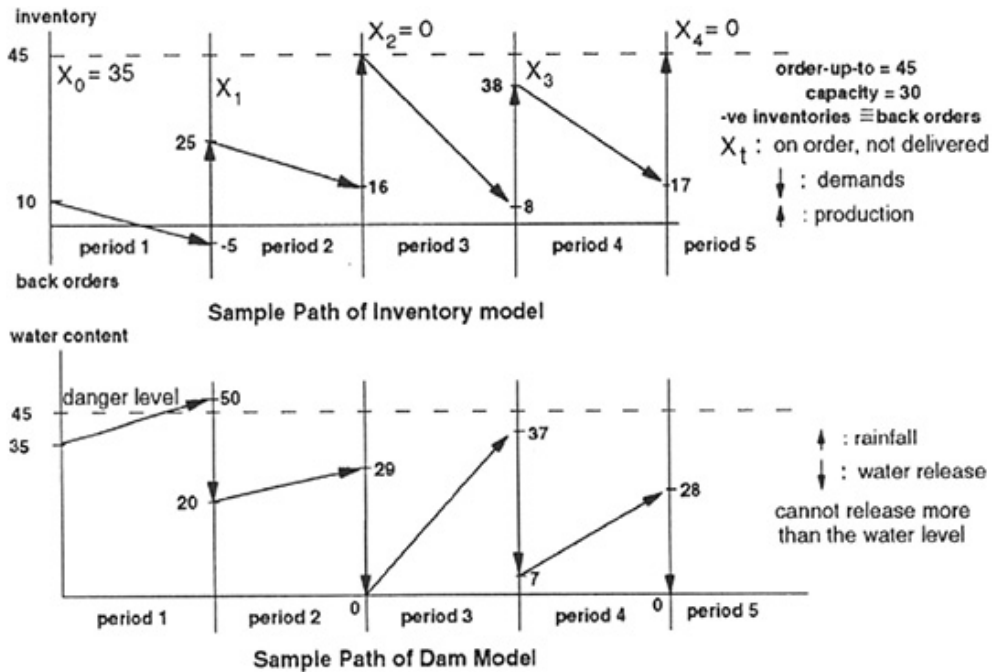
Relying on Federgruen and Zipkin's (1986a, 1986b) results that the optimal policy structure is modified base-stock, Tayur (1993) illustrates how the optimal base-stock level may be calculated using the concept of a shortfall. The shortfall is a distinct concept from the backlog. The *backlog* is the amount a customer has demanded but the firm has not provided. A *shortfall* represents the amount a firm could not produce due to capacity limits. Now the model does not track the evolution of the inventory level but the evolution of the shortfall, $S_n = \max(0, S_{n-1} + D_n - K)$, where S_n is the shortfall at the end of period n . (Since the spirit of taking the shortfall problem to the infinite-horizon, we have retained the period index n as counting forwards, rather than using period t elsewhere in the chapter counting backward.)

As seen in Figure 3.2, the connection between the inventory model (upper panel) and the dam model (lower panel) becomes clear, where (i) the shortfall level in the inventory model is the water level in the dam model, (ii) the demand outflow in the inventory model is the inflow of rainfall in the dam model, (iii) the production capacity preventing stock buildup in the inventory model is the limit of water release in the dam model, and (iv) reaching the base-stock level in the inventory model is an empty dam in the dam model. The process $\{S_n, n = 1, 2, 3, \dots\}$ is a Markov chain with a steady-state distribution $G(x)$. With a rainfall distribution $F(x)$ (also demand distribution), then the optimal order up-to level z satisfies

$$(G * F)(z) = \frac{p}{p + h} \tag{3.5}$$

where $*$ denotes convolution. The only condition required for the existence of G is that the mean demand is below capacity K . Prabhu (1965) establishes the existence and uniqueness of G for all discrete and Erlang demand distributions; for others, approximations may be utilized.

Equation (3.5) is derived in Tayur (1993) by observing that a single-stage capacitated system is equivalent to an infinite-stage uncapacitated serial system with single-period lead times, and installation 1 has a base-stock level of z and installations j have echelon base-stock levels of $z + (j - 1)K$ for all $j > 1$. By establishing the shortfall of each system is identical to



Source: Kamesam & Tayur, 1993.

Figure 3.2 Equivalence between inventory and dam models

each other and establishing the infinite-horizon cost in terms of shortfall, they differentiate with respect to z and solve to obtain Equation (3.5).

An alternative approach to finding the parameters of the policy is to search for them by evaluating the costs of different combinations of parameters. Due to a lack of closed-form solutions, the cost can be found by simulation, which is computationally expensive. Infinitesimal Perturbation Analysis (IPA), described by Fu (1994), can accelerate the search, by simulating the gradient of the objective function, rather than the value itself. This method has been applied to many systems with capacity constraints, as described below in Section 3.3.1.2. IPA involves establishing estimators of the gradients of the state variables, extending it to the performance measure (e.g., value function), and using these to derive estimates of gradients of the measure of interest (e.g., base-stock levels).

3.2.5 Capacitated Systems with Fixed Cost

Wijngaard (1972) was the earliest to explicitly consider applying a capacity limit to a dynamic stochastic inventory system (with fixed costs). He intuitively conjectured that the optimal policy should be the (s, S) policy with the amount ordered limited by the capacity, attempting to get as close to S as possible. However, he shows that this intuitive policy fails to be optimal. Shaoxiang and Lambrecht (1996) partially characterized the optimal policy by defining the

X – Y band structure. Specifically, if the inventory level is below X , order to full capacity, and if the inventory level is above Y , order nothing. The values of X and Y apply for any horizon length beyond one period. What happens between X and Y varies from problem to problem. Gallego and Scheller-Wolf (2000) further distilled the optimal policy but still not completely, showing there are now four inventory regions which are characterized (in increasing inventory order): (1) order capacity; (2) either (a) order at least up to s' or (b) no order or order capacity; (3) no order or at least to up to s' ; (4) no order. This is not a full characterization since there are, the authors note, possibly a number of intervals between s and s' where is it optimal to start and stop ordering, among other issues. Shaoxiang (2004) considered an infinite-horizon version of the model and establishes that under a concept labeled (C,K) -convexity, the difference between the X and Y is no greater than the capacity amount. The finite-horizon value function and optimal policy are shown to converge in the infinite-horizon. Also, a tighter bound on Y is derived. A counterexample to the posited (s,S) policy is shown to be preserved. Lastly, a numerical study shows that an (s,S) policy cannot be better than 11% over optimal, suggesting there is considerable value in further pursuing the optimal policy structure between X and Y .

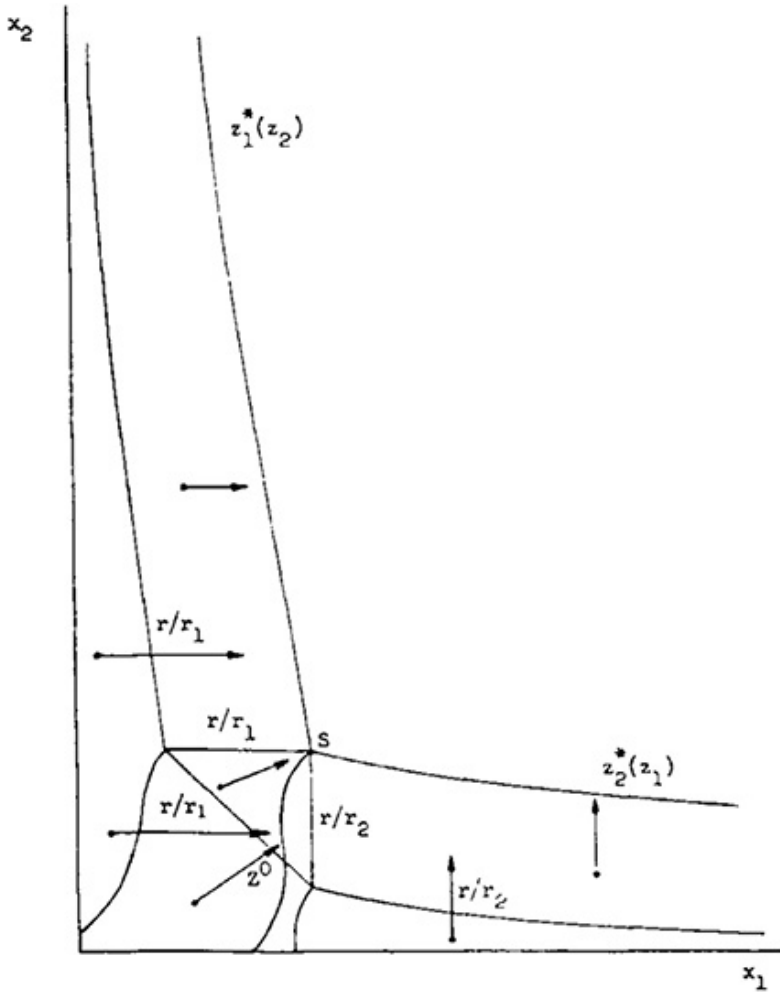
3.2.6 Multiple Items with Joint Capacity Limit

Evans (1967) was among the first to incorporate joint capacity constraints for multiple products. Evans shows the existence of an optimal policy for multiple products and describes its structure, depicted in Figure 3.3. It is illustrated for two products (1 and 2). In a nutshell, if the initial inventories (x_1 and x_2) are below the target level (S), then it is optimal to order up to the target. However, if the target cannot be reached (due to limited shared capacity r , and r_i is the amount of capacity needed for one unit of product i), then one uses a line (indicated by vector Z^0) that best balances the inventories of two products and optimal policy is to move to this line (and then along the line). Finally, if one of the initial inventories is above the target, then the other item is ordered up-to some well-defined lower level ($z_i^*(z_{-i})$). These optimal actions are illustrated in Figure 3.3.

DeCroix and Arreola-Risa (1998) extend the result for two products to an infinite-horizon setting and explicitly describe the policy when products are identical. Nahmias and Schmidt (1984) in a one-period setting numerically evaluate the performance of various heuristics for producing multiple products using one joint capacity. Glasserman (1996) also considers the infinite-horizon problem, although in continuous time to approximately minimize holding and backorder costs. The problem is effectively simplified by dedicating a fraction of the total capacity to each of the products. Atalı and Özer (2012) also consider the multi-item problem with both lower and upper bounds (akin to production smoothing) with a common intermediate product produced upstream of differentiated products, and develop a well-performing heuristic.

3.2.7 Capacity Investment

Capacity investment was investigated early, but under simplified assumptions. Usually, a predictable demand was assumed with the rate of demand increasing linearly. This type of approximation of reality allowed the focus to be on the timing of capacity additions and was easily translatable into EOQ models (Manne, 1967; Yaged, 1973). Simple extensions of these models include a nonlinear increase of demand, single capacity serving multiple demands, as well as building an inventory buffer in order to delay the next capacity investment. These were described in Luss (1982) who reviews early literature in this domain and provides several



Note: The arrows indicate optimal actions.

Source: Evans, 1967.

Figure 3.3 Structure of Evans's (1967) optimal policy

translations of deterministic demand models into dynamic programming and into network flow problems. Luss (1982) also lists many applications. Later Rajagopalan (1998) describes capacity expansion and equipment replacement models. Song et al. (2020) provides some further perspective on research dealing with capacity.

3.2.8 Reactive and Proactive Uses of Capacity

Capacity can be used in several creative ways enabling one to (a) react to better information, possibly with shorter lead time but higher cost, or (b) proactively by getting some advance information and starting to produce earlier.

The concept of reactive capacity was analyzed by Fisher and Raman (1996) and is closely related to the general idea of two lead times, shorter more expensive and longer cheaper one. Fukuda (1964) provides structural results for such model, deriving a centralized policy for ordering from two uncapacitated sources with lead times that differ by one period. The optimal policy is based on two thresholds, one for each source of product. The logic of Fukuda (1964) can be easily expanded to two capacitated sources, but cannot be expanded to cases where the difference between lead times is larger than one period. Fisher and Raman (1996) and Fisher et al. (1997) suggest assigning a role to the capacity, that is splitting regular capacity into regular capacity (with the original lead time) and reactive capacity (with shorter lead time). They evaluate the best split of capacities and analyze which products and quantities should be produced using regular capacity vs. using reactive capacity. Iyer and Bergen (1997) consider a related set-up when demand distribution and order can be adjusted at a later time and they examine the role of arrangements (wholesale price, service level) between a manufacturer and supplier in the modified settings, while Krishnan and Kapuściński (2010) point to a potential downside of these arrangements.

To make the best use of capacity, several papers advocate collecting advance information about demand. Advance demand information can be simply used to gradually build products, as refined forecasts arrive. This idea is exploited in multiple papers; see literature reviews in Hu et al. (2003) and Wang and Toktay (2008). This includes later papers such as Boyacı and Özer (2010) that combine collecting advance information with making capacity decisions.

A creative use of advance selling may also be used to take advantage of customers not knowing (rather than knowing) their preferences in advance, see e.g., Yu et al. (2015), who consider setting different prices (premium or discount) for advance and regular capacity depending on the degree of homogeneity of demand.

3.2.9 Continuous Time

Continuous time models deal with capacity rate mostly in two ways: (a) assuming constant demand rate, but allowing for other uncertainties or (b) focusing on uncertain demand.

In the first group, typically the focus is on disruptions. Kimemia and Gershwin (1983) and Bielecki and Kumar (1988) are representative examples, where both the capacity rate and demand rate are constant. They focus on disruptions in the availability of capacity and buffer these disruptions with inventory. Assuming a base-stock policy, they notice that the shortfall is independent of the order up-to level and search for the optimal order up-to level.

The second group of papers also considers continuous time, but instead employs queueing theory, where both demand and production have a certain expected “rate,” but their realizations are uncertain. Not surprisingly, the same problems can be modeled in a discrete time setting, within a dynamic-programming framework, or in a continuous time setting using a queueing framework. Each of the approaches has its own benefits: both approaches allow to derive the structure and properties of the optimal policy. Usually, the queueing theory approach imposes fairly strong assumptions such as stationary demand with exponential or deterministic inter-arrival and processing times. For example, Veatch and Wein (1994) model a make-to-stock tandem queue where the controls are the means of the exponential servers with Poisson arrivals, examining a variety of production control mechanisms, which are discussed further in Section 3.3.1. The restrictions in queueing models often come with benefits, such as some of the results can be more precise and some can be expressed in a closed form.

A good example is the area of inventory rationing. This stream of papers assumes multiple classes of customers with different revenues or penalties. At nearly the same time, Evans (1968), Kaplan (1969), and Topkis (1968) considered an uncapacitated setting in a periodic (dynamic-programming) setting and showed that, if demand is generated by multiple classes of customers, there is a rationing level such that inventory below that level should be offered only to higher-class demand (with a higher backlog cost or higher lost-sales cost). Evans stated this result in a periodic setting for two classes of demand and lost sales, Topkis has shown it for N classes of demand and lost sales, while Kaplan assumed two classes of demand and backlogging.

These results have been extended both in a queueing framework and in a discrete time dynamic-programming framework with capacity. In a queueing framework Ha (1997a, 1997b, 2000), and then later Zhao et al. (2005) showed existence of rationing levels, confirming that the same structure holds as in the original papers. Due to the queueing framework, these papers assumed a certain processing rate, which is effectively a capacity constraint. The queueing framework permits the calculation of the cost of the operating policy. Deshpande et al. (2003) considered two classes with stochastic demand and showed that a static rationing level remains optimal. Véricourt et al. (2001, 2002) show techniques to calculate the rationing levels.

In a dynamic-programming framework, Kapuściński (1996, Chapter 4) and Katircioglu (1996) explicitly impose capacity constraints and show the same results.

Queueing, without inventory considerations, allows for a different types of problems to be modeled and often focuses on intersection of inventory with other decisions. For example, Lederer and Li (1997) consider a capacity decision and pricing in a queueing framework.

Similarly, while joint inventory and capacity decisions can be easily considered in a dynamic-programming framework, Bradley and Glynn (2002) consider them in a queueing framework, using a $GI/M/1$ model, they establish an approximation that estimates best capacity decisions.

3.3 DIFFERENT SYSTEM CONFIGURATIONS

In this section, we examine the effect of capacity upon inventory in system configurations beyond the single installation. The bulk of previous work has focused on serial systems, a building block of more complicated networks, which we discuss in Section 3.3.1. In Section 3.3.2 we discuss results in assembly systems.

When considering a network of connected inventory installations, a number of questions naturally arise:

- What are optimal inventory policies?
- How are orders at one installation affected by the limited capacities at other installations?
- What role does the location of the most limiting capacity play in relation to the network structure?
- In practice, where would the most limiting capacity be located?

It is fair to characterize that knowledge of the single-installation systems is more complete than for multi-echelon systems.

3.3.1 Serial Systems

The serial system is a foundational network in multi-echelon inventory theory, à la Clark and Scarf (1960), so examining how the optimal policy is changed with the imposition of capacity limits is natural. A serial system is where one installation upstream provides inventory to one installation downstream, until the lowermost installation sells it to a (stochastic) market. Typical assumptions are that holding costs increase closer to the market and stockout costs are only incurred at the lowest installation. Clark and Scarf (1960) establish the optimality of the *echelon base-stock* policy, where each installation attempts to raise its echelon inventory (sum of local and all downstream inventory) to a desired target. Characterizing the optimal policy of the general serial capacitated system (N installations, general lead times, general capacity configurations) has been elusive. However, many special cases have been examined in the literature. We label the installations $1, 2, \dots, N$ in increasing distance from the market demand. K_i will denote the capacity limit at installation i and L_i will denote the lead time between installations $i + 1$ and i . Using this notation, we summarize the key known results:

- $N = 2, K_1 \geq K_2, L_2 = 2$: Base-stock is not optimal (Speck and van der Wal, 1991)
- $N = 2, K_1 \leq K_2, L_2 = 2$: Two-tier base-stock policies are optimal (Janakiraman & Muckstadt, 2009)
- $N = 2, K_1 \leq K_2, L_2 = 1$: The optimal policy is modified echelon base stock (Parker & Kapuściński, 2004)
- $N, K_N < \infty$: The optimal policy is modified base stock at installation N and echelon base stock at installations $1, \dots, N - 1$ (Parker & Kapuściński, 2004)
- $N, K_{N-1} \leq K_N < \infty, L_N = 1$: The optimal policy is modified echelon base stock at installation $N - 1$ and N and echelon base stock at installations $1, \dots, N - 2$ (Parker & Kapuściński, 2004)
- $N, K_{i-1} = K_i \forall i$: There exists a bound on the number of base-stock tiers at each installation (Janakiraman & Muckstadt, 2009)
- $N, K_{i-1} \geq K_i \forall i$: Under assumed base stocks, shortfalls are related to single-installation shortfalls (Huh et al., 2010)
- $N, K_{i-1} \geq K_i \forall i$: Algorithms and bounds for finding base-stock levels (Huh et al., 2016)
- $N, K_1 \leq K_i \forall i$: Under elevated inventories, the optimal inventory is branching echelon base stock (Angelus & Zhu, 2017)

3.3.1.1 Optimality results

The earliest examination of serial systems with capacity limits observed odd behavior and attempted to explain it. The “push-ahead” effect noted by Speck and van der Wal (1991) is the observation that a system will move additional stock from the upper installation to the lower installation beyond what appears to be the target base-stock level at the retailer. Let x_1, x_2 , and x_3 represent inventories at installations 1 and 2 and installation 2’s pipeline inventory. Consider their example: $K_1 = K_2 = 20, L_1 = 0, L_2 = 2, (x_3, x_2, x_1) = (20, 20, 1)$ and demands vary from 17 to 23. The retailer’s base-stock target is 20 so 19 units are shipped and 21 units are at the supplier at the start of the next period. If at most 20 units are sold in that period, all is fine, but if more than 20 are sold, the system is “handicapped” by K_1 : only 20 items can be shipped while at least 21 items are needed to reach the target base stock. They argue they would have been better off shipping 20 units rather than 19 when they had ample supply. This behavior is

what they label the “push-ahead effect.” In essence, the system is balancing (1) the additional holding cost at the retailer multiplied by the probability that demand is sufficiently low that the inventory will not be consumed, against (2) the probability that demand will be larger than retailer capacity and the supplier will not be able to move that quantity of goods at that time. Thus, the supplier theorizes that the expected payoff is superior to send the additional units when he has ample supply rather than be caught with a retailer shortage that cannot be remedied in a single shipment. Clearly, the two-period lead time upstream of the supplier plays a role in establishing the on-hand and pipeline inventory sum ($x_2 + x_3$) by which the push-ahead is executed or not, so the calculated tradeoff would presumably be effected over multiple periods if a large demand (relative to capacity) is realized. This suggests the effect may occur for a longer lead times, too. This naturally would then extend to multiple installations rather than merely longer lead times; this correspondence between lead times and installations was formalized by Glasserman and Tayur (1994).

Janakiraman and Muckstadt (2009) addressed this precise setting.¹ They demonstrate the optimality of a “two-tier base-stock” policy when $x_2 + x_3 \leq 2K$ (where $K := K_1 = K_2$). The two-tier base-stock policy is defined as having two base-stock levels at each installation, triggered by the initial inventories. A broader class of policies labeled *monotone policies* is defined and the optimal policy resides within this class. The interpretation to demonstrate compatibility with the examples above is that echelon 1 will order up-to one of two targets depending on the sum of the supplier’s installation and pipeline inventory.

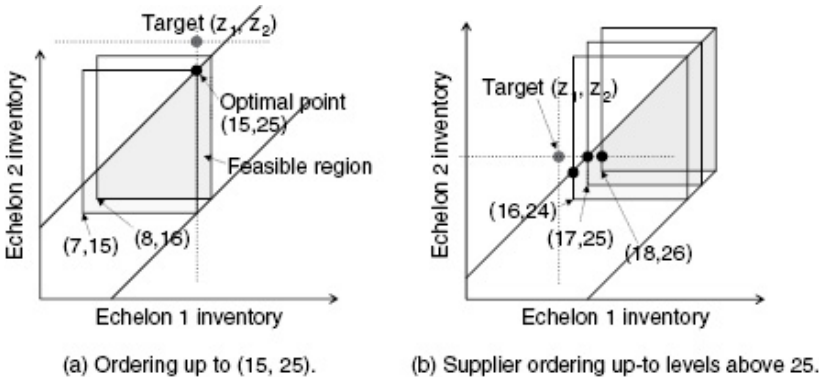
When the lead time upstream of the supplier is one period ($L_2 = 1$) and $K_1 \leq K_2$, the optimal policy is considerably simpler, and is named the *modified echelon base-stock* policy or *MEBS*. Before describing the optimal MEBS policy, Parker and Kapuściński (2004) motivate the setting with an example showing that the echelon base-stock policy is not optimal, as seen in the final two columns of Table 3.1 (all examples illustrated are for initial supplier installation inventory $x^2 = 8$ and $K_1 = 10$). A number of the examples from Table 3.1 are illustrated in

Table 3.1 Parker and Kapuściński (2004) counter-example

Initial installation inventory		Initial echelon inventory		Installation orders		Ending echelon inventory	
x^1	x^2	X^1	X^2	a^1	a^2	Y^1	Y^2
5	8	5	13	8	10	13	23
6	8	6	14	8	10	14	24
7	8	7	15	8	10	15	25
8	8	8	16	7	9	15	25
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	8	15	23	0	2	15	25
16	8	16	24	0	2	16	26
17	8	17	25	0	2	17	27
18	8	18	26	0	1	18	27
19	8	19	27	0	0	19	27
20	8	20	28	0	0	20	28

Figure 3.4, where the rectangles represent the feasible region for starting echelon inventories (X^1, X^2) at the lower left-hand corner and the solid black dot represents the optimal order up-to levels (Y^1, Y^2) . In this example, it appears that $(15, 25)$ is an echelon base-stock level from any initial inventory levels $x^1 = X^1 \leq 15$, as seen in Figure 3.4a. However, as the retailer's initial inventory (x^1 or X^1) starts as 16, the supplier surprisingly orders up to $Y^2 = 26$ (order quantity $a^2 = 2$) rather than 25, thus contradicting the posited $(15, 25)$ base-stock level. This is further contradicted when $x^1 = X^1 = 17$ when the supplier orders up to echelon $Y^2 = 27$ rather than 25 or 26. These behaviors are illustrated in Figure 3.4b. Interestingly, when the initial inventory is $(X^1, X^2) = (18, 26)$, the system orders up to $(18, 27)$, so seemingly the supplier is targeting another echelon base stock at 27. This seeming contradiction to the classic echelon base-stock policy is reconciled by recognizing that the supplier will never stock more than K_1 units of inventory, since the retailer will never order more than K_1 units. Thus, the inventories will naturally be limited to a *band* defined as $\{(X^1, X^2) | X^1 \leq X^2 \leq X^1 + K_1\}$. As illustrated in Figure 3.4, the new feasible region is the intersection of the previous rectangle and this band. Parker and Kapuściński (2004) show that the system's value function decomposes along the echelon decision variables within the band, and the band itself is an absorbing subset of the state space (if initial inventories fall outside this band, the supplier's inventory will be consumed through natural operation until it does not exceed K_1). Thus, while there may be a desired inventory target (z_1, z_2) outside the band, illustrated as $(z_1, z_2) = (15, 27)$ in Figure 3.4, it will never be reached but its influence is felt by its projection upon the band. Specifically, the retailer will order up-to 15 if starting below 15 and order nothing if above 15. And the supplier will order up to $15 + K_1 = 25$ if the retailer's initial inventory is below 15 and similarly aim to get as close to 27 as possible, limited by the augmented feasible region, by tracing the upper edge of the band until 27 is reached.

Veatch and Wein (1994) use a continuous time tandem queue model with a make-to-stock setting. They examine a variety of production control mechanisms (base stock, kanban, fixed



Note: The rectangles represent the feasible regions for various initial echelon inventories (lower left-hand corner). The solid black circles represent the optimal echelon order up-to levels. The solid gray circle represents the target echelon order up-to levels $(z_1, z_2) = (15, 27)$ for the MEBS policy.

Source: Parker and Kapuściński (2004).

Figure 3.4 The modified echelon base-stock behavior for $K_1 = 10$ and $x^2 = 8$

buffer), where the controls are the capacity rates of each installation. They demonstrate that a number of the mechanisms could be optimal under some extreme circumstances but that base stock will never be optimal. They show the following scenarios can be optimal: no inventory in the system; no finished goods inventory (implying non-idling at the supplier); and non-idling at the retailer. Through numerical examples, they show that capping inventory at the higher stage may improve the performance, which is consistent with the optimal policy derived from Parker and Kapuściński (2004).

For channels longer than $N > 2$, the optimality results are fewer. Janakiraman and Muckstadt (2009) establish a more general result for the N -stage system. Specifically, they extend their monotone results to the longer channel and establish an upper bound on the number of base-stock levels at each echelon, $2^{M_{N+1}-M_2}$, or for the entire channel, $N \cdot 2^{M_{N+1}-M_2}$, where $M_n = 1 + L_1 + L_2 + \dots + L_{n-1}$ denotes the location of stage n by summing stage lead times. Angelus and Zhu (2017) consider a scenario where the inventory levels are elevated at every echelon. Under this restriction, the value function decomposes along the echelon inventories, and the resulting optimal policy is *branching echelon base stock* (BEBS). BEBS is a state-dependent order up-to policy where the number of parameters needed to characterize the replenishment decision grows (“branches”) for each additional echelon going downstream.

Special cases in Parker and Kapuściński (2004) demonstrate how their results can be applied to longer channels. They demonstrate that if only the highest installation is capacity limited ($K_N < \infty$) then the optimal policy is modified base stock at installation N and echelon base stock for all lower installations. Also, if the two uppermost installations of an N -stage system are capacity limited such that $K_{N-1} \leq K_N$ and $L_N = 1$, then the optimal policy is MEBS for installations $N-1$ and N and echelon base stock for the lower installations. All of these latter results rely on the decomposition of the value functions into unidimensional value functions with accompanying convexity.

3.3.1.2 Stability and computational time

The remaining work on the N -stage capacity-limited serial system is mostly focused on properties of the system such as stability and convexity. Glasserman and Tayur (1994) explore the stability of this serial system under the assumption of a base-stock policy. Even though Janakiraman and Muckstadt (2009) demonstrate that a single target base-stock level policy is not optimal, it leads to insights into how the system behaves. They suggest these base-stock policies are attractive due to their simplicity, their optimality under some conditions, and the difficulty in establishing the optimal behavior in the general capacitated system. They establish that, as long as the expected demand is smaller than all the capacities, the system is *stable*. Stability in this situation is where the shortfalls (distance from reaching the base-stock levels) are finite, or less formally that the system can supply materials more quickly than demand can consume them. Specifically, they show the echelon shortfall distributions are stable and stationary, regardless of the initial inventories and the system is Harris ergodic, proved by showing that two initial inventories will *couple* in finite time. Lastly, they show how lead times may be replaced with installations with zero lead times, verifying the earlier Harris ergodic results apply to systems with positive lead times, too. Huh et al. (2010) similarly examine stability properties of the serial capacitated system acting under echelon base-stock policies. They establish a relationship between the shortfalls in the serial system and the shortfalls of a single installation, subsequently strengthening the earlier stability results, finding the shortfalls are Harris ergodic, and tightening the regeneration time results of Glasserman and Tayur (1994).

In a parallel paper, Glasserman and Tayur (1995) numerically explore these systems using a gradient simulation technique known as Infinitesimal Perturbation Analysis. This examines the sensitivity of the base-stock levels to changes in system parameters, again adopting a base-stock policy as the operating paradigm due to its simplicity. Under a variety of criteria (average cost, discounted cost), they formalize the conditions for the existence of unbiased estimators of derivatives with respect to the base-stock levels and demands. Some of their observations include: (1) the optimal costs and base-stock levels drop sharply and then level off as capacity increases; (2) optimal costs and base-stock levels increase as demand variance increases; (3) the increase in optimal cost as capacity decreases is larger at higher variance; (4) the Type-1 service level approximates the Newsvendor critical fractile; (5) the flatness around the optimal cost implies some robustness around the base-stock levels; and (6) the observed value-added at higher installations seems modest.

In their third paper, Glasserman and Tayur (1996) develop an approximation approach to calculate the echelon base-stock levels within that class of policies, using an extension of the shortfall approach of Tayur (1993). Using a test-bed of 72 problems, they establish the average relative cost error using their approximate base-stock levels relative to those found using IPA in Glasserman and Tayur (1995) is 1.9%. Since their intention is to assess the quality of the approximation technique within these base-stock policies, they do not establish the error relative to optimality, particularly since the optimal cost of these cases is exceedingly difficult to find. Therefore, their average relative error to optimal will be greater than 1.9%.

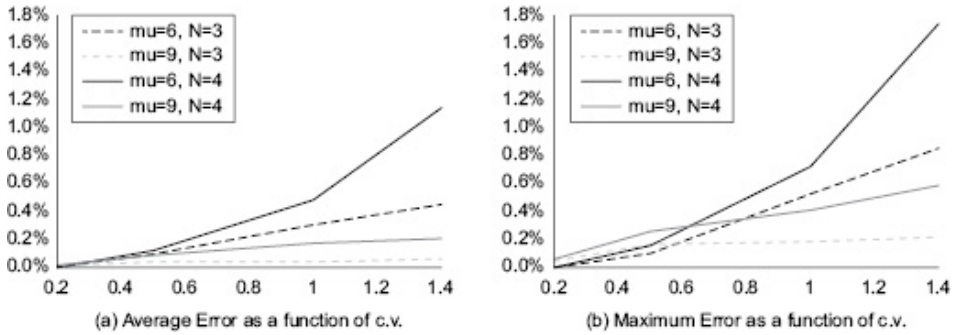
Huh et al. (2016) develop three heuristic policies based on the echelon base-stock policy, adapted from the multi-echelon literature (Federgruen & Zipkin, 1984; Shang & Song, 2003), and several bounds on performance. For light-tailed demand distributions, the algorithms are shown to be asymptotically optimal as the unit stockout cost approaches 1. In numerical studies, they compare the best heuristics with the best-performing lower bound. In the two-stage model, the best heuristic performed on average 1.1% higher than the cost of the best echelon base-stock policy and never more than 2.7% in the 75 experiment test-bed. For the four-echelon example, the average and maximum were 1.3% and 4.0%. They also report that their best-performing lower bound is on average 0.5% away from the cost of the best base-stock policy for $N = 2$ and 0.6% for $N = 4$.

In comparison, Kapuściński and Parker (2022) use the modified echelon base-stock policy over a similar demand variance and find an average error (in a “cube” around the target base stocks) of 0.1% and an average of the maximum errors (over the state space) of 0.23% for $N = 3$ stages, and 0.22% and 0.44% for $N = 4$, respectively, compared with optimality, in a test-bed of 240 cases. These errors are further described in Table 3.2, which shows the

Table 3.2 Count of cases from 240 for $N = 3$ and $N = 4$

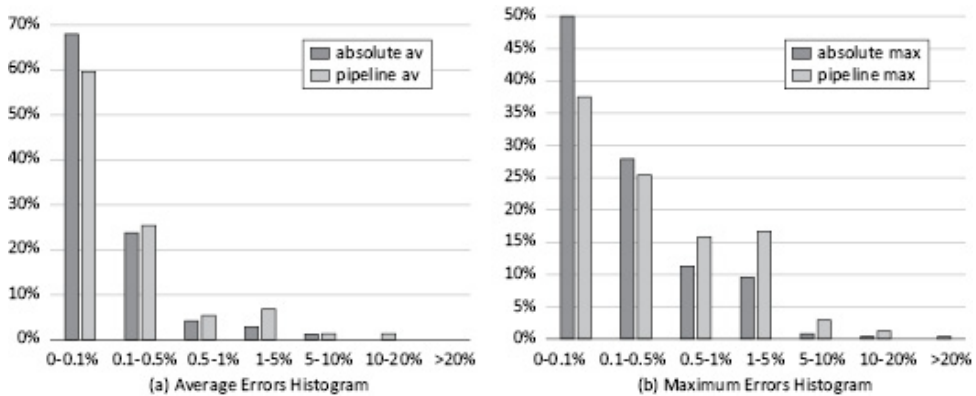
	$N = 3$		$N = 4$	
	Average error	Maximum error	Average error	Maximum error
Below 0.1%	204	150	163	120
Below 0.5%	229	178	220	187
Below 1%	233	230	230	214

Source: Kapuściński and Parker (2022).



Source: Kapuściński and Parker (2022).

Figure 3.5 Average of Average Errors and Maximum Errors as a function of the coefficient of variation



Source: Kapuściński and Parker (2022).

Figure 3.6 Average and Maximum Error Histograms for Absolute and Pipeline-Adjusted Errors

distribution of cases by error threshold, illustrating that large errors are an exception while tiny errors are common. Figure 3.5 shows how these average and maximum errors increase as N increases from 3 to 4, all errors increase as CV increases, but are worst at the low demand (μ) and high variance combination, precisely where capacity is of less consequence. The errors are lowest when utilization and variance are highest, which is where the effects of capacity are most critical. Further, the average and maximum “pipeline errors” (where the unavoidable pipeline lead time costs are removed from the MEBS and optimal horizon costs) are also calculated, mostly to isolate the effect of the MEBS policy absent other elements common to both policies, with the absolute and pipeline errors displayed in the histograms of Figure 3.6. Although the pipeline error calculation marginally increases the numbers, a similarly small

error pattern is observed, where the vast majority of average and maximum errors are much smaller than 1%. The summary suggests that while the optimal policies may be elusive in analysis and calculation, MEBS is a parsimonious base-stock policy which outperforms the conventional echelon base-stock policies.

3.3.2 Assembly Systems

Bollapragada et al. (2004) illustrates the difficulty of any form of capacity for multi-stage systems. Since finding the best base-stock policy (which is not optimal itself) is difficult, the paper considers a multi-stage assembly system with uncertain capacity. While the optimal policy is unknown, the paper proposes to replace the optimal policy with service-level requirements for each stage. They show that formulation can well approximate the one which considers actual base-stock levels in assembly systems (which is not an optimal policy).

Rosling's (1989) landmark result shows that an assembly system without capacity limits may collapse to a Clark and Scarf (1960) system under a condition known as long-term balance in the infinite-horizon. This collapse is achieved by ordering the assembly points on the basis of their echelon lead times and permits the calculation of optimal echelon base-stock levels in the equivalent serial system. Angelus and Zhu (2013) consider capacity-limited special cases of Rosling's general assembly system. They primarily investigate a system with a single assembly point at the most downstream stage, sometimes referred to as the *component assembly system*, admitting multiple components and multi-period lead times between successive stages. The overarching objective is to establish conditions under which Rosling's remarkable result may be duplicated for capacity-limited systems. First, they illustrate two scenarios which would prohibit such a collapse. The first of these is a situation where an upstream component has a larger lead time than a parallel component, leading to lower stocking level for the latter thus causing an imbalance in the echelon inventory stocks. The second situation is where the bottleneck capacity is at an upstream location for a component. The optimal stocking for this location is higher than for the parallel components at the same stage, again leading to an imbalance in the echelon stocking levels. Under either scenario, the imbalance prevents the component assembly system from being reduced to an equivalent serial system. Locating the (weakly) bottleneck capacity at the assembly point and limiting upstream lead times to one period are sufficient conditions to ensure the component assembly system collapses to a serial system, and the optimal policy will be balanced. When considering a system with subassemblies, akin to Rosling's (1989) model, with the imposition of a cost allocation assumption this more general assembly system delivers an optimal policy that is balanced and capacitated, indicating that it likewise may collapse to an equivalent serial system.

Kamesam and Tayur (1993) develop a series of algorithms based on the concept of the shortfall from target base-stock levels, under a global assumption of operation under base-stock policies. They progressively show the equivalence between a number of systems, starting with an uncapacitated assembly system and an uncapacitated serial system (à la Rosling, 1989), including algorithms for calculating echelon base-stock levels and service levels. Then duplicating the results of Tayur (1993) they show the equivalence of a single-stage capacitated system and an infinite-stage uncapacitated serial system, utilizing the dam water level correspondence to the inventory level models through the shortfalls, culminating in Equation (3.5). They then extend the analysis to a multi-stage capacitated system, showing a similar result where $K_N < \infty$ while all other stages are uncapacitated. The next step shows the equivalence

between a serial capacitated system and an uncapacitated component assembly system (i.e., assembly at the final node only). This approach, in turn, can be applied to a capacitated assembly system, where each capacitated node is translated into a series of infinite-stage uncapacitated serial systems, other than the highest node.

Huh and Janakiraman (2010) investigate a variety of properties of capacitated assembly systems, assuming echelon base-stock policies. They acknowledge this policy is not optimal but recognize it as popular among practitioners and researchers (Glasserman & Tayur, 1994, assume something similar, for example). Under this assumption, they show a number of properties of the assembly model.

3.4 APPLICATIONS AND METHODS

In this section, we describe the different methodologies (Section 3.4.1) and varying applications (Sections 3.4.2–3.4.4) where capacity plays a prominent role in inventory management.

3.4.1 Methodological

There are various methods used to analyze capacity-limited inventory systems. Below we list them with representative papers:

- dynamic programming: many papers including Federgruen and Zipkin (1986b), Parker and Kapuściński (2004)
- shortfall/stability: Tayur (1993), Glasserman and Tayur (1994), Huh et al. (2010)
- queueing analysis: Tayur (1993), Veatch and Wein (1994)
- single unit approach: Janakiraman and Muckstadt (2009)
- Infinitesimal Perturbation Analysis: Glasserman and Tayur (1995), Kapuściński and Tayur (1998), Gavarnini et al. (1999)
- infinite-horizon: Federgruen and Zipkin (1986a, 1986b), Parker and Kapuściński (2004, 2011), among others
- Markov games: Parker and Kapuściński (2011)

3.4.2 Information

The interactions between information and capacity have inspired several directions of inquiry. The information aspect has been considered with respect to future demand, and with respect to inventory levels. We consider three groups of papers:

1. The first group of papers considers the value of information in a fully decentralized setting and some papers focus on how capacity constraints may influence this value.
2. The second group looks at the minimum information that allows divisions of the same company to run in an integrated manner, when the company is capacity constrained.
3. The third group studies advance demand information when facility has limited capacity.

In the first group, multiple papers attempted to identify the settings when information about inventory in a different place in the supply chain is critical. Chen (1998) focused on a serial

supply chain, Cachon and Fisher (2000) focused on distribution systems, while Gavirneni et al. (1999) focused on a single chain (manufacturer and supplier) where the supplier has a capacity constraint. The two elements, ordering in batches and the capacity of the supplier, play the key role. Consistent with other papers (e.g., Kapuściński & Parker, 2022) decentralized decision-making benefits from knowing about inventory positions in other locations. However, tight capacity decreases the drawbacks of decentralized decision-making. The effects of capacity may be quite significant. In a similar spirit, Gavirneni (2002) examines the effect of one-period advanced information and shows that a big portion of the benefits of the value of information can be gained by knowing the order one period in advance (without knowing the inventory positions at all times)

Within the second group, Kapuściński and Parker (2022) explore the necessity of globally known information to execute the MEBS. In uncapacitated systems, Axsäter and Rosling (1993) illustrate that Clark and Scarf's (1960) echelon base-stock policy may be implemented with local information alone. At first blush, in capacity-limited systems it would seem very necessary to know inventory levels throughout the channel as this is how the policy is defined and since the random demands are "censored" by the retailer's capacity level; that is, the realized demands are not fully passed on through the retailer's orders to her supplier. Kapuściński and Parker (2022) demonstrate that *sufficient* information is passed upstream through the channel, for installations using only *local information* can mimic the actions of the MEBS policy. As noted earlier, MEBS is optimal under limited circumstances (two installations, tighter capacity at the retailer, one-period lead time at the supplier). This result contributes a considerable layer of realism to the operation of MEBS: an installation may operate using local information, not requiring knowledge of inventory levels elsewhere. Of all the information in a multi-echelon system (economic parameters, policy parameters, inventory levels), inventory levels tend to be the most dynamic and so assuming constant knowledge of these levels throughout the channel is unrealistic. MEBS is optimal under limited circumstances but numerical experiments demonstrate that MEBS performs superbly for channels longer than two installations, details of which are described in Section 3.3.1.2. Kapuściński and Parker (2022) also show this local information result is preserved for channels where the most limiting capacity resides in locations other than the retailer and the system operates a MEBS-like policy.

The third group looks at the structure of policies when advance information is available. Özer and Wei (2004) consider inventory control with limited capacity, when some advanced information (forecast about future demand) is available. A similar result is described for uncertain capacity by Hu et al. (2003).

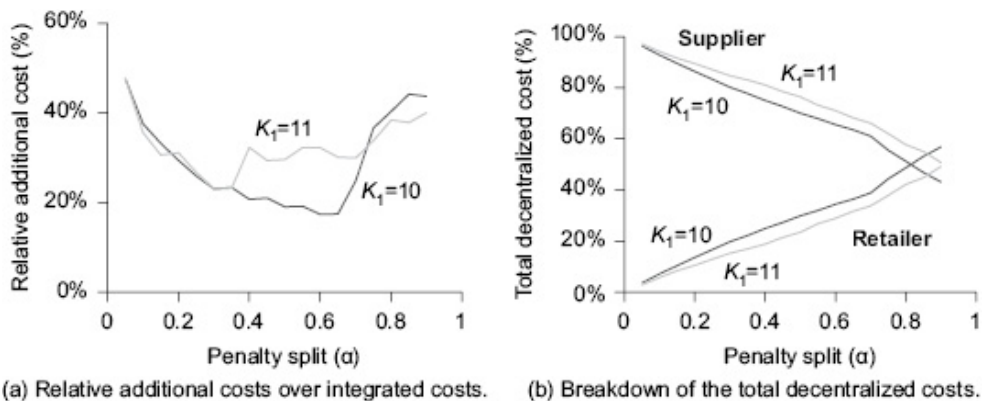
3.4.3 Vendor-Managed Inventory

The literature on Vendor-Managed Inventory identifies a number of factors where putting the decision rights about retailer inventory into supplier's (vendor's) hands may actually benefit all parties. The considered factors include better forecasting (Aviv & Federgruen, 1998), coordination of delivery (Çetinkaya & Lee, 2000), or better use of limited production (capacity). Coordination of delivery and limited production bear some similarities. Çetinkaya and Lee allow one truck to deliver goods to multiple retailers effectively deciding how to use limited capacity to satisfy their needs. Fry et al. (2001) consider a facility that produces a limited quantity, that needs to be sent to the retailer as needed, and limits how much of the product can be sent to the retailer when supply gets limited.

3.4.4 Competition

There has been limited research where competition has been injected into an inventory setting with capacity limits. Parker and Kapuściński (2011) address a two-stage serial system, analogous to Parker and Kapuściński (2004), with the key difference that the supplier and retailer are independently owned and operated. Cachon and Zipkin (1999) considered a similar channel absent capacity constraints and established the equilibrium solution, borrowing from Federgruen and Zipkin's (1984) infinite-horizon closed-form characterization of Clark and Scarf's (1960) decomposition. As already noted, there is no closed-form solution of the infinite-horizon decomposition solution of Parker and Kapuściński (2004), so Parker and Kapuściński (2011) engage a Markov game in a finite-horizon model where each firm incurs local holding costs but shares the stockout cost at the retailer, as per Pasternack (1985) and Cachon and Zipkin (1999). A Markov game is a dynamic game where each (discrete) period depends on state variables. The methodology is proof by induction albeit with the additional challenge that the existence of an equilibrium must be shown in every period. If the equilibrium is unique, the value functions and the associated equilibrium policy are well formed. If there are multiple equilibria, a determination must be made as to whether any may be selected with justification.² Parker and Kapuściński (2011) demonstrate the equilibria are Pareto improving, so a single equilibrium is selected from a continuum of equilibria, validating the value functions.

The equilibrium policy is of a modified echelon base-stock form, as in Parker and Kapuściński (2004), so comparisons between the centralized and decentralized solutions are made. The "band" restriction established by Parker and Kapuściński (2004) applies to the competitive scenario, too. The decentralized echelon base-stock levels were lower than those of the integrated system, consistent with double marginalization. The decentralized MEBS policy appears to perform well for a range of parameters. Perhaps the most interesting result is that there are examples where a tighter retailer capacity can yield greater efficiency (closer to the integrated solution). As seen in Figure 3.7a, the decentralized system costs (relative to first



Source: Parker and Kapuściński (2011).

Figure 3.7 The effect of the penalty split (α) for $K_1 = 10$ and $K_1 = 11$

best) are lower for $K_1 = 10$ than for $K_1 = 11$ for some values of α (the retailer's portion of the stockout cost). One interpretation of this observation is that the tighter capacity binds the two firms closer together when they would otherwise wish to act in diverging directions. When the system decentralized costs are separated to the two firms, as seen in Figure 3.7b, it seems clear the supplier's share of system costs reduces for the tighter capacity whereas the retailer's share of the costs increases. The explanation for this could be that when the capacity is less tight the retailer is not constraining the system as much as before and she should not carry as much of the cost burden.

3.5 CONCLUDING REMARKS

Despite the long history of research into inventory management going back to the 1950s, the research related to capacity-limited inventory management is much more recent, with the landmark results of Federgruen and Zipkin (1986a, 1986b) published in the 1980s. Since that time, our understanding of the effect of capacity upon inventory has developed substantially, extending to various system configurations and economic circumstances. Moreover, as economies and organizations become more competitive and emphasize running more “lean,” appreciating the relationship between inventory and capacity has never been more important.

So far the breakthroughs in characterizing optimal results has been limited to single installations under some cost assumptions, or limited forms of various network configurations. Results such as Janakiraman and Muckstadt (2009) suggest the likelihood of parsimonious optimal policy structure is low, as the optimal policy for serial systems, for example, appear to evade easy characterization. This suggests the more promising directions will involve the development of approximation policies and demonstrating their performance.

NOTES

1. They use a “single-unit, single-customer” approach, described more deeply in Muharremoglu, Geng, and Yang (Chapter 6 in this volume).
2. See Olsen and Parker (2014) for a pertinent discussion of these issues in the inventory context.

REFERENCES

- Allon, G., & Zeevi, A. (2011). A note on the relationship among capacity, pricing, and inventory in a make-to-stock system. *Production and Operations Management*, 20(1), 143–151.
- Angelus, A., & Zhu, W. (2013). On the structure of capacitated assembly systems. *Operations Research Letters*, 41, 19–26.
- Angelus, A., & Zhu, W. (2017). Looking upstream: Optimal policies for a class of capacitated multi-stage inventory systems. *Production and Operations Management*, 26(11), 2071–2088.
- Atalı, A., & Özer, Ö. (2012). Stochastic multi-item inventory systems with markov-modulated demands and production quantity requirements. *Probability in the Engineering and Informational Sciences*, 26(2), 263–293.
- Aviv, Y., & Federgruen, A. (1998). The operational benefits of information sharing and vendor managed inventory (VMI) programs. *Olin School of Business Working Paper*.

- Axsäter, S., & Rosling, K. (1993). Installation vs. echelon stock policies for multilevel inventory control. *Management Science*, 39(10), 1274–1280.
- Bielecki, T., & Kumar, P. R. (1988). Optimality of zero-inventory policies for unreliable manufacturing systems. *Operations Research*, 36(4), 532–541.
- Bollapragada, R., Rao, U. S., & Zhang, J. (2004). Managing inventory and supply performance in assembly systems with random supply capacity and demand. *Management Science*, 50(12), 1729–1743.
- Boyacı, T., & Özer, O. (2010). Information acquisition for capacity planning via pricing and advance selling: When to stop and act? *Operations Research*, 58(5), 1328–1349.
- Bradley, J. R., & Glynn, P. W. (2002). Managing capacity and inventory jointly in manufacturing systems. *Management Science*, 48(2), 273–288.
- Cachon, G. P., & Fisher, M. (2000). Supply chain inventory management and the value of shared information. *Management Science*, 46(8), 1032–1048.
- Cachon, G. P., & Zipkin, P. H. (1999). Competitive and cooperative inventory policies in a two-stage supply chain. *Management Science*, 45(7), 936–953.
- Cetinkaya, S., & Lee, C.-Y. (2000). Stock replenishment and shipment scheduling for vendor-managed inventory systems. *Management Science*, 46(2), 217–232.
- Chan, E. W., & Muckstadt, J. A. (1999). The effects of load smoothing on inventory levels in a capacitated production and inventory system. Tech. rep., Cornell University working paper. 41 pages.
- Chen, F. (1998). Echelon reorder points, installation reorder points, and the value of centralized demand information. *Management Science*, 44(12), S221–S234.
- Ciarallo, F. W., Akella, R., & Morton, T. E. (1994). A periodic review, production planning model with uncertain capacity and uncertain demand – Optimality of extended myopic policies. *Management Science*, 40(3), 320–332.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- de Véricourt, F., Karaesmen, F., & Dallery, Y. (2001). Assessing the benefits of different stock-allocation policies for a make-to-stock production system. *Manufacturing & Service Operations Management*, 3(2), 105–121.
- de Véricourt, F., Karaesmen, F., & Dallery, Y. (2002). Optimal stock allocation for a capacitated supply system. *Management Science*, 48(11), 1486–1501.
- DeCroix, G., & Arreola-Risa, A. (1998). Optimal production and inventory policy for multiple products under resource constraints. *Management Science*, 44(7), 950–961.
- Demirel, S., Kapuściński, R., & Yu, M. (2018). Strategic behavior of suppliers in the face of production disruptions. *Management Science*, 64(2), 533–551.
- Deshpande, V., Cohen, M. A., & Donohue, K. (2003). A threshold inventory rationing policy for service-differentiated demand classes. *Management Science*, 49(6), 683–703.
- Evans, R. V. (1967). Inventory control of a multiproduct system with a limited production resource. *Naval Research Logistics Quarterly*, 14(2), 173–184.
- Evans, R. V. (1968). Sales and restocking policies in a single item inventory system. *Management Science*, 14(7), 463–472.
- Fair, R. C. (1989). The production-smoothing model is alive and well. *Journal of Monetary Economics*, 24(3), 353–370.
- Federgruen, A., & Heching, A. (1999). Combined pricing and inventory management under uncertainty. *Operations Research*, 47(3), 454–475.
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836.
- Federgruen, A., & Zipkin, P. (1986a). An inventory model with limited production capacity and uncertain demands i. *The Average-Cost Criterion*, 11(2), 193–207.
- Federgruen, A., & Zipkin, P. (1986b). An inventory model with limited production capacity and uncertain demands ii. *The Discounted-Cost Criterion*, 11(2), 208–215.
- Fisher, M., & Raman, A. (1996). Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research*, 44(1), 87–99.
- Fisher, M., Hammond, J., Obermeyer, W., & Raman, A. (1997). Configuring a supply chain to reduce the cost of demand uncertainty. *Production and Operations Management*, 6(3), 211–225.

- Fry, M. J., Kapuściński, R., & Olsen, T. L. (2001). Coordinating production and delivery under a (z, Z)-type vendor-managed inventory contract. *Manufacturing & Service Operations Management*, 3(2), 151–173.
- Fu, M. C. (1994). Sample path derivatives for inventory systems. *Operations Research*, 42(2), 351–364.
- Fukuda, Y. (1964). Optimal policies for the inventory problem with negotiable leadtime. *Management Science*, 10(4), 690–708.
- Gallego, G., & Scheller-Wolf, A. (2000). Capacitated inventory problems with fixed order costs: Some optimal policy structure. *European Journal of Operational Research*, 126, 603–613.
- Gavirneni, S. (2002). Information flows in capacitated supply chains with fixed ordering costs. *Management Science*, 48(5), 644–651.
- Gavirneni, S., Kapuściński, R., & Tayur, S. (1999). Value of information in capacitated supply chains. *Management Science*, 45(1), 16–24.
- Glasserman, P. (1996). Allocating production capacity among multiple products. *Operations Research*, 44(5), 724–734.
- Glasserman, P., & Tayur, S. R. (1994). The stability of a capacitated, multi-echelon production-inventory system under a base-stock policy. *Operations Research*, 42(5), 913–925.
- Glasserman, P., & Tayur, S. R. (1995). Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Science*, 41(2), 263–281.
- Glasserman, P., & Tayur, S. R. (1996). A simple approximation for a multistage capacitated production-inventory system. *Naval Research Logistics*, 43, 41–58.
- Ha, A. Y. (1997a). Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43(8), 1093–1103.
- Ha, A. Y. (1997b). Optimal dynamic scheduling policy for a make-to-stock production system. *Operations Research*, 45(1), 42–53.
- Ha, A. Y. (2000). Stock rationing in an make-to-stock queue. *Management Science*, 46(1), 77–87.
- Hu, X., Duenyas, I., & Kapuściński, R. (2003). Advance demand information and safety capacity as a hedge against demand and capacity uncertainty. *Manufacturing & Service Operations Management*, 5(1), 55–58.
- Hu, X., Duenyas, I., & Kapuściński, R. (2007). Existence of coordinating transshipment prices in a two-location inventory model. *Management Science*, 53(8), 1289–1302.
- Hu, X., Duenyas, I., & Kapuściński, R. (2008). Optimal joint inventory and transshipment control under uncertain capacity. *Operations Research*, 56(4), 881–897.
- Huh, W. T., & Janakiraman, G. (2010). Base-stock policies in capacitated assembly systems: Convexity properties. *Naval Research Logistics*, 57, 109–118.
- Huh, W. T., Janakiraman, G., & Nagarajan, M. (2010). Capacitated serial inventory systems: Sample path and stability properties under base-stock policies. *Operations Research*, 58(4), 1017–1022.
- Huh, W. T., Janakiraman, G., & Nagarajan, M. (2016). Capacitated multiechelon inventory systems: Policies and bounds. *Manufacturing & Service Operations Management*, 18(4), 570–584.
- Iglehart, D., & Karlin, S. (1962). Optimal policy for dynamic inventory process with non-stationary stochastic demands. In K. J. Arrow, S. Karlin, & H. Scarf (Eds.), *Studies in applied probability and management science* (pp. 127–147). Stanford, CA: Stanford University Press.
- Iyer, A. V., & Bergen, M. E. (1997). Quick response in manufacturer-retailer channels. *Management Science*, 43(4), 559–570.
- Janakiraman, G., & Muckstadt, J. A. (2009). A decomposition approach for a class of capacitated serial systems. *Operations Research*, 57(6), 1384–1393.
- Kamesam, P. V., & Tayur, S. R. (1993). Algorithms for the analysis of multi-stage capacitated assembly systems. *IBM Research Report RC*, 8971, 45.
- Kaplan, A. (1969). Stock rationing. *Management Science*, 15(5), 260–267.
- Kapuściński, R. (1996). Analysis of capacitated manufacturing systems facing stochastic demands. Ph.D. thesis, Carnegie Mellon University.
- Kapuściński, R., & Parker, R. P. (2022). Conveying demand information in serial supply chains with capacity limits. *Operations Research*, 70(3): 1485–1505.
- Kapuściński, R., & Tayur, S. R. (1998). A capacitated production-inventory model with periodic demand. *Operations Research*, 46(6), 899–911.

- Karlin, S., & Fabens, A. (1959). A stationary inventory model with Markovian demand. *Mathematical Methods in the Social Sciences*, 159–175.
- Katircioglu, K. (1996). Dissertation title. Ph.D. thesis, University of British Columbia.
- Kimemia, J., & Gershwin, S. B. (1983). An algorithm for the computer control of a flexible manufacturing system. *AIIE Transactions*, 15(4), 353–362.
- Krane, S. D., & Braun, S. (1991). Production smoothing evidence from physical-product data. *Journal of Political Economy*, 99(3), 558–581.
- Krankel, R. M., Duenyas, I., & Kapuściński, R. (2006). Timing successive product introductions with demand diffusion and stochastic technology improvement. *Manufacturing & Service Operations Management*, 8(2), 119–135.
- Krishnan, H., Kapuściński, R., & Butz, D. A. (2010). Quick response and retailer effort. *Management Science*, 56(6), 962–977.
- Lederer, P. J., & Li, L. (1997). Pricing, production, scheduling, and delivery-time competition. *Operations Research*, 45(3), 407–420.
- Liu, X., & Tu, Y. (2008). Production planning with limited inventory capacity and allowed stockout. *International Journal of Production Economics*, 111(1), 180–191.
- Luss, H. (1982). Operations research and capacity expansion problems: A survey. *Operations Research*, 30(5), 907–947.
- Manne, A. S. (1967). *Investments for capacity expansion: Size, location and time-phasing*. MIT Press.
- Mieghem, J. V. (1998). Investment strategies for flexible resources. *Management Science*, 44(8), 1071–1078.
- Muharremoglu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, 56(5), 1089–1103.
- Muharremoglu, A., Geng, X., & Yang, N. (2022). Single-unit analysis. In *Research Handbook of Inventory Management*. Edward Elgar.
- Nahmias, S., & Schmidt, C. P. (1984). An efficient heuristic for the multi-item newsboy problem with a single constraint. *Naval Research Logistics Quarterly*, 31(3), 463–474.
- Olsen, T. L., & Parker, R. P. (2014). On Markov equilibria in dynamic inventory competition. *Operations Research*, 62(2), 332–344.
- Özer, Ö., & Wei, W. (2004). Inventory control with limited capacity and advance demand information. *Operations Research*, 52(6), 988–1000.
- Parker, R. P., & Kapuściński, R. (2004). Optimal policies for a capacitated two-echelon inventory system. *Operations Research*, 52(5), 739–755.
- Parker, R. P., & Kapuściński, R. (2011). Managing a noncooperative supply chain with limited capacity. *Operations Research*, 59(4), 866–881.
- Pasternack, B. A. (1985). Optimal pricing and return policies for perishable commodities. *Marketing Science*, 4(2), 166–176.
- Prabhu, N. U. (1965). *Queues and inventories*. Wiley, John.
- Rajagopalan, S. (1998). Capacity expansion and equipment replacement: A unified approach. *Operations Research*, 46(6), 846–857.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579.
- Shang, K. H., & Song, J.-S. (2003). Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science*, 49(5), 618–638.
- Shaoxiang, C. (2004). The infinite horizon periodic review problem with setup costs and capacity constraints: A partial characterization of the optimal policy. *Operations Research*, 52(3), 409–421.
- Shaoxiang, C., & Lambrecht, M. (1996). X-y band and modified policy. *Operations Research*, 44(6), 1013–1019.
- Song, J.-S., van Houtum, G.-J., & Mieghem, J. A. V. (2020). Capacity and inventory management: Review, trends, and projections. *Manufacturing & Service Operations Management*, 22(1), 36–46.
- Speck, C. J., & van der Wal, J. (1991). The capacitated multi-echelon inventory system with serial structure: 1. the “push ahead”-effect. *Memorandum COSOR 91-39*. Eindhoven University of Technology.
- Tayur, S. R. (1993). Computing the optimal policy for capacitated inventory models. *Communications Statistics: Stochastic Models*, 9(4), 585–598.

- Topkis, D. M. (1968). Optimal ordering and rationing policies in a nonstationary dynamic inventory model with demand classes. *Management Science*, 15(3), 160–176.
- Veatch, M. H., & Wein, L. M. (1994). Optimal control of a two-station production/inventory system. *Operations Research*, 42(2), 337–350.
- Wang, T., & Toktay, B. L. (2008). Inventory management with advance demand information and flexible delivery. *Management Science*, 54(4), 716–732.
- Wijngaard, J. (1972). An inventory problem with constrained order capacity. *Technical University Eindhoven Report 72-WSK-03*.
- Wijngaard, J. (1975). Stationary Markovian decision problems: Discrete time, general state space. Ph.D. thesis, Technical University Eindhoven.
- Yaged, B. (1973). Minimum cost routing for dynamic network models. *Networks*, 3(3), 193–224.
- Yang, J., Qi, X., & Xia, Y. (2005). A production-inventory system with Markovian capacity and outsourcing option. *Operations Research*, 53(2), 328–349.
- Yu, M., Kapuściński, R., & Ahn, H. S. (2015). Advance selling: Effects of interdependent consumer valuations and seller's capacity. *Management Science*, 61(9), 2100–2117.
- Zhao, H., Deshpande, V., & Ryan, J. K. (2005). Inventory sharing and rationing in decentralized dealer networks. *Management Science*, 51(4), 531–547.
- Zipkin, P. H. (2000). *Foundations of inventory management*. McGraw Hill.

4. Generalizations of the Clark–Scarf model and analysis

Alexandar Angelus

4.1 INTRODUCTION

More than 60 years have passed since the publication of the classic Clark and Scarf (1960) multiechelon inventory paper. This paper was the first to solve the problem of determining optimal stocking levels in a centralized, multi-period, multi-stage inventory system with stochastic demand. The seminal contribution of Clark and Scarf was to identify the structure of the optimal inventory policy that allows the objective cost function for this problem, whose dimensionality equals the number of stages in the system, to be expressed as a sum of component cost functions that each depend on only a single state variable. Owing to the convex nature of the cost functions involved, this decomposition of the objective function resolved the curse of dimensionality inherent in the problem, thus rendering it analytically and numerically tractable. Clark and Scarf’s solution procedure relied on a change of variables for the problem from installation stocks (i.e., the amount of stock at each stage) to “echelon stocks”, where each echelon stock represents the sum of stock at a particular stage plus stock in-transit to, or on hand, at all stages downstream of it. In that manner, the term “echelon” has become synonymous with “stage”. The resulting additive decomposition of a multi-dimensional objective cost function into single-dimensional component functions is referred to as *the Clark–Scarf decomposition*.

Since its appearance, the landmark Clark and Scarf (1960) paper has stimulated numerous extensions, generalizations, and applications of their original model. The field of study initiated by their work, referred to as *multiechelon inventory theory*, remains an active research area today. One reason for this is the pervasiveness of multi-stage inventory systems in the contemporary global economy. When supply chains are set up to distribute a finished product over large geographical areas, local stocking points close to the customers in different areas are needed. (In Section 4.2.3, such systems are referred to as *logistics supply chains*.) These local stocking points (also known as *installations*) may be replenished from a distribution center close to the production facility. In production-focused systems, referred to later in this chapter as *product-transforming supply chains*, stocks of raw materials and/or components are transformed into finished products through multiple discrete stages. Both types of supply chains can be modeled as multiechelon systems. For that reason, the management of multiechelon inventory systems has become a crucial part of supply chain operations in practice and supply chain research in the literature. A recent review of stochastic multiechelon inventory models by de Kok et al. (2018) provides a classification of some 400 research papers on the topic, a notable percentage of which was published in the last 20 years. More recent research in this area has been motivated by innovations in information technologies (e.g., RFID, advance demand information systems) and supply chain practices (e.g., drop-shipping,

reverse logistics, secondary markets) that have dramatically increased the technical possibilities for improved matching of supply with demand.

The early work in multiechelon inventory theory was predominantly concerned with exact models and deriving structural properties of optimal inventory policies. As the complexity of the models analyzed in the literature grew over time, establishing the Clark–Scarf decomposition became increasingly difficult, so that subsequent research focused mostly on algorithms for decision support tools and approximate and numerical studies. The multiechelon problems addressed in the literature recently, however, demonstrate a resurgence of interest in structural properties and the development of more sophisticated analytical tools needed to derive them. This chapter reviews some of those more recent papers whose primary, though not necessarily exclusive, emphasis is on deriving the structure of optimal policies.

Each paper discussed in this chapter represents a generalization of the original Clark–Scarf model along one or more of the following two broad dimensions: (1) *extra decisions* allowed at each installation (e.g., Lawson & Porteus, 2000; Angelus & Özer, 2021); and (2) *more general network structure* of the supply chain (e.g., Angelus & Porteus, 2008; Angelus & Özer, 2016). At the same time, every paper reviewed in this section has the following basic features in common with the Clark and Scarf (1960) model: (1) discrete-time; (2) finite-horizon of T periods in length; (3) salvage value function that is linear in the state variables; and (4) stochastic customer demand at the most downstream installation in the system.

The ultimate objective of each generalization of the Clark–Scarf model addressed herein is to identify the structure of the optimal policy that allows for the Clark–Scarf decomposition of the objective cost function. While this objective is not always achievable due to the complexity of the systems being considered, the efforts invested in *trying* to reach that objective have yielded novel results, insights, and solution approaches, while opening new research avenues. Even the structure of the Clark–Scarf decomposition, as the original method of analysis, has seen itself generalized in some recent work in this area (e.g., Shen et al., 2022).

The remainder of this chapter is organized as follows. Section 4.2 presents the generalizations of the Clark and Scarf model that facilitate the matching of supply with (stochastic) demand by allowing for additional product flows at each installation. One example of such additional product flows in the system is the expediting of inventory items, introduced to the literature by Lawson and Porteus (2000). Expedited orders help the system better deal with excess demand by making it possible for inventory to flow downstream faster than would otherwise be possible under the regularly scheduled orders. On the other hand, the reverse flow of products considered in Angelus and Özer (2021) facilitates dealing with excess inventory by returning it back upstream where it is cheaper to hold. The lateral flow of products out of the system discussed in Angelus (2011) represents another supply chain strategy for managing excess inventory by allowing its disposal at any installation in the system.

Section 4.3 addresses more general network structures in the form of assembly systems that represent supply chains in which multiple components are assembled at multiple installations into subassemblies. The final assembly into a finished, end-product takes place at the most downstream installation, where the finished products are used to satisfy stochastic customer demand. The challenge of managing assembly supply chains lies in the very large number of state and decision variables required to describe such systems, and the resulting curse of dimensionality that renders the numerical solution of assembly systems next to impossible, even for very small-scale supply chains. Rosling (1989) was the first paper to show how an assembly system generalization can be reduced to an equivalent Clark and Scarf multiechelon

model and thus freed of the curse of dimensionality. The subsequent papers in this area (e.g., DeCroix & Zipkin, 2005; Angelus & Porteus, 2008; Angelus & Zhu, 2013; DeCroix, 2013) address more complex supply chain effects and product dynamics in assembly systems. Some of those papers have also had to develop novel methods of analysis to derive similar equivalence results. In this section, those methods are discussed in some detail. Finally, Section 4.4 provides concluding remarks and discussion of some open research problems in multiechelon inventory theory.

4.2 GENERALIZATIONS TO MULTIPLE FLOWS OF PRODUCT

This section addresses those generalizations of the Clark and Scarf (1960) model that, in addition to their original regular flow of orders, also allow for expedited, lateral, and/or reverse flows of products in the system.

4.2.1 Systems with Expediting

Expediting of inventory is a common and important supply chain practice in industry. A majority of managers from European divisions of medium- to large-sized manufacturing companies resort to expediting to avoid back orders (Özsen & Thonemann, 2014). Expediting has also become an indispensable supply chain capability for online retailers. In general, expediting products represents a strategy for dealing with excess demand in the system, as it allows the products to flow through the supply chain toward the point of demand faster than would otherwise be possible.

A key distinction in the literature on inventory problems with expediting is whether expediting is allowed into *every* installation or *only* into the most downstream one, as those two expediting capabilities end up being modeled quite differently.

4.2.1.1 Expediting at every installation

Product expediting in the context of a multiechelon system was introduced by Lawson and Porteus (2000). In their model, each of N installations in the system can initiate both regular orders (identical to those considered in the original Clark & Scarf (1960) model) and expedited orders for products from the next installation upstream. Regular-ordered items require one period to move between any two adjacent installations in the system. The expedited flow allows an item to move downstream through the entire supply chain (or any portion thereof) faster, possibly within the same period. In what follows, x_{jt} denotes on-hand inventory at installation $j > 1$ in period t (and net inventory at installation $j = 1$); X_{jt} denotes the (positive) number of units regular-ordered into installation j in period t ; and X_{jt}^E denotes the (positive) number of units expedited into installation j from installation $j + 1$, in period t . Let $\mathbf{x}_t := \{x_{1t}, \dots, x_{Nt}\}$; $\mathbf{X}_t := \{X_{1t}, \dots, X_{Nt}\}$; and $\mathbf{X}_t^E := \{X_{1t}^E, \dots, X_{Nt}^E\}$. Hence, vectors $\mathbf{X}_t$ and $\mathbf{X}_t^E$ represent the decision variables in the model.

The sequence of events in each period t is as follows: (1) units regularly ordered in the previous period arrive at each installation; (2) inventory state $\mathbf{x}_t$ is observed; (3) expedited orders $\mathbf{X}_t^E$ are placed, starting at the most upstream installation and moving downstream (a received expedited order at each installation becomes available to be expedited further down the supply chain within the same period); (4) regular orders $\mathbf{X}_t$ are placed; (5) customer demand is

observed and satisfied by the available stock; (6) costs are incurred; and (7) regular orders depart their installations of origin. This sequence of events defines the feasible constraints as $X_{jt}^E \leq x_{j+1,t} + X_{j+1,t}^E$ and $X_{jt} \leq x_{j+1,t} + X_{j+1,t}^E - X_{jt}^E$, for $j \in [1, N-1]$ at each installation j .

The state transition equations thus become

$$x_{j,t+1} = \begin{cases} x_{1t} + X_{1t}^E + X_{1t} - D_t & \text{if } j = 1, \\ x_{jt} + X_{jt}^E - X_{j-1,t}^E + X_{jt} - X_{j-1,t} & \text{if } j > 1. \end{cases}$$

As in the original Clark and Scarf model, there is unit holding cost H_j charged on inventory at installation j in each period, and the backlogging cost p is charged to each unit of unsatisfied demand backlogged in each period. Each unit regular-ordered into installation j incurs positive cost k_j , and each unit expedited into installation j incurs positive cost k_j^E . Expedited orders cost more than regular ones; thus, $k_j^E \geq k_j$ for all j . This formulation of the expediting function captures a variety of supply chain settings: from those in which inventory can be expedited from the outside supplier all the way to the most downstream installation within a single period (i.e., the practice of “drop-shipping”, observed in a number of industries), to those in which an expedited product may take several periods (though still fewer than through regular ordering) to do the same, to those in which no expediting can take place whatsoever, and all other supply chain settings in between. Based on the assumed unit cost structure, the one-period cost function is given by

$$\gamma_t(x_{1t} + X_{1t}^E) + \sum_{j=1}^N (k_j^E X_{jt}^E + k_j X_{jt}) + \sum_{j=2}^N H_j (x_{jt} - X_{j-1,t}^E + X_{jt}^E),$$

where $\gamma_t := (p + H_1) + \mathbb{E}[(D_t - x)^+]$, with the expectation being taken over the random demand variable D_t . The problem can now be reformulated using the following echelon variables: $y_{jt} := x_{1t} + \dots + x_{jt}$; $Y_{jt}^E := y_{jt} + X_{jt}^E$; and $Y_{jt} := Y_{jt}^E + X_{jt}$. The one-period cost function now becomes

$$\gamma_t(Y_{1t}^E) + \sum_{j=1}^N (k_j^E - k_j) Y_{jt}^E + \sum_{j=1}^N (h_j + k_j) Y_{jt} - \sum_{j=1}^N k_j^E y_{jt}.$$

Let $\mathbf{y}_t := \{y_{1t}, \dots, y_{Nt}\}$, $\mathbf{Y}_t := \{Y_{1t}, \dots, Y_{Nt}\}$, and $\mathbf{Y}_t^E := \{Y_{1t}^E, \dots, Y_{Nt}^E\}$. Thus, $\mathbf{Y}_t^E$ and $\mathbf{Y}_t$ are the new decision variables of the model. The transition equations become $y_{j,t+1} = Y_{jt} - D_t$. The feasible decision set for the multiechelon inventory problem with expediting, denoted as $\mathbb{D}_{EXP}(\mathbf{y}_t)$, can be written as:

$$\mathbb{D}_{EXP}(\mathbf{y}_t) = \{\mathbf{Y}_t^E, \mathbf{Y}_t \mid y_{jt} \leq Y_{jt}^E \leq Y_{jt} \leq Y_{j+1,t}^E, 1 \leq j \leq N-1\}.$$

For this multiechelon problem with expediting, let $F_t^E(\mathbf{y}_t)$ denote the minimum expected present value of the costs over periods t through T , as of the beginning of period t , as a function of the echelon state $\mathbf{y}_t$. Thus, $F_t^E(\mathbf{y}_t)$ is the objective cost function for the problem. The resulting optimality equations are:

$$F_t^E(\mathbf{y}_t) = -\sum_{j=1}^N k_j^E y_{jt} + \min_{\mathbf{Y}_t^E, \mathbf{Y}_t \in \mathbb{D}_{EXP}} \left[\gamma_t(Y_{1t}^E) + \sum_{j=1}^N (k_j^E - k_j) Y_{jt}^E + \sum_{j=1}^N (h_j + k_j) Y_{jt} + \alpha \mathbf{E}[F_{t+1}^E(\mathbf{Y}_t - D_t)] \right].$$

Theorem 4.1 (Lawson & Porteus, 2000) *In each period t , there exist base-stock levels (S_{jt}, S_{jt}^E) at each echelon j such that the policy $(\mathbf{Y}^E, \mathbf{Y}_t)$ given recursively by*

$$\begin{aligned} Y_{jt}^E &= \begin{cases} y_{jt} \vee [S_{jt}^E \wedge Y_{j+1,t}^E] & \text{if } j < N; \\ S_{Nt}^E \vee y_{Nt} & \text{if } j = N. \end{cases} \\ Y_{jt} &= \begin{cases} Y_{jt}^E \vee [S_{jt} \wedge Y_{j+1,t}^E] & \text{if } j < N; \\ S_{Nt} \vee Y_{Nt}^E & \text{if } j = N. \end{cases} \end{aligned} \quad (4.1)$$

is optimal for $F_t^E(\mathbf{y}_t)$ and achieves its Clark–Scarf decomposition.

The policy described in Theorem 4.1 is referred to as a *top-down echelon base-stock policy* because expedited decisions are made starting at the most upstream echelon, and then moving downstream one echelon at a time while taking, at each echelon, the previously made upstream decisions as given. Once all expedited order decisions have been made, regular-order decisions at each echelon j are made next, taking the optimal expediting decisions as given lower and upper bounds on the feasible decisions at that echelon, and selecting the closest feasible decisions to the corresponding base-stock levels S_{jt} . Because this optimal policy achieves the Clark–Scarf decomposition of the objective cost function, it eliminates the curse of dimensionality and renders the problem analytically and numerically tractable.

4.2.1.2 Expediting into the most downstream installation

A different set of models has been developed to deal with supply chains in which expediting is allowed only into the most downstream installation, as it is non-trivial to prevent expediting between intermediate installations using the model of Lawson and Porteus (2000). The first such model can be found in Kim et al. (2015).

Similar to Kim et al. (2015), Shen et al. (2022) also consider a multiechelon model in which expediting takes place from any upstream installation directly to (only) the most downstream installation in the system. In their model they also allow the backlogs of customer demand to be categorized into $\tau + 2$ classes according to how delayed a customer order is. The backlog class i , for $0 \leq i \leq \tau$, refers to orders that are $\tau + i$ periods late, while any backlog that is not fulfilled by its due date is collected into class -1 . If any backlog of class $i + 1$ is not fulfilled in a period, it transits to class i at the end of that period, while the unfulfilled class -1 backlogs remain in class -1 . For each class j backlog, there is a unit penalty cost p_j per period, where p_{-1} represents the loss of customer satisfaction arising from late delivery, with $p_{-1} \geq \dots \geq p_{\tau-1} \geq p_{\tau} = 0$. Thus,

while each installation $j > 1$ can be interpreted to place only regular orders, installation 1 places both a regular order from installation 2 and expedited orders from all upstream installations. Installation 1 also decides on how to allocate the arriving inventory to $\tau + 2$ backlog classes.

The unit regular-order cost is k_j at installation j , while the cost of expediting from installation j to installation 1 is given by $k_j^E + \sum_{l=1}^{j-1} k_l$. It is assumed that $k_{j+1}^E - k_j^E \geq k_j^E - k_{j-1}^E$ for all j .

Shen et al. (2022) first show that it is optimal to allocate inventory sequentially to demand to tardier customer orders first and expedite inventory from lower installations first. Further, if a positive amount of inventory is expedited from upstream installations to installation 1 in period t , then the total amount of demand fulfilled in period t must be equal to the amount of expedited inventory plus the on-hand inventory at installation 1. Shen et al. (2022) then introduce a property referred to as *decomposable of degree 2* to describe a multi-variable function that can be expressed as a sum of convex cost component functions that each vary with at most two state variables. They establish that the objective cost function for this inventory problem with expediting is decomposable of degree 2 and that the optimal policy is an echelon base-stock policy that depends on the vector of backlog classes in each period.

In a recent paper, Gong and Wang (2021) introduce two new results pertaining to the preservation of additive convexity in multi-stage inventory problems. Using those results, they show that, when there is only a single backlog class in the multiechelon system studied by Shen et al. (2022), the objective cost function is Clark–Scarf decomposable under the optimal policy.

4.2.2 Systems with Lateral Flows of Product

Next, consider a system in which each of N installations is allowed to initiate a lateral flow of products out of the system, as well as the regular flow of orders from the next installation upstream. This lateral flow of products can be interpreted as either a disposal of excess stock in the system or as a secondary market sale of excess inventory. Secondary markets are becoming increasingly important as the means of dealing with the surplus of stock, which can accumulate in the supply chain for a variety of reasons, including volatility and random shifts in demand, the bullwhip effect, and inadequate information and forecasting systems. This is because secondary markets make it possible for supply chains to reduce excess inventory and the associated inventory-holding costs, while also receiving revenue from those markets for their excess inventory.

In what follows, X_{jt} will continue to denote a regular-flow order from installation $j + 1$ into installation j in period t , while $\mathbf{X}_t := \{X_{1t}, \dots, X_{Nt}\}$ is the regular-order schedule in period t . Let X_{jt}^D be the (positive) number of units disposed of at installation j in period t . Define $\mathbf{X}_t^D := \{X_{1t}^D, \dots, X_{Nt}^D\}$.

Figure 4.1 displays states, decisions, and product flows in this system with the lateral flow of products out of the system in the form of secondary market sales.

Random customer demands are allowed to be modulated by an exogenous Markov chain in order to capture the nonstationary nature of customer demands. Markov-modulated demand has often been used in inventory theory to model the influence of an uncertain external environment (Song & Zipkin, 1993; Chen & Song, 2001; Angelus & Porteus, 2008). Thus, there exists a countable Markov chain $\{\omega_t\}$ such that an exogenous state ω_t , determined independently of any decisions, can impact the demand distribution in each period t . I will use $D_t(\omega)$ to highlight that demand D_t depends on Markov state ω .

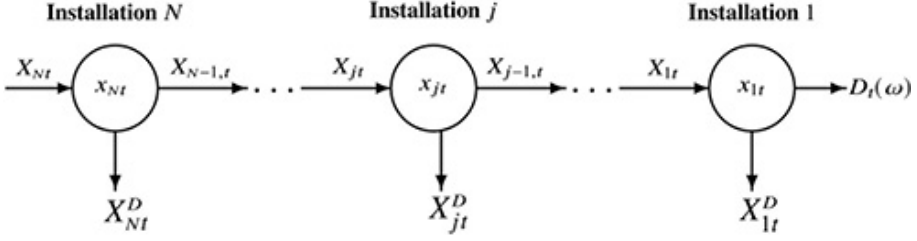


Figure 4.1 Inventory states, decisions, and product flows in a system with secondary market sales

The sequence of events in each period t is as follows: (1) units regular-ordered in the previous period arrive; (2) states x_t and ω_t are observed; (3) the decisions $(\mathbf{X}_t^D, \mathbf{X}_t)$ are made; (4) all units sold in the secondary market are removed from the supply chain; (5) customer demand is realized and satisfied from available inventory; (6) costs and revenues are incurred; and (7) regular orders depart their installations of origin. The feasible constraints become: $X_{1t}^D \leq [x_{1t}]^+$ and $X_{jt}^D \leq x_{jt}$ for $2 \leq j \leq N$; and $X_{jt} \leq x_{j+1,t} - X_{j+1,t}^D$, for $1 \leq j \leq N-1$. The state transition equations are:

$$x_{j,t+1} = \begin{cases} x_{1t} - X_{1t}^D + X_{1t} - D_t(\omega) & \text{for } j=1, \\ x_{jt} - X_{jt}^D + X_{jt} - X_{j-1,t} & \text{for } j \in [1, N]. \end{cases} \quad (4.2)$$

In addition to the customary inventory-holding and backlogging costs, each unit regular-ordered into installation j incurs positive cost k_j , while each unit sold in the secondary market from installation j generates revenue r_j . To prevent speculative ordering, it is common to assume that $k_j \geq r_j - r_{j+1}$. The expected one-period costs in period t can then be expressed as:

$$\gamma_t(\omega, x_{1t} - X_{1t}^D) + \sum_{j=1}^N (k_j X_{jt} - r_j X_{jt}^D) + \sum_{j=2}^N H_j(x_{jt} - X_{jt}^D),$$

where $\gamma_t(\omega, x_{1t} - X_{1t}^D) := (p + H_1) \mathbf{E}_\omega[(D_t(\omega) - x)^+]$, and $\mathbf{E}_\omega$ indicates that the expectation is taken over the random demand variable $D_t(\omega)$ for a given Markov state ω .

The problem can be reformulated using the following echelon variables:

$$y_{jt} := x_{1t} + x_{2t} + \dots + x_{jt}; \quad Y_{jt}^D := y_{jt} - \sum_{i=1}^j X_{it}^D; \quad Y_{jt} := Y_{jt}^D + X_{jt}.$$

Let $\mathbf{y}_t := \{y_{1t}, \dots, y_{Nt}\}$, $\mathbf{Y}_t := \{Y_{1t}, \dots, Y_{Nt}\}$ and $\mathbf{Y}_t^D := \{Y_{1t}^D, \dots, Y_{Nt}^D\}$. The feasible decisions set $\mathbb{D}_{\text{SEC}}(\mathbf{y}_t)$ can then be expressed as:

$$\mathbb{D}_{\text{SEC}}(\mathbf{y}_t) = \left\{ \mathbf{Y}_t^D, \mathbf{Y}_t \mid Y_{1t}^D \leq y_{1t}; \quad Y_{jt}^D \leq Y_{jt} \leq Y_{j+1,t}^D \leq Y_{jt}^D + y_{j+1,t} - y_{jt}; \quad 0 \leq j \leq L-1 \right\}.$$

Let $F_t^D(\omega, \mathbf{y}_t)$ denote the objective cost function for this multiechelon inventory problem with stock disposals, as a function of the Markov state ω and echelon state $\mathbf{y}_t$. The optimality equations for this problem can then be expressed as:

$$F_t^D(\omega, \mathbf{y}_t) = - \sum_{j=1}^N \Delta_j y_{jt} + \min_{\mathbf{Y}_t^D, \mathbf{Y}_t \in \mathbb{D}_{SEC}(\mathbf{y}_t)} v_t^D(\omega, \mathbf{Y}_t^D, \mathbf{Y}_t), \quad (4.3)$$

where $\Delta_j := r_j - r_{j+1}$, and

$$\begin{aligned} v_t^D(\omega, \mathbf{Y}_t^D, \mathbf{Y}_t) &:= \gamma_t(\omega, Y_{1t}^D) \\ &+ \sum_{j=1}^N [(\Delta_j - k_j + h_j) Y_{jt}^D + k_j Y_{jt}] + \alpha \mathbf{E}_\omega[F_{t+1}(\omega_{t+1}, \mathbf{Y}_t - D_t(\omega))], \end{aligned}$$

where the expectation is taken over both D_t and ω_{t+1} , given ω .

Theorem 4.2 (Angelus, 2011) *In each period t , optimal decisions $(\hat{Y}_{jt}^D, \hat{Y}_{jt})$ for the problem given in Equation (4.3) are defined at each echelon j recursively, by means of order-up-to levels S_{jt}^D and S_{jt} that are functions of ω and $\mathbf{y}_t$ as follows:*

$$\begin{aligned} \hat{Y}_{jt}^D(\omega, \mathbf{y}_t) &= \begin{cases} y_{1t} \wedge S_{1t}^D(\omega, \mathbf{y}_t) & \text{for } j = 1; \\ \hat{Y}_{j-1,t}^D(\omega, \mathbf{y}_t) \vee [(\hat{Y}_{j-1,t}^D(\omega) + y_{jt} - y_{j-1,t}) \wedge S_{jt}^D(\omega, \mathbf{y}_t)] & \text{for } j \in [2, N]; \end{cases} \\ \hat{Y}_{jt}(\omega, \mathbf{y}_t) &= \hat{Y}_{jt}^D(\omega, \mathbf{y}_t) \vee [\hat{Y}_{j+1,t}^D(\omega, \mathbf{y}_t) \wedge S_{jt}(\omega, \mathbf{y}_t)] \quad \text{for } j \in [1, N-1]. \end{aligned}$$

Due to the multi-dimensional nature of the constraints in the feasible set $\mathbb{D}_{SEC}(\mathbf{y}_t)$ and the dependence of the optimal order-up-to levels on the entire echelon state $\mathbf{y}_t$, this optimal policy does not, in general, achieve the Clark–Scarf decomposition of the objective cost function $F_t^D(\omega, \mathbf{y}_t)$. As a result, the curse of dimensionality inherent in the objective cost function for this problem remains largely unattenuated.

To address this curse of dimensionality, Angelus (2011) identifies a class of disposal policies that: (i) allow inventory disposal at each installation; and (ii) are determined by base-stock levels that are independent of the (echelon) inventory state, and (iii) achieve the Clark–Scarf decomposition of the objective cost function.

Definition 4.1 (Angelus, 2011) A disposal decisions vector $\mathbf{X}_t^D$ in period t represents a *disposal saturation policy* if there exists an installation k_t in period t such that $X_{jt}^D = 0$ for all $j > k_t$ if $k_t > 1$ and $X_{jt} = 0$ for all $j > k_t$ if $k_t < N$.

Thus, in each period t , a disposal saturation policy disposes of all inventory upstream of some installation k_t , and does not dispose of any inventory downstream of that installation (so that downstream of installation k_t only the regular flow of orders is allowed).

Theorem 4.3 (Angelus, 2011) *For every state ω and period t , there exists a disposal saturation policy specified by a set of ordered base-stock levels that achieves the Clark–Scarf*

decomposition of $F_t^D(\omega, \cdot)$. The resulting optimal regular-order policy follows an echelon base-stock policy.

In particular, the policy described in Theorem 4.3 is specified by a set of ordered echelon base-stock levels $S_{1t}^D(\omega) \geq S_{2t}^D(\omega) \geq \dots \geq S_{Nt}^D(\omega)$ recursively as follows:

$$\hat{Y}_{jt}^D(\omega) = \begin{cases} y_{1t} \wedge S_{1t}^D(\omega) & \text{for } j = 1; \\ \hat{Y}_{j-1,t}^D(\omega) \vee [\hat{Y}_{j-1,t}^D(\omega) + y_{jt} - y_{j-1,t}] \wedge S_{jt}^D(\omega) & \text{for } j \in [2, N], \end{cases}$$

In contrast to the top-down echelon base-stock policy of Lawson and Porteus (2000), this policy is implemented from the bottom up. This is because, in implementing this policy, the first inventory disposal decision made is at echelon 1, followed up by the inventory disposal decision at echelon 2, and so on. Angelus (2011) reports that the resulting performance error for this heuristic policy is less than 1%.

Gong and Wang (2021) address a similar multi-stage inventory problem with disposals and show that, when those disposals are not allowed at the most downstream installation, the objective cost function for the problem is Clark–Scarf decomposable.

4.2.3 Systems with Reverse Logistics

In contrast to forward logistics which deals with the downstream flow of products from the point of origin to the point of consumption, reverse logistics refers to principles and practices for managing the upstream flow of surplus inventory in the form of material, goods, or equipment back through the supply chain (for reuse, resale, or disposal). The value of reverse logistics currently exceeds \$200 billion in the United States alone. There is a growing recognition in both research and practice of the importance of reverse logistics for managing the mismatch between supply and demand.

In this section, we focus on reverse logistics in the form of reverse product flows that originate *within* the supply chain itself and represent a strategy for managing excess inventory (referred to as “overstock inventory” in industry), rather than on those necessitated by customers’ returns of products *into* the supply chain. (The chapter “Inventory models with returns and remanufacturing” in this handbook reviews recent research on inventory systems with customer returns of products. See also DeCroix et al., 2005; DeCroix & Zipkin, 2005).

4.2.3.1 Reverse logistics in logistics supply chains

In what follows, each of N installations in the system can initiate regular orders (identical to those in the original Clark and Scarf (1960) model), expedited orders *from* the next installation upstream (identical to those in Lawson & Porteus, 2000), and reverse orders *into* the next installation upstream.

Let X_{jt}^R denote the (positive) number of units sent upstream by reverse order from installation j into installation $j+1$ in period t . Thus, expedited and regular orders are flowing downstream while reverse orders are flowing upstream. The vector $\mathbf{X}_t^R$ of all reverse orders $\{X_{jt}^R\}$ is the *reverse-order schedule*. Random customer demands are once again modulated by an exogenous Markov chain. The sequence of events in each period t is as follows: (1) units regular- and reverse-ordered in the previous period arrive; (2) inventory state x_t and Markov state ω_t are observed; (3) expedited orders $\mathbf{X}_t^E$ are placed, starting at the most upstream installation

and moving downstream; (4) reverse orders $\mathbf{X}_t^R$ are placed and depart their installation of origin; (5) regular orders $\mathbf{X}_t$ are placed; (6) customer demand is realized and satisfied from available inventory; (7) costs are incurred; and (8) regular orders depart their origins.

The feasible constraints are: for expedited orders, $X_{jt}^E \leq x_{j+1,t} + X_{j+1,t}^E$; for reverse orders, $X_{1t}^R \leq [x_{1t} + X_{1t}^E]^+$ and $X_{jt}^R \leq x_{jt} + X_{jt}^E - X_{j-1,t}^E$ for $j > 1$; and, for regular orders, $X_{jt} \leq x_{j+1,t} - X_{j+1,t}^R + X_{j+1,t}^E - X_{jt}^E$. The state transition equations are:

$$x_{j,t+1} = \begin{cases} x_{1t} - X_{1t}^R + X_{1t}^E + X_{1t} - D_t(\omega) & \text{for } j = 1, \\ x_{jt} - X_{jt}^R + X_{j-1,t}^R + X_{jt}^E - X_{j-1,t}^E + X_{jt} - X_{j-1,t} & \text{for } j \in [2, N]. \end{cases}$$

A unit regular-ordered into installation j incurs (positive) cost k_j ; a unit expedited into installation j incurs (positive) cost k_j^E . Each unit reverse-ordered into installation $j+1$ from installation j incurs cost k_j^R , which is positive at installations 1 through $N-1$. Because excess inventory at installation N can potentially be resold to a local distributor (so that excess inventory that exits the supply chain at installation N can generate revenue), k_N^R is allowed to be negative. To prohibit the bringing of inventory into the supply chain at installation N solely for the purpose of reselling it back upstream in the next period (rather than to satisfy downstream customer demand), $k_N \geq -k_N^R$ is needed. (Otherwise, a supply chain could make unlimited profits by ordering inventory into location N and then immediately reselling it.)

The unit holding cost H_j is charged on each item of inventory at installation j in each period. This cost is composed of the financial and physical holding costs. The financial holding cost is the opportunity cost related to the value of the finished product itself, whereas the physical holding cost is associated with maintaining a unit of inventory and is therefore location-specific. Thus, if any physical transformation of the product were to take place at installation $j-1$, the unit holding cost at installation j for those items reversed-ordered would necessarily be different from those items regular-ordered into installation j (and which have not yet made it to installation $j-1$). Having the unit holding cost be identical for all items at a particular installation, regardless of their trajectory through the system, implies that items that arrive there by upstream, reverse-order flow are physically indistinguishable from those items that arrive into that installation by the downstream flow of regular orders; that is, there is no transformation of the product taking place anywhere in a supply chain (other than possibly at the most upstream installation). The objective of such supply chains, in which there is no transformation of the product, is the transportation of the finished product closer to the customer. In this chapter, such a supply chain is referred to as a *logistics supply chain*. In a logistics supply chain, it is not necessary to keep track of the trajectory of each item, but rather only of its location in any period t . With the reverse flow of orders, the one-period cost function for such a system becomes:

$$\begin{aligned} \gamma_t(\omega, x_{1t} + X_{1t}^E - X_{1t}^R)^+ &+ \sum_{j=1}^N (k_j^E X_{jt}^E + k_j^R X_{jt}^R + k_j X_{jt}) \\ &+ \sum_{j=2}^N H_j (x_{jt} - X_{j-1,t}^E + X_{jt}^E - X_{jt}^R + X_{j-1,t}^R). \end{aligned}$$

The problem can be reformulated using:

$$y_{jt} := x_{1t} + x_{2t} + \cdots + x_{jt}; \quad Y_{jt}^E := y_{jt} + X_{jt}^E;$$

$$Y_{jt}^R := Y_{jt}^E - X_{jt}^R; \quad Y_{jt} := Y_{jt}^R + X_{jt}.$$

In addition to the already defined $\mathbf{y}_t$, $\mathbf{Y}_t$, and $\mathbf{Y}_t^E$, let $\mathbf{Y}_t^R := \{Y_{1t}^R, \dots, Y_{Nt}^R\}$. Using $Y_{N+1}^E := \infty$ and $Y_{N+1}^R := \infty$, the new feasible decision set $\mathbb{D}_{MULT}(\mathbf{y}_t)$ for this problem with regular, expedited, and reverse flows of products can now be written as:

$$\mathbb{D}_{MULT}(\mathbf{y}_t) = \left\{ \mathbf{Y}_t^E, \mathbf{Y}_t^R, \mathbf{Y}_t \left| \begin{array}{ll} y_{jt} \leq Y_{jt}^E \leq Y_{j+1,t}^E & \text{for } j \in [1, N]; \\ [Y_{1t}^E]^- \leq Y_{1t}^R \leq Y_{1t}^E; Y_{j-1,t}^E \leq Y_{jt}^R \leq Y_{jt}^E & \text{for } j \in [2, N]; \\ Y_{jt}^R \leq Y_{jt} \leq Y_{j+1,t}^R - (Y_{jt}^E - Y_{jt}^R) & \text{for } j \in [1, N]. \end{array} \right. \right.$$

Angelus and Özer (2021) prove that either $X_{jt}^E = X_{jt}^R = 0$ or $X_{jt}^R = 0$ (or both) in any period t . Hence, the feasible set $\mathbb{D}_{MULT}(\mathbf{y}_t)$ can be reduced to $\mathbb{D}_{MULT}^*(\mathbf{y}_t)$ given by:

$$\mathbb{D}_{MULT}^*(\mathbf{y}_t) = \left\{ \mathbf{Y}_t^E, \mathbf{Y}_t^R, \mathbf{Y}_t \left| \begin{array}{ll} y_{jt} \leq Y_{jt}^E \leq Y_{j+1,t}^E & \text{for } j \in [1, N]; \\ [Y_{1t}^E]^- \leq Y_{1t}^R \leq Y_{1t}^E; Y_{j-1,t}^E \leq Y_{jt}^R \leq Y_{jt}^E & \text{for } j \in [2, N]; \\ Y_{jt}^R \leq Y_{jt} \leq Y_{j+1,t}^R & \text{for } j \in [1, N]. \end{array} \right. \right.$$

For this multiechelon inventory problem with reverse logistics, let $F_t^R(\omega, \mathbf{y}_t)$ denote the minimum expected present net value of the costs over periods t through T , as of the beginning of period t , as a function of the Markov state ω and echelon state $\mathbf{y}_t$. The optimality equations for this multiechelon inventory problem become

$$F_t^R(\omega, \mathbf{y}_t) = - \sum_{j=1}^N k_j^E y_{jt} + \min_{\mathbf{Y}_t^E, \mathbf{Y}_t^R, \mathbf{Y}_t \in \mathbb{D}_{MULT}^*(\mathbf{y}_t)} v_t(\omega, \mathbf{Y}_t^E, \mathbf{Y}_t^R, \mathbf{Y}_t) \quad (4.4)$$

where

$$v_t^R(\omega, \mathbf{Y}_t^E, \mathbf{Y}_t^R, \mathbf{Y}_t) := \gamma_t(\omega, Y_{1t}^R) + \sum_{j=1}^N \left[(k_j^E + k_j^R) Y_{jt}^E + k_j Y_{jt} - (k_j^R + k_j - h_j) Y_{jt}^R \right]$$

$$+ \alpha \mathbf{E}_\omega [F_{t+1}(\omega_{t+1}, \mathbf{Y}_t - D_t(\omega))]$$

Theorem 4.4 (Angelus & Özer, 2021) *Optimal decisions $(\hat{Y}_{jt}^E(\omega), \hat{Y}_{jt}^R(\omega), \hat{Y}_{jt}(\omega))$ for the problem given in Equation (4.4) are defined, at each echelon j , by means of the base-stock levels $S_{jt}^E(\omega)$, $S_{jt}^R(\omega)$, and $S_{jt}(\omega)$, as follows:*

$$\hat{Y}_{jt}^E(\omega) = \begin{cases} y_{jt} \vee [S_{jt}^E(\omega) \wedge \hat{Y}_{j+1,t}^E] & \text{for } j \in [1, N-1]; \\ y_{Nt} \vee S_{Nt}^E(\omega) & \text{for } j = N. \end{cases}$$

$$\hat{Y}_{jt}^R(\omega) = \begin{cases} [\hat{Y}_{1t}^E]^- \vee [S_{1t}^R(\omega) \wedge \hat{Y}_{1t}^E] & \text{for } j = 1; \\ \hat{Y}_{j-1,t}^E \vee [S_{jt}^R(\omega) \wedge \hat{Y}_{jt}^E] & \text{for } j \in [2, N]. \end{cases}$$

$$\hat{Y}_{jt}(\omega) = \begin{cases} \hat{Y}_{jt}^R \vee [S_{jt}(\omega) \wedge \hat{Y}_{j+1,t}^R] & \text{for } j \in [1, N-1]; \\ \hat{Y}_{Nt}^R \vee S_{Nt}(\omega) & \text{for } j = N. \end{cases}$$

Further, this policy achieves the Clark–Scarf decomposition of $F_t^R(\omega, \mathbf{y}_t)$.

By achieving the Clark–Scarf decomposition of the objective cost function, the policy described in Theorem 4.4 resolves the curse of dimensionality of the problem.

4.2.3.2 Reverse logistics in product-transforming supply chains

Next, we address the problem of regular and reverse flows of products in *product-transforming supply chains*, in which a unit of product is allowed to be physically transformed at each location as it moves downstream in the system. In such supply chains, an item that has reached a certain location will generally have different physical characteristics (and be more valuable) than an item that has not yet reached that location. Hence, in a product-transforming supply chain, different items at the same installation j may be at different stages of completion. This is because some of the items at a particular installation might have arrived there by the reverse flow from downstream installations, in which case their completion will be further along relative to those items that have arrived there by regular or expedited orders from upstream installation. Therefore, the unit inventory-holding cost for any two items at the same installation may also be different. As a result, in a product-transforming supply chain with reverse logistics, it becomes necessary to keep track, in each period, of both the location (i.e., installation) and stage of completion of each unit of inventory. As a result, the number of state and decision variables in a product-transforming supply chain increases with the square of the number of installations in the system, rendering the analysis of such systems considerably more involved. Further, in contrast to logistics supply systems, in a product-transforming supply chain there exist multiple replenishment opportunities at each installation, each with its own distinct unit cost and impact on the inventory state. This is because there may be up to $j+1$ types of inventory located at installation $j+1$, so that there can be up to $j+1$ different types of regular orders that can be placed by installation j in each period t , each with its own cost and feasible constraints.

Let i denote an item's stage of completion, which represents *the most downstream installation* reached up to that point. Thus, the inventory at installation j may consist of items at any stage of completion i , for $i \in [1, j]$, since a product from any installation i , $i < j$, where it reached the stage of completion i , may have been returned upstream into installation j . As a result, to correctly account for different unit holding and ordering costs for items at different stages of completion at the same installation, in this section, I will use the following state and decision variables:

x_{ijt} : On-hand inventory of items at installation j that are at stage of completion i at the beginning of period t , where $j \in [1, \dots, N]$ and $i \in [1, \dots, j]$;

X_{ijt}^R : number of units at installation j and stage of completion i reverse-ordered into installation $j+1$ in period t , where $j \in [1, \dots, N]$ and $i \in [1, \dots, j]$;

X_{ijt} : number of units at installation $j+1$ and stage of completion i reverse-ordered into installation j in period t , where $j \in [1, \dots, N]$ and $i \in [1, \dots, j+1]$ for $j < N$.

At the most upstream installation, only untransformed items (i.e., $i=N$) are brought *into* the supply chain; thus, when $j=N$ only regular orders with $i=N$ are allowed into the system. It follows that, in a product-transforming supply chain with reverse orders, the number of state and decision variables is of order N^2 .

Define $\mathbf{x}_t := \{x_{ijt}\}$, $\mathbf{X}_t^R := \{X_{ijt}^R\}$, and $\mathbf{X}_t := \{X_{ijt}\}$. Let H_{ij} be the unit inventory-holding cost for an item at stage of completion i at installation j , and k_{ij}^R be the unit cost of returning an item at stage of completion i from installation j to installation $j+1$. Let k_{ij} be the unit cost of regular ordering an item currently at stage of completion i from installation $j+1$ into installation j . The one-period cost function for the product-transforming supply chain with regular and reverse flows becomes:

$$\gamma_t(\omega, x_{11t} - X_{11t}^R) + \sum_{j=1}^N \sum_{i=1}^j (k_{ij}^R X_{ijt}^R + H_{ij} x_{ijt}) + \sum_{j=1}^{N-1} \sum_{i=1}^{j+1} k_{ij} X_{ijt} + k_{NN} X_{NNt}.$$

For this reverse logistics problem, let $\bar{F}_t(\omega, \mathbf{x}_t)$ denote the minimum expected net present value of the costs over periods t through T , as of the beginning of period t , as a function of the Markov state ω and inventory state $\mathbf{x}_t$. Hence, $\bar{F}_t(\omega, \mathbf{x}_t)$ is the objective cost function for the problem. The optimality equations become:

$$\bar{F}_t(\omega, \mathbf{x}_t) := \min_{\mathbf{X}_t^R, \mathbf{X}_t \in \bar{\mathbb{X}}(\mathbf{x}_t)} \bar{V}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t),$$

where

$$\begin{aligned} \bar{V}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t) = & \gamma_t(\omega, x_{11t} - X_{11t}^R) + \sum_{j=1}^N \sum_{i=1}^j (k_{ij}^R X_{ijt}^R + H_{ij} x_{ijt}) \\ & + \sum_{j=1}^{N-1} \sum_{i=1}^{j+1} k_{ij} X_{ijt} + k_{NN} X_{NNt} + \alpha \mathbf{E}_\omega [\bar{F}_{t+1}(\omega_{t+1}, \mathbf{x}_{t+1})], \end{aligned} \quad (4.5)$$

The state transitions that generate $\mathbf{x}_{t+1}$, the inventory state next period, are:

$$x_{ij,t+1} = \begin{cases} x_{11t} + X_{11t} + X_{21t} - X_{11t}^R - D_t(\omega) & \text{for } j = 1, \\ x_{ijt} + X_{ijt} - X_{i,j-1,t} - X_{ijt}^R + X_{i,j-1,t}^R & \text{for } j \in [2, N-1] \text{ and } i \in [1, j-1]; \\ x_{ijt} + X_{ijt} + X_{j+1,jt} - X_{j,j-1,t} - X_{ijt}^R & \text{for } j \in [2, N-1] \text{ and } i = j; \\ x_{iNt} - X_{i,N-1,t} - X_{iNt}^R + X_{i,N-1,t}^R & \text{for } j = N \text{ and } i \in [1, j-1]; \\ x_{NNt} + X_{NNt} - X_{N,N-1,t} - X_{NNt}^R & \text{for } j = N \text{ and } i = N. \end{cases}$$

The feasible set $\mathbb{D}_{\mathbf{x}}(\mathbf{x}_t)$ for this product-transforming supply chain problem is

$$\mathbb{D}_{\mathbf{x}}(\mathbf{x}_t) = \left\{ \mathbf{X}_t^R, \mathbf{X}_t \left| \begin{array}{ll} X_{11t}^R \in [0, [x_{11t}]^+], X_{ijt}^R \in [0, x_{ijt}] & \text{for } j \in [2, N] \text{ and } i \in [1, j]; \\ X_{ijt} \in [0, x_{i,j+1,t} - X_{i,j+1,t}^R] & \text{for } j \in [1, N-1], \text{ and } i \in [1, j+1]; \end{array} \right. \right.$$

4.2.3.3 A lower-bounding policy

We first derive a lower-bounding policy for this reverse logistics problem.

For each $j \in [1, N]$, let $\tilde{k}_j^R := \min_{i,i \in [1,j]} k_{ij}^R$ and $\tilde{H}_j := \min_{i,i \in [1,j]} H_{ij}$; for each $j \in [1, N-1]$, let $\tilde{k}_j := \min_{i,i \in [1,j+1]} k_{ij}$ with $\tilde{k}_N := k_{NN}$.

Next, let $F_t^L(\omega, \mathbf{x}_t)$ be defined as:

$$\begin{aligned} \bar{F}_t^L(\omega, \mathbf{x}_t) &= \min_{\mathbf{x}_t^R, \mathbf{x}_t \in \mathbb{D}_{\mathbf{x}}(\mathbf{x}_t)} \left\{ \gamma_t(\omega, x_{11t} - X_{11t}^R) + \sum_{j=1}^N \sum_{i=1}^j (\tilde{k}_j^R X_{ijt}^R + \tilde{H}_j x_{ijt}) \right. \\ &\quad \left. + \sum_{j=1}^{N-1} \sum_{i=1}^{j+1} \tilde{k}_j X_{ijt} + \tilde{k}_N X_{NNt} + \alpha \mathbf{E}_{\omega}[\bar{F}_{t+1}^L(\omega_{t+1}, \mathbf{x}_{t+1})] \right\} \\ &= \min_{\mathbf{x}_t^R, \mathbf{x}_t \in \mathbb{D}_{\mathbf{x}}(\mathbf{x}_t)} \left\{ \gamma_t(\omega, x_{11t} - X_{11t}^R) + \sum_{j=1}^N \left[\tilde{k}_j^R \left(\sum_{i=1}^j X_{ijt}^R \right) + \tilde{H}_j \left(\sum_{j=1}^N x_{ijt} \right) \right] \right. \\ &\quad \left. + \sum_{j=1}^{N-1} \tilde{k}_j \left(\sum_{i=1}^{j+1} X_{ijt} \right) + \tilde{k}_N X_{NNt} + \alpha \mathbf{E}_{\omega}[\bar{F}_{t+1}^L(\omega_{t+1}, \mathbf{x}_{t+1})] \right\}. \end{aligned}$$

It follows from the definitions of $\tilde{k}_j$ and $\tilde{k}_j^R$ that $\bar{F}_t^L(\omega, \mathbf{x}_t) \leq \bar{F}_t(\omega, \mathbf{x}_t)$ for any ω and $\mathbf{x}_t$; thus, $\bar{F}_t^L$ is a lower bound for $\bar{F}_t$. For each installation $j \in [1, N]$, define:

$$\begin{aligned} \tilde{x}_{jt} &:= \sum_{j=1}^N x_{ijt}; \quad \tilde{X}_{ijt}^R := \sum_{i=1}^j X_{ijt}^R; \quad \tilde{X}_{Nt} := X_{NNt} \quad \text{and} \\ \tilde{X}_{jt} &:= \sum_{i=1}^{j+1} X_{ijt} \text{ for } j \in [1, N-1]. \end{aligned}$$

Using those state and decisions variables, the optimality equations for $\bar{F}_t^L$ become:

$$\begin{aligned} \bar{F}_t^L(\omega, \tilde{\mathbf{x}}_t) &= \min_{\tilde{\mathbf{x}}_t^R, \tilde{\mathbf{x}}_t \in \mathbb{D}_{\tilde{\mathbf{x}}}^L(\tilde{\mathbf{x}}_t)} \left\{ \gamma_t(\omega, \tilde{x}_{1t} - \tilde{X}_{1t}^R) + \sum_{j=1}^N (\tilde{k}_j^R \tilde{X}_{jt}^R + \tilde{k}_j \tilde{X}_{jt} + \tilde{H}_j \tilde{x}_{jt}) \right. \\ &\quad \left. + \alpha \mathbf{E}_{\omega}[\bar{F}_{t+1}^L(\omega_{t+1}, \tilde{\mathbf{x}}_{t+1})] \right\}, \end{aligned} \tag{4.6}$$

where the feasible decision set $\mathbb{D}_{\mathbf{x}}^L(\tilde{\mathbf{x}}_t)$ is now given by:

$$\mathbb{D}_{\mathbf{x}}^L(\tilde{\mathbf{x}}_t) = \left\{ \tilde{X}_t^R, \tilde{X}_t \left| \begin{array}{ll} \tilde{X}_{1t}^R \in [0, [x_{1t}]^+], \tilde{X}_{jt}^R \in [0, x_{jt}] & \text{for } j \in [2, N]; \\ \tilde{X}_{jt}^R \in [0, x_{j+1,t} - X_{j+1,t}^R] & \text{for } j \in [1, N-1]. \end{array} \right. \right.$$

We now introduce the following set of echelon variables:

$$\tilde{y}_{jt} := \tilde{x}_{1t} + \cdots + \tilde{x}_{jt}; \quad \tilde{Y}_{jt}^R := \tilde{y}_{jt} - \tilde{X}_{jt}^R; \quad \tilde{Y}_{jt} := \tilde{Y}_{jt}^R + \tilde{X}_{jt}.$$

The dynamic program in Equation (4.6) can now be transformed into the following one:

$$\begin{aligned} \tilde{f}_t^L(\omega, \tilde{\mathbf{y}}_t) = & \sum_{j=1}^N (\tilde{k}_j^R + \tilde{k}_j) y_{jt} + \min_{\tilde{\mathbf{Y}}_t^R, \tilde{\mathbf{Y}}_t \in \bar{\mathbb{Y}}^L(\tilde{\mathbf{y}}_t)} \left\{ \gamma_t(\omega, \tilde{\mathbf{Y}}_t^R) + \sum_{j=1}^N (\tilde{k}_j \tilde{Y}_{jt} - (\tilde{k}_j^R + \tilde{k}_j) \tilde{Y}_{jt}^R) \right. \\ & \left. + \alpha \mathbf{E}_{\omega}[\tilde{f}_{t+1}^L(\omega_{t+1}, \tilde{\mathbf{Y}}_t - D_t(\omega))] \right\}, \end{aligned} \quad (4.7)$$

where the feasible decision set is given by:

$$\mathbb{D}_{\mathbf{y}}^L(\tilde{\mathbf{y}}_t) = \left\{ \tilde{\mathbf{Y}}_t^R, \tilde{\mathbf{Y}}_t \left| \begin{array}{ll} \tilde{Y}_{1t}^R \in [[\tilde{y}_{1t}]^-, \tilde{y}_{1t}], \tilde{Y}_{jt}^R \in [\tilde{y}_{j-1,t}, \tilde{y}_{jt}] & \text{for } j \in [2, N]; \\ \tilde{Y}_{jt}^R \in [\tilde{Y}_{jt}^R, \tilde{Y}_{j+1,t}^R - (\tilde{y}_{jt} - \tilde{Y}_{jt}^R)] & \text{for } j \in [1, N]. \end{array} \right. \right.$$

Theorem 4.5 (Angelus & Özer, 2021) *For each period t , state ω , and echelon j there exist base-stock levels $\tilde{S}_{jt}^R(\omega)$ and $\tilde{S}_{jt}(\omega)$, where $\tilde{S}_{jt}^R(\omega) \geq \tilde{S}_{jt}(\omega)$, such that the optimal decisions for the lower-bounding problem defined in Equation (4.7) are given by:*

$$\begin{aligned} \tilde{Y}_{jt}^{*R}(\omega) &= \begin{cases} [\tilde{y}_{1t}]^- \vee [\tilde{S}_{1t}^R(\omega) \wedge \tilde{y}_{1t}] & \text{for } j = 1; \\ y_{j-1,t} \vee [\tilde{S}_{jt}^R(\omega) \wedge \tilde{y}_{jt}] & \text{for } j \in [2, N]; \end{cases} \\ \tilde{Y}_{jt}^*(\omega) &= \begin{cases} \tilde{Y}_{jt}^{*R} \vee [\tilde{S}_{jt}(\omega) \wedge \tilde{Y}_{j+1,t}^{*R}] & \text{for } j \in [1, N-1]; \\ \tilde{Y}_{Nt}^{*R}, \vee \tilde{S}_{Nt}(\omega) & \text{for } j = N. \end{cases} \end{aligned}$$

Further, this policy achieves the Clark–Scarf decomposition of $\tilde{f}_t^L(\omega, \cdot)$ for every ω .

4.2.3.4 Heuristic policy

Having derived a lower bound on $\bar{F}_t$, the next step is to identify an effective heuristic policy. For that purpose, we introduce the following definitions.

For $j \in [1, N]$ and $i \in [1, j]$, define $h_{ij} := H_{ij} - H_{i,j+1}$ (with $h_{NN0} := H_{NN}$); for $j \in [1, N-1]$, let $\Delta H_j := H_{jj} - H_{j+1,j+1}$. Then, $\bar{V}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t)$ can be written as:

$$\begin{aligned}\bar{V}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t) &= \sum_{j=1}^N \sum_{i=1}^j (1 + \alpha) H_{ij} x_{ijt} + \bar{U}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t) \\ &\quad + \mathbf{E}_\omega \left[\min_{\mathbf{x}_{t+1}^R, \mathbf{X}_{t+1} \in \mathbb{D}_\mathbf{x}(\mathbf{x}_{t+1})} \alpha \bar{G}_{t+2}(\omega, \mathbf{x}_{t+1}, \mathbf{X}_{t+1}^R, \mathbf{X}_{t+1}) \right],\end{aligned}$$

where

$$\begin{aligned}\bar{U}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t) &= \gamma_t(\omega, x_{11t} - X_{11t}^R) + \sum_{j=1}^N \sum_{i=1}^j (k_{ij}^R - \alpha h_{ij}) X_{ijt}^R \\ &\quad + \sum_{j=1}^{N-1} \sum_{i=1}^j (k_{ij} + \alpha h_{ij}) X_{ijt} + \sum_{j=1}^{N-1} (k_{j+1,j} + \alpha \Delta H_j) X_{j+1,jt} \\ &\quad + (k_{NN} + \alpha H_{NN}) X_{NNt} - \alpha \mathbf{E}_\omega [D_t(w)];\end{aligned}$$

and

$$\begin{aligned}\bar{G}_{t+2}(\omega, \mathbf{x}_{t+1}, \mathbf{X}_{t+1}^R, \mathbf{X}_{t+1}) &= \gamma_{t+1}(\omega_{t+1}, x_{11,t+1} - X_{11,t+1}^R) + \sum_{j=1}^N \sum_{i=1}^j k_{ij}^R X_{ij,t+1}^R \\ &\quad + \sum_{j=1}^{N-1} \sum_{i=1}^{j+1} k_{ij} X_{ij,t+1} + k_{NN} X_{NN,t+1} + \alpha \mathbf{E}_{\omega_{t+1}} [\bar{F}_{t+2}(\omega_{t+2}, \mathbf{x}_{t+2})].\end{aligned}$$

Thus, $\bar{V}_t(\omega, \mathbf{x}_t, \mathbf{X}_t^R, \mathbf{X}_t)$ can be decomposed into two component functions plus a holding cost term, where $\bar{U}_t$ includes all cost terms that depend explicitly on either a regular order X_{ijt} or a reverse order X_{ijt}^R in period t . Hence, $\bar{U}_t$ captures all the direct costs incurred by any decision made in period t (where those costs are incurred in either period t or period $t+1$). The function $\bar{G}_{t+2}$, on the other hand, contains no terms that depend on any decisions made in period t ; the decisions $\mathbf{X}_t^R$ and $\mathbf{X}_t$ enter $\bar{G}_{t+2}(\omega, \mathbf{x}_{t+1}, \mathbf{X}_{t+1}^R, \mathbf{X}_{t+1})$ only through their impact on the inventory state $\mathbf{x}_{t+1}$.

The impact of X_{ijt}^R on $\bar{U}_t$ is $\delta_{ij}^R X_{ijt}^R$, where $\delta_{ij}^R := k_{ij}^R - \alpha h_{ij}$ for $j \in [1, N]$ and $i \in [1, j]$. The impact of X_{ijt} on $\bar{U}_t$ is $\delta_{ij} X_{ijt}$, where $\delta_{ij} := k_{ij} + \alpha h_{ij}$ for $j \in [1, N-1]$ and $i \in [1, j]$, and $\delta_{ij} := k_{ij} + \alpha \Delta H_j$ if $i = j+1$. Thus, δ_{ij}^R and δ_{ij} represent the cost impact of unit reverse and regular orders on $\bar{U}_t$.

At each installation $j \in [1, N]$, define $\hat{y}_{1t} = x_{11t}$; $\hat{y}_{jt} = \hat{y}_{j-1,t} + \sum_{i=1}^j x_{ijt}$. For each installation $j \in [2, N]$, let indices $i_1(j), i_2(j), \dots, i_j(j)$ be defined recursively as:

$$i_1(j) := \arg \min_{i \in [1, j]} \delta_{ij}^R; \quad i_k(j) = \arg \min_{i \in [1, j], i \notin \{i_1(j), \dots, i_{k-1}(j)\}} \delta_{ij}^R.$$

For each $j \in [1, N-1]$, let indices $u_1(j), \dots, u_{j+1}(j)$ be defined recursively as:

$$u_1(j) = \arg \min_{i \in [1, j+1]} \delta_{ij}; \quad u_k(j) = \arg \min_{i \in [1, j+1], i \notin \{u_1(j), \dots, u_{k-1}(j)\}} \delta_{ij}.$$

Indices $i_1(j), \dots, i_j(j)$ represent the ordering of reverse-order cost impact factors δ_{ij}^R from the smallest to the largest at each installation j . Indices $t_1(j), \dots, t_{j+1}(j)$ denote the equivalent ordering of regular-order cost impact factors δ_{ij} .

The idea of the proposed heuristic policy is to minimize the impact of ordering decisions on $\bar{U}_t$. Using $\{\tilde{S}_{jt}^R(\omega)\}$ and $\{\tilde{S}_{jt}(\omega)\}$ as the base-stock levels of the optimal policy for $\tilde{f}_t^L(\omega, \cdot)$, the proposed heuristic policy specifies reverse-flow and regular-flow ordering decisions $\{X_{ijt}^R\}$ and $\{X_{ijt}\}$ as follows:

- (a) If $\hat{y}_{jt} > \tilde{S}_{jt}^R(\omega)$ then $X_{11t}^R = \hat{y}_{1t} - \tilde{S}_{1t}^R(\omega)$ and X_{ijt}^R for $j \in [2, N]$ and $i \in [1, j]$ is given recursively by:

$$X_{i_1(j),jt}^R = \min \left[x_{i_1(j),jt}, y_{jt} - \tilde{S}_{jt}^R(\omega) \right];$$

$$X_{i_k(j),jt}^R = \min \left[x_{i_k(j),jt}, \hat{y}_{jt} - \tilde{S}_{jt}^R(\omega) - \sum_{q=1}^{k-1} X_{i_q(j),jt}^R \right].$$

- (b) If $\hat{y}_{jt} < \tilde{S}_{jt}(\omega)$, then: $X_{NN,t} = \tilde{S}_{Nt}(\omega) - y_{Nt}$; and X_{ijt} for $j < N$ and $i \in [1, j+1]$ are given recursively by:

$$X_{i_1(j),jt} = \min \left[x_{i_1(j),j+1,t}, \tilde{S}_{jt}(\omega) - y_{jt} \right];$$

$$X_{i_k(j),jt} = \min \left[x_{i_k(j),j+1,t}, \tilde{S}_{jt}(\omega) - \hat{y}_{jt} - \sum_{q=1}^{k-1} X_{i_q(j),jt} \right].$$

- (c) If $\tilde{S}_{jt}^R(\omega) \leq \hat{y}_{jt} \leq \tilde{S}_{jt}(\omega)$, then $X_{ijt}^R = 0$, for all $i \in [1, j]$ and $X_{ijt} = 0$ for all $i \in [1, j+1]$.

At each installation j , the proposed heuristic policy: (1) determines if either a regular order or reverse order (or neither) is required by looking at the difference between $\tilde{S}_{jt}(\omega)$ and $\hat{y}_{jt}$ for regular orders, and between $\hat{y}_{jt}$ and $\tilde{S}_{jt}^R(\omega)$ for reverse orders; (2) fulfills each such order, starting with the one at the stage of completion i that has the least cost impact on $\bar{U}_i$; and (3) if needed, continues to complete the order by resorting to items at the stage of completion with the next least cost impact on $\bar{U}_i$, and so on. In that manner, this heuristic policy can access, at each installation j , every feasible replenishment option for both reverse and regular orders, and determine the most desirable ones among them. Angelus and Özer (2021) report that the performance gap of this heuristic policy relative to the lower-bounding policy averages 4.72% across a range of model parameters and supply chain lengths, with its maximum value at 6.80% and its minimum value at 1.76%.

4.3 ASSEMBLY SYSTEMS

In this section, we address generalizations of the Clark and Scarf analysis to *assembly systems* that denote production/inventory networks in which components acquired from outside

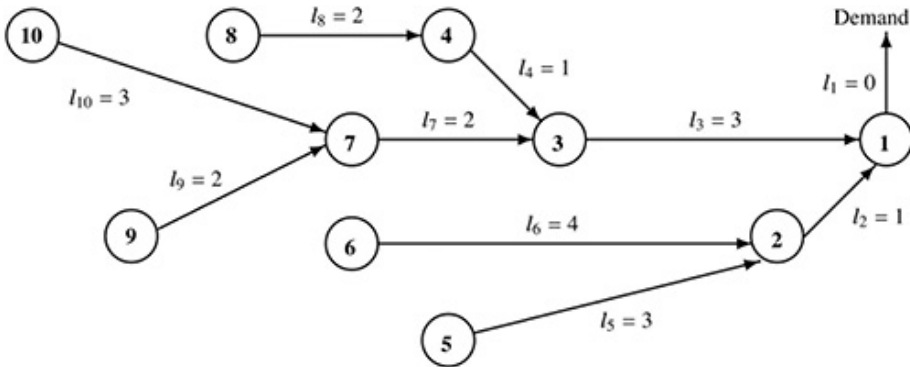
suppliers are assembled, typically in several stages, into subassemblies and then, finally, into a single, finished, end-product used to satisfy (stochastic) customer demand. Assembly networks thus represent trees with the node at the root as the finished end-product, and other nodes representing either partially assembled units (subassemblies) or externally acquired components. Also, the units are defined so that exactly one unit of each item is required for the finished end-product. Figure 4.2 depicts a generic assembly system with ten nodes.

Each node $k = 1, 2, \dots, N$ in an assembly system such as the one in Figure 4.2 is referred to as “(sub)assembly”. For each k , let l_k be the *incremental lead time* required to complete subassembly k , where subassembly 1 is the finished product. Let $s(k)$ be the unique *immediate successor* subassembly to subassembly k . Let $L_1 := l_1 = 0$, and, for each $k > 1$, let $L_k := l_k + L_{s(k)}$ be the *lead time* for subassembly k . Let $P(k)$ be the set of *immediate predecessor* subassemblies of subassembly k . The set of *components* of such a system is the set of subassemblies that have no predecessor subassemblies. Let $\mathcal{A}(k)$ denote the set of components required in the composition of subassembly k . In Figure 4.2, for example, $s(6) = 2$ and $L_9 = 7$ ($= l_9 + l_7 + l_3 + l_1$), while $P(1) = \{2, 3\}$, $P(2) = \{5, 6\}$, $P(3) = \{4, 7\}$, $P(4) = \{8\}$, $P(7) = \{9, 10\}$, $P(5) = P(6) = P(8) = P(9) = P(10) = \emptyset$, and $\mathcal{A}(1) = \{5, 6, 8, 9, 10\}$, $\mathcal{A}(2) = \{5, 6\}$, $\mathcal{A}(3) = \{8, 9, 10\}$, $\mathcal{A}(4) = \{8\}$, $\mathcal{A}(7) = \{9, 10\}$, and $\mathcal{A}(k) = \{k\}$, for $k = 5, 6, 8, 9, 10$. We refer to such a system as a *general assembly system*.

Assembly systems have inherently very large state and decision spaces because of the need to keep track of inventory at a large number of locations (i.e., nodes) in the system, and make decisions pertaining to each one of those in each period. Because of the resulting curse of dimensionality, the literature in this field has been focused on developing approaches to reduce the difficulty of managing such systems. Historically, the research on assembly systems began with the analysis of stationary systems in which costs and demand distributions do not change over time.

4.3.1 Stationary Assembly Systems

Rosling (1989) was the first to consider an assembly-system generalization of the original Clark and Scarf (1960) paper, with the regular flow of items moving downstream, holding costs at each node and backlogging cost at the most downstream node. Rosling assumes



Source: Adapted from Angelus and Ozer (2016).

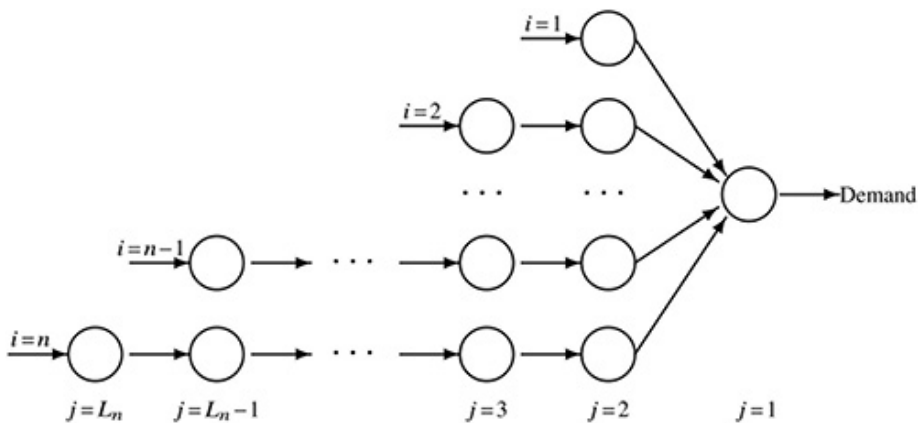
Figure 4.2 An assembly system

stationary costs and customer demands in an infinite-horizon, discounted-cost setting. His unit of analysis is an item, which represents either a subassembly in which multiple preceding items are assembled into another item that continues to flow downstream toward the final assembly. Rosling makes use of the “long-run balance” under which all echelon inventory positions for an item closer to the final assembly are lower than corresponding echelon inventory positions for an item farther from the final assembly. He shows that, if the initial state of the system satisfies the same balanced condition, the optimal policy for the assembly system can be reduced to that of an equivalent Clark and Scarf series system, in which all items equally distant from the final assembly are aggregated together.

Rosling’s approach has subsequently been used by DeCroix and Zipkin (2005) to address a stationary assembly system with uncertain product (and component) returns from customers. They describe the item-recovery pattern and restrictions on the inventory policy under which an equivalent series system is shown to exist. DeCroix (2013) makes use of Rosling’s approach in considering a stationary, infinite-horizon assembly system subject to random supply disruptions. He shows how such a system can be simplified by replacing some subsystems with a series structure.

4.3.2 Nonstationary Assembly Systems

Rosling’s approach requires that costs be stationary, and he argues that “the series interpretation generally cannot be expected to carry over when holding costs or production costs are nonstationary”. Since many assembly systems encountered in practice exhibit nonstationary parameters and finite-horizon settings, Angelus and Porteus (2008) develop a method to analyze an assembly system with nonstationary costs (and demands) in a finite-horizon setting. Their approach is based on the analysis of a *component (assembly) system* that represents an assembly system in which every subassembly, except for the finished end-product, has exactly one predecessor subassembly. Thus, components are assembled only once, at the most downstream node. Figure 4.3 depicts a generic component system with n components.



Source: Adapted from Angelus and Ozer (2016).

Figure 4.3 Component assembly system

Each component i , $i \in (1, 2, \dots, n)$ requires processing through L_i stages before becoming a part of the finished end-product, with each stage taking one period. Without loss of generality, components can be ordered so that $L_1 \leq L_2 \leq \dots \leq L_n$. Each unit of the finished product requires exactly one of each component.

Let x_{ijt} denote the number of units of component i at stage j at the beginning of period t for each i and j ($j = 1, \dots, L_i$), (where x_{i1t} is allowed to be negative, because any backlog for the final product is subtracted from the inventory level). The number of completed units available to satisfy customer demand in any given period t is given by $\min(x_{11t}, x_{21t}, \dots, x_{n1t})$.

Let X_{ijt} denote as the (nonnegative) amount of component i (regular) ordered from stage $j+1$ into stage j in period t . The feasible constraints are then $X_{ijt} \leq x_{i,j+1,t}$ for all i, j and t . The state transition equations are as follows:

$$x_{ij,t+1} = \begin{cases} x_{i1t} + X_{i1t} - D_t, & \text{for } j = 1 \\ x_{ijt} - X_{i,j-1,t} + X_{ijt}, & \text{for } j \in [2, L_i]. \end{cases}$$

Let $y_{ijt} := x_{i1t}^i + \dots + x_{ijt}^i$ and $Y_{ijt} := y_{ijt} + X_{ijt}$. The level of the finished product in period t becomes $\min(y_{11t}, \dots, y_{n1t})$. The state transitions then become $y_{i,j,t+1} = Y_{ijt} - D_t$. The feasible decision set $\mathbb{D}_{COMP}(\mathbf{y}_t)$ for the problem becomes:

$$\mathbb{D}_{COMP}(\mathbf{y}_t) := \{\mathbf{Y}_t \mid y_{ijt} \leq Y_{ijt} \leq y_{i,j+1,t}, 1 \leq j \leq L_i, 1 \leq i \leq n\},$$

The cost to order a unit of component i from stage j into stage $j+1$ in period t is k_{ijt} , while the cost to hold a unit of component i at stage j in period t is H_{ijt} . The one-period cost function is

$$\gamma_t(\min_i y_{i1t}) + \sum_{i=1}^n \sum_{j=1}^{L_i} [k_{ijt} Y_{ijt} - (k_{ijt} - h_{ijt}) y_{ijt}],$$

where $h_{ijt} := H_{ijt} - H_{i,j+1,t}$. Let $F_t^A(\mathbf{y}_t)$ be the objective cost function for the problem. The optimality equations for this system can be expressed as

$$F_t^A(\mathbf{y}_t) = \max_{\mathbf{Y}_t \in \mathbb{D}_{COMP}(\mathbf{y}_t)} \left\{ \gamma_t(\min_i y_{i1t}) + \sum_{i=1}^n \sum_{j=1}^{L_i} c_{ijt} Y_{ijt} + \alpha \mathbf{E}[F_{t+1}^A(\mathbf{Y}_t - D_t)] \right\}, \quad (4.8)$$

where $c_{ijt} = k_{ijt} - \alpha k_{ij,t+1} + \alpha h_{ij,t+1}$. To prevent a component from being moved downstream solely for the purpose of reducing costs, rather than to get closer to customer demand, we assume $c_{ijt} \geq 0$ for each component i , stage j , and period t .

Let $L = L_n$ be the longest component lead time. For each $j = 0, 1, \dots, L-1$, let $\mathcal{I}(j)$ be the set of components with a lead time strictly greater than j . These are the components for which stage j quantities and decisions apply, and are therefore called the *relevant components at stage j* .

Definition 4.2 An echelon state $\mathbf{y}_t$ is *balanced* if, for every stage j , $y_{ijt} = y_{kjt}$ for all $i, k \in \mathcal{I}(j)$. Thus, a balanced echelon state $\mathbf{y}_t$ has, at each stage, exactly the same number of units of each relevant component at that stage. (Similarly, $\mathbf{Y}_t$ is balanced if, for every stage j , $Y_{ijt} = Y_{kjt}$ for all $i, k \in \mathcal{I}(j)$).

Theorem 4.6 (Angelus & Porteus, 2008) *If the terminal value function for the problem given in Equation (4.8) is such that that unmatched components have no value (i.e., only balanced sets of components are valuable) and the system starts out balanced in period 1, then there exists a balanced optimal policy in every subsequent period.*

Hence, if the system starts out balanced in period $t=1$, it is optimal to keep it balanced throughout the time horizon. Accordingly, instead of managing the different components separately, it becomes possible to manage those at the same stage as a kit: The kit for stage j has one each of every component in $\mathcal{I}(j)$. The echelon quantity y_{ijt} in period t for each component i relevant at stage j can be represented by a single variable, z_{jt} ; that is, $y_{ijt} = z_{jt}$ for every $i \in \mathcal{I}(j)$. Let z_t denote the L vector $\{z_{1t}, \dots, z_{Lt}, t\}$. Similarly, Y_{ijt} is the same for every $i \in \mathcal{I}(j)$, and can therefore be denoted by a single variable, Z_{jt} . Let Z_t denote the L vector $\{Z_{1t}, \dots, Z_{Lt}, t\}$.

Next, define the new cost parameters, c_{jt} , for each j and t as $\tilde{c}_{jt} := \sum_{i \in \mathcal{I}(j)} c_{ijt}$. This results in an equivalent variant of the Clark and Scarf (1960) series model, with the following optimality equations, for each t and z_t :

$$f_t^A(z_t) := \max_{\substack{z_{jt} \leq Z_{jt} \leq z_{j+1,t} \\ (1 \leq j \leq L)}} \left\{ \gamma_t(z_{1t}) - \sum_{j=1}^L \tilde{c}_{jt} Z_{jt} + \alpha \mathbf{E}[f_{t+1}^A(\mathbf{Z}_t - D_t)] \right\}, \quad (4.9)$$

where $z_{L+1,t} := \infty$. It follows from Clark and Scarf (1960) that the resulting optimal policy is an echelon base-stock policy and that $f_t(z_t)$ is Clark–Scarf decomposable.

To extend their results to an assembly system with subassemblies, such as the one in Figure 4.2, Angelus and Porteus (2008) first assume that the only extra costs incurred in a system with subassemblies, in addition to the costs already introduced, are the assembly costs. Then, they relax the constraint(s) that there must be exactly the right number of each component ready when any (sub)assembly takes place (other than the final assembly at stage 1). Relaxing those constraints provides a system in which the performance is at least as good as in the original one. Further, this relaxed system has exactly the same constraints as the component assembly system already solved. Therefore, a general assembly system with subassemblies (with relaxed component-matching constraints) can be treated as an equivalent component assembly system. Thus, with relaxed component-matching constraints, the system with subassemblies shown in Figure 4.2, for example, becomes structurally equivalent to a five-component assembly system (components 5, 6, 8, 9, and 10 in the original system) with component lead times of 4, 5, 6, 7, and 8 periods, respectively.

Next, if there exists an allocation of assembly costs to its components that satisfies the required conditions, then, by the results obtained for a component system, there exists an optimal policy for the equivalent component system that is balanced. Since it is balanced, this policy is feasible in the original constrained system (i.e., a system with subassemblies) because every balanced policy satisfies the component-matching constraints. Because this policy is optimal for the equivalent component system, and feasible for the original constrained system with subassemblies, then it must be optimal for that original system with subassemblies.

To design such a cost allocation scheme, let c_{kt}^A be the discounted present value of the costs related to assembling/purchasing/transforming assembly k , evaluated at the beginning of period t . Let β_{ikt} denote the portion of the assembly cost for assembly k that is allocated to component i , for each $i \in \mathcal{A}(k)$ in period t (e.g., in Figure 4.2, allocate the assembly costs of

assembly 1 equally to each required component, so that $\beta_{51t} = \beta_{61t} = \beta_{81t} = \beta_{91t} = \beta_{10,1t} = 1/5$. In general, full allocations are required, so that $\sum_{i \in \mathcal{A}(k)} \beta_{ikt} = 1$ for each k and t . Hence, each component i is allocated $\beta_{ikt} c_{kt}^A$, for each k for which $i \in \mathcal{A}(k)$. If assembly k is initiated at the beginning of period t , it will require one unit of component i for each $i \in \mathcal{A}(k)$, and the amount $\beta_{ikt} c_{kt}^A$ will be allocated to each such component i for stage L_k in period t . How this allocation of costs would look for the system in Figure 4.2 is shown in Table 4.1. Component 6 is shown as bearing the full cost of assembly 6 – it is the only component needed for that particular assembly. The cost of assembly 7, by contrast, is shared among components 9 and 10.

Let $\mathbb{L}(j) := \{k \mid L_k = j\}$ be the set of assemblies with lead time L_k equal to j . Let $\theta_{ijt} := \sum_{k \in \mathbb{L}(j)} \beta_{ikt} c_{kt}^A$ for each i and j .

It can be shown that, if for each component i at stage j in each period t there exists an allocation of assembly costs such that $\theta_{ijt} \geq \alpha \theta_{ij,t+1}$, then the optimal policy is balanced in every period. If a policy is balanced, then, when an allocation of costs satisfies the condition $\theta_{ijt} \geq \alpha \theta_{ij,t+1}$, the actual costs incurred in a general assembly system with subassemblies are the same as those captured by the equivalent component model. Since components can be independently managed in a component assembly system, then the optimal solution of the equivalent component system is at least as good as the best one in the original (i.e., general assembly) system. Because the optimal policy for the equivalent component system is balanced, that policy is feasible in the original system with subassemblies, because, being balanced, it satisfies the component-matching constraints. Thus, this balanced policy must be optimal for the original general assembly system with subassemblies.

The approach to analyzing assembly systems introduced in Angelus and Porteus (2008) has subsequently been applied to *capacitated assembly systems*, in which each regular order X_{ijt} , placed in period t for component i to be brought into stage j is limited by a capacity constraints K_{ij} , so that the feasibility constraints become $X_{ijt} \leq \min[x_{i,j+1,t}, K_{ij}]$. Angelus and Zhu (2013) identify conditions that allow such a capacitated assembly system to be reduced to an equivalent series one. The first of those conditions requires that the bottleneck capacity in the system be located at the most downstream node, where the end-product is used to satisfy customer demand. The second condition rules out the presence of so-called “flow-through” stages in the system, in which it is not possible to hold any inventory of a particular component. Under those conditions, a capacitated assembly system with subassemblies can be reduced to an

Table 4.1 Allocation of assembly costs to components and stages

i	L_i	$k=1$	2	3	4	5	6	7	8	9	10
8	4	$\beta_{51} c_1^A$	$\beta_{52} c_2^A$			c_5^A					
5											
66	5	$\beta_{61} c_1^A$	$\beta_{62} c_2^A$				c_6^A				
8	6	$\beta_{81} c_1^A$		$\beta_{83} c_3^A$	c_4^A				c_8^A		
9	7	$\beta_{91} c_1^A$		$\beta_{93} c_3^A$				$\beta_{97} c_7^A$		c_9^A	
10											
1010	8	$\beta_{10,1} c_1^A$		$\beta_{10,3} c_3^A$				$\beta_{10,7} c_7^A$			$c_{10}^A c_{10}^A$

Source: Adapted from Angelus and Ozer (2016).

equivalent series system. In the absence of those conditions, Angelus and Zhu (2012) observe the existence of two phenomena, stockpiling and stock-withholding, that unbalance a capacitated assembly system and prevent its reduction to an equivalent series system.

Angelus and Özer (2016) consider a nonstationary assembly system with advance demand information and both expedited and regular flows of products at each node. The stochastic demand seen during period t is of the form $D_t = (D_{t,t}, \dots, D_{t,t+N})$, where $D_{t,s}$ is the demand observed in period t for delivery in period s , where $s \in [t, \dots, t+N]$ and N is the longest available delivery time offered to the customer. Thus, at the beginning of any period t , the observed demand to be fulfilled in a future period s is $O_{t,s} = \sum_{q=s-N}^{t-1} D_{q,s}$, and the available demand information is the N dimensional vector $\tilde{O}_t := (O_{t,t}, O_{t,t+1}, \dots, O_{t,t+N-1})$. The actual demand to be satisfied in period t is $O_{t,t} = \sum_{q=t-N}^{t-1} D_{q,t}$. By the end of period t , the pending demand (i.e., to be satisfied in period t) is $O_{t+1,t} = O_{t,t} + D_{t,t}$. The process D_t is allowed to depend on $\tilde{O}_t$, which is known at the beginning of period t and before D_t is realized. Angelus and Özer (2016) also start out by considering a component system, with x_{ijt} as the number of units of component i at stage j at the beginning of period t for each i and j ($j = 1, \dots, L_i$). While X_{ijt} represents the amount of component i (regular) ordered from stage $j+1$ into stage j in period t , X_{ijt}^E denotes the amount of component i expedited from stage $j+1$ into stage j in period t . The state transition equations for each component i are given by:

$$x_{ij,t+1} = \begin{cases} x_{il,t} + X_{il,t}^E + X_{il,t} - O_{t+1,t} & \text{for } j = 1, \\ x_{ijt} + X_{ijt}^E - X_{i,j-1,t}^E + X_{ijt} - X_{i,j-1,t} & \text{for } j \in [2, L_i]. \end{cases}$$

One-period costs in period t can be expressed as:

$$\begin{aligned} & \gamma_t(\tilde{O}_t, \min_i (x_{il,t} + X_{il,t}^E + X_{il,t})) \\ & + \sum_{i=1}^n \sum_{j=1}^{L_i} [(k_{ijt}^E + h_{ijt})X_{ijt}^E + (k_{ijt} + h_{ijt})X_{ijt} + H_{ijt}x_{ijt}], \end{aligned}$$

where $\gamma_t(\tilde{O}_t, x) := \mathbf{E}_{D_{t,t}|\tilde{O}_t} [p_t(O_{t+1,t} - x)^+ + H_t^A(x - O_{t+1,t})^+]$. The expectation is taken over $D_{t,t}$, given $\tilde{O}_t$ to account for the correlation across periods.

Let $\mathcal{F}_t(\tilde{O}_t, x_t)$ denote the minimum expected present value of the costs over periods t through T for this assembly system, as of the beginning of period t , given that the $\tilde{O}_t$ and x_t . The optimality equations become:

$$\mathcal{F}_t(\tilde{O}_t, x_t) = \min_{\mathbf{x}_t^E, \mathbf{x}_t \in \mathbb{D}_A(x_t)} \mathcal{G}_t(\tilde{O}_t, x_t, \mathbf{x}_t^E, \mathbf{x}_t),$$

where

$$\begin{aligned} \mathcal{G}_t(\tilde{O}_t, x_t, \mathbf{x}_t^E, \mathbf{x}_t) &= \gamma_t(\tilde{O}_t, \min_i (x_{il,t} + X_{il,t}^E + X_{il,t})) + \\ & \sum_{i=1}^n \sum_{j=1}^{L_i} [(k_{ijt}^E + h_{ijt})X_{ijt}^E + (k_{ijt} + h_{ijt})X_{ijt} + H_{ijt}x_{ijt}] + \alpha \mathbf{E}_{D_{t+1}|\tilde{O}_t} [\mathcal{F}_{t+1}(\tilde{O}_{t+1}, x_{t+1})], \end{aligned}$$

where the expectation is with respect to $D_t = \{D_{t,t}, \dots, D_{t,t+N}\}$ given $\tilde{O}_t$, and

$$\mathbb{D}_A(x_t) = \{\mathbf{X}_t^E, \mathbf{X}_t \geq 0 \mid X_{ijt}^E \leq x_{i,j+1,t} + X_{i,j+1,t}^E; \\ X_{ijt} \leq x_{i,j+1,t} + X_{i,j+1,t}^E - X_{ijt}^E \text{ for } 1 \leq i \leq n, 1 \leq j < L_i\}.$$

The optimality equations given for $\mathcal{F}_t(\tilde{O}_t, x_t)$ bear a severe curse of dimensionality: the state space has $\sum_{i=1}^n L_i$ inventory dimensions plus N advance demand dimensions. Further, optimal order quantities X_{ijt}^E and X_{ijt} for each component i and stage j depend on all the variables in the state space. Due to the presence of two flows of components in the system, the regular flow and the expedited flow, and the nonstationary cost parameters in the model neither the classic methods of Rosling (1989) for collapsing the state space nor the balance-inducing approach of Angelus and Porteus (2008) are conducive to solving the problem. Therefore, to deal with the curse of dimensionality, Angelus and Özer (2016) introduce a method for showing that the optimal policy for the underlying component system is balanced. In contrast to other approaches, their approach does not make use of echelon variables, but rather of local properties of installation state and decision variables: x_t , $\mathbf{X}_t^E$, and $\mathbf{X}_t$, respectively. Those installation decision variables are then analyzed recursively at each node in the system, starting with the most downstream node 1.

Theorem 4.7 (Angelus & Özer, 2016) *If a component system with advance demand and expediting starts out balanced in period 1, then optimal order schedules $(\mathbf{X}_t^E, \mathbf{X}_t)$ are balanced in every period $t = 1, \dots, T$. As a result, the inventory state x_t is balanced in every period $t = 1, \dots, T$.*

Hence, a component assembly system with advance demand information and expediting of inventory can be reduced to an equivalent series system. Angelus and Özer (2016) then show that the optimal policy for the resulting series system represents is a double-tiered, echelon base-stock policy that depends on vector $\tilde{O}_t$ of advance demands. This form of the optimal policy allows the system to be decomposed into a nested sequence of solvable convex subproblems.

4.4 OPEN RESEARCH PROBLEMS

We conclude this review by describing some of the open research problems in multiechelon inventory theory that are of relevance to both theory and practice. One such problem concerns systems with secondary market purchases of products that represent lateral flows of stock *into* the system (in contrast to the secondary market sales addressed in Section 4.2.2, which represent lateral flows of products *out* of the system). Secondary market purchases can thus be viewed as emergency orders at each installation used to address excess demand in the system. Secondary market purchases in multi-stage inventory systems can also be interpreted as a form of reactive capacity whose role in multi-period, multi-stage systems remains unexplored, despite a vast literature on reactive capacity in single-period, newsvendor-like settings. Such lateral flows of products into the system at each installation also generalize dual-sourcing, single-stage systems (see the chapter “Dual-sourcing, dual-mode dynamic stochastic inventory models” in this handbook) to multi-sourcing, multi-stage systems with stochastic demand.

There also remain a number of open problems in product-transforming supply chains. While there now exists a viable heuristic policy for managing reverse logistics in such systems, many questions still remain. For example, it is not known if there exists an effective policy that achieves the Clark–Scarf decomposition of the objective function. Interestingly, it is even not clear how best to define echelon inventory in such systems, as there exist different ways to aggregate inventory across different locations and stages of completion. (For example, it is possible to formulate echelon variables by aggregating installation stocks either across all stages of completion at a given location, or across all locations for a given stage of completion.) Further progress on these questions would be worthwhile as product-transforming supply chains are very common in practice and their complexity is not well understood.

Over the past 20 years, research on assembly systems has provided different approaches and important insights regarding how to optimally manage such systems. A number of interesting directions for future research have also been put forth. One such direction involves assembly systems with multiple flows of items, including regular, expedited, lateral, and reverse flows of product. Even though reverse logistics and secondary market sales represent capabilities frequently encountered in assembly systems found in practice, very little is presently known about how to manage (optimally or heuristically) those capabilities in either logistics or product-transforming supply chains with assembly structures. Consequently, making some progress in that direction would have both theoretical and practical impact.

REFERENCES

- Angelus, A. (2011). A multiechelon inventory problem with secondary market sales. *Management Science*, 57(12), 2145–2162.
- Angelus, A., & Porteus, E. (2008). An asset assembly problem. *Operations Research*, 56(3), 665–680.
- Angelus, A., & Zhu, W. (2013). On the structure of capacitated assembly systems. *Operations Research Letters*, 41(1), 19–26.
- Angelus, A., & Özer, Ö. (2016). Knowledge you can act on: Optimal policies for assembly systems with expediting and advance demand information. *Operations Research*, 64(6), 1338–1371.
- Angelus, A., & Özer, Ö. (2021). When variability trumps volatility: Optimal control and value of reverse logistics in supply chains with multiple flows of product. *Manufacturing & Service Operations Management*, 23(5), 1175–1195.
- Chen, F., & Song, J.-S. (2001). Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research*, 49(2), 226–234.
- Clark, A. H., & Scarf, H. (1960). Optimal policies for the multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- de Kok, T., Grob, C., Laumanns, M., Minner, S., Rambau, J., & Schade, K. (2018). A typology and literature review on stochastic multi-echelon inventory models. *European Journal of Operational Research*, 269(3), 955–983.
- DeCroix, G. (2006). Optimal policy for a multiechelon inventory system with remanufacturing. *Operations Research*, 54(3), 532–543.
- DeCroix, G., Song, J.-S., & Zipkin, P. (2005). A series system with returns: Stationary analysis. *Operations Research*, 53(2), 350–362.
- DeCroix, G. A. (2013). Inventory management for an assembly system subject to supply disruptions. *Management Science*, 59(9), 2079–2092.
- DeCroix, G. A., & Zipkin, P. H. (2005). Inventory management for an assembly system with product or component returns. *Management Science*, 51(8), 1250–1265.
- Gong, X., & Wang, T. (2021). Preservation of additive convexity and its applications in stochastic optimization problems. *Operations Research*, 69(4), 1015–1024.

- Kim, C., Klabjan, D., & Simchi-Levi, D. (2015). Optimal expediting policies for a serial inventory system with stochastic lead time. *Production and Operations Management*, 24(10), 1524–1536.
- Lawson, D. G., & Porteus, E. L. (2000). Multistage inventory management with expediting. *Operations Research*, 48(6), 878–893.
- Özsen, R., & Thonemann, U. W. (2015). Determining optimal parameters for expediting policies. *Manufacturing & Service Operations Management*, 17(1), 120–133.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579.
- Shen, X., Yu, Y., & Song, J.-S. J. (2022). Optimal policies for a multi-echelon inventory problem with service time target and expediting. *Manufacturing & Service Operations Management*.
- Song, J.-S., & Zipkin, P. (1993). Inventory control in a fluctuating demand environment. *Operations Research*, 41(2), 351–370.

5. Single-stage approximations of multi-echelon inventory models

Kevin H. Shang, Jing-Sheng Jeannette Song, and Sean X. Zhou

5.1 INTRODUCTION

5.1.1 Overview

For supply-chain systems with complex network structures, the forms of system-optimal inventory-control policies are often not known. Even for systems with known optimal policies, finding a system-wide optimal solution often requires solving interrelated, recursive cost functions between stages across time (see reviews by Axsäter (1993) and Federgruen (1993)). This is particularly a concern when the supply chain has different levels of information integration. In addition, a supply chain is composed of many firms, each with its own interests. These firms may not be willing to implement the optimal solution without appropriate incentives. One way to mitigate these difficulties is to design simple and effective heuristic policies.

This chapter reviews some of the recent developments of single-stage-based heuristics for multi-stage (also known as multi-echelon) inventory systems. These heuristics not only simplify computation and implementation, but also lead to simple coordination mechanisms that help the system achieve a near-optimal performance (see, Shang et al. (2009)). These heuristics also have great potential for classroom teaching and for solving complicated interdisciplinary issues that occur in supply chains. For example, in Chapter 16 of this book, supply-chain models with cash flows are discussed. The single-stage-based heuristics may be derived given the similarity of the structure of the optimal policy.

Multi-echelon inventory models can be classified by using the following attributes: deterministic vs. stochastic demand and local vs. centralized information. In this chapter, we focus on inventory systems with stochastic demand and centralized information. Unfilled demand is fully backlogged. The objective is to minimize the expected long-run average cost per period. We also review some finite-horizon or infinite-horizon models that minimize the total discounted expected cost. In addition, we discuss several extensions including models with expedited supply and systems with local information. The reader is referred to Maxwell and Muckstadt (1985), Roundy (1985), and Federgruen and Zheng (1993) for models with deterministic demand.

There are three basic configurations for multi-echelon inventory models: series, assembly, and distribution systems; see Figure 5.1. We focus primarily on series systems and briefly mention the key results for the other two systems in Section 5.4. Typical cost components for multi-echelon systems include inventory holding cost, demand backorder cost, and fixed order cost. Holding cost is the cost incurred for carrying excess inventory. Backorder cost is the penalty cost incurred for inventory shortage. Fixed costs, such as shipping costs, are costs associated with each inventory replenishment. They are usually assumed to be a constant and independent of the order quantity.

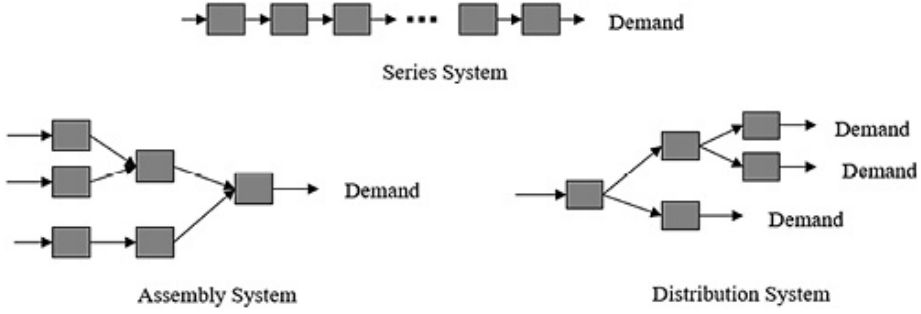


Figure 5.1 Three basic supply-chain configurations

5.2 SERIES SYSTEMS WITHOUT FIXED ORDER COSTS

We first introduce a series system. Consider a continuous-review, series inventory system with N stages. Material flows from stage N to stage $N-1$, $N-1$ to $N-2$, etc., until stage 1, where random customer demand occurs. There is a constant lead time L_j for stage j . Let $L_{[i,j]} = \sum_{k=i}^j L_k$ be the sum of lead times from stage i to stage j . Demand follows a compound Poisson process $\mathbf{D} = \{D(t), t \geq 0\}$, where $D(t)$ is cumulative demand in the time interval $(0, t]$. The arrival rate is λ and the batch size d follows a distribution $\phi(d)$. There is a linear holding cost rate h_j for each unit of on-hand inventory held at stage j . We call h_j the local holding cost rate where $h_1 \geq h_2 \geq \dots \geq h_N$. This assumption is practical because the holding cost rate is often associated with the inventory value and the monotonicity represents a value-added process of a supply chain. Define the echelon holding cost rate $h_j = h_j - h_{j+1}$ for $j = 1, \dots, N-1$ and $h_N = h_N$.

The same system configuration can be studied under the periodic-review scheme. Typically, there are two cost criteria studies under a periodic-review system, i.e., average cost and discounted cost. The former cost is considered for an infinite-horizon model whereas the latter is considered for both finite- and infinite-horizon models. The difference between continuous-review and periodic-review systems is the timing of inventory replenishment and the cost assessment, which will be explained later. In addition, for a periodic-review model with the discounted cost criterion, we often need to consider the inventory order cost c_j per unit at each stage j .

To facilitate the subsequent discussion, let us define a notation to represent the system under the average cost and the discounted cost criteria, respectively.

$$\text{Average Cost Criterion: } \mathbb{S}[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$$

$$\text{Discounted Cost Criterion: } \mathbb{S}[N, (c_i, h_i, L_i)_{i=1}^N, b, \mathbf{D}].$$

The rest of the section is organized as follows. Section 5.2.1 considers a special case with $N = 1$, i.e., a single-stage inventory system, to set the stage for the subsequent discussion; Section 5.2.2 considers series systems with stationary demand; Section 5.2.3 discusses series systems with nonstationary demand; Section 5.2.4 considers expediting options in addition to regular replenishment decisions.

5.2.1 Single-Stage Base-Stock Policy

Consider a continuous-review, single-stage system, i.e., $\mathbb{S}[1, (h, L), b, \mathbf{D}]$, in which the demand follows stationary compound Poisson process $\mathbf{D}$. It is known that a base-stock policy with base-stock level s^* is optimal. That is, the inventory manager monitors the inventory position (= inventory on order + inventory on hand – backorders) continuously. Whenever the inventory position is below s^* , an order is placed to bring the inventory position back to s^* . Otherwise, no orders are placed.

Denote $D[L] = D(t + L) - D(t)$ to be the lead-time demand, with cumulative distribution $F(\cdot)$. Let $(x)^+ = \max\{0, x\}$, $(x)^- = \max\{0, -x\}$. For any given stationary base-stock policy with base-stock level y , the steady-state on-hand inventory is $I = (y - D[L])^+$ and the steady-state number of backorders is $B = (y - D[L])^-$. Then, the long-run average inventory-backorder cost is

$$G(y) = E[hI + bB] = E[h(y - D[L])^+ + b(y - D[L])^-]. \quad (5.1)$$

Thus, the optimal base-stock level s^* minimizes Equation (5.1) over y , and we have

$$s^* = F^{-1}\left(\frac{b}{b+h}\right), \quad (5.2)$$

where $F^{-1}(\theta) = \min\{y \mid F(y) \geq \theta\}$, $\theta \in [0, 1]$ is the inverse of F . The cost expression in Equation (5.1) has exactly the same format as the newsvendor model with $D(L)$ replaced by the single-period demand D , h the overage cost, b the underage cost, and y the order quantity. The solution to Equation (5.2) corresponds to the optimal order quantity.

For the periodic-review system $\mathbb{S}[1, (h, L), b, \mathbf{D}]$ with independent and identically distributed (IID) demand, the optimal policy is also a base-stock policy, which works as follows: at the beginning of each period, an order is placed to bring the inventory position to the base-stock level s^* if the inventory position is lower than s^* ; do not order, otherwise. Because the inventory holding and backorder costs occur at the end of a period, the length of the lead time should include one additional review period, i.e., $(L + 1)$ periods. This is the only difference from the perspective of cost evaluation. The optimal base-stock level is the same as Equation (5.2) (with the adjustment of the demand during lead time). For the finite-horizon, discounted cost model $\mathbb{S}[1, (c, h, L), b, \mathbf{D}]$, let the discount rate be $0 < \alpha \leq 1$. Assuming that the end-of-horizon inventory can be returned with the order cost refunded (i.e., free returns), one can show that a myopic policy with a base-stock level shown in Equation (5.3) is optimal:

$$s^* = F^{-1}\left(\frac{\alpha^{L+1}b - c(1-\alpha)}{\alpha^{L+1}(b+h)}\right). \quad (5.3)$$

Without the free-return assumption, the myopic solution is an upper bound to the optimal solution for the finite-horizon model (see Zipkin (2000, pp. 378–379)). For the infinite-horizon, discounted cost model, it turns out that Equation (5.3) is the exact optimal base-stock level. The intuition is that, under an infinite-horizon, the end-of-horizon effect vanishes, so the system can be viewed as a free-return system.

5.2.2 Multi-Stage Echelon Base-Stock Policies

5.2.2.1 Exact optimal policy

We provide some key results for $\mathbb{S}[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$, and the discussion will focus on the continuous-review system. The analysis for the periodic-review system with IID demand is essentially the same except for the one-period adjustment for the demand during lead time. Clark and Scarf (1960) show that the echelon base-stock policies are optimal for a finite-horizon problem. For an infinite-horizon problem, Federgruen and Zipkin (1984) and Chen and Zheng (1994) prove the optimality of echelon-base-stock policies for both discounted and average cost criteria, respectively.

To facilitate our discussion, we define inventory variables below: At time t , let

$IN_j(t)$ = local net inventory at stage j ;

$IO_j(t)$ = inventory on order for stage j ;

$IT_j(t)$ = inventory in transit to stage j and $IT_0(t) = 0$;

$IN_j(t)$ = echelon net inventory at stage $j = \sum_{i=1}^j (IN_i(t) + IT_{i-1}(t))$;

$IOP_j(t)$ = echelon inventory on-order position at stage $j = IO_j(t) + IN_j(t)$;

$IP_j(t)$ = echelon inventory in-transit position at stage j

The difference between $IOP_j(t)$ and $IP_j(t)$ is the number of outstanding orders for stage j (i.e., orders not yet filled by stage $j+1$). The inventory cost is determined by the $IN_j(t)$ values. We use IN_j , IO_j , IN_j , and IOP_j to denote these inventory variables in steady state.

The echelon base-stock policy $s = (s_1, s_2, \dots, s_N)$ is implemented as follows: Stage j monitors $IOP_j(t)$ continuously, and will place an order of size $(s_j - IOP_j(t))$ if $IOP_j(t) < s_j$. Denote I_j as the local on-hand inventory at stage j and B the number of backorders at stage 1. We have $IN_j = IN_j - IOP_{j-1}$, $j \geq 2$, $I_j = [IN_j]^+$, and $B = [IN_1]^-$. Let $D_j = D[L_j]$ denote the demand during the lead time for stage j . The long-run average system-wide cost is

$$\begin{aligned} C(s) &= \mathbb{E} \left[\sum_{i=1}^N h_i I_i + bB + \sum_{i=2}^N h_i D_{i-1} \right] \\ &= \mathbb{E} \left[\sum_{i=1}^N h_i IN_i + (b + h_1)B \right]. \end{aligned} \tag{5.4}$$

Here we adopt the conventional cost accounting that the pipeline inventory incurs the holding cost of stage i before it arrives at stage $i-1$.

To evaluate the cost, we only need to characterize IN_j for all j . This process starts from stage N , $N-1$, sequentially until stage 1. More specifically, $IP_N = IOP_N$, which is s_N . For $j = N, \dots, 1$, $IN_j = IP_j - D_j$, and for $j = N-1, \dots, 1$, $IP_j = \min\{IN_{j+1}, s_j\}$. After we characterize all IN_j , we can use Equation (5.4) to find the average cost per unit of time.

Below we introduce a bottom-up recursion to obtain the average optimal cost based on Chen and Zheng (1994). Let

$$\begin{aligned} G_1(y) &= E[h_1(y - D_1) + (b + h_1')(y - D_1)^-], \\ s_1^* &= \arg \min_y G_1(y). \end{aligned} \quad (5.5)$$

For $j = 2, \dots, N$,

$$\begin{aligned} G_j(y) &= E[h_j(y - D_j) + G_{j-1}(\min\{y - D_j, s_{j-1}^*\})], \\ s_j^* &= \arg \min_y G_j(y), \end{aligned} \quad (5.6)$$

where $(s_1^*, \dots, s_N^*)$ are optimal echelon base-stock levels. The optimal cost is $G_N(s_N^*) = \min_{\mathbf{s}} C(\mathbf{s})$.

Note that the above recursion can be used to evaluate any echelon base-stock policies by replacing s_j^* with any echelon base-stock level s_j in the recursion while omitting the optimization step. With a slight modification, the above algorithm can also solve the optimal echelon base-stock levels for discounted cost criterion; see, e.g., Chao and Zhou (2007).

5.2.2.2 Single-stage bounds and heuristics

Shang and Song (2003) develop simple bounds for the optimal echelon base-stock levels based on solving a series of independent single-stage systems. The proposition below is the key idea of the construction of the solution bounds. Let $h_{[i,j]} = h_i + h_{i+1} + \dots + h_j$ for $i < j$.

Proposition 5.1 *For each $j \geq 2$ and conditioning on $IP_j = y$, assuming the echelon base-stock policy s_i^* at stage i , $i < j$ is employed, let $I_i'(y)$ denote the corresponding local on-hand inventory at stage i , $i = 1, \dots, j$, and $B(y)$ the number of backorders at stage 1. Then $G_j(y)$ is the average inventory holding and backorder cost for echelon- j . In particular, $G_j(y) = G_j^e(y) + \tau_j$, with*

$$G_j^e(y) = E[h_j I_j'(y) + h_{[j-1,j]} I_{j-1}'(y) + \dots + h_{[1,j]} I_1'(y) + (b + h_{[j+1,N]}) B(y)], \quad (5.7)$$

τ_j = the average in-transit inventory holding cost in echelon- j

$$= \sum_{i=2}^j (h_{[i,j]}) E[D_{i-1}] = \sum_{i=2}^j h_i E[\tilde{D}_{i-1}], \quad (5.8)$$

where $\tilde{D}_j = \sum_{i=1}^j D[L_i] = D[L_{(1,j)}]$.

Compared with Equation (5.4), Proposition 5.1 implies that $G_j^e(y)$ is the cost for a j -stage series system $\mathbb{S}[j, (h_i, L_i)_{i=1}^j, b_j, \mathbf{D}]$ under the echelon policy s_i^* for $i < j$ where the local holding cost rate is $h_{[i,j]}$ for stage i , $i \leq j$, and backorder cost rate $b_j = b + h_{[j+1,N]}$ for stage 1. For convenience, we term this system as *echelon- j* .

Now, for any y , according to Equation (5.7), if we replace the different holding cost rates at different stages in echelon- j with a single common value, then there would be no incentive to carry inventories at stage i , $2 \leq i \leq j$, and the echelon- j system would collapse into a single-stage system with the total lead time $L_{[1,j]}$ and base-stock level y . By setting this single common value to the minimum holding cost rate $h_j^\ell \equiv h_j$, we then obtain a single-stage lower-bound system $\mathbb{S}[1, (h_j^\ell, L_{[1,j]}), b_j, \mathbf{D}]$, whose cost function is denoted by $G_j^\ell(y)$. Define $F_j(\cdot)$ as the cumulative distribution function for $\tilde{D}_j$. Let

$$s_j'' = \operatorname{argmin} G_j^\ell(\cdot) = F_j^{-1}\left(\frac{b_j}{b_j + h_j^\ell}\right).$$

Similarly, by setting the single holding-cost-rate value to the maximum holding cost rate $h_j^u \equiv h_{[1,j]}$ in echelon- j , we obtain a single-stage upper-bound system $\mathbb{S}[1, (h_j^u, L_{[1,j]}), b_j, \mathbf{D}]$, whose cost function is denoted by $G_j^u(y)$. Let

$$s_j^\ell = \operatorname{argmin} G_j^u(\cdot) = F_j^{-1}\left(\frac{b_j}{b_j + h_j^u}\right).$$

Theorem 5.1 For any given j and y , (a) $G_j^\ell(y) + \tau_j \leq G_j(y) \leq G_j^u(y) + \tau_j$. (b) $s_j^\ell \leq s_j^* \leq s_j''$. All the inequalities become equalities for $j = 1$.

In general, there are two ways of constructing approximations for the optimal base-stock levels. (i) According to Theorem 5.1, any convex combination of s_j^ℓ and s_j'' can be used to approximate s_j^* . (ii) We can replace the coefficients of $I_i(y)$ in Proposition 5.1 with a single convex combination of them to obtain a base-stock system (similar to the bounding systems) and use its solution as the approximate solution. In principle, a numerical experiment can be carried out to identify effective weights. Shang and Song (2003, 2007) find that using a weight equal to 0.5 leads to an effective solution in most cases.

5.2.2.3 Service-level constrained model

There is an alternative way to consider inventory shortage cost by setting a target service level or fill rate in our context. See, for example, Lee and Billington (1995), Ettl et al. (2000), and Graves and Willems (2000, 2003). For the series model, the problem can be formulated as follows: For an echelon base-stock policy $\mathbf{s}$,

$$R(\mathbf{s}) = \text{fill rate under } \mathbf{s} = P(IN_1 > 0),$$

$$H(\mathbf{s}) = \text{average holding cost under } \mathbf{s} = E\left[\sum_{j=1}^N h_j IN_j + h_1^-(IN_1)^-\right].$$

Let β be a pre-determined fill rate. The problem is to solve

$$(SC) \quad \min_{\mathbf{s}} H(\mathbf{s}) \quad s.t. \quad R(\mathbf{s}) \geq \beta.$$

Shang and Song (2006) show that the above single-stage bounds can be applied to the service-level constrained model by using an imputed backorder cost parameter $b = \beta h_1' / (1 - \beta)$.

5.2.3 Nonstationary Demand

For many products, the demand process is nonstationary due to seasonality, weather, and technological or economic factors.

5.2.3.1 Independent but nonidentically distributed demand

Clark and Scarf (1960) study a periodic-review, finite-horizon system in which the demand distributions across periods are independent but may not be identical. The objective is to minimize the expected total discounted cost of the system. They show that under this demand model, a time-varying echelon base-stock policy $\{s_j^*(t), j = 1, \dots, N, t = 1, \dots, T\}$ is optimal, where T is the length of the planning horizon, $s_j^*(t)$ is the optimal echelon base-stock level for stage j in period t . Despite its simple structure, the computation of the policy parameters is complex, requiring solving multiple functional equations recursively.

To simplify the computation, Shang (2012) develops a single-stage heuristic for the optimal policy of this model. The idea is similar to that of Shang and Song (2003): for each echelon- j system, we construct a lower- and an upper-bound system such that $s_j^*(t)$ is bounded by the solutions of these bounding systems in each period. Specifically, the upper-bound system, which generates the lower-bound solution $s_j^\ell(t)$, is constructed by requiring stage $i (< j)$ to always order up to stage $i+1$'s echelon net inventory in each period. By doing so, an inventory unit ordered by stage j will arrive at stage 1 in $L_{[1,j]}$ periods. In other words, stage i , $i = 2, 3, \dots, j$ can be viewed as an in-transit point and the echelon- j system collapses into a single-stage system with a lead time $L_{[1,j]}$. The holding and backorder costs are h_j^u and b_j , respectively. To compute the resulting unit purchase cost c_j^u , consider the total cost incurred for a unit ordered by stage j . This unit incurs an order cost when it arrives at each of its downstream stages and a holding cost in each period before it arrives at stage 1. That is,

$$c_j^u = c_j + \sum_{i=2}^j \alpha^{L_{[i,j]}} (c_{i-1} + h_i).$$

The resulting upper-bound system is $\mathbb{S}[1, (c_j^u, h_j^u, L_{[1,j]}), b_j, \mathbf{D}]$, and optimal solution in each period t , $s_j^u(t)$, can be obtained by solving a one-dimensional dynamic program.

The lower-bound system can be constructed with a similar idea: by regulating $h_i = 0$ and $c_i = 0$ for $i < j$, there is no incentive to keep the inventory at stage $i = j, j-1, \dots, 2$, and the series system collapses into a single-stage system $\mathbb{S}[1, (c_j^\ell, h_j^\ell, L_{[1,j]}), b_j, \mathbf{D}]$, where

$$c_j^\ell = c_j + \alpha^{L_{[1,j]}} h_j.$$

Let the optimal solution for the lower-bound system be $s_j^u(t)$ and x_j be the initial inventory position before ordering for stage j . Shang (2012) shows the following results.

Theorem 5.2

- (1) If $x_{j-1} < s_{j-1}^*(t)$, $s_j^*(t) \geq s_j^\ell(t)$ for $t > L_{[1,j]}$, where $s_j^\ell(t)$ is the solution obtained from $\mathbb{S}[1, (c_j^u, h_j^u, L_{[1,j]}), b_j, \mathbf{D}]$.
- (2) $s_j^u(t) \geq s_j^*(t)$ for $t > L[1, j]$, where $s_j^u(t)$ is the solution obtained from $\mathbb{S}[1, (c_j^\ell, h_j^\ell, L_{[1,j]}), b_j, \mathbf{D}]$.

To further simplify the computation for $s_j^\ell(t)$ and $s_j^u(t)$, Shang (2012) suggests using a myopic solution in Equation (5.3) to approximate these solutions. A numerical study shows that the myopic solutions are quite effective.

5.2.3.2 Time-series demand model

Dong and Lee (2003) extend Clark and Scarf (1960) to an infinite-horizon problem with a time-series demand model, known as the Martingale model of forecast evolution (MMFE). The MMFE characterizes the demand forecast evolution by incorporating both past demands and other influential factors. It includes an ARMA (autoregressive and moving average) process with a minimum mean-square error forecast as a special case. Dong and Lee (2003) show that the optimal policy for this system is a state-dependent echelon base-stock policy $\{s_j^*(\mathbf{D}), j = 1, \dots, N\}$, where the state $\mathbf{D}$ is the demand forecast made before making the ordering decision. The computation of the policy parameters is very complex, if not prohibitive, due to the continuous state. As a consequence, the authors develop a simple solution $\hat{s}_j(\mathbf{D})$ to approximate $s_j^*(\mathbf{D})$. The computation of $\hat{s}_j(\mathbf{D})$ is based on a simple approximation of the induced penalty function for echelon- j . For AR(1) demand process, $s_j(\mathbf{D})$ can be obtained in closed form, which sheds light on the effect of the primitive model parameters. Under certain assumptions, such as costless returns of excess inventory to the upstream stage, the optimal myopic base-stock levels $s_j^*(\mathbf{D})$ are always attainable, $\hat{s}_j(\mathbf{D})$ is a lower bound to $s_j^*(\mathbf{D})$, and the resulting system cost is an upper bound to the optimal system cost. For IID demand, the solution lower bound and cost upper bound coincide with those in Chao and Zhou (2007) and Shang and Song (2003) (by setting $\alpha = 1$ in Dong and Lee (2003)).

5.2.3.3 Markov-modulated demand

Chen and Song (2001) extend Clark and Scarf (1960) to another time-correlated demand model – Markov-modulated demand (MMD) – where the demand distribution of each period depends on a “world” state of an underlying Markov chain capturing the evolution of environmental factors that influence the demand, such as weather, consumer taste, seasonality, product life cycles, and economic conditions. Specifically, let $k = 1, \dots, \mathcal{K}$ be the state of the Markov chain $\{M(t), t = 1, 2, \dots\}$. Suppose that in period t , $M(t) = k$, and the demand in period t follows a distribution $F_k(\cdot)$. The independent but nonidentical demand in §5.2.3.1 is a special case of this model with T states, and the transition probability from state t to $t+1$ is one and zero otherwise. The MMD demand process is time-correlated through the Markov chain, which is external to the demand process. By contrast, under MMFE in §5.2.3.2, the demand in each period explicitly depends on the demand history. Chen and Song (2001) show that a world-state-dependent echelon base-stock policy $\{s_j^*(k), j = 1, \dots, N, k = 1, \dots, \mathcal{K}\}$ is optimal

that minimizes the long-run average cost. That is, if in period t , $M(t) = k$, then echelon- j 's optimal action is to follow a base-stock policy with base-stock level $s_j^*(k)$. They also provide an iterative algorithm to compute the optimal policy.

Muharremoglu and Tsitsiklis (2008) show that the state-dependent echelon base-stock policy continues to be optimal for more general systems with (exogenous and sequential) stochastic lead times and discrete demand processes (including MMD). They decompose the problem into a series of single-unit single-demand problems, and develop a dynamic programming algorithm for computing the optimal policy.

To simplify the computational complexity of these exact algorithms, Chen et al. (2017) develop simple-to-compute bounds and heuristics on optimal echelon base-stock levels based on derivative bounds on the optimality equations. When the state space of the Markov chain degenerates to a singleton, the MMD demand process becomes IID and their bounds reduce to those in Shang and Song (2003). For the special case of independent but nonidentical demand, their solution upper bounds are the same as those in Shang (2012), and the single-stage lower bounding system is constructed in a similar way as in Shang and Song (2003) and Shang (2012). Their solution lower bounds, however, differ from those in Shang (2012), because the derivative-based approach does not assume the optimal echelon base-stock levels to be achievable in each period. In general, this assumption does not hold; see Song and Zipkin (1993) and Song and Zipkin (1996), thus the "lower bound" obtained using this assumption may fail to bound the optimal solution under MMD, as illustrated in the numerical studies in Chen et al. (2017).

Computing the derivative-based solution lower bound involves optimization over demand state permutations, which can be challenging if there is a large number of states. To simplify computation, Chen et al. (2017) develop easy-to-compute heuristic solutions that only require evaluating a linear combination of derivatives of the newsvendor cost functions of each demand state, which also makes the solutions more intuitive.

Abhyankar and Graves (2001) design a heuristic policy for a two-stage system with MMD, assuming the upstream stage can always fulfill orders from the downstream stage. With this assumption, the two-stage system is decoupled into two single-stage systems, with the downstream stage operating under a state-dependent installation base-stock policy, and the upstream stage operating under a static installation base-stock policy.

5.2.4 Model with Expediting

In this section, we consider an infinite-horizon, periodic-review series system where each stage has two ordering decisions: expedited order and regular order. For each stage $1 \leq i \leq N$, the lead time L_i^r for regular orders is 1, and L_i^e for expedited orders is 0 (the analysis and results can be extended to $L_i^e = l_i$ and $L_i^r = l_i + 1$ for a general nonnegative l_i ; see Zhou and Chao (2009) for details). The assumption that the lead-time difference between a regular and an expedited order must be 1 appears restrictive, but relaxing it makes the problem too complicated to yield an optimal control policy that is analytically solvable. This is because the resulting state space for each stage has to be augmented to include the pipeline inventory scheduled to arrive in future periods. In this case, it is known that, even for a single-stage system $N = 1$, the optimal policy is complicated and state-dependent (Whittemore and Saunders, 1977). For convenience, let $\mathbf{L}_i = (L_i^e, L_i^r)$.

For each stage i , the unit expedited and regular order cost from stage $i + 1$ is k_i^E and k_i^R , respectively, with $k_i^E > k_i^R$. The sequence of events is as follows. First, at the beginning of every period, each stage receives the regular order placed in the previous period. Second, starting from stage N , each stage places expedited and regular orders sequentially. Specifically, stage N first places its expedited order from the outside supplier and receives it immediately. It then places a regular order which will be delivered at the beginning of the next period (note that stage N 's expedited order is immediately available to satisfy the order from stage $N - 1$). Stage $N - 1$ then decides its expedited and regular orders from stage N . Again, the expedited order stage $N - 1$ received can be used to satisfy the order from stage $N - 2$. This top-down ordering process continues until stage 1 places its expedited and regular orders from stage 2. Finally, demand is realized during the period at stage 1 and all costs are incurred at the end of the period. The objective is to minimize the expected total discounted cost over an infinite planning horizon.

Lawson and Porteus (2000) show that the optimal policy of this problem is of an echelon base-stock type with two echelon base-stock levels at each stage, and the optimal base-stock levels can be computed through a nested recursive algorithm. First, let $c_i^E = k_i^E - k_i^R + h_i$ and $c_i^R = \alpha k_i^E - k_i^R$. Because $k_i^E > k_i^R$ and $h_i \geq 0$, c_i^E is positive. In addition, we assume $c_i^R \geq 0$. If this is not the case, then the regular shipping mode will never be used and the model reduces to the Clark–Scarf model with a single supply mode between stages. Let $\mathbf{c}_i = (c_i^E, c_i^R)$. Then we can denote the series system with expediting as $\mathbb{S}^e[N, (\mathbf{c}_i, h_i, \mathbf{L}_i)_{i=1}^N, b, \mathbf{D}]$.

We now present the computational algorithm for the optimal base-stock levels. Denote $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$. Define $G_i^E(y) = c_i^E y + (h_i + b)E[(y - D)^-]$, which is convex with minimizer s_i^E . Let, for $i = 1, \dots, N$,

$$G_i^R(y) = -c_i^R y + G_i^E(y \wedge s_i^E) + \alpha E[G_i^E((y - D) \vee s_i^E)], \quad (5.9)$$

$$s_i^R = \arg \min_y G_i^R(y), \quad (5.10)$$

and, for $i = 1, \dots, N - 1$,

$$G_{i+1}^E(y) = c_{i+1}^E y + G_i^R(y \wedge s_i^R), \quad (5.11)$$

$$s_{i+1}^E = \arg \min_y G_{i+1}^E(y), \quad (5.12)$$

where both $G_i^E(\cdot)$ and $G_i^R(\cdot)$ are univariate convex functions. Note that the algorithm above does not directly lead to the minimum expected total discounted cost for the system, which can be obtained through successive approximation from the dynamic program (see the dynamic program formulation in Lawson and Porteus (2000)). The optimal top-down echelon base-stock policy (Lawson and Porteus (2000)) works as follows. Starting from stage N , each stage tries to raise its echelon net inventory and position to expedited base-stock level s_i^E and regular base-stock level s_i^R , respectively, taking upstream decisions as given and ignoring downstream decisions. Let $(\mathbf{s}^E, \mathbf{s}^R)$ denote the vectors of optimal base-stock levels.

The main idea of deriving the lower (upper, respectively) bounds to the optimal base-stock levels is to find simple upper (lower, respectively) bounds of the first order derivative of G_i^E and G_i^R (see, Chao and Zhou (2007); Zhou and Chao (2009)). Since s_1^E is known in a closed

form, we shall only develop bounds for s_i^E , $i \geq 2$, and for s_i^R , $i \geq 1$. Let $c_0^R = 0$. In the following, for ease of presentation, let $\bar{F}$ denote the complementary CDF of one-period demand and $\bar{F}_i$ denote the complementary CDF of i -period demand.

Theorem 5.3 For $i = 1, \dots, N$, the lower bounds for s_i^E and s_i^R are, respectively,

$$s_i^{E,\ell} = \bar{F}^{-1} \left(\frac{\sum_{j=1}^i \alpha^{i-j} (c_j^E - c_{j-1}^R)}{\alpha^{i-1} (h_1 + b)} \right), \quad (5.13)$$

and

$$s_i^{R,\ell} = \bar{F}^{-1} \left(\frac{-c_i^R + \sum_{j=1}^i \alpha^{i-j+1} (c_j^E - c_{j-1}^R)}{\alpha^i (h_1 + b)} \right). \quad (5.14)$$

We next present a set of newsvendor-type upper bounds, which are obtained by constructing lower-bound functions for $(G_i^E(y))'$ and $(G_i^R(y))'$.

Theorem 5.4 For $i = 1, \dots, N$, the upper bounds for s_i^E and s_i^R are, respectively:

$$s_i^{E,u} = \bar{F}_i^{-1} \left(\frac{c_i^E - c_{i-1}^R + \alpha c_{i-1}^E}{h_1 + b - \sum_{j=1}^{i-2} (\alpha c_j^E - c_j^R)} \right); \quad (5.15)$$

and

$$s_i^{R,u} = \bar{F}_{i+1}^{-1} \left(\frac{\alpha c_i^E - c_i^R}{\alpha \left(h_1 + b - \sum_{j=1}^{i-1} (\alpha c_j^E - c_j^R) \right)} \right). \quad (5.16)$$

Note that Zhou and Chao (2009) derive other sets of simple lower and upper bounds for (s^E, s^R) . Based on these simple bounds, we can develop a heuristic by using weighted averages of these solution bounds. Zhou and Chao (2009) demonstrate that a heuristic with equal weights of lower and upper bounds performs quite well numerically.

5.3 SERIES SYSTEMS WITH FIXED ORDER COSTS

We now consider the system with fixed order costs. The system configuration and primitives are the same as those in Section 5.2. In addition, there is a fixed order cost k_j incurred for each inventory replenishment at stage j . When fixed order costs are not negligible, a firm may not order inventory at every review point as it will incur too much ordering cost. If a

firm has installed an automatic replenishment system, which is able to review the inventory position continuously, an echelon (r, q) policy can be implemented to replenish inventory. On the other hand, many supply-chain firms replenish inventory according to a fixed schedule and echelon (s, T) policies are ideal for shipment coordination. We only focus on the average cost criterion. In Section 5.3.1 we set the stage by reviewing the basic results for the single-stage system. We consider (r, q) policies in Section 5.3.2 and (s, T) policies in Section 5.3.3 for multi-stage systems. For each of these two policies, we further consider two scenarios, depending on whether the batch size q or reorder interval T is fixed. If these variables are fixed, the total order cost per period becomes a constant and we shall omit it in the total cost function.

We continue to define notations to represent these four systems.

The (r, q) model with fixed batch sizes: $\mathbb{S}_Q[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$;

The general (r, q) model: $\mathbb{S}_Q[N, (h_i, k_i, L_i)_{i=1}^N, b, \mathbf{D}]$;

The (s, T) model with fixed reorder intervals: $\mathbb{S}_T[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$;

The general (s, T) model: $\mathbb{S}_T[N, (h_i, k_i, L_i)_{i=1}^N, b, \mathbf{D}]$.

5.3.1 SINGLE-STAGE (r, q) AND (s, T) POLICIES

We consider the special case with $N = 1$. When fixed order cost $k > 0$, it is known that an (r, q) policy is optimal for continuous-review systems with Poisson demand, and a (s, S) policy is optimal for periodic-review systems (or continuous-review system with compound Poisson demand). However, due to its ease of coordinating production batches between stages, the (r, q) policy is often adopted in practice.

For the $\mathbb{S}_Q[1, (h, k, L), b, \mathbf{D}]$ system under an (r, q) policy, the manager monitors the inventory position continuously and places an order of size q whenever the inventory position reaches the reorder point r . For the compound Poisson demand case, an (r, q) policy is replaced with the (r, nq) policy, i.e., placing an order of an integer multiple of batch size q so that the inventory position is between $r + 1$ and $r + q$. The analysis for the (r, nq) policy is essentially the same as that for the (r, q) policy (see, e.g., Shang (2008)). Clearly, a continuous-review base-stock policy with base-stock levels s is a special case of the (r, q) policy by setting $q = 1$ and $s = r + 1$.

The inventory order position is uniformly distributed over $[r + 1, r + q]$ and the total average cost per unit of time under an (r, q) policy can be expressed as follows:

$$C(r, q) = \frac{k\lambda}{q} + \frac{\sum_{y=1}^q G(r+y)}{q}. \quad (5.17)$$

The first term in Equation (5.17) is the average fixed order cost per unit time (note that if under a (r, nq) policy, there could be potentially two types of fixed cost, one is incurred per batch, the other is incurred per order) and the second term is the average inventory holding cost and backorder cost per unit time. The cost function $C(r, q)$ is jointly convex in r and q , so we can obtain the optimal policy by first fixing q and obtaining the best reorder point:

$$r(q) = \arg \min_r \left\{ \sum_{y=1}^q G(r+y) \right\} \quad (5.18)$$

Then, one can obtain the optimal order quantity q^* and the optimal reorder point $r^* = r(q^*)$. See Federgruen and Zheng (1992) for an efficient algorithm.

For the $\mathbb{S}_T[1, (h, k, L), b, \mathbf{D}]$ system under an (s, T) policy, the inventory manager reviews the inventory position every T period and places orders, if needed, to keep the inventory position at the base-stock level s . It is clear that if $T = 1$, the (s, T) policy reduces to the conventional base-stock policy for periodic-review systems. The total average cost per unit of time under a (s, T) policy can be expressed as follows (see, e.g., Liu and Song (2012)).

$$C(s, T) = \frac{k}{T} + \frac{\sum_{\tau=1}^T G(s, \tau)}{T}. \quad (5.19)$$

where

$$G(y, \tau) = E[h(y - D[L + \tau])^+ + b(y - D[L + \tau])^-]. \quad (5.20)$$

Here $G(y, \tau)$ is the expected inventory holding and demand backlogging cost at the end of period $L + \tau$ given the current inventory position is y . Here, $C(s, T)$ is not jointly convex. Liu and Song (2012) provide an efficient algorithm to compute the optimal s^* and T^* .

5.3.2 MULTI-STAGE (R, Q) POLICIES

We consider the $\mathbb{S}_Q[N, (h_i, k_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system with Poisson demand with demand rate λ . The echelon (r, q) policy is implemented as follows: Stage j monitors $IOP_j(t)$ continuously, and will place an order of size q_j when $IOP_j(t)$ reaches the reorder point r_j . There is a fixed order cost k_j for each stage j whenever an order is placed. It is often assumed that the batch sizes satisfy integer-ratio relations, i.e., $q_j = m_j q_{j-1}$, where $q_j, m_j \in \mathbb{N}$ and $\mathbb{N}$ is the set of positive integers, $j = 2, \dots, N$, for coordination of orders and ease of analysis.

Let $\mathbf{r} = (r_1, \dots, r_N)$ and $\mathbf{q} = (q_1, \dots, q_N)$. The total average cost per unit of time can be expressed as follows (see Chen (2000)).

$$\begin{aligned} C(\mathbf{r}, \mathbf{q}) &= \sum_{i=1}^N \frac{k_i \lambda}{q_i} + E\left[\sum_{i=1}^N h_{[i, N]} I_i' + bB + \sum_{i=2}^N h_{[i, N]} D_{i-1}\right] \\ &= \sum_{i=1}^N \frac{k_i \lambda}{q_i} + E\left[\sum_{i=1}^N h_i IN_i + (b + h_1)B\right]. \end{aligned} \quad (5.21)$$

The first term in Equation (5.21) is the average fixed order cost per unit of time, and the second term is the average inventory holding cost and backorder cost per unit of time.

To evaluate the cost, we only need to characterize IN_j for all j . This process starts from stage N , $N-1$, sequentially until stage 1. More specifically, $IP_N = IOP_N$, which is uniformly

distributed over $\{r_N + 1, \dots, r_N + q_N\}$. For $j = N, \dots, 1$, $IN_j = IP_j - D_j$, and for $j = N - 1, \dots, 1$, $IP_j = O_j[IN_{j+1}]$, where

$$O_j[x] = \begin{cases} x, & \text{if } x \leq r_j, \\ x - mq_j, & \text{otherwise,} \end{cases} \quad (5.22)$$

with m being the largest integer such that $x - mq_j > r_j$. After we characterize all IN_j , we can use Equation (5.21) to find the average cost per unit of time.

We can also derive a similar bottom-up recursion to evaluate the average inventory holding and backorder costs. The idea is similar to that of the base-stock system. Recall $G_1(y)$ in Equation (5.5). For $j = 2, \dots, N$,

$$G_j(y) = E[h_j(y - D_j) + G_{j-1}(O_{j-1}[y - D_j])]. \quad (5.23)$$

Then,

$$C(\mathbf{r}, \mathbf{q}) = \sum_{i=1}^N (k_i \lambda) / q_i + \sum_{x=1}^{q_N} G_N(r_N + x) / q_N. \quad (5.24)$$

Note that when $q_j = 1, \forall j$, the recursion in Equation (5.23) reduces to Equation (5.6).

For fixed batch sizes $\mathbf{q}$, the optimal reorder point r_j^* can be obtained recursively.

First, $r_1^* = \arg \min_y \bar{G}_1(y) \stackrel{\text{def}}{=} \sum_{x=1}^{q_1} G_1(y + x)$. For $j = 2, \dots, N$, suppose r_{j-1}^* is known.

Substituting r_{j-1}^* for r_{j-1} in the O_{j-1} function in Equation (5.23), the optimal reorder point is

$$r_j^* = \arg \min_y \bar{G}_j(y) \stackrel{\text{def}}{=} \sum_{x=1}^{q_j} G_j(y + x).$$

To find the optimal batch sizes $\mathbf{q}^*$, Chen and Zheng (1998) develop lower and upper bounds to the total cost function. These cost bounds are a sum of N separable cost functions of batch sizes. With these results, they obtain the bounds for the optimal batch size for each stage. The optimal batch sizes can then be found via enumeration within the bounds. See Shang and Zhou (2010) for an alternative algorithm.

5.3.2.1 Bounds for reorder points with fixed batch sizes

We consider the $\mathbb{S}_{\mathbb{Q}}[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system, i.e., the batch sizes are fixed. The recursions Equation (5.23) are similar to those in the base-stock policy. We observe that for each stage j , the optimal reorder point r_j^* does not depend on the decisions at upstream stages. To determine r_j^* , the echelon- j manager only needs to know b , $h_{[1, N]}$, and the parameters within his/her echelon: $(r_i^*, q_i, D[L_i], h_i)$, for $i < j$. Following the same idea, we can charge the minimum holding cost rate h_j and the largest holding cost rate $h_{[1, j]}$ to construct a lower-bound and an upper-bound system for echelon- j , respectively. The resulting cost functions are $G_j^\ell(\cdot)$ and $G_j^u(\cdot)$, and $G_j^\ell(\cdot) \leq G_j(\cdot) \leq G_j^u(\cdot)$.

Recall that stage j follows a (r_j, q_j) policy. With ample supply for echelon- j , IOP_j is uniformly distributed over $\{r_j + 1, \dots, r_j + q_j\}$. Thus, the average inventory holding and backorder cost for echelon- j takes an expectation over IOP_j and we can derive the following:

$$\frac{1}{q_j} \sum_{x=1}^{q_j} G_j^\ell(y+x) + \underline{\tau}_j \leq \frac{1}{q_j} \sum_{x=1}^{q_j} G_j(y+x) \leq \frac{1}{q_j} \sum_{x=1}^{q_j} G_j^u(y+x) + \bar{\tau}_j.$$

Thus, the echelon- j cost is bounded by the costs of $\mathbb{S}_{\mathbb{Q}}[1, (h_j^\ell, L_{[1,j]}), b_j, \mathbf{D}]$ and $\mathbb{S}_{\mathbb{Q}}[1, (h_j^u, L_{[1,j]}), b_j, \mathbf{D}]$.

Define $r_j^u = \operatorname{argmin}_y \left\{ \sum_{x=1}^{q_j} G_j^u(y+x) \right\}$, and $r_j^\ell = \operatorname{argmin}_y \left\{ \sum_{x=1}^{q_j} G_j^\ell(y+x) \right\}$.

Theorem 5.5 For $j = 1, \dots, N$, (1) $G_j^u(y) + \tau_j \geq G_j(y) \geq G_j^\ell(y) + \tau_j$, for all y . (2) $r_j^\ell \leq r_j^* \leq r_j^u$. When $j = 1$, the above inequalities reduce to equalities.

Again, one can use the weighted-average approaches on either the reorder points or holding cost rates to obtain an approximation of r_j^* .

5.3.2.2 Simple solution for batch sizes

We now consider the $\mathbb{S}_{\mathbb{Q}}[N, (h_i, k_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system. Chen and Zheng (1998) construct cost bounds to replace the average total cost in the objective function. Their idea is to construct upper and lower bounds for the *induced penalty cost* function (i.e., the penalty cost charged to the upstream stage for not being able to fulfill the downstream orders) for each stage j . These penalty cost-bound functions only depend on a stage's batch size and are independent of its downstream's batch sizes. With the stage's holding cost, along with the constructed induced penalty cost bounds, they develop cost bounds for the total cost. These cost bounds are a sum of separable functions of batch sizes. Thus, a standard clustering algorithm (e.g., Maxwell and Muckstadt (1985)) can be applied to find the optimal solution for these cost-bound problems. They propose heuristics by converting these optimal solutions to power-of-two batch sizes.

Shang (2008) proposes a simpler heuristic that employs the same two steps as the heuristic for the deterministic model. The key idea is that the single-stage lower-bound function $G_j^\ell(\cdot)$ can be used to generate an effective batch size q_j . Intuitively, for the single-stage (r, q) system, the optimal batch size q^* is determined by the "shape" of the inventory holding and backorder cost function, as well as the fixed order cost k (Zheng (1992)). Shang (2008) first decouples the total system cost for given batch sizes with the corresponding optimal reorder points into each stage and shows that the lower-bound function $G_j^\ell(\cdot)$ has a similar shape as that of the stage cost function. Consequently, he replaces the stage cost function with $G_j^\ell(q_j)$ and derives a simple heuristic that includes two steps: *clustering* and *minimization*.

In the clustering step, the stages are grouped into disjoint clusters $\{c(1), c(2), \dots, c(M)\}$ according to cost ratios. Let $\mathcal{S} = \{1, 2, \dots, N\}$. For any $i, j \in \mathcal{S}$ with $i \leq j$, the set $\{i, i+1, \dots, j\}$ is called a cluster. These consecutive stages in each cluster will use the same batch size. Specifically, define

$$h[m] = \sum_{i \in c(m)} h_i, \quad h^-[m] = \sum_{i \in c(m)} h_{[i, N]}, \quad \text{and} \quad k[m] = \sum_{i \in c(m)} k_i.$$

These clusters satisfy the following two conditions:

- (i) $k[1]/h[1] < \dots < k[M]/h[M]$,
- (ii) for each cluster $c(m) = \{l_1, \dots, l_2\}$, there does not exist an l with $l_1 \leq l < l_2$ so that $k[m^-]/h[m^-] < k[m^+]/h[m^+]$, where $c(m^-) = \{l_1, \dots, l\}$ and $c(m^+) = \{l+1, \dots, l_2\}$.

Through a two-dimensional diagram suggested by Zipkin (2000), the above clusters $\{c(1), c(2), \dots, c(M)\}$ can be conveniently identified.

In the minimization step, a single-stage problem is solved for each cluster $c(m)$ sequentially, starting with $m = 1$. In each problem, the solution of batch size $Q_{c(m)}$ is restricted to be an integer multiple of $Q_{c(m-1)}$, $m \geq 2$. $Q_{c(m)}$ is the solution to the following problem:

$$\min_Q \left\{ \frac{\lambda k[m] + \sum_{x=1}^Q G_{c(m)}^\ell(r^\ell(Q) + x)}{Q} \right\}, \quad (5.25)$$

$$\text{s.t. } Q = qQ_{c(m-1)}, \quad q \in \mathfrak{T}^+, \quad m > 1,$$

where

$$r^\ell(Q) = \arg \min_y \left\{ \sum_{x=1}^Q G_{c(m)}^\ell(y + x) \right\}, \quad G_{c(m)}^\ell(y) = \sum_{i \in c(m)} G_i^\ell(y),$$

and $G_j^\ell(\cdot)$ is the cost function of $\mathbb{S}[1, (h_j^\ell, L_{[1,j]}), b_j, \mathbf{D}]$. Then, set $q_i^a = Q_{c(m)}$ for $i \in c(m)$; $(q_1^a, \dots, q_N^a)$ are the heuristic batch sizes. One can apply the recursion in Equation (5.23) to find the corresponding heuristic reorder points $(r_1^a, \dots, r_N^a)$ or the single-stage approximations for the reorder points.

5.3.3 Multi-Stage (s, T) Policies

This section considers a multi-stage system with (s, T) policies. We assume that demands in different periods are IID and integer-valued. Let λ denote the mean of the one-period demand. Section 5.3.3.1 assumes that the reorder intervals are fixed and studies the $\mathbb{S}_{\mathbb{T}}[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system, and we present heuristics for the optimal base-stock levels. Section 5.3.3.2 assumes that the reorder intervals are decision variables and studies the $\mathbb{S}_{\mathbb{T}}[N, (h_i, k_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system. We provide an approximation for the optimal reorder intervals by solving single-stage (s, T) models.

An echelon (s, T) policy is operated as follows: Stage j orders at the beginning of every T_j period. If the echelon inventory order position is less than an echelon base-stock level s_j , the stage orders to bring the inventory order position back to s_j . Clearly, when $T_j = 1$, the (s, T) policy reduces to the periodic-review base-stock policy. We refer to these T_j -th periods as order periods, and T_j as the reorder interval of stage j . The reorder intervals follow integer-ratio relations: $T_{j+1} = n_j T_j$, where $T_j, n_j \in \mathbb{N}$, $j = 1, \dots, N-1$. We assume that all shipments are synchronized. That is, a downstream stage, whenever possible, places an order when its upstream stage receives a shipment. (A synchronized shipping policy dominates a non-synchronized one; see Chao and Zhou (2009).) The objective is to find the (s, T) policy such that the average total cost per period is minimized.

We now discuss how to evaluate the average total cost per period under the (s, T) policy. This total cost includes two parts, the average inventory holding and backorder cost (inventory-related cost) per period and the average review cost per period. We first show how to evaluate the former.

Consider the dynamics of the echelon inventory variables under the (s, T) policies. Suppose that stage N places an order at the beginning of an order period t . Define a *cycle* for stage j , $j = 1, \dots, N$, with respect to t as a time interval that includes periods $t + L_{[j, N]} + \tau$, $\tau = 0, \dots, T_N - 1$. As we shall see below, this order will directly or indirectly determine IN_j^- and IN_j within stage j 's cycle (as here we consider periodic-review systems, IN_j^- denote the echelon net inventory before a period and IN_j the ending echelon net inventory of a period). Since the system repeats itself when stage N places an order in every T_N periods, it is a regenerative process with a cycle length of T_N periods. Thus, the long-run average inventory-related cost per period is equal to the expected inventory-related cost incurred in the cycle divided by T_N . Since the expected cost is determined by IN_j , we show below how to derive IN_j within the cycle.

We start from stage N . Let $D[t_1, t_2]$ denote the total demand in periods $t_1, t_1 + 1, \dots, t_2 - 1$ and $D[t_1, t_2]$ denote the total demand in periods $t_1, t_1 + 1, \dots, t_2$. Suppose that stage N orders at the beginning of an order period t . Because stage N has ample supply, $IP_N(t) = IOP_N(t) = s_N$. This order will arrive at stage N in period $t + L_N$. Since there will be no other order periods until period $t + T_N$, for $\tau = 0, \dots, T_N - 1$,

$$IN_N^-(t + L_N + \tau) = s_N - D[t, t + L_N + \tau),$$

and $IN_N(t + L_N + \tau) = IN_N^-(t + L_N + \tau) - D$.

Now consider stage $j = N - 1, N - 2, \dots, 1$ sequentially. Define $\lfloor a \rfloor$ as a roundoff operator, which returns the greatest integer less than or equal to a , a real number. Let $\text{mod}_x(y)$ be an operator that returns the remainder of y divided by x , $x \in \mathbb{N}$, and $y \in \{0, \mathbb{N}\}$. According to the synchronized replenishment rule, stage j will order in periods $t + L_{[j+1, N]} + \lfloor \tau / T_j \rfloor T_j$, for $\tau = 0, \dots, T_N - 1$. IP_j is determined jointly by $IOP_j = s_j$ and stage $j + 1$'s net echelon net inventory IN_{j+1}^- . That is, for $\tau = 0, \dots, T_N - 1$,

$$\begin{aligned} IP_j \left(t + L_{[j+1, N]} + \left\lfloor \frac{\tau}{T_j} \right\rfloor T_j \right) \\ = \min \left\{ IN_{j+1}^- \left(t + L_{[j+1, N]} + \left\lfloor \frac{\tau}{T_j} \right\rfloor T_j \right), s_j \right\}. \end{aligned} \quad (5.26)$$

Equation (5.26) means that if stage $j + 1$ has sufficient stock such that $IN_{j+1}^- > s_j$, $IP_j = s_j$. Otherwise, stage $j + 1$ will ship as much as possible, in which case $IP_j = IN_{j+1}^-$.

The IP_j in the order periods will further determine IN_j^- and IN_j within periods $t + L_{[j, N]} + \tau$, $\tau = 0, \dots, T_N - 1$:

$$\begin{aligned} IN_j^-(t + L_{[j, N]} + \tau) &= IP_j \left(t + L_{[j+1, N]} + \left\lfloor \frac{\tau}{T_j} \right\rfloor T_j \right) \\ &\quad - D \left[t + L_{[j+1, N]} + \left\lfloor \frac{\tau}{T_j} \right\rfloor T_j, t + L_{[j, N]} + \left\lfloor \frac{\tau}{T_j} \right\rfloor T_j + \text{mod}_{T_j}(\tau) \right), \end{aligned} \quad (5.27)$$

and $IN_j(t + L_{[j, N]} + \tau) = IN_j^-(t + L_{[j, N]} + \tau) - D$.

We write $IN_j(\tau)$ to represent $IN_j(t + L_{[j,N]} + \tau)$ in steady state. The long-run average inventory-related costs per period are equal to

$$G(\mathbf{s}, \mathbf{T}) = \frac{1}{T_N} \mathbb{E} \left[\sum_{\tau=0}^{T_N-1} \left(\sum_{j=1}^N h_j IN_j(\tau) + (b + h_1) [IN_1(\tau)]^- \right) \right], \quad (5.28)$$

where $\mathbf{s} = (s_1, \dots, s_N)$ and $\mathbf{T} = (T_1, \dots, T_N)$.

Below we provide a similar bottom-up recursion to conveniently evaluate $G(\mathbf{s}, \mathbf{T})$. The idea behind this scheme is similar to that for the (r, q) policy: At each iteration, we evaluate the average inventory-related costs for echelon- j , referred to as $G_j(y, \mathbf{T}_j)$ where $\mathbf{T}_j = (T_1, T_2, \dots, T_j)$, providing that stage j 's echelon inventory order position $IOP_j(t)$ is equal to y and its downstream stage $i (< j)$ follows an (s, T) policy with parameters (s_i, T_i) .

Proposition 5.2 *Define*

$$G_1(y, \mathbf{T}_1) = \frac{1}{T_1} \left(\sum_{\tau=0}^{T_1-1} \mathbb{E} [h_1 (y - D[L_1 + \tau]) + (b + h_{1,N}) (y - D[L_1 + \tau])^-] \right). \quad (5.29)$$

For $j = 2, \dots, N$, define recursively

$$\begin{aligned} G_j(y, \mathbf{T}_j) = & \frac{1}{T_j} \sum_{\tau=0}^{T_j-1} \mathbb{E} [h_j (y - D[L_j + \tau]) \\ & + G_{j-1} \left(\min \left\{ s_{j-1}, \left(y - D \left[L_j + \left\lfloor \frac{\tau}{T_{j-1}} \right\rfloor T_{j-1} \right) \right\} \right) \right]. \end{aligned} \quad (5.30)$$

Then, $G(\mathbf{s}, \mathbf{T}) = G_N(s_N, \mathbf{T}_N)$.

We next determine the average fixed costs per period. For stage j , the review cost k_j is incurred for every T_j . So the average cost per period is k_j / T_j . The total cost per period is

$$C(\mathbf{s}, \mathbf{T}) = \sum_{j=1}^N \left(\frac{k_j}{T_j} \right) + G(\mathbf{s}, \mathbf{T}). \quad (5.31)$$

Similar to the (r, q) policy, for fixed reorder intervals $\mathbf{T}$, the optimal base-stock levels can be obtained recursively. First, let $s_1^* = \arg \min_y G_1(y, \mathbf{T}_1)$. For $j = 2, \dots, N$, suppose s_{j-1}^* is known, we substitute s_{j-1}^* for s_{j-1} in Equation (5.30), and let $s_j^* = \arg \min_y G_j(y, \mathbf{T}_j)$. Then, $(s_1^*, \dots, s_N^*)$ are the optimal base-stock levels. Note that $G_j(y, \mathbf{T}_j)$ are convex in y for all j .

Finding the optimal reorder intervals is more complicated. Shang and Zhou (2010) decompose the total cost function using an induced penalty cost function, i.e., the penalty cost charged to an upstream stage if the stage cannot fulfill its downstream stage's order. Then,

they develop bounds to the induced penalty cost function by regulating downstream reorder intervals. With these steps, they obtain solution bounds for the optimal reorder intervals. An enumeration within the solution bounds yields the optimal reorder intervals.

5.3.3.1 Bounds for base-stock levels with fixed reorder intervals

We consider the $\mathbb{S}_{\mathbb{T}}[N, (h_i, L_i)_{i=1}^N, b, \mathbf{D}]$ system, and develop bounds for the optimal base-stock levels $(s_1^*, \dots, s_N^*)$ when $\mathbf{T}$ is fixed. By setting the holding cost to the minimum and maximum value within the echelon, one can derive the lower-bound system $\mathbb{S}_{\mathbb{T}}[1, (h_j, L_{[1,j]}), b_j, \mathbf{D}]$ and the upper-bound system $\mathbb{S}_{\mathbb{T}}[1, (h_{[1,j]}, L_{[1,j]}), b_j, \mathbf{D}]$. For a given inventory position y at the order epoch, let the resulting cost function be $G_j^\ell(y, T_j)$ and $G_j^u(y, T_j)$, respectively, and define $s_j^\ell = \arg \min_y G_j^\ell(y, T_j)$ and $s_j^u = \arg \min_y G_j^u(y, T_j)$.

Proposition 5.3 For $j = 1, \dots, N$, (1) $G_j^\ell(y, T_j) + \tau_j \leq G_j(y, \mathbf{T}_j) \leq G_j^u(y, T_j) + \tau_j$. (2) $s_j^\ell \leq s_j^* \leq s_j^u$.

Similar to the (r, q) policy, an effective heuristic for the base-stock levels can be found by weighted averaging s_j^ℓ and s_j^u or h_j and $h_{[1,j]}$.

5.3.3.2 Simple solution for reorder intervals

Shang and Zhou (2010) provide a heuristic for optimal reorder intervals by solving the sum of independent cost-bound functions. They show that when the downstream stages $i < j$ employ the stage j 's reorder interval T_j , the resulting induced penalty cost function is a lower bound to the original one calculated from the system where stage i uses policy $(s_i^*, \mathbf{T}_i)$, $i < j$. Similarly, they show that when the downstream stages $i < j$ use the smallest reorder interval, i.e., $T_i = 1$, the resulting induced penalty cost function is an upper bound to the original one. Using these cost-bound functions, they construct upper and lower bounds for the cost of each stage. They propose solving the sum of these cost-bound functions subject to the integer-ratio constraints to obtain heuristic solutions.

Shang et al. (2009) provide a simpler heuristic that works well in their numerical study. The key idea of the heuristic is to approximate the exact stage cost function by the single-stage cost function $G_j^\ell(y, T_j)$. More specifically, for any given T_j , one can find the best base-stock level $s_j(T_j)$. Then, define $G_j^\ell(T_j) := G_j^\ell(s_j(T_j), T_j)$, and solve the following problem:

$$\begin{aligned} \min_{\mathbf{T}} \sum_{j=1}^N (k_j \lambda_j / T_j + G_j^\ell(T_j)) \\ \text{s.t. } T_{j+1} = n_j T_j, \quad j = 1, \dots, N-1. \end{aligned}$$

Then, a similar two-step algorithm of clustering and minimization suggested for the (r, q) policy in §5.3.2.2 can generate an effective heuristic for the optimal reorder intervals.

5.4 EXTENSIONS

In this section, we discuss a few extensions. Sections 5.4.1 and 5.4.2 consider assembly and distribution systems, respectively; Section 5.4.3 discusses a series system in which each location only has its own local demand information.

5.4.1 Assembly System

An assembly system is a supply-demand network with a single end item 1, such that every other item $j > 1$ has just one successor. In contrast to a series system, an item j may have several predecessors, which compose the set $Pre(j)$. Assume to make a unit of item j requires one unit of each of all of its predecessors $i \in Pre(j)$. Customers demand only the end item and their arrival follows a compound Poisson process.

Following the earlier notation, each item j has its constant assembly lead time L_j . Inventory of item j incurs installation holding cost h'_j and unsatisfied demand of the end item incurs backlogging cost b . There are no fixed ordering costs. Similar to the series system, define the echelon inventory holding cost of item j as $h_j = h'_j - \sum_{i \in Pre(j)} h'_i$.

Rosling (1989) and Zipkin (2000) show that under certain conditions, an assembly system can be transformed into an equivalent series system and so the echelon base-stock type of policy is optimal. Moreover, the policy evaluation and optimization can be done exactly the same as those of the series system discussed previously. Consequently, we can similarly derive the single-stage bounds and heuristics for the optimal base-stock levels. We provide a brief discussion on the transformation and the conditions as follows.

Define

$\underline{L}_j$ = forward echelon lead time for item j , including j 's own lead time. This is the minimal time required to move a unit of item j to the customer. Thus

$$\underline{L}_1 = L_1$$

$$\underline{L}_i = L_i + \underline{L}_j, \quad i \in Pre(j).$$

Renumber the items so that $\underline{L}_j$ is increasing in j . Also set $\underline{L}_0 = 0$. Define

$$L_j'' = \underline{L}_j - \underline{L}_{j-1}.$$

Observe that $L_j'' \leq L_j$ for all j .

Now we define a class of echelon base-stock policies: there are policy parameters s_j while we adjust that policy to improve the balance among the inventories. It is called *balanced echelon base-stock policy* in Zipkin (2000). The best policy of this type is optimal (Rosling (1989)).

Let $IT_j(t)$ be the stock (in item j 's units) shipped in the interval $[t - L_j, t)$. Define

$$IT_j^-(t) = \text{portion of } IT_j(t) \text{ shipped in the interval } [t - L_j, t - L_j''),$$

$$IN_j^+(t) = IT_j^-(t) + IN_j(t).$$

The policy works as follows. Item J uses an ordinary echelon base-stock policy with policy parameter s_J . For item $j > 1$ we adjust the base-stock policy using the variable $IN_{j+1}^+(t)$. Specifically, we decide the quantity to ship of item j so as to bring $IP_j(t)$ as close as possible

to $\min\{s_j, IN_{j+1}^+(t)\}$. That is, if $IP_j(t^-)$ is already more than this quantity before any shipment, we ship nothing. If $IP_j(t^-)$ is less, we ship the difference, provided there is a sufficient inventory of items $i \in \text{Pre}(j)$ to do so. If one of those inventories is too small, we ship as much as possible.

Suppose that $IP_j(t) \leq IN_{j+1}^+(t)$ for all $j < J$, which is known as the *long-run balance* condition. If this holds for all t , then for all $j < J$ and $i \in \text{Pre}(j)$, $IN_i(t) \geq IN_{j+1}^+(t)$. Thus, we never have to worry about the predecessor inventories; they are guaranteed to be sufficient. We need only compare $IP_j(t)$ to $\min\{s_j, IN_{j+1}^+(t)\}$.

We now describe the equivalent series system under the long-run balance condition. It has J stages, the same demand process and echelon holding cost rate h_j , and demand backlogging cost b as the assembly system, but lead times L_j'' . Note that its IN_j is precisely the IN_j^+ of the original assembly system. We can then apply the policy evaluation algorithm of a series system to this series system, using the policy parameters $\mathbf{s}$. The average cost, in the assembly system's terms, is

$$E\left[\sum_{j=1}^J h_j IN_j^+ + (b + h_1') [IN_1^+]^- \right] - \sum_{j=1}^J h_j E[D(L_j - L_j'')]]$$

By applying the optimization algorithm in the previous section to this equivalent series system, the resulting base-stock levels $\mathbf{s}^*$ is optimal for the original assembly system. Hence, we can apply the same idea of deriving lower and upper bounds of the optimal base-stock levels.

5.4.2 Distribution System

The distribution system is reversed to the assembly system, where each stage has one predecessor and multiple successors. Our subsequent discussion will focus on a one-warehouse, multi-retailer system in which the demands are IID Poisson processes between retailers. Let 0 be the warehouse index and $j = 1, \dots, N$ represents N retailers. So h_0 denotes the echelon inventory holding cost at the warehouse while h_j and b_j are the inventory holding and back-order cost at retailer j , respectively.

For any echelon base-stock level $\mathbf{s} = (s_0, s_1, \dots, s_N)$, the average total cost per period can be expressed

$$\begin{aligned} G(\mathbf{s}) &= E[h_0 I_0 + h_0 \sum_{j>0} IT_j + \sum_{j>0} (h_j I_j + b_j B_j)] \\ &= E[h_0 IN_0 + \sum_{j>0} (h_j IN_j + (b + h_0 + h_j) B_j)], \end{aligned} \tag{5.32}$$

where B_j is the backorder at retailer j . We refer the reader to Chapter 8.6 of Zipkin (2000) for the derivation of these variables.

The complication of the optimal replenishment policy is that, in addition to determining order quantities, the decision-maker has to determine the optimal allocation policy when the warehouse does not have enough inventory to satisfy the retailer's orders. Even if the warehouse has sufficient inventory, it may be optimal to reserve some inventory for future orders

from retailers with higher shortage costs, i.e., inventory rationing. Clark and Scarf (1960) show that under the so-called *balanced* assumption, i.e., the on-hand inventory can be instantaneously transferred between retailers as needed, the myopic allocation rule is optimal. The myopic allocation rule specifies that the warehouse only needs to decide the best inventory allocation between retailers for the current period without considering the future. In fact, Clark and Scarf show that under the balanced assumption, the echelon base-stock policy along with the myopic allocation is optimal. (The resulting cost is often used as a lower bound to examine the effectiveness of heuristics in the literature.) Thus, under the assumption, one can apply a bottom-up recursion to obtain the optimal base-stock level similar to that of the series system. The resulting base-stock levels, along with the myopic allocation rule, are an effective heuristic for the distribution system. Although the balanced assumption can simplify the optimization step when solving for the echelon base-stock level for the warehouse, it requires solving an allocation decision at the warehouse.

Gallego et al. (2007) propose aggregation and decomposition methods to further simplify the above process by eliminating the allocation step to obtain the optimal base-stock levels. The aggregation method is to aggregate the retailer sites into a single-stage system by choosing one holding cost rate and one backorder cost rate for all retailers. Gallego et al. (2007) choose the smallest holding and backorder cost rates to construct a lower bound to the optimal cost. When the cost rates are identical, the retailer sites can be consolidated into a single-stage system with the lead-time demand equal to the sum of all lead-time demands of retailers. Consequently, the distribution system becomes a two-stage series system.

The decomposition method is to create dedicated inventory at the warehouse that replenishes each retailer's inventory. Thus, there is no inventory pooling at the warehouse, and the system can be decomposed into N two-stage systems with the original problem parameters. After calculating stage 2's base-stock level of each series system, the warehouse's base-stock level is simply the sum of these stage 2's base-stock levels.

In short, with the aggregation and decomposition ideas, one can transform a distribution system into either a two-stage series system or N two-stage series systems. As a result, the single-stage approximations proposed in Section 5.2.2 can be applied accordingly. In fact, Rong et al. (2017) use the same decomposition idea and apply the single-stage approximation to generate the base-stock levels in their decomposition-aggregation heuristic. They also propose the recursive optimization heuristic, which applies a bottom-up approach that sequentially solves the optimal base-stock level for each stage (without considering the allocation decision).

5.4.3 Local Information

So far, we assume that the demand information is centralized. However, in practice, a supply chain may have different levels of information integration. We now consider a series system where each location only knows its local demand information. The local demand for a stage is the order placed by its immediate downstream stage. We provide a summary in the literature for some known results under local information. Chen (2000) shows that there is a one-to-one correspondence between the echelon and local base-stock policies. Axsäter and Rosling (1993) show that the local (r, q) policy is a special case of the echelon (r, q) policy (and that the local policy is therefore suboptimal). Chen (1998) provides an algorithm to search for the optimal local reorder points when batch sizes are fixed. Shang et al. (2009) provide an approach

to obtain the optimal batch sizes for the local (r, q) policy. It is conceivable that the simple solution bounds can be developed for the local (r, q) policies.

Shang et al. (2010) characterize the dynamics of key inventory variables under the local (s, T) policy. These dynamics lead to a simple, bottom-up recursion that can evaluate a given local (s, T) policy. More specifically, by converting the local inventory variables into echelon ones, the local policy can be evaluated as if it were an echelon (s, T) policy with modified system lead times. This evaluation procedure can be used further to find the optimal local base-stock levels for given reorder intervals. Unlike the (r, q) policy, the local (s, T) policy is not a special case of the echelon (s, T) policy. Nevertheless, we show that the optimal echelon (s, T) policy always dominates the optimal local one.¹

5.5 SUMMARY AND FUTURE RESEARCH

This chapter summarizes recent developments in single-stage-based heuristics for multi-echelon inventory models. We show that a decision-maker can solve a series of single-stage problems to obtain effective solutions. These single-stage heuristics not only simplify the computation, but also help implementation.

We lay out a few important future research directions. First, the excess demand in the reviewed inventory models is assumed to be fully backlogged. It would be useful to extend the idea to the lost-sales inventory models. The lost-sales models share similar technical challenges as the dual-sourcing models with arbitrary lead times. The connection between the dual-sourcing model and the lost-sales model for a single-stage system has been established (Sheopuri et al. (2010)). However, our understanding of multi-echelon inventory systems with lost sales (or expediting with arbitrary lead times) is still very limited. We hope its optimality analysis and a similar single-stage-based heuristics can be developed for the lost-sales model.

Second, the current single-stage-based heuristics have been successfully applied to serial systems (as pointed out in Section 5.4.1, under certain conditions, an assembly system can be transformed into an equivalent series system) and distribution systems with (s, T) policies. However, a supply chain is an intricate network with complicated system structures. For example, some parts of a supply chain may be serial types of networks and some distribution types. In the deterministic demand network model, the power-of-two solution has been proven to guarantee a worst-case cost bound. It will be important, though quite challenging, to develop a simple heuristic for general supply-chain networks with stochastic demand.

Third, multi-echelon models lay the foundation for studying supply-chain problems. However, there are relevant business activities related to operations that are not specifically incorporated into the models. For example, a supply chain includes material, information, and financial flows, and these flows are often entangled with each other. The multi-echelon literature has been focusing on the first two flows. Nevertheless, the 2008 financial crisis demonstrates a need for jointly studying these three flows together. It is conceivable that such a joint problem will be extremely difficult to analyze. Thus, one would hope that the single-stage-based heuristics can help simplify the analysis of these joint problems.

Joint inventory and dynamic pricing models in the literature are for single-stage systems, e.g., Federgruen and Heching (1999). The impact of pricing policy on the inventory decision has been well investigated for the single-stage system. It will be interesting to study the impact of the pricing policy on inventory decisions in the supply chain. The idea of constructing these

single-stage heuristics reviewed in this chapter may be useful for constructing simple joint inventory and pricing policies for a multi-echelon system.

Finally, there is a growing trend in developing and analyzing learning algorithms, both online and offline, for inventory systems with unknown demand distribution. So far most of the work is focused on single-stage systems (e.g., Levi et al. (2007), Huh et al. (2009)). Recently, Zhang et al. (2020) study sample-based approximation algorithms for series systems with unknown demand distribution and derive sample size upper bounds to guarantee the performance of sample-based optimal solutions. An interesting direction is on deriving the analytical performance of the single-stage approximations and their possible applications in systems where demand distribution is unknown.

NOTE

1. An echelon policy may not always dominate a local one. With this result, Shang and Zhou (2010) provide a method to search for the optimal reorder intervals for the local (s, T) policy. We conjecture that the simple solutions can be developed for the local (s, T) policies as well.

REFERENCES

- Abhyankar, H. S., & Graves, S. C. (2001). Creating an inventory hedge for Markov-modulated poisson demand: An application and model. *Manufacturing & Service Operations Management*, 3(4), 306–320.
- Axsäter, S. (1993). Continuous review policies for multi-level inventory systems with stochastic demand. In *Logistics of Production and Inventory*, Edited by S.C Graves, A.H.G. Rinnooy Kan and P.H. Zipkin, 4, 175–197.
- Axsäter, S., & Rosling, K. (1993). Notes: Installation vs. echelon stock policies for multilevel inventory control. *Management Science*, 39(10), 1274–1280. <https://doi.org/10.1287/mnsc.39.10.1274>.
- Chao, X., & Zhou, S. X. (2007). Probabilistic solution and bounds for serial inventory systems with discounted and average costs. *Naval Research Logistics*, 54(6), 623–631. <https://doi.org/10.1002/nav.20234>.
- Chao, X., & Zhou, S. X. (2009). Optimal policy for a multiechelon inventory system with batch ordering and fixed replenishment intervals. *Operations Research*, 57(2), 377–390. <https://doi.org/10.1287/opre.1080.0636>.
- Chen, F. (1998). Echelon reorder points, installation reorder points, and the value of centralized demand information. *Management Science*, 44(12-part-2), S221–S234. <https://doi.org/10.1287/mnsc.44.12.s221>.
- Chen, F. (2000). Optimal policies for multi-echelon inventory problems with batch ordering. *Operations Research*, 48(3), 376–389. <https://doi.org/10.1287/opre.48.3.376.12427>.
- Chen, F., & Song, J.-S. (2001). Optimal policies for multiechelon inventory problems with markov-modulated demand. *Operations Research*, 49(2), 226–234.
- Chen, F., & Zheng, Y.-S. (1994). Lower bounds for multi-echelon stochastic inventory systems. *Management Science*, 40(11), 1426–1443. <https://doi.org/10.1287/mnsc.40.11.1426>.
- Chen, F., & Zheng, Y.-S. (1998). Near-optimal echelon-stock (r, nQ) policies in multistage serial systems. *Operations Research*, 46(4), 592–602. <https://doi.org/10.1287/opre.46.4.592>.
- Chen, L., Song, J.-S., & Zhang, Y. (2017). Serial inventory systems with markov-modulated demand: Derivative bounds, asymptotic analysis, and insights. *Operations Research*, 65(5), 1231–1249.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490. <https://doi.org/10.1287/mnsc.6.4.475>.
- Dong, L., & Lee, H. L. (2003). Optimal policies and approximations for a serial multiechelon inventory system with time-correlated demand. *Operations Research*, 51(6), 969–980. <https://doi.org/10.1287/opre.51.6.969.24920>.

- Ettl, M., Feigin, G. E., Lin, G. Y., & Yao, D. D. (2000). A supply network model with base-stock control and service requirements. *Operations Research*, 48(2), 216–232. <https://doi.org/10.1287/opre.48.2.216.12376>.
- Federgruen, A., & Zheng, Y.-S. (1993). Optimal power-of-two replenishment strategies in capacitated general production/distribution networks. *Management Science*, 39(6), 710–727. <https://doi.org/10.1287/mnsc.39.6.710>.
- Federgruen, A. (1993). Centralized planning models for multi-echelon inventory systems under uncertainty. in *Logistics of Production and Inventory*, Edited by S.C Graves, A.H.G. Rinnooy Kan and P.H. Zipkin, 4, 133–173.
- Federgruen, A., & Heching, A. (1999). Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3), 454–475. <https://doi.org/10.1287/opre.47.3.454>.
- Federgruen, A., & Zheng, Y.-S. (1992). An efficient algorithm for computing an optimal (r, q) policy in continuous review stochastic inventory systems. *Operations Research*, 40, 808–813.
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836. <https://doi.org/10.1287/opre.32.4.818>.
- Gallego, G., Özalp Özer, & Zipkin, P. (2007). Bounds, heuristics, and approximations for distribution systems. *Operations Research*, 55(3), 503–517. <https://doi.org/10.1287/opre.1060.0373>.
- Graves, S. C., & Willems, S. P. (2000). Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management*, 2(1), 68–83. <https://doi.org/10.1287/msom.2.1.68.23267>.
- Graves, S. C., & Willems, S. P. (2003). Supply chain design: Safety stock placement and supply chain configuration. *Supply Chain Management: Design, Coordination and Operation*, 95–132. [https://doi.org/10.1016/S0927-0507\(03\)11003-1](https://doi.org/10.1016/S0927-0507(03)11003-1).
- Huh, W. T., Janakiraman, G., Muckstadt, J. A., & Rusmevichientong, P. (2009). An adaptive algorithm for finding the optimal base-stock policy in lost sales inventory systems with censored demand. *Mathematics of Operations Research*, 34(2), 397–416. <https://doi.org/10.1287/moor.1080.0367>.
- Lawson, D. G., & Porteus, E. L. (2000). Multistage inventory management with expediting. *Operations Research*, 48(6), 878–893. <https://doi.org/10.1287/opre.48.6.878.12399>.
- Lee, H. L., & Billington, C. (1995). The evolution of supply-chain-management models and practice at hewlett-packard. *Interfaces*, 25(5), 42–63. <https://doi.org/10.1287/inte.25.5.42>.
- Levi, R., Roundy, R. O., & Shmoys, D. B. (2007). Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 821–839. <https://doi.org/10.1287/moor.1070.0272>.
- Liu, F., & Song, J.-S. (2012). Good and bad news about the (s, t) policy. *Manufacturing & Service Operations Management*, 14, 42–49. <https://doi.org/10.1287/msom.1110.0353>.
- Maxwell, W. L., & Muckstadt, J. A. (1985). Establishing consistent and realistic reorder intervals in production-distribution systems. *Operations Research*, 33(6), 1316–1341. <https://doi.org/10.1287/opre.33.6.1316>.
- Muharremoglu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, 56(5), 1089–1103.
- Rong, Y., Atan, Z., & Snyder, L. (2017). Heuristics for base-stock levels in multi-echelon distribution networks. *Production and Operations Management*, 26(9), 1760–1777.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579. <https://doi.org/10.1287/opre.37.4.565>.
- Roundy, R. (1985). 98%-effective integer-ratio lot-sizing for one-warehouse multi-retailer systems. *Management Science*, 31(11), 1416–1430. <https://doi.org/10.1287/mnsc.31.11.1416>.
- Shang, K. H. (2008). Note: A simple heuristic for serial inventory systems with fixed order costs. *Operations Research*, 56(4), 1039–1043. <https://doi.org/10.1287/opre.1080.0547>.
- Shang, K. H. (2012). Single-stage approximations for optimal policies in serial inventory systems with nonstationary demand. *Manufacturing & Service Operations Management*, 14(3), 414–422. <https://doi.org/10.1287/msom.1110.0373>.
- Shang, K. H., & Song, J.-S. (2003). Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science*, 49(5), 618–638. <https://doi.org/10.1287/mnsc.49.5.618.15147>.
- Shang, K. H., & Song, J.-S. (2006). A closed-form approximation for serial inventory systems and its application to system design. *Manufacturing & Service Operations Management*, 8(4), 394–406. <https://doi.org/10.1287/msom.1060.0114>.

- Shang, K. H., & Song, J.-S. (2007). Serial supply chains with economies of scale: Bounds and approximations. *Operations Research*, 55(5), 843–853. <https://doi.org/10.1287/opre.1070.0406>.
- Shang, K. H., Song, J.-S., & Zipkin, P. H. (2009). Coordination mechanisms in decentralized serial inventory systems with batch ordering. *Management Science*, 55(4), 685–695. <https://doi.org/10.1287/mnsc.1080.0981>.
- Shang, K. H., & Zhou, S. X. (2009). A simple heuristic for echelon policies in serial supply chains. *Operations Research Letters*, 37(6), 433–437. <https://doi.org/10.1016/j.orl.2009.08.003>.
- Shang, K. H., & Zhou, S. X. (2010). Optimal and heuristic echelon (r, nQ, t) policies in serial inventory systems with fixed costs. *Operations Research*, 58(2), 414–427. <https://doi.org/10.1287/opre.1090.0734>.
- Shang, K. H., Zhou, S. X., & van Houtum, G.-J. (2010). Improving supply chain performance: Real-time demand information and flexible deliveries. *Manufacturing & Service Operations Management*, 12(3), 430–448. <https://doi.org/10.1287/msom.1090.0277>.
- Sheopuri, A., Janakiraman, G., & Seshadri, S. (2010). New policies for the stochastic inventory control problem with two supply sources. *Operations Research*, 58(3), 734–745. <https://doi.org/10.1287/opre.1090.0799>.
- Song, J.-S., & Zipkin, P. (1993). Inventory control in a fluctuating demand environment. *Operations Research*, 41(2), 351–370.
- Song, J.-S., & Zipkin, P. H. (1996). Managing inventory with the prospect of obsolescence. *Operations Research*, 44(1), 215–222.
- Whittmore, A. S., & Saunders, S. C. (1977). Optimal inventory under stochastic demand with two supply options. *SIAM Journal on Applied Mathematics*, 32(2), 293–305. <https://doi.org/10.1137/0132023>.
- Zhang, K., Gao, X., Wang, Z., & Zhou, S. X. (2020). Sampling-based approximation for serial multi-echelon inventory system. Working Paper, Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China.
- Zheng, Y.-S. (1992). On properties of stochastic inventory systems. *Management Science*, 38(1), 87–103. <https://doi.org/10.1287/mnsc.38.1.87>.
- Zhou, S. X., & Chao, X. (2009). Newsvendor bounds and heuristics for serial supply chains with regular and expedited shipping. *Naval Research Logistics*. <https://doi.org/10.1002/nav.20388>.
- Zipkin, P. (2000). *Foundations of inventory management*. McGraw Hill.

6. Single-unit analysis

Alp Muharremoglu, Xin Geng, and Nan Yang

6.1 INTRODUCTION

In many inventory systems, units of goods go through multiple stages. These may be different stages in a supply chain or different operations in a production facility. The control of such multi-echelon inventory systems has undoubtedly been a fascinating research topic and has attracted numerous scholars. Ever since the seminal work by Clark and Scarf (1960), the traditional approach in multi-echelon inventory theory is to decompose the original system into a series of one-stage subsystems in a dynamic program recursion; this approach has been very fruitful (see, e.g., Federgruen and Zipkin, 1984; Chen and Zheng, 1994; Chen and Song, 2001).

We, on the other hand, view the inventory management problem from a very different perspective, which results in a different way of decomposing the original problem. This approach, called *single-unit analysis*, is based on the observation that, in a multi-echelon inventory system with backlog, one unit of final product (including all its components) is paired with one unit of customer demand. Applying this analysis technique, many important multi-echelon inventory problems could be decomposed into much simpler and manageable subproblems, each consisting of one unit–customer pair. While the arrival time of the customer is governed by the demand distribution, the goal is therefore to move the unit through the system so that the trade-off between holding cost and backlog cost is optimized. Since only one unit and one customer are present in this subproblem, it is relatively easy to structurally characterize and efficiently compute its optimal policy. Then, the original problem can also be solved accordingly.

The origin of single-unit analysis dates back to the 1990s. It was pioneered by Axsäter (1990), by making the observation that any particular unit ordered is used to fill a particular demand, and then he matched this unit with “its demand” to evaluate their associated expected cost. Using this approach, he developed an efficient method to evaluate the cost of a given base-stock policy for a two-echelon distribution system in continuous time. Two of Axsäter’s subsequent works continue to apply the approach to similar settings: Axsäter (1993a) extends this technique to systems with batch ordering and Axsäter (1993b) studies a distribution system with a particular allocation rule in discrete time. Zipkin (1991) also applied this technique to evaluating base-stock policies in a multi-echelon inventory system with compound-Poisson demands, which is further extended by Song and Zipkin (1992) and Song and Zipkin (1996a) to Markov-modulated Poisson demands. In addition, some other early works that have used the single-unit approach, e.g., Katircioglu and Atkin (1998) and Achy-Brou (2001), mainly focus on the derivation of the optimal policy structure.

To demonstrate the broad application of single-unit analysis and show its effectiveness in solving optimal inventory control problems, we review a stream of papers that use the single-unit method in various different settings. These works concentrate on serial systems and assembly systems. The two types of systems are in general compatible with the single-unit analysis.

Muharremoglu and Tsitsiklis (2008) considered a serial system with Markov-modulated demand and lead times, and conducted a formal single-unit analysis to prove the optimality of state-dependent echelon base-stock policies. Several extensions to this problem (still for serial systems) have been investigated as well, including a problem with expediting options (Muharremoglu, 2002; Muharremoglu and Tsitsiklis, 2003), batch ordering (Muharremoglu, 2002), and a particular class of capacitated serial systems (Janakiraman and Muckstadt, 2009). Applying the decomposition technique to an assembly system, Chen and Muharremoglu (2014) extend Rosling's (1989) model to characterize the optimal policy for a more general setting and provide a milder optimality condition.

Even if the original problem cannot be decomposed, single-unit analysis can still be applied to provide useful insights, develop approximation policies, and suggest novel ideas about how to solve the problem. Two papers are notable in this regard. In a serial system with general (exogenous) stochastic lead times, Muharremoglu and Yang (2010) provide a method to evaluate base-stock policies by linking the original problem to a closely related single-unit single-customer problem, even though the decomposition technique does not go through in this case. In a single-product assemble-to-order system with stochastic lead times, Muharremoglu et al. (2021) take the single-unit approach to propose a polynomial time algorithm to evaluate the performance of a component base-stock policy.

Lastly, single-unit analysis has been applied to solve some context-specific inventory problems. For example, Martínez-de-Albéniz and Lago (2010) and Berling and Martínez-de-Albéniz (2011) both study single-echelon systems with stochastic demand, and, in particular, consider stochastic selling prices and procurement costs, respectively. Moreover, Berling and Martínez-de-Albéniz (2016) consider a continuous-stage serial supply chain and focus on transportation speed optimization. All these papers take advantage of the single-unit decomposition approach in solving inventory control problems.

In Section 6.2, we introduce some preliminary concepts through a toy example. Section 6.3 presents the main results from papers that study serial systems, whereas Section 6.4 focuses on the application of single-unit analysis in assembly systems. Section 6.5 discusses several related papers that consider context-specific inventory problems. Finally, Section 6.6 summarizes and concludes this chapter.

6.2 A TOY EXAMPLE ON A SERIAL SYSTEM

Before we review the applications of single-unit analysis in different models, we first use this approach to establish the form of an optimal policy for a simple multi-echelon inventory control problem. Through showing this toy example, we lay out some preliminary concepts and definitions, which not only reveal the fundamental ideas of this approach, but also facilitate the understanding of the works to be reviewed. All definitions in this section are adopted from Muharremoglu and Tsitsiklis (2008).

6.2.1 Problem Description

Consider a discrete-time serial system with M stages, indexed by $1, \dots, M$ from downstream to upstream. Stage j ($j = 1, \dots, M-1$) replenishes its inventory by ordering from stage $j+1$; stage M receives stocks from an outside supplier (with unlimited stock), which for exposition

ease is referred to as stage $M + 1$. There is a lead time L_j for orders flowing from stage $j + 1$ to stage j , and in this simple example we assume that $L_j = 1$ for all $j = 1, \dots, M$. Section 6.2.4 will discuss the extension to general lead times, which is fairly straightforward. In this example, we look at a finite planning horizon indexed as $t = 1, 2, \dots, T$. In period t , customer demand D_t is a random variable with known distribution and occurs only at stage 1. Any demand that is not immediately satisfied is backlogged. Demands are assumed to be i.i.d. across periods. Finally, ordering, holding, and backorder costs are all linear, with rates $c_j \geq 0$, $h_j > 0$ for all stages $j = 1, \dots, M$ (let $c_{M+1} = h_{M+1} = 0$), and $b > 0$. Pipeline inventory is charged holding cost corresponding to the destination echelon. The objective is to minimize the expected total cost.

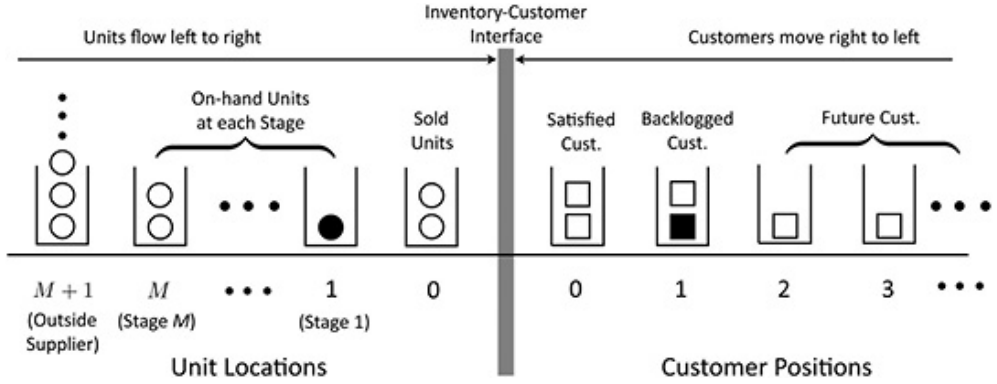
The fundamental idea of single-unit analysis is to view an inventory system as a collection of unit–customer pairs, where each unit is “destined” to be given to a unique customer (before any control is applied). To formulate the problem based on this idea, some basic concepts and definitions are in order.

First, we define the location of a unit. In this example, there are $M + 2$ locations for units to flow through. Every stage where a unit can be stored, including the outside supplier, constitutes a location with the same index as its stage index. In addition, any sold (consumed) unit has location 0, which is clearly an artificial location. We create this location so that all units are kept in the system for ease of numbering them. By our model setup, there are countably infinitely many units at location $M + 1$, but only finitely many units at any other location. Hence, we may index all units in sequence based on their locations in the system at time 1, breaking ties arbitrarily. Let z_t^i be the location of the i^{th} unit at time t ($i = 1, \dots, \infty; t = 1, \dots, T$).

Second, we define the position of a customer. In each period, a quantity of demand arrives. We treat each unit of demand as a distinguishable object, and refer to each unit of demand as “a customer”. Then, we can index all customers in sequence according to their arrival times (those arriving together are numbered in any order). Note that, even though we do not know the exact arrival times of future customers, we can still talk about a “next customer”, a “second to next customer”, etc. Customers who have been satisfied have position 0 (these are the matched demand by units at location 0). Arrived but unsatisfied customers have position 1. All future customers are sequentially assigned positions 2, 3, ..., in the same order as they are indexed. Let y_t^i be the position of the i^{th} customer at time t ($i = 1, \dots, \infty; t = 1, \dots, T$).

We remark that the key idea behind the concepts of unit location and customer position is to establish an indexing convention over a countably infinite number of unit–customer pairs. Although different papers have used slightly different definitions, they all achieve this goal and are equivalent through relabeling. Figure 6.1 provides a visual illustration of the above definitions.

The sequence of events during period t is as follows. (1) The units that were ordered in the last period (recall that lead time is assumed to be one) arrive at each stage $j = 1, 2, \dots, M$, to the extent that they are available at the upstream stage. (2) Order is placed from stage j to stage $j + 1$, $j = 1, 2, \dots, M$. (3) Demand realizes and the customers’ positions are updated according to the demand D_t . That is, customers $2, 3, \dots, D_t + 1$ arrive to position 1; all other customers move forward to position 1 with D_t steps. (4) Units on hand at stage 1 (i.e., location 1) are matched with arrived customers (those with position 1) until at least one side ends up empty. (5) Ordering, holding, and backorder costs are charged. Note that Figure 6.1 depicts the moment between event (3) and event (4) during period t .



Note: The time point depicted by the figure is at the end of event (3) and the beginning of event (4) during period t . Specifically, the order placed in $t-1$ has arrived at each stage, the new order for period t has been submitted, and the current demand of two customers has arrived to position 1, waiting to be paired with units at location 1. At this moment, the unit and customer in black are to be paired and are going to location/position 0 at the end of event (4), and there will be a backlog of 1 unit.

Figure 6.1 *Illustration of the unit locations and customer positions in the example of the serial system*

6.2.2 Policy Classification and Problem Decomposition

Based on the above problem formulation, we introduce a few classes of policies and formally describe the problem decomposition, which could greatly facilitate the characterization of the optimal policy. First, we define a control policy for the problem. A control policy $\pi = (\mathbf{u}_1, \dots, \mathbf{u}_T)$ dictates the hold/release decisions for all units in every period, given the units' locations and customers' positions. Specifically, at time t , given $(z_t^i, y_t^i)_{i=1}^\infty$, the control $\mathbf{u}_t = \{u_t^i\}_{i=1}^\infty$ where $u_t^i = 0$ means "hold unit i " and $u_t^i = 1$ means "release unit i ". Obviously, for units at location $j = M+1, M, \dots, 2$, the hold/release control decision corresponds to the ordering decision at stage $j-1$; i.e., if stage j places an order of quantity q , then the controls of q units held at location $j+1$ would be "release". Finally, the control for the units at location 1 is governed by the demand process as previously described, and the control for the units at location 0 is irrelevant.

Note that in our model setup, the units released at any location can be any units at that location. Hence, we will focus on a class of policies that always release the units with the lowest indexes and thus units can never overtake each other:

Monotonic policies: State $(z_t^i, y_t^i)_{i=1}^\infty$ at time t is called monotonic if and only if $z_t^i \leq z_t^j$ for units i and j where $i < j$. A policy is monotonic if it guarantees that a monotonic state at time t always results in a monotonic state at time $t+1$.

For simplicity, we assume that the system always starts from a monotonic state. In fact, we can easily see that the class of monotonic policies is optimal, because, for any non-monotonic policy, one could copy that policy in the number of released units but choose those with the lowest indexes, resulting in a monotonic policy and the same cost. Moreover, since customers arrive in the order of their indexes, we can further consider a class of policies named committed policies:

Committed policies: A policy is committed if customer i can be served only by unit i , and unit i can be received only by customer i . This means that, when a committed policy is used, the paired units and customers must be those with the lowest indexes at location 1 and position 1, respectively.

A direct corollary from the above definitions is that every monotonic policy is also a committed policy; therefore, the class of committed policies is also optimal. Furthermore, such commitment requirement for pairing units with customers makes the problem decomposable to independent subsystems, each consisting of one unit–customer pair with index i . Next, we consider the collection of subsystems and show that the class of decoupled policies defined below is optimal for the overall system:

Decoupled policies: A policy is decoupled if it can be represented in terms of a series of mappings, each pertaining to the control problem of a unit–customer pair. For example, the control policy u_t^i , which should be a function of the entire state $\{(z_t^i, y_t^i)\}_{i=1}^\infty$, is a decoupled policy if it can simply be written as:

$$u_t^i = \hat{u}_t^i(z_t^i, y_t^i), \quad \forall i, t.$$

To see why the class of decoupled policies is optimal, we consider the following facts. First, due to the linear cost structure, the total cost incurred by the overall system is the sum of the costs incurred by every subsystem. Second, when combining any control policy for each subsystem together, we obtain a feasible policy for the overall system; hence, the optimal expected total cost for the overall system serves as a lower bound for the sum of optimal expected total cost for all subsystems. Third, since every monotonic and committed policy for the overall system corresponds to a committed policy for each subsystem, the sum of the optimal expected total cost for all subsystems must be a lower bound for the optimal expected total cost of the overall system. This is because the monotonic and committed policy is optimal. Therefore, the previous two facts together indicate that the two optimal expected total costs should be the same. Lastly, each unit–customer subsystem can be operated independently, meaning that the decision of whether or not to release a unit is isolated and only depends on state variables of the corresponding unit–customer pair. Thus, since the subsystems are operationally independent, they can be managed identically, independently, and optimally to obtain the optimal policy for the overall system. It is worth noting that, although operationally independent, the subsystems are connected stochastically through the demands.

6.2.3 Optimal Policies

Since the overall system is now decomposed into much smaller and simpler subsystems, we may first characterize the optimal control structure for a subsystem. Define $U_t^*(y, z) \subseteq \{0, 1\}$ as the set of optimal decisions for the single-unit single-customer subsystem at time t when the unit is at location $z = 2, 3, \dots, M+1$ (these are the locations where a control decision is needed) and the customer is at position y . Then, the following lemma, whose proof can be found in Muharremoglu and Tsitsiklis (2008), implies that the subsystem is optimally controlled by a threshold-type policy that releases a unit if the corresponding customer is near enough.

Lemma 6.1 *If $U_t^*(y', z) = \{1\}$, then $1 \in U_t^*(y, z)$ for every $y < y'$.*

Lemma 6.1 allows us to develop the threshold position $y_t^*(z) := \max_y \{1 \in U_t^*(y, z)\}$ for every location z , so that the control $\hat{u}_t = 1$ for units at location z if and only if $y_t \leq y_t^*(z)$ is optimal for the subsystem.

Moving back to the overall problem, we can simply apply the above decoupled policy to every subsystem to obtain the optimal overall policy. Not surprisingly, the resulting policy is the state-dependent echelon base-stock policy. Specifically, the optimal policy for the problem in our example is for every stage $j = 1, 2, \dots, M$ to order enough inventory to raise the inventory position to $S_j^* := y_i^*(j+1) - 1$. Note that it is very intuitive to show this statement through the single-unit argument. The above optimal decoupled policy $\hat{u}_t$ releases a certain amount of units at stage $j+1$ to serve the customers who have already arrived (with positions 0 and 1) and the next $y_i^*(j+1) - 1$ future customers (with positions $2, 3, \dots, y_i^*(j+1)$). Equivalently, it tries to raise the inventory position at stage j to $y_i^*(j+1) - 1$. We remark that, compared to using the traditional dynamic programming approach, the above single-unit analysis can be simpler and more intuitive.

6.2.4 Extension to General Lead Times

Recall that we have assumed that the lead time at every stage is deterministic and equals one. This is the reason why in our model the stage indexes are the same as the non-zero location indexes. The above single-unit approach can be easily extended to the case with arbitrary deterministic lead times. While the concept of customer position is defined in exactly the same way as aforementioned, the definition of unit location is slightly modified. In particular, in between consecutive stages j and $j+1$, we insert $L_j - 1$ locations where L_j is the order lead time from stage $j+1$ to stage j ($j = 1, \dots, M$). Then, locations are labeled in the same manner, starting from location 0 for the consumed units, all the way to location $\sum_j L_j + 1$ where the outside supplier is. Apparently, the location index for each stage j will not be j in general, and the control decision for units at the locations denoting the in-transit periods is always “release”. However, our previous analysis is unaffected by this change.

Moreover, the above single-unit analysis can be further extended to admit stochastic lead times with bounded support subject to the restriction that orders do not cross; i.e., the order placed earlier will arrive earlier. Keeping everything else the same, we insert $l_j - 1$ locations between stages $j+1$ and j , where l_j is the upper bound of the random lead time L_j . Hence, the movement of units along the pipeline locations is dictated by the stochastic lead time processes. Note that, for general stochastic lead times that permit order crossing, single-unit analysis can still be applied to derive keen insights; but the problem may not be decomposable any more (see Section 6.3.2).

6.3 APPLICATIONS IN MANAGING SERIAL SYSTEMS

This section focuses on the application of single-unit analysis in inventory control problems for serial systems. Consider an M -stage serial system as described in Section 6.2.1, except that the order lead times between stages are general. Besides, the objective is to minimize the expected total cost in either finite-horizon or infinite-horizon settings. The papers we review in this section fall into two categories. The first set of papers assumes deterministic lead times, but includes additional practical considerations. The second assumes stochastic lead times, which can make the inventory control problem considerably more difficult.

6.3.1 Serial Systems with Deterministic Lead Times and Additional Considerations

In this subsection, we review some papers that apply single-unit analysis to inventory control problems while taking into account additional practical considerations. Since the primary

objective of these papers is to solve the problems in some special serial systems, they focus on deterministic lead times.

6.3.1.1 Costly expediting options

We start with a serial system in which orders could be expedited at some cost. In this setting, demand in each period is assumed to be i.i.d., and the order cost c_j is generalized so that units can be sent from any stage j ($j = 2, \dots, M$) to any downstream stage $i < j$, by incurring a cost c_{ji} . The lead time for any order, regardless of its origin and destination, is assumed to be one period. This is the model studied by Muharremoglu and Tsitsiklis (2003). We review how they use single-unit analysis to decompose the problem and characterize the optimal policy.

The main idea is similar to Muharremoglu and Tsitsiklis (2008), i.e., a unit and a customer are paired, and the overall problem can be decomposed into a series of single-unit single-customer subproblems. Moreover, different unit–customer pairs can be controlled identically and independently of each other. After the structure of the subproblem is examined, the link between the optimal subproblem policies and the overall problem is established, so the optimal policy for the overall problem can be characterized.

Consider a subproblem with a single unit–customer pair. The state of the subproblem is (z, y) , i.e., the unit location z and customer position y . Since the lead time is assumed to be one, the location z would correspond to the actual stages. The control variable u in this setting is not simply 0 and 1, but contains all destination stages including the current stage; i.e., $u(z) \in \{z, z-1, \dots, 1\}$ where $u(z) = z$ means to hold the unit. In the presence of costly expediting options, the decomposability result depends on an additional assumption of the expediting costs. Specifically, the costs are assumed to be supermodular: For any stages $i < j \leq k < l$, $c_{li} + c_{kj} \geq c_{lj} + c_{ki}$. This assumption precludes order crossing for optimization purposes. Indeed, any decision that results in order overtaking can be replaced with another that has at most the same cost and preserves the order sequence. Therefore, under this assumption, one can just focus on monotonic and decoupled policies for the control problem in the single-unit single-customer subsystem.

The resulting policies are not as simple as echelon base-stock policies, but can be summarized by $(M+1)M/2$ threshold values for each time period. These values are called extended base-stock levels. Particularly, in each time period t , for any $j > i \geq 1$, there exists an S_t^{ji} that depends on both the stage from which units are to be shipped as well as the stage that is the intended destination. Given these threshold values S_t^{ji} , the control decision u_t as a function of (z_t, y_t) is determined in the following manner. The control variable u_t is a non-decreasing step function in y_t and thresholds S_t^{ji} are exactly the points where the function increases. Between these step points, the function is constant. Hence, the subproblem policy partitions the customer position axis into segments that determine the stage the unit will be shipped to if the corresponding customer's position falls into that segment. Specifically, suppose at time t the customer is at position y_t and the unit is at the location $z_t = j > 1$. Then the decision $u_t(z_t)$ is made so that $u_t(z_t) = j$ if $y_t > S_t^{j,j-1}$ and $u_t(z_t) = i \in \{1, \dots, j-1\}$ if $S_t^{j,i-1} < y_t \leq S_t^{ji}$.

Since the overall problem is optimally solved by decoupled policies, the above threshold policy for the single-unit single-customer subproblem can be translated into an optimal control for the overall problem. Particularly, given the thresholds S_t^{ji} ($j > i \geq 1$), the quantity to be shipped from stage j to stage i at time t is given by

$$\left(\min \{I_t^j, S_t^{ji}\} - \max \{I_t^{j-1}, S_t^{j,i-1}\} \right)^+,$$

where I_t^j represents the echelon inventory position at stage j at time t . This is called *extended echelon base-stock policies* and is proven to be optimal by Muharremoglu and Tsitsiklis (2003).

Theorem 6.1 (Muharremoglu and Tsitsiklis, 2003) *The set of extended echelon base-stock policies is optimal for the finite-horizon model and the infinite-horizon model with either discounted cost or average cost criterion.*

The echelon base-stock policy is simply a special case where all thresholds S_t^{ji} equal $-\infty$ except for $i = j - 1$, so units can only move to the immediate downstream stage.

6.3.1.2 Batch ordering constraint

Next, we turn to look at a serial system with a batch ordering constraint. In this setting, the inventory at stage j is replenished by placing an order for units stored at stage $j + 1$; but the ordering quantity is restricted to be a multiple of some batch size Q_j . Along the stages in the serial system, the batch sizes are assumed to be nested, i.e., $Q_{j+1} = n_j Q_j$ for some integer n_j ($j = 1, \dots, M - 1$). Lastly, demands are assumed to be i.i.d. random variables, whereas order lead times are all one period. This model is studied by Muharremoglu (2002, Chapter 5).

The analysis is, again, based on the decomposition idea. In this setting, when formulating the overall problem using unit locations z_t^i and customer positions y_t^i , the batch ordering requirement should be met by the control decision u_t^i on each unit. That is, $|\{i \mid z_t^i = j + 1, u_t^i = 1\}|$ should be a multiple of Q_j . Using the single-unit formulation, it is easy to see that monotonic and committed policies are optimal. Decoupled policies, however, are clearly not optimal, because coupling among units naturally arises due to the batch size constraint. Moreover, each subproblem involves a single batch of units paired with the customers who are destined to receive those units. Note that these customers are consecutive in arriving sequence, and therefore the position of the last customer contains enough information about all others' positions. Unlike the decomposition in Muharremoglu and Tsitsiklis (2008), the subproblems in this setting are not all identical, but depend on the batch sizes at each stage.

Define a k -unit subproblem as one in which there are k units, including the ones at the outside supplier, and k customers, following the same dynamics as the overall problem. Assume that the system starts with the inventory at stage j being a multiple of Q_{j-1} for all $j = 2, \dots, M$. Due to the nested order batch sizes constraint, the above property will be preserved throughout the horizon. Let $\beta^j Q_{j-1}$ be the number of units at stage j at the beginning of the horizon, assuming $Q_0 = 1$. Clearly, β^j 's are all non-negative integers. Then, the overall problem can be decomposed into a series of smaller subproblems, β^j of which are Q_{j-1} -unit subproblems, for $j = 1, \dots, M$, and the rest are Q_M -unit subproblems.

Now that the overall problem can be decomposed into smaller subproblems, each subproblem can be analyzed to obtain the optimal control of the specific single batch of units. In fact, an optimal control can be determined by considering only the location of the batch and the position of the paired customers. Specifically, for every stage $j \geq 1$ and time t , there exists an optimal policy for the Q_j -unit subproblem that ships a batch of size Q_j to stage j if and only if the position of the last customer associated with this batch is below a certain threshold $\bar{y}_t^j$.

This optimal policy for the subproblem can be translated to the set of echelon (R, nQ) policies for the overall problem. Under this policy, if the echelon inventory level of stage j at time

t falls below a reorder point R_t^j , some number of batches of size Q_j are ordered to bring the echelon inventory level back above R_t^j . This policy is shown by Chen (2000) to be optimal when the objective is to minimize the long-run average cost. By the same argument as used in Section 6.2.3, the aforementioned policy equivalently tries to set the echelon inventory position at stage j to a level above $\bar{y}_t^j - 1$. Therefore, it is the same as an (R, nQ) policy with the reorder point $R_t^j = \bar{y}_t^j - 1$ at stage j in period t . Formally, the main result is stated as follows.

Theorem 6.2 (Muharremoglu, 2002, Chapter 5) *The set of (R, nQ) policies is optimal for the finite-horizon model and the infinite-horizon model with discounted cost criterion.*

As such, Muharremoglu (2002, Chapter 5) uses the single-unit analysis method to extend the result in Chen (2000) to broader types of planning horizons and objectives.

6.3.1.3 Limited echelon capacity

Finally, we review the work by Janakiraman and Muckstadt (2009), who use single-unit analysis to study a class of capacitated serial systems. In their main model, a two-echelon inventory system, in which both echelons have an identical ordering/production capacity limit Θ , is considered. The demand is Markov-modulated (same as Muharremoglu and Tsitsiklis, 2008) but the lead time is deterministic. For the case that the upstream lead time is exactly one period, Parker and Kapuściński (2004) have proven the optimality of a modified echelon base-stock policy. As one of the primary contributions, Janakiraman and Muckstadt (2009) utilizes a decomposition approach to extend the policy and the optimality result to the case where the lead time is two periods.

First, from the view of single-unit analysis, the problem for an M -stage system can be formulated in terms of the unit locations z_t^i and customer positions y_t^i , along with the control variable u_t^i for each unit located at the actual stages. Assume orders shipped from stage j to stage $j-1$ experience a lead time L_{j-1} ; then $\bar{L}_j = 1 + L_1 + \dots + L_{j-1}$ is the location index of the actual stage j for $j = 2, \dots, M+1$ (the location for stage 1 is, by definition, 1). Note that the order placed by stage $j-1$ should not exceed the available stock at stage j or the ordering/production capacity limit Θ . Hence, the release/hold decision u_t^i for unit i should comply with the capacity limit constraint, i.e., $1\{i \mid u_t^i(z_t^i) = 1\} \leq \min\{\Theta, 1\{i \mid z_t^i = \bar{L}_j\}\}$ for $j = 2, \dots, M+1$. A policy satisfying the above condition is called a feasible policy. With this formulation, it is straightforward to see that the search for optimal policies can be safely restricted to the class of feasible and monotonic policies.

Note that the decomposition defined by Janakiraman and Muckstadt (2009) is not like the ones previously reviewed, because the resulting subsystem is not single-unit/single-customer, but is unit capacity. Specifically, the decomposition entails Θ subsystems with unit ordering/production capacity. Subsystem k ($1 \leq k \leq \Theta$) has a countably infinite number of unit-customer pairs with indexes (pertaining to the overall system) $k, k + \Theta, k + 2\Theta, \dots$. Every stage in each subsystem has a capacity limit of releasing at most one unit. Intuitively, there is a natural connection between the series of units $i, i + \Theta, \dots$: Under monotonic policies, unit $i + \Theta$ can be affected by the capacity constraint at stage j at time t if unit i has not been released from stage $j+1$ yet. Therefore, following the same dynamic as in the overall system, these units collectively act as if every stage can release at most one unit at a time. This subsystem is simpler to analyze for this reason. Although being different from the single-unit single-customer decomposition, this decomposable capacitated system can still be optimally controlled by decoupled policy, which can be shown using the standard single-unit argument:

Theorem 6.3 (Janakiraman and Muckstadt, 2009) *When subsystem k is managed optimally using a decoupled policy in periods $t, t+1, \dots, T$, for all $1 \leq k \leq \Theta$, the resulting policy is optimal for the entire system.*

Note that this is true in spite of the fact that the demand processes of the subsystems are not independent. Applying the above results to a two-echelon system, Janakiraman and Muckstadt (2009) characterize the optimal policies in more detail. In particular, when the upstream lead time is one period, the decomposition concept developed above can be used to prove the same result shown by Parker and Kapuściński (2004), whose proof is based on dynamic programming. More importantly, when the upstream lead time is two periods, Janakiraman and Muckstadt (2009) characterize the structure of the optimal policy. Consider the first unit, indexed by i , located at stage 3 (outside supplier). Since the upstream lead time is two periods, the location of unit i is $z^i = L_1 + 3$. Note that location $L_1 + 2$ is the pipeline between stage 3 and stage 2 and location $L_1 + 1$ is stage 2. Hence, how many units are in these two locations reveals critical information for managing the subsystem. In fact, along with the position of customer i and the Markov chain that modulates the demand process, this information is sufficient to determine the optimal control for this subsystem at time t , which turns out to be a threshold-type policy. Such a threshold policy for the subsystems can help infer the structure of the optimal policy for the overall system, which constitutes the main contribution of Janakiraman and Muckstadt (2009), i.e., showing the optimality of the so-called “two-tier base-stock” policy for the two-period lead time case.

6.3.2 Serial Systems with Stochastic Lead Times

There are many papers in the inventory management literature that study stochastic lead times, with an early reference in the book by Hadley and Whitten (1963). They characterize the long-run behavior of a single-stage system under the assumption that lead times are i.i.d. and that orders arrive in the same sequence they were placed. These two assumptions can facilitate the analysis, but do not generally hold simultaneously. Hence, studies on stochastic lead times following Hadley and Whitten (1963) generally focus on one of the two assumptions. On the one hand, the lead times are taken to be i.i.d. (thus allowing order crossing) in some papers such as Zalkind (1978), Robinson et al. (2001), and Song and Zipkin (1996b). On the other hand, lead time processes that preclude order crossing are also common in the literature; lead times introduced by Kaplan (1970) and exogenous sequential lead time studied by Zipkin (1986) are two examples. As shall be seen in the following, Muharremoglu and Tsitsiklis (2008) adopt Kaplan’s lead time assumption and link the uncertainty to a modulating Markov chain, whereas Muharremoglu and Yang (2010) assume a general class of exogenous lead times that includes all aforementioned lead times.

6.3.2.1 Markov-modulated lead times

We now review the main results from the paper by Muharremoglu and Tsitsiklis (2008). This paper formalizes the method of single-unit analysis, uses the method to decompose the problem, and proves the optimality of a class of state-dependent echelon base-stock policies in a serial system where the uncertainty in demand and lead times arises from an exogenous Markov chain $\{\xi_t\}_{t=1}^{\infty}$.

One of the reasons why the problem is decomposable is that the order lead times are assumed to be such that an order cannot arrive at its destination before an earlier order does. Therefore, the lead time model precludes order crossing and ensures some independence of order arrivals, and thus the unit–customer pair under single-unit analysis could be decoupled from each other.

Using the concepts of unit location and customer position from single-unit analysis, Muharremoglu and Tsitsiklis (2008) propose a class of state-dependent echelon base-stock policies and show its optimality. Let $A = \{a_1, a_2, \dots, a_{M+1}\}$ be the set of unit locations where a_j corresponds to the location of stage j ; and let $A' = A \setminus \{a_1\}$. In addition, for any location $z > 0$, let $v(z) = \max \{j \mid j \in A \text{ and } j \leq z\}$. That is, starting from location z and going in the downstream direction, $v(z)$ is the first location that is an actual stage. Then, a policy is a state-dependent echelon base-stock policy if for every t and every realization of the Markov chain $\xi_t = \xi$, there exists a value $S_t^{v(z-1)}(\xi)$ such that, given the locations of units, z^i , and positions of customers, y^i , the following equation holds for every location $z \in A'$:

$$\begin{aligned} & \mathbb{I}\{i \mid z^i = z, u_t^i = 1\} \\ &= \min \left\{ \left[S_t^{v(z-1)}(\xi) - \left(\mathbb{I}\{i \mid 1 \leq z^i \leq z-1\} - \mathbb{I}\{i \mid y^i = 1\} \right) \right]^+, \mathbb{I}\{i \mid z^i = z\} \right\}. \end{aligned} \quad (6.1)$$

On the left-hand side, it is the control decision at location $z \in A'$, which corresponds to stage $2, 3, \dots, M$, of how many units to release. On the right-hand side, the term $S_t^{v(z-1)}(\xi)$ is the base-stock level for location $v(z-1)$ (i.e., the immediate downstream stage); the term $\mathbb{I}\{i \mid 1 \leq z^i \leq z-1\} - \mathbb{I}\{i \mid y^i = 1\}$ represents the echelon inventory position at $v(z-1)$, as every demand from customers with position $y^i = 1$ is backlogged; and the last term $\mathbb{I}\{i \mid z^i = z\}$ is the total number of units at location z . Therefore, such a policy operates in the following way: For every actual stage except stage 1, the echelon inventory position at the next downstream stage is calculated and enough units (to the extent of availability) are released to raise this number to a target value. The echelon inventory position is the total number of units at locations $1, \dots, v(z-1)$, plus the units in transit toward location $v(z-1)$ (i.e., units at locations $v(z-1)+1, \dots, z-1$), minus the backlogged demand.

Despite being more complicated than the toy example in Section 6.2, the serial system with Markov-modulated lead time and demand processes studied in Muharremoglu and Tsitsiklis (2008) can still be analyzed using the same single-unit method. In fact, the set of policies described in Equation (6.1) is able to effectively replicate any monotonic and decoupled policies on monotonic states (see these definitions in Section 6.2.2). By focusing on the single-unit single-customer subsystem and the movement of this unit–customer pair, the following is shown by Muharremoglu and Tsitsiklis (2008):

Proposition 6.1 (Muharremoglu and Tsitsiklis, 2008) *Suppose that π is a monotonic and decoupled policy. Then, there exists a state-dependent echelon base-stock policy that agrees with π at every monotonic state.*

Using similar arguments as seen in Section 6.2.3, one can prove the optimality of the monotonic and decoupled policies for the subsystem, and therefore the entire system. As such,

the set of state-dependent echelon base-stock policies is optimal for the overall problem. Moreover, this optimality result holds for finite planning horizon and infinite horizon with either discounted cost criterion or long-run average cost criterion, as stated below.

Theorem 6.4 (Muharremoglu and Tsitsiklis, 2008) *The set of state-dependent echelon base-stock policies is optimal for the finite-horizon model, the infinite-horizon model with discounted cost criterion, and the infinite-horizon model with average cost criterion.*

It is worth mentioning that the analysis for different types of planning horizon and optimality criteria follows the same basic single-unit idea and thus remains largely the same. Furthermore, for an infinite-horizon problem with an average cost criterion, the average cost can be written in a form that explicitly reflects the decomposition technique developed here. Specifically, the optimal average cost, C^* , equals the product of the expected demand, $\bar{d}$, and the minimized cost, C , associated with a particular unit–customer pair; i.e., $C^* = \bar{d}C$. Lastly, the single-unit decomposition approach can give rise to efficient algorithms as well, because the base-stock levels can be calculated by simply computing an optimal policy for a single-unit single-customer subproblem. Indeed, Muharremoglu and Tsitsiklis (2008) provide several algorithms that are easy to understand and to implement.

6.3.2.2 Exogenous lead times

The work of Muharremoglu and Yang (2010) studies single and multi-stage serial systems with a quite general class of stochastic lead times, referred to as exogenous lead times. The lead time model does not make assumptions about order crossing, so the paper actually covers both the sequential and the non-sequential lead times. A lead time process $(L_1(t), \dots, L_M(t))$ is exogenous if the lead time process $L_j(t)$ of stage j is independent of the lead time processes of all other stages and independent of the demand process. The demand in each time period is assumed to be i.i.d. random variables with expectation $E[D_i] = \bar{d}$. Within the class of base-stock policies, the objective is to determine the optimal base-stock levels, i.e., $\mathbf{s} = (s_1, \dots, s_M) \geq 0$, so that the infinite-horizon average cost $C(\mathbf{s})$ is minimized.

The main contribution of Muharremoglu and Yang (2010) is in providing a method to determine the optimal base-stock levels and evaluate the cost of a given base-stock policy. The single-unit analysis is at the core of its methodology. However, unlike Muharremoglu and Tsitsiklis (2008), the original problem is *not* decomposable and the optimal policy is *not* decoupled. Indeed, orders may overtake each other for exogenous lead times and the movement of unit–customer pairs would thus get entangled. Nevertheless, Muharremoglu and Yang (2010) still apply single-unit analysis to derive insightful structural results. Specifically, the total cost for the entire system is shown to be the product of the expected demand and the average cost of a unit–customer pair; i.e., $C(\mathbf{s}) = \bar{d} \times E[\zeta]$, where ζ is a random variable representing the cost incurred while moving a unit to match with the corresponding customer.

For a sequential lead time, the long-run cost of a typical unit–customer pair can be further simplified by solving a related single-unit problem. Suppose that the steady-state distribution of the lead time process exists and is denoted by $\mathbf{L}^{ss}$. Consider a version of the single-unit problem where the lead time has the same distribution as $\mathbf{L}^{ss}$. Given the position of the corresponding customer at the initial time, y_0 , the total cost of the base-stock $\mathbf{s}$ for matching this unit and the customer is denoted by $J(y_0, \mathbf{s}, \mathbf{L}^{ss})$. Then, the total average

cost associated with the single-unit problem, $C^s(\mathbf{s}, \mathbf{L}^{ss})$, can be calculated by taking the limit $C^s(\mathbf{s}, \mathbf{L}^{ss}) = \lim_{y_0 \rightarrow \infty} J(y_0, \mathbf{s}, \mathbf{L}^{ss})$. As such, this cost can be used to find the average cost for the overall system.

Proposition 6.2 (Muharremoglu and Yang, 2010) *If the lead time process for the original problem is sequential, then $C(\mathbf{s}) = \bar{d}C^s(\mathbf{s}, \mathbf{L}^{ss})$ for all $\mathbf{s}$.*

This finding is analogous to the result of Muharremoglu and Tsitsiklis (2008), where the system can be decomposed into independently managed subsystems. By contrast, the result by Muharremoglu and Yang (2010) is achieved through algebra, in spite of the fact that the original problem is not decomposable in general.

For systems with general exogenous (non-sequential) lead times, the exact average cost is difficult to compute because $C^s(\mathbf{s}, \mathbf{L}^{ss})$ cannot represent the true cost of the single-unit system due to possible order crossing. In light of this difficulty, Muharremoglu and Yang (2010) take some additional steps to solve the problem. First, the *order-based ordered* lead time $\hat{L}_j(t)$ is defined as the duration between the t^{th} order release from stage $j+1$ and the t^{th} order arrival at stage j . Obviously, these orders do not have to be the same order because orders may not arrive in the same sequence they were released. Second, define $V_j(t)$ to be the number of outstanding orders between stages $j+1$ and j at time t . Let $\hat{\mathbf{L}}^{ss}$ and $\mathbf{V}^{ss}$ be the steady-state random vector of the process $(\hat{L}_1, \dots, \hat{L}_{M-1})$ and $(V_1, \dots, V_{M-1})$, respectively. Then, Muharremoglu and Yang (2010) make a couple of critical observations: (1) $\hat{\mathbf{L}}^{ss}$ and $\mathbf{V}^{ss}$ have the same distribution; and (2) the infinite-horizon average cost of a single-stage problem depends on the lead time process only through the distribution of $\mathbf{V}^{ss}$, or, equivalently, the distribution of the steady-state order-based ordered lead time process $\hat{\mathbf{L}}^{ss}$.

Next, Muharremoglu and Yang (2010) relate the original problem to the single-unit problem by going through an intermediate surrogate problem, which is described as follows. Everything else being identical, the surrogate problem simply replaces the lead time process $L_j(t)$ in the original system with the order-based ordered lead time $\hat{L}_j(t)$. See Table 2 in Muharremoglu and Yang (2010, p. 119) for an illustrative example regarding the transformation from the original problem to its surrogate problem.

Hence, although the original lead time process is not sequential, the lead time process for the surrogate problem is. As such, the above Proposition 6.2 (Muharremoglu and Yang, 2010) may be applied to evaluate the average cost for the surrogate problem; let $C'(\mathbf{s})$ be this cost. First, for single-stage systems, $C(\mathbf{s})$ and $C'(\mathbf{s})$ are proven to be equal; i.e., the surrogate problem is an exact representation (in terms of average cost) of the original problem in the single-stage case.

Proposition 6.3 (Muharremoglu and Yang, 2010) *For a single-stage problem, $C(s) = C'(s) = \bar{d}C^s(s, \hat{\mathbf{L}}^{ss})$ for all base-stock level s .*

This result is derived from the aforementioned critical observations that $C(s)$ depends on the lead time process only through the distribution of $\mathbf{V}^{ss}$, which has the same distribution as $\hat{\mathbf{L}}^{ss}$. Indeed, in the single-stage case, the supplier would send the order quantity that equals to the i.i.d. demand. Second, for multi-stage systems, $C(s) = C'(s)$ would hold only under certain conditions. Intuitively, the order sizes are no longer realizations of the i.i.d. demand process

in the multi-stage case; rather, they are dependent on the inventory availability at upstream stages. As shown by Muharremoglu and Yang (2010), the single-unit method is exact for a multi-stage problem in four cases: (i) when orders do not cross, or (ii) when order crossing occurs only at the most upstream stage, or (iii) when the difference between base-stock levels of consecutive stages is sufficiently large, or (iv) in two-stage systems with deterministic upstream lead time, when the base-stock levels are the same. Finally, if the above conditions do not hold, then $C'(s)$ serves as a good approximation, producing near-optimal base-stock levels and close cost estimates.

6.4 APPLICATIONS IN MANAGING ASSEMBLY SYSTEMS

In this section, we review works that showcase the application of single-unit analysis in managing assembly systems. Consider a periodically reviewed assembly system with stochastic demand. In this system, components are ordered from outside suppliers, transferred and combined to form subassemblies at some intermediate stages, and finally assembled together into a finished product. Suppose that every time an assembly takes place, one component or subassembly can be made into only one subassembly or final product.

Moreover, unlike serial systems, there are multiple flow units (components, subassemblies, and final products) in an assembly system. However, in terms of the final finished product, there is only a single kind to give to customers. Therefore, the fundamental idea of matching one product with one customer using single-unit analysis still works for assembly systems, although tracing one product requires tracing all its components. To unify terminology, the word “unit” in the single-unit analysis (e.g., “unit–customer pair”) for assembly systems refers to the final product.

In the following, we review two papers. The first paper applies a single-unit method to analyze a general assembly system and obtains the single-unit single-customer decomposition. The second paper focuses on a special assembly system, i.e., a single-product assemble-to-order system, and utilizes single-unit analysis to develop an efficient algorithm for policy evaluation and optimization.

6.4.1 General Assembly Systems

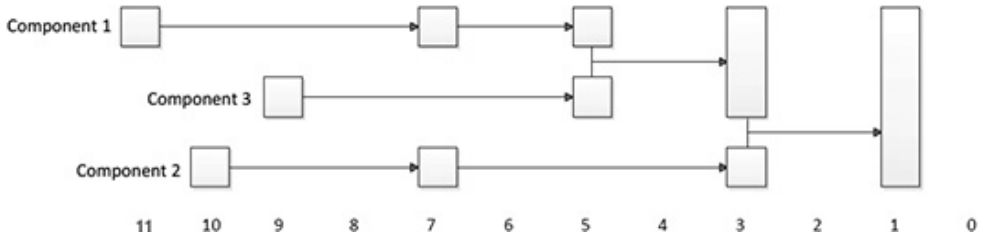
The paper by Chen and Muharremoglu (2014) studies an assembly system with a general tree structure and assumes deterministic lead times and stochastic demand. The optimal inventory control policy for this system in the long run has been given by Rosling (1989). However, depending on the starting state, Rosling’s policy could be suboptimal in the short run. The main contribution of Chen and Muharremoglu (2014), therefore, is a characterization of the optimal policy structure for a general assembly system with an arbitrary starting state. As an important first step toward deriving this result, Chen and Muharremoglu (2014) decompose the problem into a set of single-unit single-customer subproblems.

Let the number of components be K . Since each component will go through the system, let the total lead time of this process for component k be L_k . The general idea of defining the unit location is exactly the same as laid out in Section 6.2.1, e.g., a unit at location 0 is already sold to a customer, and the location with the highest index represents the outside supplier. However, there are some notable differences from the serial systems. The details are illustrated by an

example in Figure 6.2. In particular, here in the assembly setting, a unit means a set of all components, and thus the location of a unit is a vector of the locations of all components of the unit. Therefore, the system dynamics could be described in terms of component locations. For instance, once component 1 and component 3 reach location 5, they are ready to be assembled, and it takes two time periods until the subassembly is ready and arrives at location 3. It is important to note that, from that point on, those two particular components (1 and 3) will always have the same location at any time.

The single-unit single-customer subproblem could then be formulated as follows. Each component k is indexed from location 0 to location L_k , breaking ties arbitrarily. The location of unit i at time t in the assembly system is a vector $\mathbf{z}_t^i = (z_t^1, \dots, z_t^K)$, where z_t^k is the location of the i^{th} component k at time t . The position of customer i , y_t^i , is defined in the same way as described in Section 6.2.1. In addition, the control decision at time t , $\mathbf{u}_t^i$ is also a vector, with each entry u_t^k representing a hold/release action for component k . In an assembly system, the control variable $\mathbf{u}_t^i$ should satisfy an additional constraint: As soon as the assembly of two components is initiated, the two components should either both move or both stay. For the cost, let $h_t^k(l)$ be the per-unit holding cost of component k at location l in period t , and let b be the per-unit backlogging cost. Then, the objective is to determine the control variable $\mathbf{u}_t^i$ so that the total cost of the subsystem is minimized across a finite planning horizon.

This problem can be solved by relating it to a corresponding serial system, constructed in the following way. Consider a single-unit single-customer subproblem for the assembly system with initial state $(\mathbf{z}_0, y_0)$ (superscript i omitted). For each component k , let $B(k)$ represent the set of stocking locations for component k , either as itself or part of a subassembly/final product. Then, the stocking locations of the corresponding serial system are $B_{\text{ser}} = \cup_k \{l \leq z_0^k, l \in B(k)\}$, which are the stocking locations of the assembly system where at least one component can still be stored in a later period. The holding cost at location l at time t is $\hat{h}_t(l) := \sum_{k=1}^K h_t^k(\min\{z_0^k\})$, for $l = 0, \dots, w$, where $w := \max\{l | l \in B_{\text{ser}}\}$ is the total lead time of the serial system. The backlog costs and demand process remain the same. Note that the structure of the corresponding serial system depends only on the initial state $(\mathbf{z}_0, y_0)$. According to Muharremoglu and



Note: Total lead times of the three components are $L_1 = 10$, $L_2 = 9$, and $L_3 = 8$, respectively. Rectangles are actual storing stages. Vertical segment represents an assembling operation. Horizontal arrows are on the same scale as the delays between stages, which is reflected by the numbers at the bottom. These numbers are the locations in this system. This figure is adapted from Chen and Muharremoglu (2014).

Figure 6.2 Illustration of the locations in a three-component assembly system

Tsitsiklis (2008), such a single-unit single-customer serial system can be optimally controlled by an echelon threshold policy. Specifically, for any stocking location $l \in B_{ser} \setminus \{0\}$, there is a threshold s_t^l , such that a unit at l is released if the customer is close enough, i.e., $y_t \leq s_t^l$. Consider the following control of a threshold form. Let $l_t = \max_r z_r$ be the location of the furthest component(s). Release a component k from its stocking location z_t^k if and only if it is one of the furthest components (i.e., $z_t^k = l_t$) and $y_t \leq s_t^l(\mathbf{z}_0)$ (here $l = l_t$ is changing dynamically).

Theorem 6.5 (Chen and Muharremoglu, 2014) *The above-defined threshold policy is optimal for a single-unit single-customer subsystem.*

Note that the threshold s_t^l from the corresponding serial system is written as a function of the initial state. Hence, the single-unit assembly problem with *any* initial state $(\mathbf{z}_0, y_0)$ can be solved.

Finally, observe that the series of single-unit single-customer subproblems can be separately and independently solved to obtain the optimal policy for the entire system. Therefore, the above threshold policy for each unit–customer pair provides clues for finding the overall optimal policy. The optimal policy for an arbitrary initial state, called *dynamic balanced echelon base-stock policy*, is one where the base-stock level s_t^n is dynamically evolving, in contrast to Rosling’s result (1989) where base-stock levels are constant. In addition, the evolution depends on the total demand observed so far; i.e., $s_t^n = s_t^n(\hat{D}_t)$ where $\hat{D}_t = D_0 + \dots + D_{t-1}$ is the cumulative demand up to period $t-1$. For exposition brevity, readers are referred to Definition 7 in Chen and Muharremoglu (2014) for the detailed functional form of $s_t^n(\hat{D}_t)$. Here, we simply state the main result:

Theorem 6.6 (Chen and Muharremoglu, 2014) *There exists an optimal dynamic balanced echelon base-stock policy for the overall assembly system.*

6.4.2 Single-Product Assemble-To-Order Systems

Next, we review the paper by Muharremoglu et al. (2021). This paper studies a single-product assemble-to-order (ATO) system, which can be viewed as a special assembly system. In particular, there is only one assembling operation and that is to produce the final product from all components. The paper employs a discrete-time model and assumes K components, each of which is procured from an outside supplier. The stochastic lead time process of all components, $(L_1(t), \dots, L_K(t))$, is assumed to be exogenous (defined in Muharremoglu and Yang, 2010). The demands over time are i.i.d. random variables. The paper focuses on component base-stock policies and tackles the problem of policy evaluation and optimization. Since the outstanding orders of components are all initiated by a common demand, the computational challenge is thus to calculate the distribution of the minimum of K correlated random variables. Suppose that the system is managed via a component base-stock policy, $\mathbf{s} = (s_1, \dots, s_K)$. Without loss of generality, let $s_1 \geq s_2 \geq \dots \geq s_K \geq 0$. The objective is to evaluate the policy performance under the infinite-horizon average cost criterion. Given the holding costs and backlog cost, the system’s long-run average cost is straightforward to derive, and it turns out that the expected backlog, denoted by $E[B]$, is the key quantity to compute. This is done by applying the following single-unit technique.

Consider a typical unit–customer pair. The single unit consists of K components that are destined to be assembled into a final product. The single customer is assumed to be in position y_t at time t . The main idea is to examine the movement of this unit–customer pair and characterize the release times and lead times of the components, as well as the arrival time of the customer.

As the demand in period t , D_t , realizes, the focal customer's position becomes smaller. Since a component base-stock policy is enforced, an important implication is that component k for this customer is released when the customer's position becomes s_k or less. Assume that, at time 0, $y_0 > s_1$ (the customer is initially in the distant future). As time progresses, the position of this customer will sequentially cross the base-stock levels. Let r_k be the time the threshold s_k is crossed; i.e., $y_{r_k-1} > s_k \geq y_{r_k}$. At this moment, component k is released. By assumption, it is easy to see $r_K \geq r_{K-1} \geq \dots \geq r_1$.

The release times and lead times of components together decide the ready time of the final product. To be specific, component k will arrive at time $r_k + L_k(r_k)$, and the final product will be ready for assembly when the latest component arrives, which is $\max_{1 \leq k \leq K} \{r_k + L_k(r_k)\}$. Moreover, the focal customer will eventually arrive in the system and demand the final product, by which time the position of the customer will become no more than 1. Denote this arrival time by $\tau = \min\{t : y_t \leq 1\}$.

Now, one can write the policy performance for the above single-unit single-customer system. Let $\alpha = \max_{1 \leq k \leq K} r_k + L_k(r_k) - r_K$ be the elapsed time between the moment when component K was released and the ready time of the final product. Let $\gamma = \tau - r_K$ be the elapsed time between the moment when component K was released and the arrival time of the customer. Hence, if $\alpha > \gamma$, then the product is not ready before the customer arrives, and the customer will be backlogged; otherwise the customer is immediately satisfied. As a result, the amount of time this typical customer is backlogged is $(\alpha - \gamma)^+$. Since all unit–customer pairs can be treated in the exactly same way, the above variables essentially have the same distribution for all customers. Thus, $E[(\alpha - \gamma)^+]$ is the expected backlog for any arriving customer in the distant future. As such, Muharremoglu et al. (2021) obtains one of the main results.

Proposition 6.4 (Muharremoglu et al., 2021) *The expected backlog for the overall system is given by*

$$E[B] = E[(\alpha - \gamma)^+] \cdot E[D]. \quad (6.2)$$

In the above, $E[D]$ is the average number of customers arriving in each period.

To apply Equation (6.2), the distributions of α and γ must be computed. The two random variables are generally correlated and depend on many factors such as the demand process. To facilitate the analysis, two auxiliary variables are identified: The overshoot $\omega := s_K - y_{r_n}$ and the undershoot $\upsilon := y_{r_n-1} - s_K$. As a result, α and γ can be nicely decoupled by conditioning on ω and υ . Thus, Equation (6.2) can be rewritten as a multiplication of conditional probabilities, which links the above single-unit analysis back to the overall policy evaluation.

Finally, the random variable υ can greatly reduce the computational burden in finding the performance measure. Instead of memorizing each demand realization in every past period, the algorithm only needs to keep a record of the sum of the past demands with the aid of υ . Therefore, the curse of dimensionality is avoided, and an algorithm with polynomial complexity can be developed. The paper by Muharremoglu et al. (2021) is the first to propose an

efficient algorithm for the performance evaluation of base-stock policies in a single-product ATO system with stochastic lead times, and the key idea is generated by single-unit analysis on the system.

6.5 OTHER RELATED MODELS

The single-unit technique can also facilitate inventory management for some complex systems with context-specific features. The decomposition approach usually leads to more tractable analysis and more insightful results. In this section, we review three papers of this kind. All of these papers consider serial inventory systems; however, each of them takes into account some particular situation that gives rise to complicated issues.

Martínez-de-Albéniz and Lago (2010) investigate a periodic-review single-echelon system with constant lead time but non-stationary, correlated, stochastic demand and cost. In addition, the selling price of the product forms a non-increasing stochastic process. Operated under a base-stock policy, the objective is to minimize the expected cost by balancing the holding cost and the backlog cost. The difficulty of this control problem lies in the non-stationarity of the relevant stochastic process, which may yield a high-dimensional state space in the dynamic program formulation. The paper utilizes the single-unit approach to formulate a related optimization problem for a typical unit–customer pair, and provides sufficient conditions for myopic policy, which has a closed-form expression, to be optimal.

Consider a standard periodic-review single-echelon system with backordering in an infinite-horizon setting. The product is sourced from an outside supplier and the order lead time is a constant L period. To set up for the single-unit analysis, the paper starts by showing that, when prices, $\{P_i\}$, are non-increasing over time, the i^{th} customer must be served with the i^{th} unit under any optimal policy. Thus, the overall problem can be formulated unit by unit. For the i^{th} unit–customer pair, let t_i be the (only) time that unit i is ordered, and T_i be the time customer i arrives. Note that all information about the demand process is embedded in the arrival process T_i . The unit is ready when the order is delivered after the lead time, meaning that the time when the customer pays for the unit is $T'_i := \max\{T_i, t_i + L\}$. The overall inventory control problem is to determine the order times $\{t_i\}$ for all i so that the net present value (with discount rate α) of the profit is maximized. Since costs and revenue are defined to be linear, the total profit has a linear structure, which warrants the single-unit decomposition. Specifically, one only needs to decide the order time t_i for unit i , independently from all other units, to maximize the profit of fulfilling customer i by unit i . The decomposition is shown below:

$$E_{\mathcal{I}_0} \left\{ \sum_{t=1}^{\infty} \alpha^t \left[\sum_{i=1}^{\infty} \mathcal{G}(t, t_i, T_i, T'_i) \right] \right\} \stackrel{\text{Decomposed}}{\Rightarrow} E_{\mathcal{I}_0} \left\{ \sum_{t=1}^{\infty} \alpha^t \mathcal{G}(t, t_i, T_i, T'_i) \right\}.$$

The function $\mathcal{G}$ depends on the time period, the movement of the unit–customer pair, and the control; of course, it also depends on the selling price process $\{P_i\}$, the procurement cost process, and the per-unit holding and backlog costs. The expectation in the objective function is conditioning on the initial information set $\mathcal{I}_0$; and the evolution of all system stochastic processes is determined by the past and present information up to time t .

Compared to the standard inventory control approach, where the inventory level and the demand forecast for all future periods constitute the state space, the single-unit single-customer decomposition focuses on a series of smaller problems, where the arrival time T_i of customer i is the only required information. As Martínez-de-Albéniz and Lago (2010) show, this simpler formulation leads to analytical formulas for each i and generates insights into the possible optimality of myopic policies. Note that similar results are also derived by Yu and Benjaafar (2009) using the same approach.

Berling and Martínez-de-Albéniz (2011) study a similar setting as in Martínez-de-Albéniz and Lago (2010), except that they focus on a continuous-review system with Poisson demand. Moreover, they do not consider the selling price of the product, but assume that the procurement price forms a particular continuous stochastic process such as the Brownian motion and Ornstein–Uhlenbeck process. The main contribution of Berling and Martínez-de-Albéniz (2011) is an explicit characterization of the optimal price-dependent base-stock policy in terms of a series of threshold prices. This is accomplished by using the single-unit decomposition technique.

Consider a single-echelon continuous-review system with a constant lead time L and Poisson demand of rate λ . There is a per-unit procurement price, which is stochastic; moreover, holding cost and backlog cost are linear with a constant rate. The objective is to minimize the expected discounted cost with the continuous discount rate α . Let t_i and T_i be, again, the order time of unit i and the arrival time of customer i , respectively. Then, the problem becomes selecting the order times for all units $\{t_i\}_{i=1}^{\infty}$ to minimize the expected discounted total cost for the entire system. The single-unit analysis approach can be applied here to decompose the objective function into an independent total cost function for a single-unit i . That is, this problem can be viewed unit by unit:

$$E \left\{ \int_{t=1}^{\infty} \left[\sum_{k=1}^{\infty} \hat{G}(t, t_i, T_i) \right] e^{-\alpha t} dt \right\} \xRightarrow{\text{Decomposed}} E \left\{ \int_{t=1}^{\infty} \hat{G}(t, t_i, T_i) e^{-\alpha t} dt \right\}$$

Again, $\hat{G}$ also depends on the procurement cost process and all other constant system parameters. As shown by Berling and Martínez-de-Albéniz (2011), the above single-unit decomposition approach leads to a novel interpretation of the base-stock policy, which is in terms of threshold prices.

Finally, the work by Berling and Martínez-de-Albéniz (2016) represents another example of applying the single-unit analysis to inventory problems in complex systems. In their paper, a serial system with continuous stages and continuous-review control is studied. Unlike the previous papers, the main management decision is the transportation speed, i.e., the order lead time. In this regard, the paper is related to the inventory literature that considers expediting (e.g., Parker and Kapuściński, 2004; Muharremoglu and Tsitsiklis, 2003).

Consider a serial system where products are moved along a continuum of locations $[0, U]$. Any $x \in [0, U]$ represents where the inventory is and a decision shall be made at this location on the speed v of the moving inventory. An operating cost $m(x, v) \geq 0$ is charged, which can be seen as the sum of holding cost and transportation cost. Demand arises at location $x = 0$ and forms a Poisson process. Taking all costs incurred in moving, holding, and possibly backlogging the inventory into account, the inventory control problem is to decide on the real-time transportation speed so that the discounted total cost is minimized. This problem

can be solved by a standard approach that involves a high-dimensional dynamic programming formulation.

Using a single-unit analysis approach, on the other hand, this optimal control problem can be decomposed into one-dimensional subproblems by tracking a single unit. Since the orders cannot cross at optimality, unmet demands are backlogged, and all costs are linear in units, one may just follow each unit from the beginning (when it enters the system at $x = U$) to the end (when it satisfies a demand at $x = 0$) and focus on optimizing the single-unit subsystem. The formulation gives rise to a much simpler control problem, from which insightful results can be derived. Here, the single-unit method facilitates the analysis by avoiding the more difficult tracking of the inventory level and its evolving distribution.

6.6 CONCLUDING REMARKS

As we have seen above, single-unit analysis serves as a useful tool in solving inventory control problems in various settings. The main underlying reason is that this analytical technique usually allows for decomposition of the original problem and decoupling of control policies. Specifically, instead of focusing on each order, each echelon inventory level, and each realized demand, single-unit analysis turns to examining a typical unit–customer pair by tracking all the cost-generating moments from the beginning to the end. Since the decomposed problem only has a single unit and a single customer, the dynamic is simple and it is often more likely to obtain a well-structured control policy. Then, the resulting policy can be translated back to apply to the original overall problem. For systems that are not decomposable, on the other hand, single-unit analysis becomes an idea-generating tool as it suggests novel insights into solving the inventory problem. Moreover, it provides ideas about developing efficient algorithms that may not be found using traditional approaches.

Basically, single-unit analysis may be used for an inventory system where a unit and a customer are paired before any control decision is made. Moreover, the original problem is decomposable if the system dynamics may be decoupled into single-unit single-customer subproblems. Therefore, despite its wide application in inventory management, the single-unit approach may not completely go through in certain settings. For example, if the costs have non-linear forms, then the cost of a subproblem becomes interdependent and the system is thus not decomposable in general; however, for certain types of non-linear cost, the system may still be decomposable after some modification of the relevant concepts (see, e.g., Federgruen and Wang, 2015). Similarly, if there are general unequal capacity limits at each stage in a serial system, then the control of subproblems would be coupled. Another example is systems with lost sales, where a unit would have to be reassigned to another customer once a demand is lost; so, the unit–customer pair is not destined in the beginning. In these situations, single-unit analysis can still help identify certain aspects that make the problem difficult and hint at possible solutions to overcome the difficulties.

Apart from the serial and assembly systems, distribution systems are also an important type of multi-echelon inventory system. For such a system, unless an allocation rule is given to dictate where a unit is shipped to, units cannot be paired with particular customers in advance. Under a specific allocation rule, single-unit analysis may be applied to independently control each unit–customer pair (see, e.g., Axsäter, 1990). However, the consideration of allocation rules can make things more complicated to apply the single-unit analysis in

general distribution systems. Overall, the above situations, where single-unit analysis may be applicable after some model reformulation or modification, indicate that this method could be employed in broader settings than discussed in this chapter. Therefore, when thinking of any potential opportunity to employ single-unit analysis, one should not be confined to a specific setting.

REFERENCES

- Achy-Brou, A. (2001). A new approach to multistage serial inventory systems. Master's thesis, Massachusetts Institute of Technology.
- Axsäter, S. (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research*, 38(1), 64–69.
- Axsäter, S. (1993a). Exact and approximate evaluation of batch-ordering policies for two-level inventory systems. *Operations Research*, 41(4), 777–785.
- Axsäter, S. (1993b). Optimization of order-up-to-s policies in two-echelon inventory systems with periodic review. *Naval Research Logistics*, 40(2), 245–253.
- Berling, P., & Martínez-de-Albéniz, V. (2011). Optimal inventory policies when purchase price and demand are stochastic. *Operations Research*, 59(1), 109–124.
- Berling, P., & Martínez-de-Albéniz, V. (2016). Dynamic speed optimization in supply chains with stochastic demand. *Transportation Science*, 50(3), 1114–1127.
- Chen, F. (2000). Optimal policies for multi-echelon inventory problems with batch ordering. *Operations Research*, 48(3), 376–389.
- Chen, F., & Song, J. S. (2001). Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research*, 49(2), 226–234.
- Chen, F., & Zheng, Y. (1994). Lower bounds for multi-echelon stochastic inventory systems. *Management Science*, 40(11), 1426–1443.
- Chen, S., & Muharremoglu, A. (2014). Optimal policies for assembly systems: Completing Rosling's characterization. Working paper. Available at SSRN 2475218.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- Federgruen, A., & Wang, M. (2015). A continuous review model with general shelf age and delay-dependent inventory costs. *Probability in the Engineering and Informational Sciences*, 29(4), 507–525.
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836.
- Hadley, G., & Whitten, T. M. (1963). *Analysis of inventory systems*. Prentice Hall.
- Janakiraman, G., & Muckstadt, J. A. (2009). A decomposition approach for a class of capacitated serial systems. *Operations Research*, 57(6), 1384–1393.
- Kaplan, R. S. (1970). A dynamic inventory model with stochastic lead times. *Management Science*, 16(7), 491–507.
- Katircioglu, K., & Atkins, D. (1998). New optimal policies for a unit demand inventory problem. Working Paper, University of British Columbia.
- Martínez-de-Albéniz, V., & Lago, A. (2010). Myopic inventory policies using individual customer arrival information. *Manufacturing & Service Operations Management*, 12(4), 663–672.
- Muharremoglu, A. (2002). A new perspective on multi-echelon inventory systems. Ph.D. thesis, Massachusetts Institute of Technology.
- Muharremoglu, A., & Tsitsiklis, J. N. (2003). Dynamic leadtime management in supply chains. Working paper, Massachusetts Institute of Technology.
- Muharremoglu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, 56(5), 1089–1103.
- Muharremoglu, A., & Yang, N. (2010). Inventory management with an exogenous supply process. *Operations Research*, 58(1), 111–129.

- Muharremoglu, A., Yang, N., & Geng, X. (2021). Single-product assemble-to-order systems with exogenous lead times. Working paper.
- Parker, R. P., & Kapuściński, R. (2004). Optimal policies for a capacitated two-echelon inventory system. *Operations Research*, 52(5), 739–755.
- Robinson, L. W., Bradley, J. R., & Thomas, L. J. (2001). Consequences of order crossover under order-up-to inventory policies. *Manufacturing & Service Operations Management*, 3(3), 175–188.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579.
- Song, J.-S., & Zipkin, P. (1992). Evaluation of base-stock policies in multiechelon inventory systems with state-dependent demands part i: State-independent policies. *Naval Research Logistics (NRL)*, 39(5), 715–728.
- Song, J.-S., & Zipkin, P. (1996a). Evaluation of base-stock policies in multiechelon inventory systems with state-dependent demands part ii: State-dependent depot policies. *Naval Research Logistics (NRL)*, 43(3), 381–396.
- Song, J.-S., & Zipkin, P. (1996b). The joint effect of leadtime variance and lot size in a parallel processing environment. *Management Science*, 42(9), 1352–1363.
- Yu, Y., & Benjaafar, S. (2009). A customer-item decomposition approach to stochastic inventory systems with correlation. Working Paper, University of Minnesota.
- Zalkind, D. (1978). Order-level inventory systems with independent stochastic leadtimes. *Management Science*, 24(13), 1384–1392.
- Zipkin, P. (1986). Stochastic leadtime in continuous-time inventory models. *Naval Research Logistics*, 33(4), 763–774.
- Zipkin, P. (1991). Evaluation of base-stock policies in multiechelon inventory systems with compound-Poisson demands. *Naval Research Logistics*, 38(3), 397–412.

7. Robust inventory management

Michael R. Wagner

7.1 INTRODUCTION

In this chapter, we present models that minimize cumulative ordering, holding, and shortage costs for a classic T -period inventory management problem for a single product with no fixed costs and backordering allowed. Demand is stochastic, not necessarily identically distributed over periods, and possibly correlated across periods. The distributions of demand are not known, though we assume we have some information about them (e.g., moment information). We apply recent advances in robust optimization by utilizing uncertainty sets that are inspired by the limit theorems of probability, such as the Central Limit Theorem (CLT), the Strong Law of Large Numbers (SLLN), and the Law of the Iterated Logarithm (LIL). A robust optimization approach, coupled with these specific uncertainty sets, leads to tractable models, including closed-form solutions in a variety of scenarios. The simple nature of these closed-form results allows a decision maker to better understand the inventory management strategy he/she implements, in contrast to the computational solution to an optimization problem.

Our basic model is static, where all ordering decisions must be made at the start of the planning horizon, and there are many applications for such a model. For instance, specifying multiple order quantities in advance occurs in some supply chain contracts, such as “advanced booking discount” (ABD) programs (Tang et al. (2004)), advanced purchase contracts (Ozer and Wei (2006)), and contracts for short lifecycle products (e.g., fashion, high technology), where there is usually a single ordering opportunity (Chen et al. (2016)). However, we also apply our static model in a rolling horizon framework, allowing the decision in each period to depend on the currently observed inventory position, which results in a dynamic model that can react to new information (e.g., demand realizations). Similar rolling horizon implementations are also studied by Mamani et al. (2017), Solyalı et al. (2016), and Wagner (2018).

7.1.1 Literature Review

We first position our chapter with respect to the broader robust inventory management literature. One of the first such models appears in Bertsimas and Thiele (2006), where the authors utilize the “budgets of uncertainty” idea from Bertsimas and Sim (2004) to address the conservatism of robust optimization. Bienstock and Özbay (2008) extend Bertsimas and Thiele (2006) in various directions. See and Sim (2010) analyze a “factor-based” model of uncertainty, which results in a second-order cone reformulation. Wagner (2010, 2011) studied similar problems from the “online optimization” perspective, where the only property known about demand is non-negativity. Ardestani-Jaafari and Delage (2016) extend the ideas from Gorissen and Hertog (2013) to analyze more general robust optimization problems involving sums of piecewise linear functions, which can be applied to inventory management. Mamani et al. (2017) study a similar problem, where the uncertainty sets are motivated by the Central

Limit Theorem, which results in closed-form solutions; this paper strongly influences our chapter. Solyalı et al. (2016) propose a new robust formulation of inventory control based on ideas of facility location. Wagner (2018) provides a continuous-time formulation of a similar problem, where the uncertainty set is motivated by the Strong Law of Large Numbers. Sinha et al. (2021) extend Mamani et al. (2017) to accommodate revenues as well as period-dependent economic parameters. Sinha et al. (2021) and Wagner (2018) also influence this chapter.

We consider both static as well as dynamic rolling horizon implementations in this chapter. Ben-Tal et al. (2004) studies an adjustable robust optimization problem, where variable values can be changed once unknown quantities are realized, which is shown to be NP-hard. Bertsimas et al. (2010) prove the optimality of policies that are affine in previous periods' demand realizations, for a general class of multi-stage robust optimization models where unknown parameters are constrained to lie in intervals. Iancu et al. (2013) extend this research by more fully characterizing the problem structures where affine policies are optimal. Mamani et al. (2017), Solyalı et al. (2016), and Wagner (2018) study rolling horizon implementations similar to that considered in this chapter; in particular, Wagner (2018) studies various rolling horizon contexts which depend on whether or not the observed demand stream is consistent with the original robust uncertainty set.

Finally, we discuss the selection of the robust uncertainty sets. In earlier robust optimization papers, the uncertainty sets were selected to be mathematically tractable, including interval, polyhedral, and ellipsoidal sets. In recent years, researchers have designed uncertainty sets that are motivated by the limit theorems of probability. Bertsimas et al. (2011) analyze queueing networks using a robust uncertainty set that is motivated by the Law of the Iterated Logarithm. Bandi and Bertsimas (2012) consider uncertainty sets motivated by the Central Limit Theorem, which are also applied in detailed investigations of option pricing (Bandi and Bertsimas (2014b)), auction design (Bandi and Bertsimas (2014a)), queueing theory (Bandi et al. (2015, 2018)), and inventory management (Mamani et al. (2017); Sinha et al. (2021)). Wagner (2018) uses uncertainty sets that are motivated by the Strong Law of Large Numbers in an inventory management context.

7.2 STATIC ROBUST INVENTORY MANAGEMENT

In this chapter, we use the following generic uncertainty set for the demand vector $(D_1, \dots, D_T)$ over T periods,

$$\Omega = \left\{ (D_1, \dots, D_T) : \underline{S}_t \leq \sum_{j=1}^t D_j \leq \bar{S}_t, \ell_t \leq D_t \leq u_t, t = 1, \dots, T \right\}, \quad (7.1)$$

where the parameters $(\underline{S}_t, \bar{S}_t, \ell_t, u_t)_t$ are used to capture the relevant structure of the various limit theorems of probability. In order to guarantee that Ω is not empty, we assume $\underline{S}_t \leq \bar{S}_t$, $0 \leq \ell_t \leq u_t$, $\underline{S}_t - \bar{S}_{t-1} \leq \bar{S}_t - \bar{S}_{t-1}$, and $[\underline{S}_t - \bar{S}_{t-1}, \bar{S}_t - \bar{S}_{t-1}] \cap [\ell_t, u_t] \neq \emptyset$, for $t = 1, \dots, T$, since $\underline{S}_t - \bar{S}_{t-1} \leq D_t = \sum_{j=1}^t D_j - \sum_{j=1}^{t-1} D_j \leq \bar{S}_t - \bar{S}_{t-1}$.

Let c denote the unit procurement cost, h the unit inventory holding cost, b the unit backordering cost, and z_t the ordering decision in period t . Our static robust inventory management problem, as a function of the uncertainty set Ω , is

$$\begin{aligned}
& \min_{(z_1, \dots, z_T) \geq 0} \sum_{t=1}^T (cz_t + m_t) \\
& \text{s.t.} \quad y_t = \sum_{i=1}^t (z_i - D_i), t = 1, \dots, T \\
& \quad m_t \geq hy_t, t = 1, \dots, T, \forall (D_1, \dots, D_T) \in \Omega \\
& \quad m_t \geq -by_t, t = 1, \dots, T, \forall (D_1, \dots, D_T) \in \Omega,
\end{aligned} \tag{7.2}$$

where y_t is the inventory position in period t , m_t captures the mismatch cost $\max\{hy_t, -by_t\}$ in period t , and we assume zero initial inventory $y_0 = 0$. In particular, since Equation (7.2) is a robust optimization model, the variables $(z_i, y_i, m_i)_i$ are feasible if and only if they satisfy the last two sets of constraints for every demand vector in Ω .

The partial-sum structure of Ω allows us to determine, in closed form, the minimum and maximum cumulative demands for the first k periods, $k = 1, \dots, T$, which we denote as

$$\underline{D}_k = \min_{(D_1, \dots, D_T) \in \Omega} \sum_{t=1}^k D_t \quad \text{and} \quad \overline{D}_k = \max_{(D_1, \dots, D_T) \in \Omega} \sum_{t=1}^k D_t, \tag{7.3}$$

respectively. Note that both $\underline{D}_k$ and $\overline{D}_k$ are increasing in k . The next lemma, whose proof can be found in Mamani et al. (2017), provides expressions for these partial sums for the uncertainty set defined in Equation (7.1).

Lemma 7.1 For $k = 1, \dots, T$,

$$\begin{aligned}
\overline{D}_k = \min & \left\{ \sum_{t=1}^k u_t, \min_{i < k} \left\{ \overline{S}_i + \sum_{t=i+1}^k \min\{u_t, \overline{S}_t - \overline{S}_{t-1}\} \right\}, \overline{S}_k, \right. \\
& \left. \min_{i > k} \left\{ \overline{S}_i - \sum_{t=k+1}^i \max\{\ell_t, \underline{S}_t - \overline{S}_{t-1}\} \right\} \right\}
\end{aligned}$$

and

$$\begin{aligned}
\underline{D}_k = \max & \left\{ \sum_{t=1}^k \ell_t, \max_{i < k} \left\{ \underline{S}_i + \sum_{t=i+1}^k \max\{\ell_t, \underline{S}_t - \overline{S}_{t-1}\} \right\}, \underline{S}_k, \right. \\
& \left. \max_{i > k} \left\{ \underline{S}_i - \sum_{t=k+1}^i \min\{u_t, \overline{S}_t - \overline{S}_{t-1}\} \right\} \right\}.
\end{aligned}$$

The next lemma provides an optimality condition for Equation (7.2), and the proof may be found in Mamani et al. (2017).

Lemma 7.2 *If $kb < c \leq (k+1)b$ for some integer $0 \leq k \leq T-1$, then the robust order quantities satisfy*

$$\sum_{j=1}^t z_j^* = \frac{b\bar{D}_t + h\underline{D}_t}{b+h}, t=1, \dots, T-k \quad \text{and} \quad z_t^* = 0, t > T-k.$$

If $Tb < c$, $z_t^* = 0$ for all t .

This lemma balances the worst-case holding and backordering costs for the first $T-k$ periods, and the entire horizon if $c \leq b$. In other words, the worst-case holding cost in period t for the smallest possible cumulative demand $\underline{D}_t$, namely $h(\sum_{j=1}^t z_j - \underline{D}_t)$, will equal the worst-case backordering cost for the largest possible cumulative demand $\bar{D}_t$, namely $b(\bar{D}_t - \sum_{j=1}^t z_j)$.

It is also insightful to note that the optimality condition from Lemma 7.2 is identical to the newsvendor solution for demand that is uniformly distributed on $[\underline{D}_t, \bar{D}_t]$ with overage and underage costs of h and b , respectively. Therefore, the derivation of the robust quantities can be viewed as a sequence of newsvendor solutions, where in the t -th application the costs up to period t are minimized for uniformly distributed demand, whose interval is determined from the uncertainty set.

Furthermore, for the $kb < c \leq (k+1)b$ case, since $(b\bar{D}_t + h\underline{D}_t)/(b+h)$ is non-negative and increasing in t , Lemma 7.2 provides a recursion for determining the robust ordering quantities:

$$z_t^* = \begin{cases} \frac{b\bar{D}_t + h\underline{D}_t}{b+h} - \sum_{j=1}^{t-1} z_j^*, & t=1, \dots, T-k, \\ 0, & t > T-k. \end{cases} \quad (7.4)$$

While this recursion is valid for all parameterizations $(\underline{S}_t, \bar{S}_t, \ell_t, u_t)_t$ of the uncertainty set Ω in Equation (7.1), we are more interested in non-recursive closed-form expressions for the ordering quantities z_t^* . In the next subsections, we discuss various parameterizations of the uncertainty set Ω , motivated by limit theorems of probability, that allow such closed-form expressions.

7.2.1 Central Limit Theorem

Consider the sum of T independent identically distributed (i.i.d.) random variables X_t , $t=1, \dots, T$, each with mean λ and standard deviation σ . In one of its simplest forms, the CLT states that

$$\lim_{T \rightarrow \infty} P \left(\frac{\sum_{t=1}^T X_t - T\lambda}{\sqrt{T}\sigma} \leq \theta \right) = \Phi(\theta), \forall \theta, \quad (7.5)$$

where Φ is the cumulative distribution function of a standard normal random variable.

While the CLT is typically presented for i.i.d. random variables, variants of the CLT can accommodate non-identically distributed random variables as well as correlated random variables. The Lindeberg CLT allows for independent *non-identically distributed* random variables; more details about this extension can be found in Feller (1968), Chapter X, Section 5. There also exist variants of the CLT that allow for correlation; for instance, see Billingsley (2012), Chapter 5, Section 27 for technical details.

We assume that the demand in period t , D_t , is a random variable with mean λ_t and standard deviation σ_t ; we assume that these statistics are known, but the distributions are not. We also assume that the demand covariance matrix Σ is known. The partial sum $\sum_{t=1}^k d_t$ has mean $\sum_{t=1}^k \lambda_t$ and standard deviation $\sqrt{\mathbf{e}_k' \Sigma_k \mathbf{e}_k}$, where Σ_k is the first k rows and columns of Ω and $\mathbf{e}_k$ is the k -dimensional vector of all ones. Motivated by the CLT, we let $\underline{S}_t = \sum_{j=1}^t \lambda_j - \Gamma_t \sqrt{\mathbf{e}_t' \Sigma_t \mathbf{e}_t}$, $\bar{S}_t = \sum_{j=1}^t \lambda_j + \Gamma_t \sqrt{\mathbf{e}_t' \Sigma_t \mathbf{e}_t}$, $\ell_t = \max\{\lambda_t - \Gamma_t \sigma_t, 0\}$ and $u_t = \lambda_t + \hat{\Gamma}_t \sigma_t$, so that our uncertainty set becomes

$$\Omega_{CLT} = \left\{ (D_1, \dots, D_T) : -\Gamma_t \leq \sum_{j=1}^t \frac{D_j - \lambda_j}{\sqrt{\mathbf{e}_t' \Sigma_t \mathbf{e}_t}} \leq \Gamma_t, \right. \\ \left. \max\{\lambda_t - \Gamma_t \sigma_t, 0\} \leq D_t \leq \lambda_t + \hat{\Gamma}_t \sigma_t, t = 1, \dots, T \right\}. \quad (7.6)$$

The constraints $D_t \in [\max\{\lambda_t - \hat{\Gamma}_t \sigma_t, 0\}, \lambda_t + \hat{\Gamma}_t \sigma_t]$ preclude extreme demand movements that are unlikely to appear in practice. Furthermore, these period-specific constraints can be viewed to hold with high probability, due to, for example, the Chebyshev inequality: $P(|D_t - \lambda_t| \geq \hat{\Gamma}_t \sigma_t) \leq 1 / \hat{\Gamma}_t^2$. The $-\Gamma_t \leq \sum_{j=1}^t (D_j - \lambda_j) / \sqrt{\mathbf{e}_t' \Sigma_t \mathbf{e}_t} \leq \Gamma_t$ constraints are motivated by the CLT. The $\Gamma_t \geq 0$ and $\hat{\Gamma}_t \geq 0$ are tunable parameters that allow adjustment of the conservatism of the robust optimization approach.

Practically speaking, the Γ_t parameters will be chosen as small constants, usually no more than 3; this is motivated by the fact that, if Z is a standard normal random variable, then $P(-3 \leq Z \leq 3) \approx 99.7\%$. The simplest parameterization is to set $\Gamma_t = \Gamma$, for all t , where $\Gamma \in [2, 3]$. Another parameterization that we explore in our chapter, which emphasizes the limit aspect of the CLT, is to set $\Gamma_t \rightarrow \infty$ for $t < T$ and $\Gamma_T \in [2, 3]$. The $\hat{\Gamma}_t$ parameters can be set using a qualitative description of the demand D_t in period t . For example, if demand in period t is assumed to be symmetric, we select $\hat{\Gamma}_t \leq \lambda_t / \sigma_t$, so that the constraint $\lambda_t - \hat{\Gamma}_t \sigma_t \leq D_t \leq \lambda_t + \hat{\Gamma}_t \sigma_t$ is symmetric around λ_t . In contrast, if the distribution of demand in period t is asymmetric, a value of $\hat{\Gamma}_t > \lambda_t / \sigma_t$ is required to have an asymmetric constraint $0 \leq D_t \leq \lambda_t + \hat{\Gamma}_t \sigma_t$; this implies that $\hat{\Gamma}_t \in (\lambda_t / \sigma_t, 3]$, which is feasible only when the coefficient of variation $\sigma_t / \lambda_t \geq 0.33$. Put differently, in this chapter we study asymmetric demand distributions where the lower bound of demand is zero. Such demand distributions arise in inventory management and are used to model the demand for “slow-moving” or “low-volume” products where demand is intermittent, such as in Chen and Yu (2005). For instance, many spare aircraft parts satisfy the bound on σ_t / λ_t and exhibit asymmetric demand patterns, earning descriptions of “lumpy” and “erratic” in the literature; for example, see Ghobbar and Friend (2003), and Williams (1984). In particular, Ghobbar and Friend (2003) study a variety of aircraft parts with coefficients of variation up to 1.28. Therefore, highly variable asymmetric demands warrant a value of $\hat{\Gamma}_t \in (\lambda_t / \sigma_t, 3]$ and the resulting asymmetric uncertainty set.

Finally, we point out that the Γ_t and $\hat{\Gamma}_t$ parameters control a robustness-conservatism trade-off in our robust optimization model. As Γ_t and $\hat{\Gamma}_t$ are increased, the model becomes more robust in that it covers more possible demand realizations in the uncertainty set. In exchange, the model is arguably more conservative, as it is accounting for a larger set of possible demand realizations. Ben-Tal et al. (2009, pp. 32–33) provide guidance on constructing robust uncertainty sets with respect to the support of the uncertain parameters, and argue that choosing an uncertainty set smaller than the support leads to good performance of a robust model.

7.2.1.1 Symmetric demand

Our first theorem, for symmetric demand ($\lambda_t - \hat{\Gamma}_t \sigma_t \geq 0, \forall t$) and whose proof may be found in Mamani et al. (2017), characterizes the robust order quantities in closed form. The subsequent discussion focuses on the $c \leq b$ case; if $c > b$, the ordering pattern is preserved, with the simple modification that ordering stops before period T , per Lemma 7.2. Furthermore, we set $\Gamma_t = \infty$ for $t < T$ to retain only a constraint on the full summation $\sum_{t=1}^T D_t$, in the asymptotic spirit of the CLT. If finite values of Γ_t , $t < T$, are required, the ordering quantities can still be determined using the recursion in Equation (7.4), though we are unable to derive closed form z_t^* for this case.

Theorem 7.1 *If $\lambda_t - \hat{\Gamma}_t \sigma_t \geq 0$ for all t and $c \leq b$, the robust order quantities are*

$$z_t^* = \begin{cases} \lambda_t + \hat{\Gamma}_t \sigma_t \left(\frac{b-h}{b+h} \right), & t \leq \kappa \\ \lambda_t + \left(\Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} + \sum_{j=\kappa+2}^T \hat{\Gamma}_j \sigma_j - \sum_{j=1}^{\kappa} \hat{\Gamma}_j \sigma_j \right) \left(\frac{b-h}{b+h} \right), & t = \kappa + 1 \\ \lambda_t - \hat{\Gamma}_t \sigma_t \left(\frac{b-h}{b+h} \right), & t > \kappa + 1, \end{cases} \quad (7.7)$$

where $\kappa = \max \left\{ t : \sum_{j=1}^t \Gamma_j \sigma_j \leq (\Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} + \sum_{j=1}^T \Gamma_j \sigma_j) / 2 \right\}$. If $kb < c \leq (k+1)b$ for some integer $1 \leq k \leq T-1$, the orders in Equation (7.7) are applied for $t = 1, \dots, T-k$ and $z_t^* = 0$ for $t = T-k+1, \dots, T$. If $Tb < c$, $z_t^* = 0$ for all t .

If $\Sigma = \sigma^2 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $\lambda_t = \lambda$ for all t , we obtain the case of i.i.d. demand with mean λ and standard deviation σ . For simplicity, we also let $\Gamma_T = \hat{\Gamma}_T = \Gamma$ for all t . The ordering quantities for i.i.d. demand are especially simple, and are presented in the following corollary.

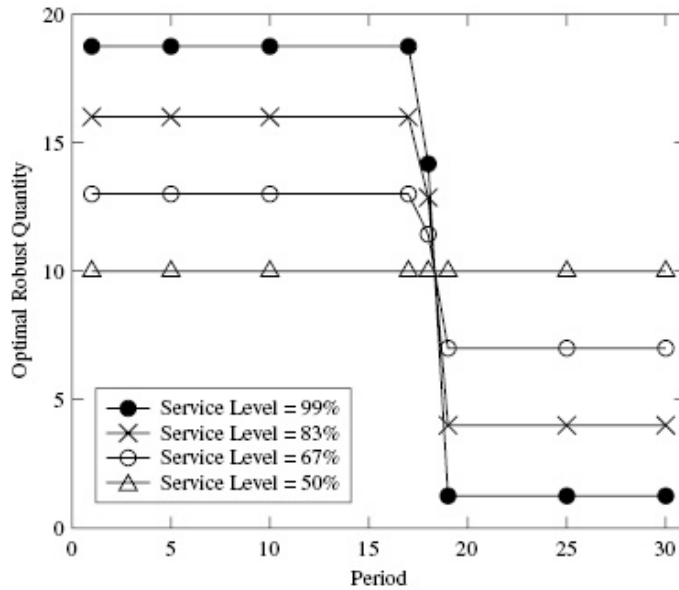
Corollary 7.1 (I.I.D. Demand) *If $\lambda - \Gamma \sigma \geq 0$ and $c \leq b$, the robust order quantities are*

$$z_t^* = \begin{cases} \lambda + \Gamma \sigma \left(\frac{b-h}{b+h} \right), & t \leq \lfloor \tau \rfloor \\ \lambda - \Gamma \sigma \left(\frac{b-h}{b+h} \right) (1-2\varepsilon), & t = \lfloor \tau \rfloor + 1 \\ \lambda - \Gamma \sigma \left(\frac{b-h}{b+h} \right), & t > \lfloor \tau \rfloor + 1, \end{cases} \quad (7.8)$$

where $\tau = (T + \sqrt{T})/2$ and $\varepsilon = \tau - \lfloor \tau \rfloor$. If $kb < c \leq (k+1)b$ for some integer $1 \leq k \leq T-1$, the orders in Equation (7.8) are applied for $t = 1, \dots, T-k$ and $z_t^* = 0$ for $t = T-k+1, \dots, T$. If $Tb < c$, $z_t^* = 0$ for all t .

The robust strategy in Corollary 7.1 orders in every period; this will not be true when $\lambda - \Gamma\sigma < 0$ (which we study in the next subsection). Note that if ordering too much is equivalent to ordering too little (i.e., $b = h$), then the robust strategy will always order the mean of demand λ , as in the classic newsvendor solution (when the median equals the mean). More generally, in Figure 7.1, we plot various ordering strategies and observe a clear pattern that depends on the relative size of b and h , conveniently represented as the service level $b/(b+h)$. For service levels that are at least 50%, the strategy first orders aggressively, until a threshold $\lfloor \tau \rfloor$, and then reduces the orders; the opposite behavior of ordering conservatively at first, and then aggressively, is observed for service levels below 50% (plots are omitted).

The threshold period induced by the value of τ , where the ordering behavior changes, can be explained as follows: When $t \leq \lfloor \tau \rfloor$, the order quantity is driven by the period-dependent constraints $D_t \in [\lambda - \Gamma\sigma, \lambda + \Gamma\sigma]$, whereas when $t > \lfloor \tau \rfloor + 1$ the order quantity is driven by the CLT constraints $-\Gamma \leq (\sum_{i=1}^T D_i - T\lambda) / (\sqrt{T}\sigma) \leq \Gamma$; when $t = \lfloor \tau \rfloor + 1$, both have an influence. Note that this threshold is independent of the values of λ , σ and Γ ; this is not true in the next subsection, where $\lambda - \Gamma\sigma < 0$ and the uncertainty set contains the zero vector.



Source: Mamani et al. (2017).

Figure 7.1 Illustration of Corollary 7.1 for $n = 30$, $\mu = 10$, $\sigma = 3$, $\Gamma = 3$ and various service levels $b/(b+h)$

7.2.1.2 Asymmetric demand

Our next theorem, for asymmetric demand ($\lambda_t - \hat{\Gamma}_t \sigma_t < 0, \forall t$) and whose proof may be found in Mamani et al. (2017), characterizes the robust order quantities in closed form. The subsequent discussion focuses on the $c \leq b$ case; if $c > b$, the ordering pattern is preserved, with the simple modification that ordering stops before period T , as per Lemma 7.2. Furthermore, we let $\Gamma_t \rightarrow \infty$ for $t < T$ to retain only a constraint on the full summation $\sum_{i=1}^T D_i$, in the asymptotic spirit of the CLT.

Theorem 7.2 *If $\lambda_t - \hat{\Gamma}_t \sigma_t < 0$ for all t and $c \leq b$, the robust order quantities are*

$$z_t^* = \begin{cases} \left(\frac{b}{b+h} \right) (\lambda_t + \hat{\Gamma}_t \sigma_t), & t \leq \kappa_1 \\ \left(\frac{b}{b+h} \right) \left(\sum_{j=1}^T \lambda_j + \Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} - \sum_{j=1}^{\kappa_1} (\lambda_j + \hat{\Gamma}_j \sigma_j) \right), & t = \kappa_1 + 1 \\ 0, & \kappa_1 + 1 < t \leq \kappa_2 \\ \left(\frac{h}{b+h} \right) \left(\sum_{j=1}^{\kappa_2+1} (\lambda_j + \hat{\Gamma}_j \sigma_j) - \Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} - \sum_{j=1}^T \hat{\Gamma}_j \sigma_j \right), & t = \kappa_2 + 1 \\ \left(\frac{h}{b+h} \right) (\lambda_t + \hat{\Gamma}_t \sigma_t), & t \geq \kappa_2 + 2, \end{cases} \quad (7.9)$$

where

$$\kappa_1 = \max \left\{ t : \sum_{j=1}^t (\lambda_j + \Gamma_j \sigma_j) \leq \sum_{j=1}^T \lambda_j + \Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} \right\}$$

and

$$\kappa_2 = \max \left\{ t : \sum_{j=1}^t (\lambda_j + \Gamma_j \sigma_j) \leq \Gamma_T \sqrt{\mathbf{e}' \Sigma \mathbf{e}} + \sum_{j=1}^T \Gamma_j \sigma_j \right\}.$$

If $kb < c \leq (k+1)b$ for some integer $1 \leq k \leq T-1$, the orders in Equation (7.9) are applied for $t = 1, \dots, T-k$ and $z_t^* = 0$ for $t = T-k+1, \dots, T$. If $Tb < c$, $z_t^* = 0$ for all t .

We first explain the interval of zero ordering, which was not present in the previous subsection for a symmetric uncertainty set. For the asymmetric uncertainty set, demand can be zero, and the worst-case cumulative demands $\underline{D}_t$ and $\underline{D}_t$ are not necessarily strictly increasing in t . Indeed, the periods with zero ordering, $t \in (\kappa_1 + 1, \kappa_2]$, correspond exactly to the periods where $\underline{D}_t$ and $\underline{D}_t$ are both constant. Therefore, matching the worst-case costs, or equivalently $\sum_{j=1}^t z_j^* = (bD_t + h\underline{D}_t) / (b+h)$ results in zero ordering.

Next, as in the previous subsection, we now focus on the case of i.i.d. demand by setting $\Sigma = \sigma \mathbf{I}$ and $\lambda_t = \lambda$ for all t ; we also let $\Gamma_T = \hat{\Gamma}_t = \Gamma$ for all t . The optimal ordering quantities are presented in the following corollary.

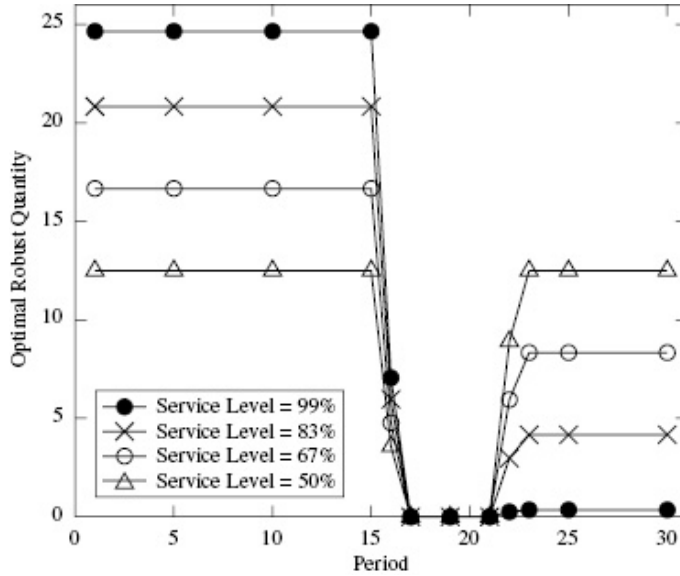
Corollary 7.2 (I.I.D. Demand) *If $\lambda - \Gamma\sigma < 0$ and $c \leq b$, the robust order quantities are*

$$z_t^* = \begin{cases} \left(\frac{b}{b+h}\right)(\lambda + \Gamma\sigma), & t \leq \lfloor \tau_1 \rfloor \\ \varepsilon \left(\frac{b}{b+h}\right)(\lambda + \Gamma\sigma), & t = \lfloor \tau_1 \rfloor + 1 \\ 0, & \lfloor \tau_1 \rfloor + 1 < t \leq \lfloor \tau_2 \rfloor \\ (1-\varepsilon) \left(\frac{h}{b+h}\right)(\lambda + \Gamma\sigma), & t = \lfloor \tau_2 \rfloor + 1 \\ \left(\frac{h}{b+h}\right)(\lambda + \Gamma\sigma), & t > \lfloor \tau_2 \rfloor + 1, \end{cases} \quad (7.10)$$

where $\tau_1 = (T\lambda + \sqrt{T}\Gamma\sigma) / (\lambda + \Gamma\sigma)$, $\tau_2 = (T + \sqrt{T})\Gamma\sigma / (\lambda + \Gamma\sigma)$, $\varepsilon = \tau_1 - \lfloor \tau_1 \rfloor$, and $\varepsilon = \tau_2 - \lfloor \tau_2 \rfloor$. If $kb < c \leq (k+1)b$ for some integer $1 \leq k \leq T-1$, the orders in Equation (7.10) are applied for $t = 1, \dots, T-k$ and $z_t^* = 0$ for $t = T-k+1, \dots, T$. If $Tb < c$, $z_t^* = 0$ for all t .

The different structure of the uncertainty set for the $\lambda - \Gamma\sigma < 0$ case results in qualitatively different ordering behavior; in particular, there is a range of periods for which no ordering takes place ($z_t^* = 0$). There are now two thresholds, τ_1 and τ_2 , that dictate this lack of ordering, rather than the single threshold τ under the case where $\lambda - \Gamma\sigma \geq 0$. It can be easily shown using basic algebra that $\tau_1 < \tau < \tau_2$. The drivers of the thresholds τ_1 and τ_2 , where the ordering behavior changes, can be explained as follows: When $t \leq \lfloor \tau_1 \rfloor$, the order quantity is driven by the period-dependent constraints $D_t \in [0, \lambda + \Gamma\sigma]$, whereas when $t > \lfloor \tau_2 \rfloor + 1$ the order quantity is driven by the CLT constraints $-\Gamma \leq (\sum_{i=1}^T D_i - T\lambda) / (\sqrt{T}\sigma) \leq \Gamma$. When $\lfloor \tau_1 \rfloor + 1 \leq t \leq \lfloor \tau_2 \rfloor + 1$, both constraints have an influence. Unlike threshold τ in the previous section, these thresholds depend on λ , σ , and Γ . Basic calculus shows that $\frac{\partial \tau_1}{\partial \lambda} > 0$ and $\frac{\partial \tau_2}{\partial \lambda} < 0$, which implies that, as the mean of demand increases, the range of intervals with zero ordering is reduced. Intuitively, this ordering strategy approaches that of the previous subsection. Conversely, we can show that $\frac{\partial \tau_1}{\partial \sigma} < 0$ and $\frac{\partial \tau_2}{\partial \sigma} > 0$, which implies that as the standard deviation of demand increases, the range of intervals with zero ordering is enlarged. Note that it is possible to have $\tau_2 > T$ and the last regime of ordering is never reached; this occurs if $T < (\Gamma\sigma / \lambda)^2$. In this scenario, ordering only takes place in early periods, which then ceases after period $\lfloor \tau_1 \rfloor + 1$. The ordering behavior when $b = h$ also differs from the $\lambda - \Gamma\sigma \geq 0$ case; apart from not ordering in the middle of the horizon, the robust strategy will order a constant amount equal to half the maximum demand $(\lambda + \Gamma\sigma) / 2$; note that this is in contrast to a newsvendor solution. However, this solution makes intuitive sense: since the costs of over and under ordering are the same, the robust strategy orders the midpoint of the range of demand $[0, \lambda + \Gamma\sigma]$.

In Figure 7.2, we plot various ordering strategies for service levels $b / (b+h) \geq 50\%$. We observe some similar patterns to that in Figure 7.1: The strategy first orders aggressively, then



Source: Mamani et al. (2017).

Figure 7.2 Illustration of Corollary 7.2 for $n = 30$, $\mu = 10$, $\sigma = 5$, $\Gamma = 3$ and various service levels $b/(b + h)$

zero ordering, finally followed by conservative ordering; the opposite behavior is observed for service levels below 50% (plots are omitted).

7.2.2 Strong Law of Large Numbers

The SLLN is another powerful probabilistic limit theorem about the average of random variables. For i.i.d. random variables $X_1, X_2, \dots, X_T$ with mean λ , the SLLN states that the average converges to the mean

$$\lim_{T \rightarrow \infty} \frac{\sum_{t=1}^T X_t}{T} = \lambda,$$

almost surely. This limit theorem motivates the following uncertainty set,

$$\Omega_{SLLN} = \left\{ (D_1, \dots, D_T) : \lambda - \varepsilon \leq \frac{\sum_{t=1}^T D_t}{T} \leq \lambda + \varepsilon, \right. \\ \left. \lambda - \delta \leq D_t \leq \lambda + \delta, t = 1, \dots, T \right\}, \quad (7.11)$$

where δ and ε are tunable parameters such that $0 \leq \varepsilon, \delta \leq \lambda$. Alternatively, we could omit the $\delta \leq \lambda$ requirement and use a lower bound $\max\{\lambda - \delta, 0\}$ for D_t , for all t , possibly allowing δ to depend on t , as in the CLT case, which would lead to new order quantities, whose derivation we leave to the reader. Intuitively, the ε parameter reflects the fact that the SLLN has not converged for a finite T , and the Ω parameter allows individual demands to fluctuate away from the mean λ . See Wagner (2018) for advice on setting these parameters.

For the case where $\delta \leq \lambda$, letting $\xi = T(\varepsilon + \delta) / 2\delta$, the following worst-case cumulative demands for Ω_{SLLN} are derived by Sinha et al. (2021), which are discrete-time analogues of continuous-time results in Wagner (2018).

Lemma 7.3

$$\underline{D}_t^{SLLN} = \begin{cases} t(\lambda - \delta), & t \leq \lfloor \xi \rfloor, \\ t\lambda - T\varepsilon - (T - t)\delta, & t > \lfloor \xi \rfloor, \end{cases}$$

$$\overline{D}_t^{SLLN} = \begin{cases} t(\lambda + \delta), & t \leq \lfloor \xi \rfloor, \\ t\lambda + T\varepsilon + (T - t)\delta, & t > \lfloor \xi \rfloor. \end{cases}$$

Applying the recursion in Equation (7.4), we obtain the following closed-form solutions for the uncertainty set Ω_{SLLN} .

Theorem 7.3

$$z_t^* = \begin{cases} \lambda + \delta \left(\frac{b - h}{b + h} \right), & t \leq \lfloor \xi \rfloor, \\ \lambda - \delta \left(\frac{b - h}{b + h} \right), & t > \lfloor \xi \rfloor, \end{cases}$$

It is interesting to note the similarity of Theorem 7.3 with Corollary 7.1. Indeed, if we set $\delta = \Gamma\sigma$ and $\varepsilon = \delta/\sqrt{T}$, the ordering strategies are (almost) identical. However, this choice of ε suggests a $O(1/\sqrt{T})$ convergence rate, whereas the SLLN has a faster rate of convergence $O(1/\sqrt{T \log \log T})$, as per the LIL, which suggests an alternative choice of ε that is a function of $1/\sqrt{T \log \log T}$. We discuss an uncertainty set motivated by the LIL in the next section.

7.2.3 Law of the Iterated Logarithm

The CLT and SLLN are obtained by scaling the sum of random variables by $\sqrt{T}$ and T , respectively. Scaling instead by $\phi(T) \triangleq \sqrt{T \log \log T}$ provides the LIL. In particular, for i.i.d. random variables X_t with mean λ and standard deviation Ω , the LIL consists of the following two results

$$\limsup_{T \rightarrow \infty} \frac{\sum_{t=1}^T X_t - T\lambda}{\sigma\sqrt{2\phi(T)}} = 1 \quad \text{and} \quad \liminf_{T \rightarrow \infty} \frac{\sum_{t=1}^T X_t - T\lambda}{\sigma\sqrt{2\phi(T)}} = -1,$$

almost surely.

A LIL-motivated uncertainty set was first used by Bertsimas et al. (2011), in the context of queueing systems, and we use the same idea to define

$$\Omega_{LIL} = \left\{ (D_1, \dots, D_T) : - (1 + \varepsilon) \leq \frac{\sum_{t=1}^T D_t - T\lambda}{\sigma\sqrt{2}\phi(T)} \leq (1 + \varepsilon), \right. \\ \left. \lambda - \delta \leq D_t \leq \lambda + \delta, t = 1, \dots, T \right\}, \quad (7.12)$$

where $\varepsilon \geq 0$ and $0 \leq \delta \leq \lambda$ are adjustable parameters. As in the SLLN case, we could omit the $\delta \leq \lambda$ restriction, and instead use a lower bound $\max\{\lambda - \delta, 0\}$ for all D_t , possibly allowing δ to depend on t , which would lead to new order quantities, whose derivation we leave to the reader. For the case where $\delta \leq \lambda$, letting $\zeta = \lfloor \frac{T\delta + (1 + \varepsilon)\sigma\sqrt{2}\phi(T)}{2} \rfloor + 1$, the following worst-case cumulative demands are derived in Sinha et al. (2021).

Lemma 7.4

$$\underline{D}_t^{LIL} = \begin{cases} t(\lambda - \delta), & t < \zeta, \\ t\lambda - (1 + \varepsilon)\sigma\sqrt{2}\phi(T) - (T - t)\delta, & t \geq \zeta, \end{cases}$$

$$\overline{D}_t^{LIL} = \begin{cases} t(\lambda + \delta), & t < \zeta, \\ t\lambda + (1 + \varepsilon)\sigma\sqrt{2}\phi(T) + (T - t)\delta, & t \geq \zeta. \end{cases}$$

Applying the recursion in Equation (7.4), we obtain the following closed-form solutions for the uncertainty set Ω_{LIL} .

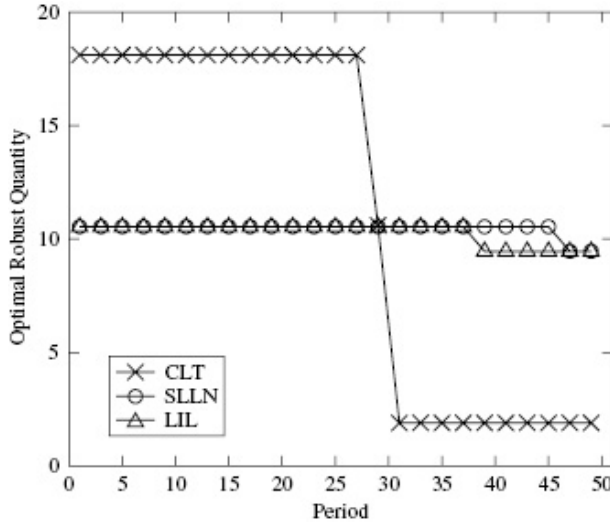
Theorem 7.4

$$z_t^* = \begin{cases} \lambda + \delta \left(\frac{b - h}{b + h} \right), & t < \zeta, \\ \lambda - \delta \left(\frac{b - h}{b + h} \right), & t \geq \zeta. \end{cases}$$

Superficially, these order quantities appear identical to that of Theorem 7.3. The ε and δ parameters indeed play similar roles in both approaches. However, the thresholds for switching between the two distinct order quantities are different, since $\zeta \neq \xi$.

7.2.4 Choice of Uncertainty Set

Which uncertainty set should be used? *A priori*, it is unclear which is best. The set Ω_{CLT} is perhaps the easiest to apply, since the parameters Γ_t and $\hat{\Gamma}_t$ are interpretable as numbers of



Note: For Theorem 7.3, we use $\varepsilon = 0.5$ and $\delta = 0.6$. For Theorem 7.4, we use $\varepsilon = 0.05$ and $\delta = 0.6$.

Figure 7.3 Comparison of the ordering strategies from Corollary 7.1 and Theorems 7.3–7.4 for $T = 50$, $\mu = 10$, $\sigma = 3$, $\Gamma = 3$ and a service level of $b/(b + h) = 95\%$

standard deviations. Of course, there is no guarantee that Ω_{CLT} will work best; it is possible that either Ω_{SLLN} or Ω_{LIL} would result in better performance, albeit with a more difficult parameter interpretation related to probabilistic convergence rates. However, since closed-form solutions exist for all three uncertainty sets, it is fairly straightforward to test the three approaches simultaneously in numerical experiments. Furthermore, the various parameters associated with the uncertainty sets could be determined using a cross-validation type of technique. We also point out that these three sets are not necessarily the only ones that could lead to closed-form solutions. Indeed, any set that is of the form of Equation (7.1) can, via Lemmas 7.1–7.2 and the recursion in Equation (7.4), lead to closed-form solutions. Perhaps even other uncertainty sets, not of the form of Equation (7.1), could lead to similar solutions. We conclude this section with Figure 7.3 that presents a comparison of the order quantities from Corollary 7.1 (CLT uncertainty set) and Theorems 7.3–7.4 (SLLN and LIL uncertainty sets, respectively).

7.3 DYNAMIC ROBUST INVENTORY MANAGEMENT

In this section, we analyze a dynamic robust model based on a rolling horizon implementation. In particular, at the beginning of period k , past demands and order quantities are known, as is the current inventory position, and this information is used to create a new problem instance for the remaining horizon, namely periods $t = k, \dots, T$. The past realized demands and order quantities are denoted by $(\hat{D}_1, \dots, \hat{D}_{k-1})$ and $(z_1^*, \dots, z_{k-1}^*)$, respectively, and the inventory position at the end of period $k-1$ is denoted by $\hat{y}_{k-1} = \sum_{t=1}^{k-1} (z_t^* - \hat{D}_t)$. Our presentation is based on the generic CLT uncertainty set in Equation (7.6) and mirrors the discussion in Mamani et al. (2017), though the SLLN and LIL uncertainty sets in Equations (7.11) and (7.12), respectively, could be used instead.

We utilize a new uncertainty set Ω_{CLT}^k in the formulation of the problem for the remaining horizon, which is defined as the intersection of the original uncertainty set Ω_{CLT} , given in Equation (7.6), with the realized demand values, or

$$\begin{aligned}\Omega_{CLT}^k &= \Omega_{CLT} \cap \{D_1 = \hat{D}_1, \dots, D_{k-1} = \hat{D}_{k-1}\} \\ &= \{(D_k, \dots, D_T) : -\Gamma_t \leq \frac{\sum_{j=1}^{k-1} \hat{D}_j + \sum_{j=k}^t D_j - \sum_{j=1}^t \lambda_j}{\sqrt{\mathbf{e}_t S_t \mathbf{e}_t}} \leq \Gamma_t, \\ &\quad \max\{\lambda_t - \hat{\Gamma}_t \sigma_t, 0\} \leq D_t \leq \lambda_t + \hat{\Gamma}_t \sigma_t, t = k, \dots, T\}.\end{aligned}$$

We assume that Ω_{CLT}^k is not empty, though Wagner (2018) discusses various projection techniques to resolve an empty set.

We next provide a *sequence* of optimization models, indexed by k , that collectively define a dynamic version of the model in Equation (7.2). These models are used to find the robust order quantity in period k , z_k^* , for $k = 1, \dots, T$, which is a function of the inventory position $\hat{y}_{k-1}$. In period $k = 1, \dots, T$, we leverage the known information by solving the following model

$$\begin{aligned}\min_{(z_k, \dots, z_T) \geq 0} \quad & \sum_{t=k}^T (cz_t + m_t) \\ \text{s.t.} \quad & y_t = y_{k-1} + \sum_{j=k}^t (z_j - D_j), t = k, \dots, T \\ & m_t \geq hy_t, t = k, \dots, T, \forall (D_k, \dots, D_T) \in \Omega_{CLT}^k \\ & m_t \geq -by_t, t = k, \dots, T, \forall (D_k, \dots, D_T) \in \Omega_{CLT}^k,\end{aligned}\tag{7.13}$$

where y_t is the inventory position in period $t \geq k$. While the model in Equation (7.13) solves for $(z_k^*, \dots, z_T^*)$, only the solution z_k^* will be implemented in period k , as the final determination of z_t^* , for $t > k$, will depend on the realized $\hat{y}_{t-1}$, and will be provided by the solution of the model indexed by t .

We next update the definitions of $\underline{D}_t$ and $\bar{D}_t$ to accommodate the known values of past demands:

$$\underline{D}_t = \min_{(D_k, \dots, D_T) \in \Omega_{CLT}^k} \sum_{j=k}^t D_j, t = k, \dots, T$$

and

$$\bar{D}_t = \max_{(D_k, \dots, D_T) \in \Omega_{CLT}^k} \sum_{j=k}^t D_j, t = k, \dots, T.\tag{7.14}$$

The intermediate problem defined in Equation (7.13) typically has a non-zero initial inventory position $\hat{y}_{k-1}$, which requires a generalization of Lemma 7.2. We present a generic version of the full problem, defined in Equation (7.2), which can be easily tailored to the subproblems indexed by k ; a formal proof can be found in Mamani et al. (2017). We first present some conditions that facilitate the presentation of the generalized lemma.

Condition 1: $\ell b < c \leq (\ell + 1)b$ for some integer $0 \leq \ell \leq T - 1$.

Condition 2: $\frac{b\bar{D}_m + h\underline{D}_m}{b+h} < y_0 \leq \frac{b\bar{D}_{m+1} + h\underline{D}_{m+1}}{b+h}$ for some integer $0 \leq m \leq T - 1$, where $\underline{D}_0 = \bar{D}_0 = 0$.

Condition 3: $m < T - \ell$.

These conditions have the following interpretations: If we begin with a positive inventory, Condition 1 determines when ordering should *stop*, by comparing the unit ordering cost c and unit backordering cost b ; Condition 2 determines when ordering should *start*, by evaluating how long the initial inventory will last to satisfy demand; Condition 3 makes sure that the start of ordering occurs before the stopping of it. If we instead begin with a backlog, ordering starts immediately, and stops according to Condition 1. These interpretations are presented mathematically in the next lemma.

Lemma 7.5 *The robust order quantities satisfy the following.*

- If $y_0 > 0$ and Conditions 1–3 hold:
 - $z_t^* = 0, t = 1, \dots, m$,
 - $\sum_{j=m+1}^t z_j^* = \frac{b\bar{D}_t + h\underline{D}_t}{b+h} - y_0, t = m+1, \dots, T - \ell$,
 - $z_t^* = 0, t = T - \ell + 1, \dots, T$.
- If $y_0 < 0$ and Condition 1 holds:
 - $\sum_{j=1}^t z_j^* = \frac{b\bar{D}_t + h\underline{D}_t}{b+h} - y_0, t = 1, \dots, T - \ell$,
 - $z_t^* = 0, t = T - \ell + 1, \dots, T$.

Otherwise, $z_t^* = 0$ for all t .

In period k , our rolling horizon model only needs to solve for z_k^* , which implies that only the minimum and maximum demands in the *first* period are needed. These are $\bar{D}_t$ and $\underline{D}_t$, from Equation (7.14), evaluated at $t = k$:

$$\underline{D}_k = \min_{(D_k, \dots, D_T) \in \Omega_{CLT}^k} D_k \quad \text{and} \quad \bar{D}_k = \max_{(D_k, \dots, D_T) \in \Omega_{CLT}^k} D_k. \quad (7.15)$$

The following theorem characterizes the optimal order quantities for the rolling horizon model, whose proof can be found in Mamani et al. (2017).

Theorem 7.5 If $c \leq b(T - k + 1)$, the robust order quantities in period $k = 1, \dots, T$ are

$$z_k^* = \max \left\{ \frac{b\bar{D}_k + h\underline{D}_k}{b + h} - y_{k-1}, 0 \right\}, \quad (7.16)$$

where

$$\begin{aligned} \bar{D}_k = \min & \left\{ \lambda_k + \Gamma_k \sigma_k, \min_{t=k, \dots, T} \left\{ \sum_{j=1}^t \lambda_j + \Gamma_t \sqrt{\mathbf{e}'_t \Sigma_t \mathbf{e}_t} - \sum_{j=1}^{k-1} D_j \right. \right. \\ & \left. \left. - \sum_{j=k+1}^t \max \left\{ \lambda_j - \Gamma_j \sigma_j, 0, \lambda_j - \Gamma_j \sqrt{\mathbf{e}'_j \Sigma_j \mathbf{e}_j} - \Gamma_{j-1} \sqrt{\mathbf{e}'_{j-1} \Sigma_{j-1} \mathbf{e}_{j-1}} \right\} \right\} \right\} \end{aligned}$$

and

$$\begin{aligned} \underline{D}_k = \max & \left\{ \max \{ \lambda_k - \Gamma_k \sigma_k, 0 \}, \max_{t=k, \dots, T} \left\{ \sum_{j=1}^t \lambda_j - \Gamma_t \sqrt{\mathbf{e}'_t \Sigma_t \mathbf{e}_t} - \sum_{j=1}^{k-1} D_j \right. \right. \\ & \left. \left. - \sum_{j=k+1}^t \min \left\{ \lambda_j + \Gamma_j \sigma_j, \lambda_j + \Gamma_j \sqrt{\mathbf{e}'_j \Sigma_j \mathbf{e}_j} + \Gamma_{j-1} \sqrt{\mathbf{e}'_{j-1} \Sigma_{j-1} \mathbf{e}_{j-1}} \right\} \right\} \right\}. \end{aligned}$$

If $c > b(T - k + 1)$, $z_k^* = 0$ for $k = 1, \dots, T$.

The order quantities in Theorem 7.5 are state-dependent base-stock policies. The base stock for period $k = 1, \dots, T$ is $\frac{b\bar{D}_k + h\underline{D}_k}{b + h}$, where the dependence on state is from the definitions of $\bar{D}_k$ and $\underline{D}_k$, which depend on the past realized demands $\hat{D}_t$ for $t < k$.

We point out that Mamani et al. (2017) demonstrate that these closed-form solutions outperform the more complicated robust models of Bertsimas et al. (2010) and Bertsimas and Thiele (2006), that do not provide closed-form solutions, when demand is not i.i.d. and has serial correlation. Furthermore, since Bertsimas and Thiele (2006) show that their model outperforms a dynamic program with a misspecified distribution, by transitivity, our model also outperforms such a dynamic program. Finally, Solyalı et al. (2016) also pursue a rolling horizon implementation, which outperforms Ben-Tal et al. (2004), Bertsimas and Thiele (2006), See and Sim (2010), and others.

7.4 CONCLUSION

There are many possible extensions to the results presented in this chapter. For instance, there could exist capacities on orders and inventories, which are treated by both Mamani et al. (2017) and Sinha et al. (2021). It is also possible to incorporate revenues, and maximize profit, instead

of minimizing cost. This extension is the primary focus of Sinha et al. (2021), and the analysis is more challenging since the objective of the robust model can be bimodal. Sinha et al. (2021) also consider period-dependent economic parameters, which adds another challenge, since the natural extension $(b_t \bar{D}_t + h_t \underline{D}_t) / (b_t + h_t)$ is no longer necessarily increasing in t , which causes complications for the recursion in Equation (7.4). One could also consider a network of inventory locations, which is analyzed by Bertsimas and Thiele (2006) using similar techniques.

The robust models in this chapter utilize a set membership description of uncertainty. It is possible to use a different variant of robust optimization to analyze these inventory management problems, namely *distributionally* robust optimization. In this approach, the uncertain parameters (e.g., demand) are realizations from some *unknown* stochastic distribution. Some information about the distribution is known, such as the mean and standard deviation, and an uncertainty set for the *distribution*, that is consistent with the known information, is utilized in the robust model. The newsvendor model has been successfully analyzed using this approach, starting with the seminal work by Scarf (1958) and continued by, for example, Lei and Wagner (2021), Natarajan et al. (2018), and Perakis and Roels (2008); other examples of this style of robust optimization research can be found in the comprehensive literature review of Natarajan et al. (2018).

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of a Neal and Jan Dempsey Faculty Fellowship.

REFERENCES

- Ardestani-Jaafari, A., & Delage, E. (2016). Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research*, 64(2), 474–494.
- Bandi, C., & Bertsimas, D. (2012). Tractable stochastic analysis in high dimensions via robust optimization. *Mathematical Programming*, 134(1), 23–70.
- Bandi, C., & Bertsimas, D. (2014a). Optimal design for multi-item auctions: A robust optimization approach. *Mathematics of Operations Research*, 39(4), 1012–1038.
- Bandi, C., & Bertsimas, D. (2014b). Robust option pricing. *European Journal of Operational Research*, 239(3), 842–853.
- Bandi, C., Bertsimas, D., & Youssef, N. (2015). Robust queueing theory. *Operations Research*, 63(3), 676–700.
- Bandi, C., Bertsimas, D., & Youssef, N. (2018). Robust transient analysis of multi-server queueing systems and feed-forward networks. *Queueing Systems*, 89(3–4), 351–413.
- Ben-Tal, A., Ghaoui, L. E., & Nemirovski, A. (2009). *Robust optimization*. Princeton University Press.
- Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2), 351–376.
- Bertsimas, D., Gamarnik, D., & Rikun, A. (2011). Performance analysis of queueing networks via robust optimization. *Operations Research*, 59(2), 455–466.
- Bertsimas, D., Iancu, D., & Parrilo, P. (2010). Optimality of affine policies in multi-stage robust optimization. *Mathematics of Operations Research*, 35(2), 363–394.
- Bertsimas, D., & Sim, M. (2004). Price of robustness. *Operations Research*, 52(1), 35–53.
- Bertsimas, D., & Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, 54(1), 150–168.
- Bienstock, D., & Özbay, N. (2008). Computing robust basestock levels. *Discrete Optimization*, 5(2), 389–414.

- Billingsley, P. (2012). *Probability and measure* (Anniversary edition ed.). Wiley.
- Chen, F., & Yu, B. (2005). Quantifying the value of leadtime information in a single-location inventory system. *Manufacturing & Service Operations Management*, 7(2), 144–151.
- Chen, S., Lee, H., & Moynadeh, K. (2016). Supply chain coordination with multiple shipments: The optimal inventory subsidizing contracts. *Operations Research*, 64(6), 1320–1337.
- Feller, W. (1968). *An introduction to probability theory and its applications, volume 1* (3rd ed.). Wiley.
- Ghobbar, A., & Friend, C. (2003). Evaluation of forecasting methods for intermittent parts demand in the field of aviation: A predictive model. *Computers and Operations Research*, 30(14), 2097–2114.
- Gorissen, B., & Hertog, D. (2013). Robust counterparts of inequalities containing sums of maxima of linear functions. *European Journal of Operational Research*, 227(1), 30–43.
- Iancu, D., Sharma, M., & Sviridenko, M. (2013). Supermodularity and affine policies in dynamic robust optimization. *Operations Research*, 61(4), 941–956.
- Lei, J., & Wagner, M. (2021). Data-driven distributionally robust newsvendor models with censored demand. Submitted.
- Mamani, H., Nassiri, S., & Wagner, M. (2017). Closed-form solutions for robust inventory management. *Management Science*, 63(5), 1625–1643.
- Natarajan, K., Sim, M., & Uichanco, J. (2018). Asymmetry and ambiguity in newsvendor models. *Management Science*, 64(7), 3146–3167.
- Ozer, O., & Wei, W. (2006). Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management Science*, 52(8), 1238–1257.
- Perakis, G., & Roels, G. (2008). Regret in the newsvendor model with partial information. *Operations Research*, 56(1), 188–203.
- Scarf, H. (1958). A min-max solution of an inventory problem. In K. Arrow, S. Karlin, & H. Scarf (Eds.), *Studies in the mathematical theory of inventory and production* (pp. 201–209). Stanford, CA: Stanford University Press.
- See, C., & Sim, M. (2010). Robust approximation to multi-period inventory management. *Operations Research*, 58(3), 583–594.
- Sinha, S., Wagner, M., & Ghate, A. (2021). A robust multi-period newsvendor model with inventory balance constraints. Submitted.
- Solyali, O., Cordeau, J., & Laporte, G. (2016). The impact of modeling on robust inventory management under demand uncertainty. *Management Science*, 62(4), 1188–1201.
- Tang, C., Rajaram, K., Alptekinoglu, A., & Ou, J. (2004). The benefits of advance booking discount programs: Model and analysis. *Management Science*, 50(4), 465–478.
- Wagner, M. (2010). Fully distribution-free profit maximization: The inventory management case. *Mathematics of Operations Research*, 35(4), 728–741.
- Wagner, M. (2011). Online lot-sizing problems with ordering, holding and shortage costs. *Operations Research Letters*, 39(2), 144–149.
- Wagner, M. (2018). Robust inventory management: An optimal control approach. *Operations Research*, 66(2), 426–447.
- Williams, T. (1984). Stock control with sporadic and slow-moving demand. *The Journal of the Operational Research Society*, 35(10), 939–948.

8. Dual-sourcing, dual-mode dynamic stochastic inventory models

Linwei Xin and Jan A. Van Mieghem

8.1 INTRODUCTION

The dual-sourcing (or dual-mode) dynamic stochastic inventory model studies the replenishment of inventory from two sources (or two transportation modes) in the presence of demand uncertainty. The classic assumption, also adopted here, is that one source (mode) is cheap, whereas the other is fast. The former then is called the regular source, whereas the latter is the express source. Dell, for example, sources the majority of its computers sold in the United States from Asia and supplements the supply as needed with nearby sourcing from Mexico. The Asian source has a longer lead time, but lower unit replenishment cost than the Mexican source. In practice, the objective of dual sourcing is to offer high availability at low cost in the presence of uncertainty, by judiciously ordering from the two sources. This objective is modeled by minimizing the total expected holding, backlogging, and procurement costs. The earliest study of such a dual-sourcing model dates back to several papers in the early 1960s, and it continues to attract academic attention because the optimal strategy largely remains unknown, except for the special case in which the lead-time difference between the two sources equals one review period. The reason is that with non-consecutive lead times, the dynamic program governing the optimal replenishment policy is multi-dimensional.

The dual-mode dynamic stochastic inventory model studies the replenishment of inventory using two transportation modes. Whether involving one or two sources, the dual-mode model is mathematically equivalent to the dual-sourcing model. The dual-mode model is relevant to the many organizations that, occasionally, expedite replenishment orders.

A vast literature studies multi-sourcing inventory models conducted over the past six decades, and we refer the interested reader to an excellent review (Minner (2003)) as well as a recent update (Svoboda et al., 2021). Our review is limited to dynamic inventory replenishment with stochastic demands, and we cover both discrete- and continuous-time models. Perhaps one key difference between our review paper and others is that our review is more technically detailed and provides self-contained proofs of several fundamental results in the dual-sourcing literature. We also highlight theory advanced recently by using asymptotic analysis.

8.2 DISCRETE-TIME MODEL

Consider a discrete-time, discrete-review inventory model in which, for any time period t , stochastic customer demand D_t is satisfied from finished-goods inventory (FGI), and excess demand is backlogged. Let I_t denote the net inventory (i.e., on-hand inventory minus back-orders) at the beginning of each period before demand realization. When I_t is positive, it is

the on-hand inventory level. When it is negative, its absolute value is the number of back-orders. In each period $t \in \mathbb{N}$, the firm can replenish its FGI from two sources or using two modes: (a) a fast source (mode) E , called express, emergency, or expedited, with deterministic lead time $L^E \in \mathbb{N}$ and per-unit replenishment cost $c^E > 0$; (b) a slow source (mode) R , called regular, with deterministic lead time $L^R \in \mathbb{N}$, where $L^R > L^E$ and per-unit replenishment cost $0 < c^R < c^E$. We assume negligible fixed order costs. For convenience, denote the following:

$$c = c^E - c^R > 0 \text{ and } L = L^R - L^E > 0.$$

Let h and b denote the per-period per-unit holding and backlog cost. The assumption that $c < bL$ is standard; otherwise, the marginal cost of procuring a unit from E instead of R exceeds the maximal benefit of avoiding a unit of backlog over L periods and single sourcing from the regular supplier is optimal.

Period t starts by placing the express and regular replenishment orders, denoted by $q_t^E \geq 0$ and $q_t^R \geq 0$, respectively. In the standard dual-sourcing model, orders are non-negative yet unrestricted or uncapacitated. Then, orders $q_{t-L^E}^E$ and $q_{t-L^R}^R$ are received and added to the finished-goods inventory. Finally, demand D_t is observed and satisfied, net inventory I_t is updated via

$$I_{t+1} = I_t + q_{t-L^E}^E + q_{t-L^R}^R - D_t,$$

and costs are incurred. For the latter without loss of generality, it is convenient to adopt the standard “accounting trick” in the inventory literature (e.g., Zipkin (2000)) where we charge in period t for orders $q_{t-L^R}^R, q_{t-L^E}^E$ placed in periods $t-L^R$ and $t-L^E$, respectively. Let C_t be the sum of the holding and backorder costs incurred in time period t , plus the ordering cost incurred for orders placed earlier:

$$C_t \triangleq c^R q_{t-L^R}^R + c^E q_{t-L^E}^E + G(I_t),$$

where $[x]^+ = \max[x, 0]$, $[x]^- = \max[-x, 0]$, and $G(x) \triangleq h[x]^+ + b[x]^-$. Note that $G(x)$ is a convex function.

We now formalize the family of admissible policies Π that determine how new orders are placed. An admissible policy π consists of a sequence of deterministic measurable functions $\{f_t^\pi, t \geq 1\}$, where each f_t^π is with domain $\mathbb{R}^{L^R+L^E+1}$ and range $\mathbb{R}_+^2$. Order $i \in \{E, R\}$ then is placed using the information encoded in the state of the system at time t :

$$q_t^i = f_{i,t}^\pi(q_{t-L^R}^R, \dots, q_{t-1}^R, q_{t-L^E}^E, \dots, q_{t-1}^E, I_t).$$

Let Π denote the family of all such admissible policies π .

To denote the dependence of the cost on the policy π , we use the notation C_t^π . Let $C(\pi)$ denote the long-run average cost incurred by policy π :

$$C(\pi) \triangleq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}[C_t^\pi].$$

The dual-sourcing problem is to find a policy π that achieves the lowest long-run average cost. Note that although finite-horizon optimality is important in practice, most existing dual-sourcing models focus on minimizing long-run average cost. We make the following two reductions to simplify our analysis through the rest of the section:

1. $c^R = 0$ such that the per-unit replenishment costs from the fast and slow sources are c and 0, respectively;
2. $L^E = 0$ such that the fast and slow lead times are 0 and L , respectively.

Note that the first reduction is without loss of generality under the long-run average cost criterion, because the total long-run average inventory sourced from R and E should equal the expected demand in a stable system. Under a discounted cost criterion, the reduction can also be without loss of generality as long as the ending salvage value is carefully defined. The second reduction is also without loss of generality, after using the notion of expedited inventory position and applying a standard state-reduction technique in the inventory literature (e.g., Xin and Goldberg (2018)).

We use $OPT(L) \triangleq \inf_{\pi \in \Pi} C(\pi)$ to denote the optimal cost. Here, we emphasize the dependence on L and suppress the dependence on other parameters.

8.2.1 Exact Optimality Result for Consecutive Lead Times: Single-Index Dual-Base-Stock Policy

With consecutive lead times $L=1$, the underlying dynamic program is one-dimensional and the state variable is the regular inventory position (i.e., the sum of the net inventory and all outstanding orders). The optimal policy thus also is a function of this single variable, or “index.” The celebrated result is that the optimal policy with consecutive lead times is a *Single-Index Dual-Base-Stock* policy. This policy is a natural generalization of the optimal single-base-stock policy under single sourcing and is simple to implement: it has two order-up-to levels S^E, S^R (one for each source) and first brings the regular inventory position (which is the single index) up to S^E by ordering from E and then brings the regular inventory position (including the newly placed express order) further up to S^R by ordering from R . The name “single index” comes from the fact that the policy only uses one state variable (i.e., regular inventory position). Later we introduce a different “dual-index” policy that uses two state variables.

Theorem 8.1 (Fukuda (1964))

When $L=1$, there exists an optimal *Single-Index Dual-Base-Stock* policy.

We provide a self-contained proof of Theorem 8.1 here. Recall that we assume throughout this chapter that $L^E=0$. Let us consider a T -period discounted problem without an ending salvage value. We proceed backward and first look at the last period T , where the cost-to-go function V_T depends on $x_T \triangleq I_T + q_{T-1}^R$, the initial net inventory in period T plus the regular order placed in period $T-1$ (received just before the last demand D_T is observed):

$$V_T(x_T) = \min_{q_T^E \geq 0} \left\{ cq_T^E + \mathbb{E} \left[G(x_T + q_T^E - D_T) \right] \right\}. \quad (8.1)$$

Define

$$S_T^E \in \arg \min_{y \in R} \{cy + \mathbb{E}[G(y - D_T)]\}. \quad (8.2)$$

Note that S_T^E exists because the objective function is convex; it is also non-negative under the assumption that $c < b$. It follows that the optimal express ordering decision in Equation (8.1) is to bring the net inventory up to S_T^E , namely, $q_T^{E,*} = (S_T^E - x_T)^+$. Hence,

$$V_T(x_T) = -cx_T + \left\{ c \max\{x_T, S_T^E\} + \mathbb{E}[G(\max\{x_T, S_T^E\} - D_T)] \right\},$$

which is convex in x_T from the definition of S_T^E in Equation (8.2). Note that the regular ordering decision does not matter in the last period, because the order will not arrive on time.

Now we go backward to period $T-1$:

$$\begin{aligned} V_{T-1}(x_{T-1}) &= \min_{q_{T-1}^R, q_{T-1}^E \geq 0} \left\{ cq_{T-1}^E + \mathbb{E}[G(x_{T-1} + q_{T-1}^E - D_{T-1})] \right. \\ &\quad \left. + \alpha V_T(x_{T-1} + q_{T-1}^E + q_{T-1}^R - D_{T-1}) \right\} \\ &= \min_{q_{T-1}^E \geq 0} \left\{ cq_{T-1}^E + \mathbb{E}[G(x_{T-1} + q_{T-1}^E - D_{T-1})] \right. \\ &\quad \left. + \alpha \min_{q_{T-1}^R \geq 0} \mathbb{E}[V_T(x_{T-1} + q_{T-1}^E + q_{T-1}^R - D_{T-1})] \right\} \end{aligned} \quad (8.3)$$

Similarly, define

$$S_{T-1}^R \in \arg \min_{y \in R} \mathbb{E}[V_T(y - D_{T-1})]. \quad (8.4)$$

It follows that the optimal regular ordering decision in Equation (8.3) is $q_{T-1}^{R,*} = (S_{T-1}^R - x_{T-1} - q_{T-1}^E)^+$. Hence,

$$\begin{aligned} V_{T-1}(x_{T-1}) &= -cx_{T-1} + \min_{q_{T-1}^E \geq 0} \left\{ c(x_{T-1} + q_{T-1}^E) \right. \\ &\quad \left. + \mathbb{E}[G(x_{T-1} + q_{T-1}^E - D_{T-1})] + \alpha V_T(\max\{x_{T-1} + q_{T-1}^E, S_{T-1}^R\} - D_{T-1}) \right\}. \end{aligned} \quad (8.5)$$

Define

$$S_{T-1}^E \in \arg \min_{y \in R} \left\{ cy + \mathbb{E}[G(y - D_{T-1}) + \alpha V_T(\max\{y, S_{T-1}^R\} - D_{T-1})] \right\}, \quad (8.6)$$

which is the optimal express order-up-to level in Equation (8.5). Define additionally

$$\begin{aligned}
U_1 &\in \arg \min_{y \in R} \left\{ cy + \mathbb{E} \left[G(y - D_{T-1}) \right] \right\}, \\
U_2 &\in \arg \min_{y \in R} \left\{ cy + \mathbb{E} \left[G(y - D_{T-1}) + \alpha V_T(y - D_{T-1}) \right] \right\}.
\end{aligned} \tag{8.7}$$

From the definitions of Equations (8.4) and (8.7), one can see that U_2 is always between S_{T-1}^R and U_1 ; namely, we have either $U_1 \leq U_2 \leq S_{T-1}^R$ or $U_1 \geq U_2 \geq S_{T-1}^R$. As a consequence,

$$S_{T-1}^E = \begin{cases} U_1 & \text{if } U_1 \leq U_2 \leq S_{T-1}^R; \\ U_2 & \text{if } U_1 \geq U_2 \geq S_{T-1}^R, \end{cases} \tag{8.8}$$

and the optimal express ordering decision in Equation (8.5) is $q_{T-1}^{E,*} = (S_{T-1}^E - x_{T-1})^+$. Here, the reason $S_{T-1}^E = U_1$ for the first case of Equation (8.8) is that the max operator in Equation (8.6) becomes a constant S_{T-1}^R under $y = U_1$. Similarly, one can conclude that $S_{T-1}^E = U_2$ for the second case of Equation (8.8).

In conclusion, the optimal policy in period $t-1$ is associated with two order-up-to levels S_{T-1}^E and S_{T-1}^R , and the policy first orders from E to bring the net inventory up to S_{T-1}^E and then orders from R to bring the inventory position up to S_{T-1}^R . Note the possibility that $S_{T-1}^E \geq S_{T-1}^R$ according to Equation (8.8). In that case, the optimal policy only uses the express source. Finally, one can see that $V_{T-1}(x_{T-1})$ is also convex in x_{T-1} .

By using a similar induction, one can eventually prove that for each $t = 1, \dots, T$, there exist S_t^R and S_t^E such that the optimal policy orders bring the net inventory and inventory position up to the two order-up-to levels, respectively. One can also use standard techniques to extend the optimality result to the infinite-horizon discounted problem (e.g., Feinberg (2016)) and the long-run average counterpart (e.g., Schäl (1993); Huh et al. (2011)), respectively. The results can be easily extended to general consecutive lead times as well, by simply replacing the function G as a convolution over multiple periods. Finally, in the special case of independent and identically distributed (iid) demands, $S_{T-1}^E = U_1 = S_T^E$ due to the fact that $S_T^E \leq S_{T-1}^R$. Hence, by induction, the optimal express order-up-to levels $\{S_t^E\}$ are stationary and are myopic solutions to Equation (8.2).

8.2.2 General (Non-Consecutive) Lead Times

With non-consecutive lead times $L > 1$, the underlying dynamic program is L -dimensional and the optimal policy is no longer a simple function of one state variable or index, but depends on the entire inventory pipeline vector (e.g., Whittemore and Saunders (1977)). Sheopuri et al. (2010) further confirm this point and show that simple order-up-to policies are not optimal for a special case of dual-sourcing systems with non-consecutive lead times. Their result is established by showing the equivalence of that special case with a single-sourcing lost-sales system for which order-up-to policies are not optimal (Karlin and Scarf (1958)). We next elaborate on this connection.

8.2.2.1 Connection to a single-sourcing lost-sales model

Sheopuri et al. (2010) establish a connection between a special dual-sourcing system with non-consecutive lead times and a single-sourcing lost-sales system, which is notoriously

challenging to optimize (see Chapter 1). The special dual-sourcing system has the following characteristics: (1) the decision q_t^E is made after D_t is observed, and this order is delivered instantly. Sheopuri et al. (2010) refer to that setting as “the express lead time of the dual-sourcing system $L^E = -1$ ”; (2) the stock-out cost of the lost-sales system is equal to the ordering-cost difference of the dual-sourcing system; and (3) the backorder cost of the dual-sourcing system is sufficiently large.

Sheopuri et al. (2010) prove that this special dual-sourcing system is equivalent to a single-sourcing lost-sales system. Note that the assumptions above (especially the third assumption) essentially imply that immediately clearing all stock-outs by expediting is optimal. The traditional dual-sourcing problem is even more challenging because the express lead time cannot be -1 , and the optimal policy might have stock-outs. Because an exact optimal policy for non-consecutive lead times seems out of reach, a considerable amount of existing research focuses on constructing and analyzing various heuristic policies. We next review several simple yet promising heuristics.

8.2.3 Single-Index and Dual-Index Policies

A natural family to investigate are the Single-Index Dual-Base-Stock policies (which are optimal for consecutive lead times and we abbreviate their name to Single-Index (SI) policies going forward). Rather than considering the entire L -dimensional state vector, SI policies only track one state variable or “index”: the sum of the net inventory and all outstanding orders (also called regular inventory position). Specifically, define

$$IP_t^R \triangleq I_t + q_{t-L}^R + \dots + q_{t-1}^R. \quad (8.9)$$

An SI policy first brings the regular inventory position up to S^E by ordering from E , and then brings the regular inventory position (including the newly placed express order) up to S^R by ordering from R :

$$q_t^E = (S^E - IP_t^R)^+, \quad q_t^R = (S^R - IP_t^R - q_t^E)^+.$$

Scheller-Wolf et al. (2007) analyze the performance of SI policies for general lead times. Although SI policies are easy to understand and implement, they are usually outperformed by *Dual-Index* policies, which are described next.

A more complex Dual-Index policy considers two measures, or “indexes,” of the entire state. The typical two-dimensional state reduction considers two particular aggregations (sums) of the state: the aforementioned regular inventory position and the expedited inventory position. The expedited (or express) inventory position is the sum of the net inventory and all the orders due to arrive in the next L^E periods (recall that we assume throughout this chapter that $L^E = 0$):

$$IP_t^E \triangleq I_t + q_{t-L}^R. \quad (8.10)$$

For the rest of the section, we review a specific but widely used family of Dual-Index policies called *Dual-Index Dual-Base-Stock* policies, which are closely related but more sophisticated than SI policies as described earlier. For notational convenience, we call them Dual-Index (DI)

policies going forward. We note that the name “Dual-Index” first appeared in Veeraraghavan and Scheller-Wolf (2008) in the discrete-time setting, and later Song and Zipkin (2009) use it to reinterpret the policy proposed by Moynadeh and Schmidt (1991) in the continuous-time setting.

Similar to an SI policy, a DI policy is associated with two order-up-to levels (S^E, S^R) satisfying $S^E \leq S^R$. The difference is that the DI policy uses both the regular and expedited inventory positions; by contrast, the SI policy uses only the regular inventory position. More specifically, in each period t , the DI policy first brings the expedited inventory position up to S^E by ordering from E and then brings the regular inventory position up to S^R by ordering from R :

$$q_t^E = (S^E - IP_t^E)^+, \quad q_t^R = (S^R - IP_t^R - q_t^E)^+. \quad (8.11)$$

In some cases, regular orders may push the expedited inventory position above S^E , causing an *overshoot*: $O_t \triangleq (IP_t^E - S^E)^+$, a distinguishing feature from a single-sourcing system under an order-up-to policy. Define $\Delta S \triangleq S^R - S^E$. One can verify that the overshoot distribution O_t only depends on ΔS and is independent of S^E (e.g., Veeraraghavan and Scheller-Wolf (2008)). We use $L=1$ to demonstrate this property. Suppose $I_1 = S^E$. Then, it is easy to see that $O_1 = 0$ and $O_{t+1} = (\Delta S - D_t)^+$ for all $t \geq 1$. This property implies that computing the optimal pair $(\Delta S, S^E)$ can be efficient. For example, in the long-run average setting, the overshoot distribution has a steady-state O_∞ such that for each ΔS , the best $S^E(\Delta S)$ (as a function of ΔS) can be computed by solving a newsvendor problem (with demand distribution $D_1 - O_\infty$). Regarding how to compute the best ΔS , although the cost function is not necessarily convex in ΔS , the numerical experiments in Veeraraghavan and Scheller-Wolf (2008) suggest that it appears to be unimodal in ΔS . Therefore, simple one-dimensional search methods (e.g., golden search) can work well.

The performance of the best DI policy can be very close to that of the optimal policy when the lead-time difference between the two suppliers is small. Indeed, it reduces to the SI policy when the lead-time difference is exactly one period (because the regular and express inventory positions coincide) such that it is optimal as reviewed in Section 8.2.1. Moreover, this family of policies includes the family of single-sourcing policies. For example, when the two base-stock levels coincide, the DI policy reduces to a single-sourcing policy from E ; when $S^E = -\infty$, the DI policy reduces to a single-sourcing policy from R . Finally, DI policies are expected to outperform SI policies given that both Li and Yu (2014) and Hua et al. (2015) show that the optimal fast (slow) order is more sensitive to soon(late)-to-receive orders.

DI policies have also been generalized to more complex base-stock type policies (e.g., Sheopuri et al. (2010); Hua et al. (2015)). For instance, Sheopuri et al. (2010) propose a heuristic under which the expedited order follows a base-stock policy and the regular order follows a *vector base-stock* policy, which originated from a single-sourcing lost-sales model. Specifically, the vector base-stock policy compares an inventory-system vector consisting of the partial sums of pipeline inventories to a fixed vector, and orders the minimum difference between the components of the system vector and those of the fixed vector. Numerical results suggest that the vector base-stock policy generally outperforms the DI and SI policies for general lead times.

Although DI policies can perform well for small lead-time differences, their performance seems to deteriorate as L grows, which leads to the study of the family of Tailored Base-Surge policies.

8.2.4 Tailored Base-Surge Policy

In this section, we review a family of simple and natural policies that are implemented in practice: *Tailored Base-Surge* (TBS) policies. These policies were first proposed and analyzed by Allon and Van Mieghem (2010), although closely related standing-order policies had been studied earlier (e.g., Rosenshine and Obee (1976)). We refer to Mini-Case 6 in Van Mieghem (2008) for more about the motivation and background of TBS policies.

We assume that demands are iid for the rest of the section. Each TBS policy π_{r,S^E} is associated with two parameters (r, S^E) : a constant order r is placed at the regular source in each period to meet a *base* level of demand, while the orders placed at the express source follow an order-up-to- S^E rule to manage demand *surges*. Specifically,

$$q_t^E = (S^E - IP_t^E)^+, \quad q_t^R = r.$$

Note that dual-sourcing inventory systems in which a constant-order policy is implemented for the regular source are equivalent to systems with constant returns, which have been well studied in the literature (e.g., Fleischmann and Kuik (2003); DeCroix et al. (2005)). Hence, we can think of an equivalent single-sourcing backlogged inventory system with demands $D-r$ and zero lead time, controlled by a base-stock policy with the base-stock level S^E . Let I_n be the net inventory in period n after replenishment but before seeing the demand. Then, from the inventory dynamics,

$$I_{n+1} = \max \left[I_n + r - D_n, S^E \right].$$

Assume that the initial net inventory $I_1 = S^E$. By using induction, it follows that

$$I_{n+1} = S^E + \max \left[\left(nr - \sum_{j=1}^n D_j \right), \dots, (2r - D_{n-1} - D_n), (r - D_n), 0 \right].$$

In the limiting case, it is straightforward to verify that the steady-state net inventory exists and has the same distribution as $S^E + I_r^\infty$ as long as $r < \mathbb{E}[D]$, where

$$I_r^\infty \triangleq \sup_{j \geq 0} \left(jr - \sum_{i=1}^j D_i \right).$$

Similar to DI policies as discussed in Section 8.2.3, the term I_r^∞ represents the overshoot under a TBS policy and is independent of S^E . In addition, it can be interpreted as the steady-state waiting time in a GI/D/1 queue with inter-arrival distribution D and deterministic processing time r . It follows that

$$C(\pi_{r,S^E}) = c(\mathbb{E}[D] - r) + \mathbb{E}\left[G\left(I_r^\infty + S^E - D\right)\right], \quad (8.12)$$

where D is independent of I_r^∞ . The term $\mathbb{E}[D] - r$ in Equation (8.12) captures the average amount of inventory ordered from the express source in the long run and $I_r^\infty + S^E - D$ represents the steady-state net inventory at the end of each period. Note that $C(\pi_{r,S^E})$ is independent of the lead-time difference L . For each r , the minimization problem $\min_{S \in \mathbb{R}} C(\pi_{r,S})$ that finds the optimal base-stock level is equivalent to a standard newsvendor problem (with demand distribution $D - I_r^\infty$). Furthermore, $\min_{S \in \mathbb{R}} C(\pi_{r,S})$ is convex in r (e.g., Janakiraman et al. (2015)). Let (r^*, S^*) be an optimal pair:

$$r^* \in \arg \min_{0 \leq r \leq \mathbb{E}[D]} \left(\min_{S \in \mathbb{R}} C(\pi_{r,S}) \right), \quad S^* \in \arg \min_{S \in \mathbb{R}} C(\pi_{r^*,S}).$$

Combining all the above, we conclude that the optimal pair can be computed efficiently by solving a one-dimensional convex program on a compact set.

In the special case in which the demand has a two-point support distribution, Janakiraman et al. (2015) establish the optimality of TBS policies. More specifically, when demand is composed of a base-demand component plus a surge-demand component, as long as the probability of the surge is sufficiently small, a TBS policy is optimal. They also show that the performance of TBS, relative to the optimal policy, improves as the lead time of the regular source increases. For general demand distributions, Xin and Goldberg (2018) prove the asymptotic optimality of TBS policies.

Theorem 8.2 (Xin and Goldberg (2018))

$$\lim_{L \rightarrow \infty} \frac{C(\pi_{r^*,S^*})}{OPT(L)} = 1.$$

Because the survey paper by Goldberg et al. (2020) provides a nice and detailed overview of the proof, we only briefly mention the two key ideas here. The first key idea is to use Jensen's inequality to obtain a lower bound on the optimal cost. More specifically, suppose that the optimal policy has a stationary measure of the net inventory and pipeline vector. By using the fact that the optimal cost is jointly convex in the net inventory and pipeline vector, one can apply Jensen's inequality to obtain a lower bound on the optimal cost by simply replacing the steady-state (random) vector with its expectation. Combined with the stationarity, the pipeline now becomes a constant vector, which connects to a constant-order policy. The second key idea in the proof is to show that the lower bound converges to the cost of a constant-order policy as $L \rightarrow \infty$. More specifically, the convergence to a long-run average backlogged model with returns (i.e., Equation (8.12)) by its discounted counterpart is established as the discount factor goes to one. The argument essentially validates the so-called Schäl conditions, and we refer interested readers to Schäl (1993) and Huh et al. (2011) for more discussion. Finally, for more recent research on asymptotic optimality of constant-order policies in other related inventory models, we refer to Goldberg et al. (2016), Xin and Goldberg (2016), Xin et al. (2017), Chen et al. (2019), Bu et al. (2020), Xin (2021c), and Xin (2021b).

The discussion thus far suggests that DI policies perform well for small L and TBS policies perform well for large L . However, a fundamental question of whether there exists a simple policy (perhaps by incorporating features from other good-performing heuristics) that performs well for all L , remains open. We answer this question and introduce the family of Capped Dual-Index policies in the following section.

8.2.5 Capped Dual-Index Policy

Given that the optimal dual-sourcing policy for stochastic demands and general lead times continues to be elusive, Sun and Van Mieghem (2019) investigate whether the optimal policy can be found when stochastic optimization is replaced by robust optimization. We refer to Chapter 7 “Robust inventory management” by Michael R. Wagner in this book for more details. Sun and Van Mieghem (2019) solve the robust optimization and call the robust-optimal policy the *Capped Dual-Index* (CDI) policy. They show that the optimal policy is a DI policy with an added constraint (or “cap”) on the slow order. Specifically, the slow orders q_i^R of (8.11) are modified to $q_i^R = \min \{ (S^R - IP_i^R - q_i^E)^+, r \}$, where r is a parameter of the policy determined by the decision-maker. This cap restricts the overshoot and also smooths the reordering over multiple periods if a sudden demand surge occurs. Here, we use the same notation r for the cap of a CDI policy as for the constant-order quantity of a TBS policy, because CDI indeed has a strong connection to TBS (see below). The authors also show numerically that the CDI policy performs admirably (in the conventional stochastic demand setting) relative to other heuristics. The key contribution of Sun and Van Mieghem (2019) is to highlight how adding a simple cap to an order can improve inventory performance in systems that are intractable with exact stochastic analysis. Recall that a Dual-Base-Stock policy is optimal for consecutive lead times (Theorem 8.1), and it is also a special case of a CDI policy (with an implicit cap $S^R - S^E$ on the slow order).

CDI policies also have strong connections to other well-performing heuristics. On one hand, as the regular order-up-to level grows (due to, e.g., a growing L) while other parameters remain the same, the CDI policy can behave like a TBS policy (because regular orders are always capped). On the other hand, as L decreases (or the cap r increases), the CDI policy can resemble a DI policy. This natural transition depending on the value of L explains CDI’s superior performance. The performance of CDI policies has also been investigated theoretically in a continuous-time model by Xin (2021a), which we review in Section 8.3.3. However, one caveat: computing the optimal (S^R, S^E, r) is more challenging due to the lack of convexity. We provide a summary of the heuristic policies reviewed for the discrete-time model in Table 8.1.

Subsequently, following the same idea of adding an order cap, Xin (2021c) studies the performance of a *capped base-stock* policy in the single-sourcing lost-sales model. Unsurprisingly, the numerical results demonstrate its superior performance, which is driven primarily by the newly added order cap that helps smooth orders and avoids chasing surging demands. Capped Base-Stock policies also are highly interpretable and easy to implement. The high-level idea that a good policy should have orders be smoother than the filled demands was indeed raised early (for example, see Section 9.6.5 of Zipkin (2000)), and similar policies were studied in related inventory models (we refer to Xin (2021c) for more discussions there).

Table 8.1 A summary of heuristic policies for the discrete-time model

Policy	Parameters
Single-Index: $q_t^E = (S^E - IP_t^R)^+$, $q_t^R = (S^R - IP_t^R - q_t^E)^+$ Pros: track only one state variable (or index); optimal for $L=1$ Cons: usually outperformed by Dual-Index policy	(S^R, S^E)
Dual-Index: $q_t^E = (S^E - IP_t^E)^+$, $q_t^R = (S^R - IP_t^R - q_t^E)^+$ Pros: usually outperform Single-Index policy; optimal for $L=1$ Cons: less effective for large L	(S^R, S^E)
Tailored Base-Surge: $q_t^E = (S^E - IP_t^E)^+$, $q_t^R = r$ Pros: asymptotically optimal as $L \rightarrow \infty$ Cons: less effective for small L	(S^E, r)
Capped Dual-Index: $q_t^E = (S^E - IP_t^E)^+$, $q_t^R = \min\left\{(S^R - IP_t^R - q_t^E)^+, r\right\}$ Pros: superior performance in general (including median L) Cons: challenging to compute the optimal parameters due to the lack of convexity	(S^R, S^E, r)

Note: IP_t^E and IP_t^R represent the expedited and regular inventory positions, respectively, as defined in Equations (8.9) and (8.10).

8.2.6 Dual Sourcing with Capacitated Sources

In this section, we review several recent results for capacitated dual-sourcing models. Federgruen et al. (2021) study discrete-time stochastic dual-sourcing systems with consecutive lead times when sources are capacitated, and find optimal dual-sourcing policies by using (C_1, C_2, K_1, K_2) -convexity. They show when the fast source is capacitated, a capped base-stock policy is optimal for the fast supplier: order-up-to the base-stock level if you can; otherwise, order the cap.

Gijsbrechts et al. (2021) extend Fukuda (1964) to a setting where the fast per-unit replenishment cost increases to mc^E if $q_t^E > k$. Here, k represents the fast supplier's base capacity and $m \geq 1$ is the multiplier paid for ordered units beyond the base capacity (e.g., capturing overtime premium). This piecewise linear fast purchasing cost models the fast supplier's volume flexibility, which increases when either k increases or m decreases. (When $m \rightarrow \infty$, it models a hard constraint for the fast supplier.) They prove that a *modified dual base-stock* policy with three base-stock levels is optimal with consecutive lead times. Fast orders follow a *modified (single) base-stock* policy, with base-stock levels $S_1^E > S_2^E$: the base capacity is used to raise the inventory position toward the higher base-stock level S_1^E ; if this action raises the inventory position above the lower base-stock level S_2^E , no overtime is used. Otherwise, the base capacity and overtime are used to raise the inventory position up to S_2^E . This policy creates a region of "inaction" where only k units are ordered from the fast supplier and additional units are "postponed" by ordering from the slow supplier. After ordering from the fast supplier, the inventory position is raised to the slow base-stock level S^R .

These exact optimality results for consecutive lead times were preceded by heuristics for general lead times that utilized order smoothing to alleviate capacity constraints: Boute and Van Mieghem (2015) use the modified (single) base-stock policy in a single-sourcing setting to motivate their *Dual-Sourcing Smoothing (DSS)* policy, a linear order-smoothing heuristic in a dual-sourcing model where both sources can have installed capacity (i.e., a capacity that incurs cost regardless of the ordered quantity). They adopt a conventional discrete-time inventory model with a linear control rule that smoothes orders and allows an exact and analytically tractable analysis of single and dual-sourcing policies under normal demand. Distinguishing features of the model are that it captures each source's lead time, capacity cost, and flexibility to work overtime. Boute and Van Mieghem (2015) use Lagrange's inversion theorem to provide exact and simple square-root-bound formulas for the strategic sourcing allocations¹ and the value of dual sourcing. The formulas provide structural insight into the impact of financial, operational, and demand parameters, and a starting point for quantitative decision-making.

Boute et al. (2021) introduce the POUT-TBS policy: a constant-order policy for the slow supplier, combined with a proportional order-up-to (POUT) policy for the fast supply. "Proportional" means only a fraction of the actual inventory deviation is corrected in each decision, deferring the remaining correction to be recovered in future decisions. Similar to DSS, the POUT-TBS policy smoothes the fast orders to avoid expensive overtime production beyond the installed fast capacity. Using a break-even analysis, the authors investigate the lead time, demand and cost characteristics that make dual sourcing with a small capacitated fast supply (also called a SpeedFactory) more desirable than complete off-shoring. They adopt Z-transforms to present exact analyses under normally distributed demand. This paper appears to be the first to provide a stochastic analysis of dual sourcing under correlated and non-stationary demand, which is prevalent in practice and is shown to significantly increase the viability of dual sourcing relative to the traditionally assumed iid demand in dual-sourcing research.

8.3 CONTINUOUS-TIME MODEL

Discrete-time inventory models correspond to discrete-review policies, which are prevalent in practice. In addition, discrete-time models allow the use of dynamic programming to study optimal policies. By contrast, continuous-time inventory models correspond to continuous-review policies, which may be easier to implement in digital environments. From a theoretical perspective, it connects naturally to queueing systems, and hence provides another classic toolbox to study inventory problems. We refer to Zipkin (2000) for more comments on the connections/differences between discrete-time and continuous-time models. In particular, this setting permits the consideration of stochastic lead times in some models, which are mostly absent in the discrete-time models.

In this section, we consider the continuous-time version of the discrete-time dual-sourcing inventory model. All parameters remain the same with the following exceptions: (1) customer demand arrives as a Poisson process with rate λ ; (2) the holding and backorder penalty costs are measured per unit of inventory per unit of time; (3) we consider general $L^E \geq 0$; and (4) we assume $h=1$ (without loss of generality by using a simple scaling argument). The objective is

to minimize the long-run average cost per unit of time. Let $C(\pi)$ denote the long-run average cost per unit of time incurred by policy π , and let OPT denote the optimal cost.

Similar to the discrete-time setting, the optimal policy is complex and we focus on three heuristics in the rest of the section, with an exception of Section 8.3.4, where we introduce some results in the form of the optimal policy in a special case. In particular, we discuss DI, TBS, and CDI policies in the following sections. The main insights are essentially the same as in the discrete-time model; specifically, DI outperforms TBS for small lead-time differences, TBS outperforms DI for large lead-time differences, and CDI performs uniformly well. However, because of the strong connection to queueing systems, conducting an asymptotic analysis in the continuous-time model to characterize exactly the threshold of the lead-time difference on when DI outperforms TBS, is possible. Similar to Table 8.1, we provide a summary of the heuristic policies reviewed for the continuous-time model in Table 8.2.

Remark 8.1 The continuous-time dual-sourcing model is closely related to but more complex than the analogous continuous-time single-sourcing lost-sales model as studied by Reiman (2004) and Xin (2021a). In particular, if $L^E=0$ and the decision-maker decides to avoid any stock-outs, the dual-sourcing model will reduce to the single-sourcing lost-sales model (i.e., the express ordering cost c corresponds to the lost-sales penalty cost in the lost-sales model). However, because L^E can be strictly positive, and having stock-outs may be better than clearing stock-outs immediately by expediting, the dual-sourcing model is more difficult to analyze. Interestingly, as we prove later, the dual-sourcing model indeed reduces to the lost-sales

Table 8.2 A summary of heuristic policies for the continuous-time model

Policy	Parameters
Dual-Index places an order whenever a demand arrives; the order comes from E iff the number of outstanding orders from R that will not arrive within time L^E reaches $S^R - S^E$	(S^R, S^E)
Tailored Base-Surge places an order from R every r^{-1} time units; only orders from E if the express inventory position falls below S^E	(S^E, r)
Capped Dual-Index the regular inventory position is reviewed r^{-1} time units after the last regular order is placed, and an order is placed if it falls below S^R ; otherwise, it waits until the regular inventory position falls below S^R and then places an order; an express order can be placed any time as long as the express inventory position falls below S^E	(S^R, S^E, r)
Performance Summary similar to the discrete-time model, DI outperforms TBS for small lead-time differences, TBS outperforms DI for large lead-time differences, and CDI performs uniformly well. Under Assumption 1, we can also characterize exactly the threshold of the lead-time difference. Namely, when both L and c are large, the threshold of lead-time difference is approximately $\alpha^* c$, where $\alpha^* \approx 0.69787$	

Note: Each order requests only one unit of inventory, due to the Poisson demand assumption.

model in the worst-case scenario, and both models have the same worst-case ratio of 1.79. We provide further discussion later.

We make the following assumption for the rest of the section.

Assumption 1 $L = \alpha c$ for some $\alpha \geq b^{-1}$.

Similar to the discrete-time model, we assume $\alpha \geq b^{-1}$ (equivalent to $bL \geq c$); otherwise, one can easily see that there exists an optimal policy sourcing exclusively from R .

As we conduct an asymptotic analysis, the above assumption essentially leads to the setting where $L, c \rightarrow \infty$, similar in spirit to the large lead-time and penalty-cost setting in the analogous single-sourcing lost-sales model by Reiman (2004) and Xin (2021a), which arguably presents the worst case. In addition, in that lost-sales model, Reiman (2004) proves that a threshold exists such that the best constant-order policy outperforms the best base-stock policy if and only if (iff) the ratio of the lead time to the penalty cost exceeds the threshold, and this threshold has a closed-form expression asymptotically. We extend this type of asymptotic analysis to the dual-sourcing setting here. The reason we scale L and c (instead of L and b) is that this combination is a direct extension of scaling the lead time and penalty cost in the single-sourcing lost-sales setting. Moreover, as we describe below, the inventory systems under a DI policy and a TBS policy have nice connections to queueing systems. The assumption of large lead-time and ordering-cost differences essentially leads to heavy-traffic systems, which makes our analysis more convenient. Furthermore, when the lead-time difference L grows, the market-equilibrium price difference c (arguably) grows as well to maintain an effective two-supplier market. We provide more discussion about the single-sourcing lost-sales model in Section 8.3.3.2.

8.3.1 Dual-Index Policy in Continuous Time

In this section, we discuss the family of DI policies in the continuous-time setting. We ignore SI policies because an SI policy reduces to a single-sourcing base-stock policy in the continuous-time model with unit demand. Each DI policy π_{S^R, S^E} is associated with two parameters (S^R, S^E) , where $S^R \geq S^E$. The policy triggers orders to keep the regular inventory position (i.e., net inventory plus all outstanding orders) at S^R and the express inventory position (i.e., net inventory plus all outstanding orders that will arrive within time L^E) at least S^E . Namely, the policy places an order of one unit either from E or R whenever a demand arrives (in practice, this type of order-up-to policy is often used for controlling the stock levels of expensive and slow-moving items). The order is placed to E iff the number of outstanding orders from R that will not arrive within time L^E reaches $S^R - S^E$. Equivalently, we define $S \triangleq S^R$ and $U \triangleq S^R - S^E \geq 0$; that is, the DI policy $\pi_{S, U}$ keeps the regular inventory position at S and orders only from E iff U outstanding orders from R will not arrive within time L^E . Note that $S \leq U$ (equivalent to $S^E \leq 0$) is possible.

The steady state of the inventory system under $\pi_{S, U}$ was first derived by Moynzadeh and Schmidt (1991). Song and Zipkin (2009) point out that the system has a nice connection to a two-station tandem queue. In this tandem queue, both stations have an infinite number of servers with deterministic processing times L and L^E , respectively. We use (N_1, N_2) to denote the numbers of busy servers at stations 1 and 2, respectively. Jobs arrive as a Poisson process with rate λ . Each job is determined to either pass through both stations (i.e., a regular order) or skip the first station and go directly to the second station (i.e., an express order). An express

order happens iff the number of busy servers at station 1 reaches U , that is, iff $N_1 = U$. After the job is completed at station 2, it exits the network, indicating the arrival of the order to the on-hand inventory. The inventory state under $\pi_{S,U}$ can be completely captured by (N_1, N_2) , which represent the numbers of outstanding orders that will not and will arrive within time L^E , respectively. Note that due to the overflow, the queueing network is not the classic Jackson network that automatically guarantees a product-form equilibrium solution structure, but it still has the product-form solution as an extension (e.g., Jackson (1963)). The steady-state distribution has the following product form:

$$\mathbb{P}(N_1 = n_1, N_2 = n_2) = \frac{1}{G(U)} \phi_1(n_1) \phi_2(n_2) \quad \text{for } 0 \leq n_1 \leq U, n_2 \geq 0, \quad (8.13)$$

where

$$\phi_1(n) \triangleq \frac{e^{-\lambda L} (\lambda L)^n}{n!}, \quad \phi_2(n) \triangleq \frac{e^{-\lambda L^E} (\lambda L^E)^n}{n!}, \quad G(U) \triangleq \sum_{n_1=0}^U \phi_1(n_1).$$

In other words, the marginal distribution N_2 is a Poisson distribution with rate λL^E , N_1 is a truncated Poisson distribution with rate λL , and N_1, N_2 are independent.

Moinzadeh and Schmidt (1991) obtain the solutions in Equation (8.13) using balance equations. Song and Zipkin (2009), after observing the connection to queueing networks, directly apply Jackson's result to obtain these solutions. Because of this connection to queueing networks, Song and Zipkin (2009) extend the dual-sourcing system to allow each source to be modeled as a Jackson network consisting of several stations. Orders placed to a particular source go through the network following a pre-determined route. In particular, the lead times at each source can be iid such that the product-form solution still holds with L and L^E replaced by their means.

We next characterize the long-run average cost of $\pi_{S,U}$. Define the following Erlang loss function:

$$B(n, a) \triangleq \frac{\frac{a^n}{n!}}{\sum_{k=0}^n \frac{a^k}{k!}}, \quad n, a > 0.$$

Then, the term $B(U, \lambda L) = \mathbb{P}(N_1 = U)$ captures the fraction of orders from E , that is, the utilization of the source E . Therefore,

$$C(\pi_{S,U}) = \lambda c B(U, \lambda L) + \mathbb{E} \left[(S - N_1 - N_2)^+ + b (N_1 + N_2 - S)^+ \right].$$

Note that the above closed-form expression under a DI policy is unique to the continuous-time setting, and no simple expression of the long-run cost function exists in the analogous discrete-time setting. Taking advantage of the closed form of $C(\pi_{S,U})$, Moinzadeh and Schmidt (1991) develop an efficient algorithm to numerically compute the best DI policy (see Song and Zipkin (2009) for a nice summary right after Equation (8.3) in their paper).

Although considerable progress has been made, analytically quantifying the gap between the best DI policy and the optimal policy is still challenging. To analyze this gap, Xin (2021a) initiates an asymptotic analysis under Assumption 1. Performing the asymptotic analysis requires understanding the limits of the Erlang loss function $B(n, a)$ and the truncated Poisson distribution N_1 (we can ignore N_2 because it does not scale with L). The former turns out to be related to the following hazard-rate function of the standard normal distribution:

$$\psi(x) \triangleq \frac{\phi(x)}{1 - \Phi(x)},$$

where $\phi(x) \triangleq (\sqrt{2\pi})^{-1} e^{-\frac{x^2}{2}}$ and $\Phi(x) \triangleq \int_{-\infty}^x \phi(y)dy$ are the probability density function (pdf) and cumulative distribution function (cdf) of the standard normal distribution, respectively. Note that $\psi(x)$ is strictly increasing and strictly convex. The following asymptotic result relates the Erlang loss function $B(n, a)$ to $\psi(x)$ (Jagerman (1974)): for each $\beta \in \mathbb{R}$,

$$\lim_{a \rightarrow \infty} \sqrt{a} B(\lfloor a + \beta \sqrt{a} \rfloor, a) = \psi(-\beta).$$

Regarding the limit of N_1 (which depends on U), one can prove that the following scaled distribution $(N_1(U_\beta) - U_\beta) / \sqrt{\lambda L}$ converges weakly to Z_β as $L \rightarrow \infty$ (Xin (2021a)), where $U_\beta \triangleq \lfloor \lambda L + \beta \sqrt{\lambda L} \rfloor$ for $\beta \in \mathbb{R}$ and Z_β is a truncated normal distribution with mean $-\beta$, variance 1, and within $(-\infty, 0]$. Combining all the above, we are able to provide a (tight) upper bound on the limit of the scaled cost incurred by the best DI policy in continuous time.

Theorem 8.3 (Xin (2021a))

Under Assumption 1,

$$\begin{aligned} \lim_{c \rightarrow \infty} \frac{\min_{s, U} C(\pi_{s, U})}{\sqrt{\lambda c}} &\leq \sqrt{\alpha} \min_{\beta, s \in \mathbb{R}} \left\{ \frac{1}{\alpha} \psi(-\beta) + \mathbb{E} \left[(s - Z_\beta)^+ + b(Z_\beta - s)^+ \right] \right\} \\ &\triangleq F_{DI}(b, \alpha). \end{aligned} \quad (8.14)$$

In addition, for each fixed α , $F_{DI}(b, \alpha)$ is non-decreasing in b .

As we illustrate below, the limit $\lim_{b \rightarrow \infty} F_{DI}(b, \alpha)$ is exactly the scaled cost incurred by the optimal base-stock policy in the related single-sourcing lost-sales model, and the limit is non-decreasing in α .

8.3.2 Tailored Base-Surge Policy in Continuous Time

In this section, we discuss the family of TBS policies in the continuous-time setting. Xin (2021a) proposes and studies this version of TBS policies. Each TBS policy π_{r, S^E} is associated with two parameters (r, S^E) . The policy places an order of one unit from R every r^{-1} time unit and only orders from E if the express inventory position (i.e., net inventory plus all outstanding orders that will arrive within time L^E) is strictly less than S^E (more precisely, exactly $S^E - 1$)

right after receiving a demand. It is slightly different from the TBS policy discussed in the discrete-time review setting, where the policy orders a constant amount of inventory (possibly multiple units) every period. The parameter r can be interpreted as the supply rate, which must be strictly smaller than the demand rate λ ; otherwise, the resulting inventory system is unstable. The steady-state express inventory position under π_{r,S^E} (denote by IP^E) is closely related to the steady-state number-in-system I_r^∞ in a D/M/1 queue, where customers arrive every r^{-1} time unit and the processing time follows an exponential distribution with mean λ^{-1} . Indeed, IP^E has the same distribution as $I_r^\infty + S^E$. It follows that the net inventory in its steady state has the same distribution as IP^E minus the express lead-time distribution (a Poisson distribution with rate λL^E). Therefore,

$$C(\pi_{r,S^E}) = c(\lambda - r) + \mathbb{E} \left[\left(I_r^\infty + S^E - P(\lambda L^E) \right)^+ + b \left(I_r^\infty + S^E - P(\lambda L^E) \right)^- \right],$$

where $P(\lambda L^E)$ is a Poisson distribution with rate λL^E and the term $(\lambda - r)$ represents the amount of inventory ordered from E captured by the difference between the demand rate λ and regular supply rate r . In addition, $C(\pi_{r,S^E})$ is independent of the slow lead time because orders arrive every r^{-1} time units from R .

Similarly, analytically quantifying the gap between the best TBS policy and the optimal policy is difficult. Xin (2021a) initiates an asymptotic analysis under Assumption 1. Performing the asymptotic analysis requires understanding the limit of the I_r^∞ . From queueing theory, it is well known that $(2\lambda\theta I_r^\infty)/(r^2\sqrt{c})$ converges weakly to an exponential distribution with mean one when $r = \lambda - \theta/\sqrt{c}$ for some $\theta > 0$. Hence, we are able to provide a (tight) upper bound on the limit of the scaled cost incurred by the best TBS policy.

Theorem 8.4 (Xin (2021a))

Under Assumption 1,

$$\lim_{c \rightarrow \infty} \frac{\min_{r,S} C(\pi_{r,S})}{\sqrt{\lambda c}} \leq \sqrt{2}.$$

8.3.3 Capped Dual-Index Policy in Continuous Time

In this section, we discuss the family of CDI policies in the continuous-time setting. Xin (2021a) proposes and studies this version of CDI policies. Recall that in the discrete-time model described earlier, each CDI policy is associated with three parameters: two base-stock levels and an additional order cap associated with R . Similarly, each CDI policy $\pi_{S^R, S^E, r}$ is associated with three parameters (S^R, S^E, r) in the continuous-time setting as well. It is similar to the DI policy π_{S^R, S^E} , and the only difference is that it has an additional cap on the regular order frequency. Specifically, r^{-1} time units after the last regular order is placed, the regular inventory position is reviewed and an order of size one is placed if it falls below S^R ; otherwise, it waits until the regular inventory position falls below S^R and then places an order. Meanwhile,

an express order can be placed at any time as long as the express inventory position falls below S^E . The family of CDI policies clearly includes the family of TBS policies and the family of DI policies. Specifically, a CDI policy reduces to a TBS policy when $S^R = \infty$ and reduces to a DI policy when $r = \infty$. Because characterizing the optimal CDI policy is difficult (even in the asymptotic setting), as a first step toward theoretically understanding its performance, Xin (2021a) proposes the following natural upper bound on the cost incurred by the best CDI policy: the minimum between the cost incurred by the best TBS policy and the cost incurred by the best DI policy.

Lemma 8.1 (Xin (2021a))

$$\min_{S^R, S^E, r} C(\pi_{S^R, S^E, r}) \leq \min_{S, U} C(\pi_{S, U}) \bigwedge \min_{r, S} C(\pi_{r, S}).$$

Xin (2021a) proves that this simple upper bound already achieves a 1.79-approximation asymptotically; that is, the ratio of this upper bound to the optimal cost is no greater than 1.79 in the interested asymptotic regime. As an immediate consequence, the best CDI policy achieves a 1.79-approximation as well. Define

$$\text{RATIO(DS)} \triangleq \sup_{\lambda, \alpha, b, L^E > 0} \limsup_{c \rightarrow \infty} \frac{\min_{S, U} C(\pi_{S, U}) \bigwedge \min_{r, S} C(\pi_{r, S})}{OPT}.$$

Theorem 8.5 (Xin (2021a))

Under Assumption 1, $\text{RATIO(DS)} \leq 1.79$.

We also briefly discuss the significance of this result. Among all the research that has been conducted on the dual-sourcing inventory model with general lead times over the past six decades, only a few papers provide performance guarantees (summarized in Table 8.3). Moreover, Theorem 3.3 is the only result with a worst-case ratio in the stochastic setting, although it is an asymptotic one. Finally, note that the cost-balancing technique used to prove the 2-approximation in the single-sourcing lost-sales model in Levi et al. (2008) cannot be directly applied to the dual sourcing (neither discrete-time nor continuous-time). The technique relies on an assumption that once an order is placed in some period, the associated expected marginal holding cost that will be incurred over the rest of the planning horizon is a function only of the realized demands over the rest of the horizon and is unaffected by any future decisions. This assumption fails in the dual-sourcing model with general lead times, because the marginal holding cost associated with a regular order can be affected by a future express order.

8.3.3.1 Proof sketch of Theorem 8.5

In this section, we provide a proof sketch of Theorem 8.5 and refer to a complete proof from Xin (2021a). We first derive a lower bound on the optimal cost by using a single-sourcing backlogged model through a coupling argument. Define

$$F_{\text{lower}}(x) \triangleq \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right) \phi \left(\Phi^{-1} \left(\frac{1}{1+x} \right) \right).$$

Table 8.3 A comparison of dual-sourcing papers with general lead times and performance guarantees

Review Frequency	Demand	Policy	Performance Guarantee
<u>Janakiraman et al. (2015)</u>			
periodic	iid two-point distribution	TBS	optimal
<u>Xin and Goldberg (2018)</u>			
periodic	iid general distribution	TBS	asymptotically optimal (large L)
<u>Sun and Van Mieghem (2019)</u>			
periodic	distribution-free robust rolling-horizon model [#]	CDI	optimal
<u>Xin (2021a)</u>			
continuous	Poisson process	CDI	1.79 (large L and c)

[#]Sun and Van Mieghem (2019) establish the optimality of CDI policies in a robust rolling-horizon model.

Lemma 8.2 (Xin (2021a))

Let $C^{*,R}(b)$ be the optimal long-run average cost per unit of time when only sourcing from R (recall that the holding- and penalty-cost parameters are 1 and b , respectively, and no ordering cost from R arises). Then, $OPT \geq C^{*,R}(c/L)$. In addition, under Assumption 1,

$$\lim_{c \rightarrow \infty} \frac{C^{*,R}\left(\frac{c}{L}\right)}{\sqrt{\lambda c}} = F_{lower}(\alpha). \quad (8.15)$$

The lower bound $C^{*,R}(c/L)$ corresponds to a single-sourcing backlogged model with an adjusted backlog penalty cost c/L (per unit of inventory per unit of time): the benefit of a shorter lead time L by using the express source at the expense of c is mitigated by the adjusted penalty cost c/L in the single-sourcing backlogged model. Regarding the asymptotic expression of $C^{*,R}(c/L)$ in Equation (8.15), similar to the analogous discrete-time setting (and perhaps not surprising), the optimality of a base-stock policy has been well established. In addition, the inventory system under a base-stock policy corresponds to an $M/D/\infty$ queue whose steady-state distribution is given by the celebrated Palm's Theorem, which says that the steady-state distribution of the number of outstanding orders (i.e., the orders already placed but not yet received) is a Poisson distribution with rate λL . It follows that the lower bound can be calculated by using the newsvendor problem below:

$$C^{*,R}\left(\frac{c}{L}\right) = \min_{S \in \mathbb{R}} \mathbb{E} \left[(S - P(\lambda L))^+ + \frac{c}{L} (S - P(\lambda L))^- \right],$$

where $P(\lambda L)$ is a Poisson distribution with rate λL . Then, a straightforward calculation leads to Equation (8.15).

Note that the lower bound in Equation (8.15) is independent of b , which could be a potential problem because we are primarily interested in the worst-case ratio. Fortunately, as demonstrated earlier in Theorems 8.3 and 8.4, the best TBS and DI policies do not diverge in b .

Combining Lemma 8.2 with Theorems 8.3 and 8.4 implies the following upper bound on the ratio:

$$\begin{aligned} \text{RATIO(DS)} &\leq \sup_{\alpha, b > 0} \frac{F_{DI}(b, \alpha) \bigwedge \sqrt{2}}{F_{lower}(\alpha)} \\ &= \sup_{\alpha > 0} \frac{\lim_{b \rightarrow \infty} F_{DI}(b, \alpha) \bigwedge \sqrt{2}}{F_{lower}(\alpha)}. \end{aligned} \quad (8.16)$$

The next step is to calculate the limit $\lim_{b \rightarrow \infty} F_{DI}(b, \alpha)$, which is a critical component of the entire proof. We essentially want to prove that the joint optimization problem over $\beta, s \in \mathbb{R}$ in Equation (8.14) converges to the same optimization problem fixing $s=0$ up front as $b \rightarrow \infty$. Doing so avoids the need to solve the non-trivial joint optimization problem explicitly for each (b, α) pair. The intuition is as follows. When c is large, the optimal DI policy has a non-positive optimal s , corresponding to the setting where it orders from E only when the express inventory position is non-positive, which is known as a “reactive” control policy in the literature; namely, the express source engages only to cover backlogs (e.g., see Allon and Van Mieghem (2010)). After L and c both go to infinity, the worst case of $\max_{b > 0} F_{DI}(b, \alpha)$ arises when $b \rightarrow \infty$ such that the optimal s converges to zero, namely, using the express supplier to clear stock-outs immediately.

Lemma 8.3 (Xin (2021a))

For each $\alpha > 0$,

$$\lim_{b \rightarrow \infty} F_{DI}(b, \alpha) = \sqrt{\alpha} \min_{\beta \in \mathbb{R}} \left\{ \beta + \left(1 + \frac{1}{\alpha} \right) \psi(-\beta) \right\} \triangleq G_{DI}(\alpha).$$

Combining (8.16) with Lemma 8.3 implies

$$\text{RATIO(DS)} \leq \sup_{\alpha > 0} \frac{G_{DI}(\alpha) \bigwedge \sqrt{2}}{F_{lower}(\alpha)}. \quad (8.17)$$

We claim that the right-hand side of Equation (8.17) is no greater than 1.79. Indeed, according to Xin (2021a), the first term $G_{DI}(\alpha)$ in the numerator is exactly the scaled cost incurred by the optimal base-stock policy in the single-sourcing lost-sales model, the second term $\sqrt{2}$ in the numerator is exactly the scaled cost incurred by the optimal constant-order policy in the lost-sales model, and the denominator $F_{lower}(\alpha)$ is also exactly the lower bound derived for the lost-sales model. Therefore, we conclude that the right-hand side of (8.17) is exactly the derived upper bound on the interested ratio for the single-sourcing lost-sales model, which is equal to 1.79 from Xin (2021a). The proof of Theorem 8.5 is completed.

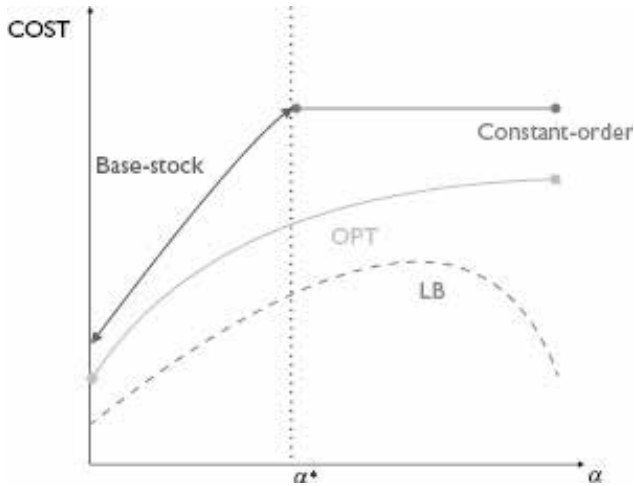
Recall that we discussed the connection between the dual-sourcing and single-sourcing lost-sales systems in the discrete-time setting in Section 8.2.2.1. Through Theorem 8.5, we build a deeper connection and prove that after both L and c go to infinity in the analogous continuous-review setting, the third assumption above (i.e., the large penalty-cost setting) indeed presents the worst-case scenario, and the worst-case ratio 1.79 is the same as the one in the single-sourcing lost-sales system.

8.3.3.2 1.79-approximation of capped base-stock policies in the single-sourcing lost-sales model

Finally, to explain ratio 1.79 of Theorem 8.5 at a high level, we briefly review the 1.79-approximation result of capped base-stock policies in the continuous-time single-sourcing lost-sales model. Similarly, customer demand arrives as a Poisson process with rate λ , the lead time is L , and the lost-sales penalty and holding costs are p and $h=1$ (without loss of generality), respectively. Under the assumption that $L=\alpha p$ for constant $\alpha>0$, which is similar in spirit to Assumption 1, Reiman (2004) conducts an asymptotic analysis under large L and p , and explicitly calculates the costs of the best base-stock and best constant-order policies. Reiman (2004) also proves that the best constant-order policy outperforms the best base-stock policy when the lead time reaches a certain threshold, and explicitly characterizes this threshold asymptotically. More specifically, there exists a critical threshold α^* such that the best constant-order policy outperforms the best base-stock policy iff $\alpha > \alpha^*$, and this critical threshold is approximately 0.69787. Building on it, Xin (2021a) proves that the minimum between the cost incurred by the best constant-order policy and the cost incurred by the best base-stock policy achieves a 1.79-approximation. As an immediate consequence, the best capped base-stock policy achieves a 1.79-approximation as well. We provide a proof sketch in Figure 8.1. Regarding whether the capped base-stock policy can still achieve a constant-ratio approximation in the non-asymptotic setting, the numerical results by Xin (2021a) are promising and encouraging. We conjecture that it still holds and leave further investigation of this conjecture as an important future research question.

8.3.4 Continuous-Time Dual Sourcing with Capacitated Sources

Continuous-time dual-sourcing from capacitated supply sources gives rise to production–inventory models where inventory can be replenished from two sources that are each modeled as single-server (or multi-server) queues. (Lead times thus become endogenous and stochastic as the flow time of a single-server queue.) A notable paper is by Song and Zipkin (2009), who offer an efficient exact evaluation tool (product form as described in Section 8.3.1) and an exact optimization method for the DI policy studied earlier by Moynzadeh and Schmidt (1991). Song and Zipkin (2009) also generalize the constant lead times considered by Moynzadeh and Schmidt (1991) to various stochastic lead times. Their generalization covers the following: (1) iid lead times, which include the constant lead times as a special case; (2) endogenous stochastic lead times, where the regular source consists of two nodes in tandem and each node is a Jackson network, whereas the express source consists only the second node; this model includes exponential or Erlang processing times (the time an order spent at each node contains both the queue and processing time), as special cases; and (3) exogenous, sequential stochastic lead times, which also include constant lead times as a special case. Song et al. (2022) further extend the study to multiple sources with expediting.



Notes: The upper bound is the minimum of the cost incurred by the best base-stock policy and the cost incurred by the best constant-order policy. We show that the former cost (the red solid curve with arrow ends) as well as the optimal cost (the yellow solid curve with square ends) grows with α (interpreted as the proxy of the lead time after scaling up both the lead time and penalty cost), but the latter cost (the blue solid curve with round ends) does not grow with α . Hence, the worst-case ratio of the upper bound to the optimal cost can only be achieved when $\alpha \in [0, \alpha^*]$. We next relax the optimal cost to the lower bound (the green dashed curve), which is similar to the lower bound in Equation (8.15), and compare the upper bound (interpreted as the cost incurred by the best base-stock policy under the lost-sales system) with the lower bound (interpreted as the cost incurred by the best base-stock policy under the analogous backlogged system with an adjusted penalty cost). Finally, we prove that the ratio of the upper bound to the lower bound is monotone in α on $[0, \alpha^*]$ such that the worst-case ratio is achieved at exactly $\alpha = \alpha^*$, resulting in the worst-case ratio of 1.79.

Figure 8.1 We provide a proof sketch of the 1.79-approximation result for the single-sourcing lost-sales model by using this graph

Song et al. (2017) consider a special case of the dual-sourcing system of Song and Zipkin (2009) in the sense that the queueing network at each node only contains one server. A regular order goes through both nodes, whereas an express order skips the first node and goes through only the second node (again, the time an order spends at each node is the sum of the queue and processing time). They characterize the optimal policy, which consists of a constant threshold and a switching curve that is a function of N_2 (the number of outstanding orders at the second node). These policy parameters determine when to order and from which source. Using this insight, they develop a heuristic policy that outperforms the DI and TBS policies in a numerical study.

Given that many of the continuous-time models are not amenable to exact analysis, one can proceed in two common ways: (1) solve the exact problem numerically or via simulation, or (2) solve an approximate problem analytically. We review two papers adopting option (2). Bradley (2004) considers a production–inventory problem in which the inventory can be replenished from in-house production or through a subcontractor. The author constructs a Brownian

approximation of the optimal control problem, assuming that the manufacturer uses a Single-Index Dual-Base-Stock policy. Allon and Van Mieghem (2010) introduce the TBS policy to analyze a capacitated dual-sourcing problem in continuous time, where capacity is also optimized. By using only a single base-stock, their replenishment policy is simpler (essentially one-dimensional) and provides greater tractability. They present performance bounds on the optimal cost and prove that economic optimization brings the system into the so-called heavy-traffic regime. The theoretical significance is that heavy traffic is not assumed, but that it is the proven result of capacity optimization. From a practical perspective, their proposition guarantees that the system converges to a tractable Brownian limiting system as the demand rate $\lambda \rightarrow \infty$. The authors provide an analytic characterization of the asymptotically optimal TBS dual-sourcing policy, including its strategic allocation, base-stock level, and the expected cost.

8.4 FUTURE RESEARCH DIRECTIONS

We close this review by providing some potentially fruitful research directions.

First, our existing knowledge of dual sourcing can be used as a tool to deal with emerging technologies and topics. For instance, Song and Zhang (2020) study a spare-parts supply chain and model it as a hybrid inventory system in which spare parts can be sourced from a remote supplier or printed on demand by a local 3D printer with limited capacity. However, once the sourcing decision is made, a part is single-sourced. A more desirable strategy is to have 3D printing serve as a backup replenishment. That is, each part is dual-sourced. One complication is that different parts share the same finite printing capacity in the multiple spare-parts setting. Nonetheless, 3D printing serves as a backup plan and presents a potential solution to part stock-outs between replenishments.

Second, dual-sourcing research may benefit from further collaboration with practitioners. For example, Peng et al. (2012) propose a dual-mode equipment-procurement framework, which has been implemented at Intel Corporation. This framework combines dual-source procurement with option contracts in three layers: a contract negotiation layer, where the firm chooses the best combination of lead time and price for each mode from the supply contract menu; a capacity reservation layer, where the firm reserves total equipment-procurement quantities from the two supply modes before the planning horizon starts; and an execution layer, where the firm orders equipment from the two supply modes based on the updated demand information. The implementation of this framework has resulted in significant dollar savings.

Third, our understanding and the practice of dual sourcing may also benefit from embracing state-of-the-art machine-learning/deep-learning techniques or from studying dual-sourcing problems with non-stationary parameters (e.g., demands). Gijsbrechts et al. (2020) take an important step in this direction. They apply the Asynchronous Advantage Actor Critic (A3C) algorithm, a type of Deep Reinforcement Learning (DRL) algorithm. They show that DRL can match the performance of SI and TBS policies (yet CDI still outperforms DRL) and other approximate dynamic programming methods, with limited changes to the tuning parameters across all studied problems. Yet, the initial tuning remains computationally burdensome. The authors therefore recommend that “generating structural policy insight or designing specialized policies that are (ideally provably) near-optimal thus remains desirable.” It would be interesting to investigate how to combine our domain knowledge in dual sourcing with state-of-the-art machine-learning techniques.

Fourth, it is important to go beyond the classic model and explore alternative dual-sourcing modeling options. For example, in the pharmaceutical industry, the selections of dual suppliers are not based on cost efficiency, but on other aspects such as supply lead-time variability, yield uncertainty, and so on. Hence, we need new models that differentiate the two suppliers along these dimensions (beyond constant lead times and costs). In fact, dual sourcing in supply-chain management literature (outside inventory) is usually linked with supplier risk.

ACKNOWLEDGMENTS

The authors sincerely thank Jeannette Song for motivating us to write this chapter. The authors are also grateful to Joren Gijsbrechts, Woonghee Tim Huh, Ganesh Janakiraman, Jeannette Song, and Li Xiao for stimulating discussions.

NOTE

1. “Strategic allocations” specify how the average total sourcing volume is allocated to both sources.

REFERENCES

- Allon, G., & Van Mieghem, J. A. (2010). Global dual sourcing: Tailored base-surge allocation to near- and offshore production. *Management Science*, 56(1), 110–124.
- Boute, R. N., Disney, S. M., Gijbrechts, J., & Van Mieghem, J. A. (2021). Dual sourcing and smoothing under non-stationary demand time series: Re-shoring with speedfactories. *Management Science*, Forthcoming.
- Boute, R. N., & Van Mieghem, J. A. (2015). Global dual sourcing and order smoothing: The impact of capacity and lead times. *Management Science*, 61(9), 2080–2099.
- Bradley, J. R. (2004). A Brownian approximation of a production-inventory system with a manufacturer that subcontracts. *Operations Research*, 52(5), 765–784.
- Bu, J., Gong, X., & Yao, D. (2020). Constant-order policies for lost-sales inventory models with random supply functions: Asymptotics and heuristic. *Operations Research*, 68(4), 1063–1073.
- Chen, X., Stolyar, A. L., & Xin, L. (2019). Asymptotic optimality of constant-order policies in joint pricing and inventory control models. Available at SSRN: <https://ssrn.com/abstract=3375203>.
- DeCroix, G., Song, J.-S., & Zipkin, P. (2005). A series system with returns: Stationary analysis. *Operations Research*, 53(2), 350–362.
- Federgruen, A., Liu, Z., & Lu, L. (2021). Dual sourcing: Creating and utilizing flexible capacities with a second supply source. Working Paper, Columbia University.
- Feinberg, E. A. (2016). Optimality conditions for inventory control. *INFORMS Tutorials in Operations Research*, 14–45.
- Fleischmann, M., & Kuik, R. (2003). On optimal inventory control with independent stochastic item returns. *European Journal of Operational Research*, 151(1), 25–37.
- Fukuda, Y. (1964). Optimal policies for the inventory problem with negotiable leadtime. *Management Science*, 10(4), 690–708.
- Gijsbrechts, J., Boute, R. N., Disney, S. M., & Van Mieghem, J. A. (2021). Volume flexibility at responsive suppliers in reshoring operations. Working paper.
- Gijsbrechts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. (2020). Can deep reinforcement learning improve inventory management? Performance on dual sourcing, lost sales and multi-echelon problems. Available at SSRN: <https://ssrn.com/abstract=3302881>.

- Goldberg, D. A., Katz-Rogozhnikov, D. A., Lu, Y., Sharma, M., & Squillante, M. S. (2016). Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. *Mathematics of Operations Research*, 41(3), 898–913.
- Goldberg, D. A., Reiman, M. I., & Wang, Q. (2020). A survey of recent progress in the asymptotic analysis of inventory systems. *Production and Operations Management*, Forthcoming.
- Hua, Z., Yu, Y., Zhang, W., & Xu, X. (2015). Structural properties of the optimal policy for dual-sourcing systems with general lead times. *IIE Transactions*, 47(8), 841–850.
- Huh, W. T., Janakiraman, G., & Nagarajan, M. (2011). Average cost single-stage inventory models: An analysis using a vanishing discount approach. *Operations Research*, 59(1), 143–155.
- Jackson, J. R. (1963). Jobshop-like queueing systems. *Management Science*, 10(1), 131–142.
- Jagerman, D. (1974). Some properties of the erlang loss function. *Bell System Technical Journal*, 53(3), 525–551.
- Janakiraman, G., Seshadri, S., & Sheopuri, A. (2015). Analysis of tailored base-surge policies in dual sourcing inventory systems. *Management Science*, 61(7), 1547–1561.
- Karlin, S., & Scarf, H. (1958). Inventory models of the arrow-harris-marschak type with time lag. In K. J. Arrow, S. Karlin, H. Scarf (Eds.), *Studies in the mathematical theory of inventory and production*, chap. 9 (pp. 155–178). Stanford University Press.
- Levi, R., Janakiraman, G., & Nagarajan, M. (2008). A 2-approximation algorithm for stochastic inventory control models with lost-sales. *Mathematics of Operations Research*, 33(2), 351–374.
- Li, Q., & Yu, P. (2014). Multimodularity and its applications in three stochastic dynamic inventory problems. *Manufacturing & Service Operations Management*, 16(3), 455–463.
- Minner, S. (2003). Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics*, 81–82, 265–279.
- Moynadeh, K., & Schmidt, C. P. (1991). An (s-1, s) inventory system with emergency orders. *Operations Research*, 39(2), 308–321.
- Peng, C., Erhun, F., Hertzler, E. F., & Kempf, K. G. (2012). Capacity planning in the semiconductor industry: Dual-mode procurement with options. *Manufacturing & Service Operations Management*, 14(2), 170–185.
- Reiman, M. I. (2004). A new simple policy for a continuous review lost-sales inventory model. Unpublished manuscript.
- Rosenshine, M., & Obee, D. (1976). Analysis of a standing order inventory system with emergency orders. *Operations Research*, 24(6), 1143–1155.
- Schäl, M. (1993). Average optimality in dynamic programming with general state space. *Mathematics of Operations Research*, 18(1), 163–172.
- Scheller-Wolf, A., Veeraraghavan, S., & van Houtum, G.-J. (2007). Effective dual sourcing with a single index policy. Working paper, Carnegie Mellon University, Pittsburgh, PA.
- Sheopuri, A., Janakiraman, G., & Seshadri, S. (2010). New policies for the stochastic inventory control problem with two supply sources. *Operations Research*, 58(3), 734–745.
- Song, J.-S., Xiao, L., Zhang, H., & Zipkin, P. (2017). Optimal policies for a dual-sourcing inventory problem with endogenous stochastic lead times. *Operations Research*, 65(2), 379–395.
- Song, J.-S., Xiao, L., Zhang, H., & Zipkin, P. (2022). Smart policies for multisource inventory systems and general tandem queues with order tracking and expediting. *Operations Research*, 70(4), 2421–2438.
- Song, J.-S., & Zhang, Y. (2020). Stock or print? impact of 3d printing on spare parts logistics. *Management Science*, 66(9), 3860–3878.
- Song, J.-S., & Zipkin, P. (2009). Inventories with multiple supply sources and networks of queues with overflow bypasses. *Management Science*, 55(3), 362–372.
- Sun, J., & Van Mieghem, J. A. (2019). Robust dual sourcing inventory management: Optimality of capped dual index policies and smoothing. *Manufacturing & Service Operations Management*, 4(21), 713–948.
- Svoboda, J., Minner, S., & Yao, M. (2021). Typology and literature review on multiple supplier inventory control models. *European Journal of Operational Research*, 293(1), 1–23.
- Van Mieghem, J. A. (2008). *Operations strategy: Principles and practice*. Dynamic Ideas.
- Veeraraghavan, S., & Scheller-Wolf, A. (2008). Now or later: A simple policy for effective dual sourcing in capacitated systems. *Operations Research*, 56(4), 850–864.

- Whittemore, A. S., & Saunders, S. C. (1977). Optimal inventory under stochastic demand with two supply options. *SIAM Journal on Applied Mathematics*, 32(2), 293–305.
- Xin, L. (2021a). 1.79-approximation algorithms for continuous review single-sourcing lost-sales and dual-sourcing inventory models. *Operations Research*, Forthcoming.
- Xin, L. (2021b). Asymptotic analysis of a remanufacturing system with non-identical lead times. Available at SSRN: <https://ssrn.com/abstract=3760906>.
- Xin, L. (2021c). Understanding the performance of capped base-stock policies in lost-sales inventory models. *Operations Research*, 69(1), 61–70.
- Xin, L., & Goldberg, D. A. (2016). Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Operations Research*, 64(6), 1556–1565.
- Xin, L., & Goldberg, D. A. (2018). Asymptotic optimality of tailored base-surge policies in dual-sourcing inventory systems. *Management Science*, 64(1), 437–452.
- Xin, L., He, L., Bewli, J., Bowman, J., Feng, H., & Qin, Z. (2017). On the performance of tailored base-surge policies: Theory and application at walmart.com. Available at SSRN: <https://ssrn.com/abstract=3090177>.
- Zipkin, P. (2000). *Fundamentals of inventory management*. McGraw Hill.

9. Assemble-to-order systems

Levi DeValve, Jing-Sheng Jeannette Song, and Yehua Wei

9.1 INTRODUCTION

Assemble-to-order (ATO) systems are a manufacturing approach to reduce inventory while increasing market responsiveness. Under ATO, firms assemble multiple types of products upon receiving customer requests, so finished goods inventory is completely eliminated. It also reduces component inventory through *common components*, each of which is shared by several products. This allows the component inventory to be flexibly allocated to products after the demands arrive, realizing the benefits of risk pooling. With advanced information and production technologies, more and more manufacturing companies, such as Dell, Lenovo, and BMW, have implemented ATO. In a broader sense, the inventory planning and order fulfillment of e-retailing companies, such as Amazon, can be viewed as an ATO system due to the presence of multi-item orders.

To fully reap the benefits of ATO systems, it is crucial to manage component inventory effectively. This entails two decisions: component inventory replenishment and common component inventory allocation. Because each product is assembled from several components, the shortage of any component can delay the product delivery. Therefore, it is desirable to have a coordinated component replenishment policy. Heterogeneous replenishment lead times complicate such coordination. Similarly, the decision of whether to allocate a unit of common component inventory to a specific product needs to trade off the revenues and waiting costs of all products that share this component. The more common components there are, the more complex the policy.

Given these challenges, the structure of the optimal inventory policy for ATO systems remains largely unknown and is surely complex. Scholars have been actively searching for efficient and effective approximate control policies. A survey by Song and Zipkin (2003) covered the works before 2003, and one by Atan et al. (2017) summarized the studies up to 2017. In addition, Goldberg et al. (2019) included some of the recent asymptotic analyses. Most of the earlier works have assumed a simple and commonly used allocation policy, and studied performance evaluation and optimization of a given class of replenishment policies. More recently, studies on the structure of optimal or near-optimal inventory policies in various ATO models have begun to emerge. These studies have also provided fresh perspectives on managing these systems. In this chapter, instead of covering the entirety of the vast ATO literature, we complement the existing surveys by outlining the critical intuition and ideas in the developments of optimal or near-optimal policies, with an emphasis on the literature that appeared in the last 15 years. We start by covering the single-period model (Section 9.2), and then move to dynamic models (Section 9.3). Finally, we discuss several future research directions (Section 9.4).

We adopt the following set of notations. The bold letters are reserved for vectors and matrices. For two vectors $\mathbf{x}$ and $\mathbf{y}$ of the same dimension, we use $\mathbf{x} \cdot \mathbf{y}$ to represent their dot/inner product. For any ATO system, we let m and n denote the number of components and the number

of products, respectively, and $\mathbb{Z}$ and $\mathbb{R}$ ($\mathbb{Z}_+$ and $\mathbb{R}_+$) denote the set of (non-negative) integers and reals, respectively. The assembly structure, commonly known as the Bill-of-Materials, is described by an m by n matrix $\mathbf{A}$, with $A_{ij} \in \mathbb{Z}_+$ representing the number of units of component (type) i needed to assemble product (type) j . For the special case of a binary matrix $\mathbf{A}$, the assembly structure can be represented by a bipartite graph where the sets of components and products form the bipartition, and $A_{ij} = 1$ denotes an edge between component i and product j . In that case, the sets $N(i)$ and $N(j)$ represent the neighbors of a component node and product node in the graph, respectively. As an example, the *M-system* is depicted in Figure 9.1, consisting of two components (represented by triangles) and three products (represented by circles).

There are three sets of cost parameters: the unit inventory ordering cost for each component i , denoted by c_i ; the unit inventory salvage/holding cost for each component i , denoted by h_i ; and the unit shortage (backorder) cost for each product j , denoted by b_j . In addition, the procurement lead time for component i is L_i .

9.2 ONE-PERIOD MODELS

In this section, we focus on a general two-stage stochastic programming (SP) formulation for one-period ATO models. As we shall discuss throughout the chapter, such models may have direct implications for more sophisticated dynamic models.

Consider the following one-period, two-stage ATO model. In the first stage, which occurs at the beginning of the period, component inventories are ordered and received in anticipation of demand. The component inventory vector is denoted by $\mathbf{y} \in \mathbb{Z}_+^m$, and we note that the procurement lead times are not considered in the one-period model. During the period, demand $\mathbf{d} \in \mathbb{Z}_+^n$ for different products is realized, where d_j represents the demand for product j . In the second stage, which occurs at the end of the period, the components are allocated to the (realized orders of) products for assembly. Let x_j denote the number of product j assembled and z_j the demand shortage of product j after assembly.

The second-stage allocation-decision problem can be then written as

$$\begin{aligned} G(\mathbf{y} \mid \mathbf{d}) &= \min_{\mathbf{x}, \mathbf{z}} \mathbf{b} \cdot \mathbf{z} + \mathbf{h} \cdot (\mathbf{y} - \mathbf{A}\mathbf{x}) \\ \text{s.t.} \quad &\mathbf{A}\mathbf{x} \leq \mathbf{y}, \\ &\mathbf{x} + \mathbf{z} = \mathbf{d}, \\ &\mathbf{x}, \mathbf{z} \in \mathbb{Z}_+^n. \end{aligned} \tag{9.1}$$

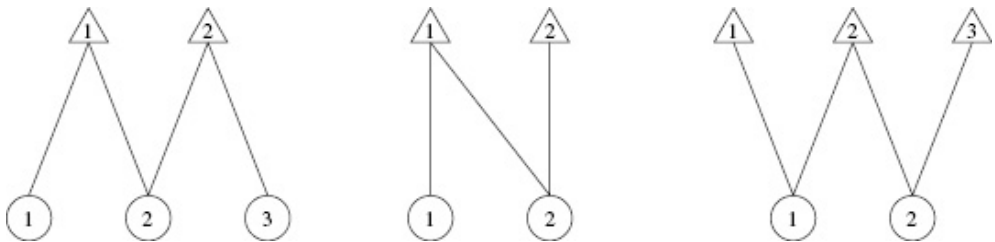


Figure 9.1 The *M-System*, *N-System*, and *W-system* (left to right)

Let D_j denote the stochastic demand for product j with integral support. Then, the first-stage replenishment-decision problem can be written as

$$\min_{\mathbf{y}} \mathbf{c} \cdot \mathbf{y} + \mathbb{E}_{\mathbf{D}}[G(\mathbf{y} \mid \mathbf{D})], \text{ s.t. } \mathbf{y} \in \mathbb{Z}_+^m. \quad (9.2)$$

Equation (9.2) is the standard one-period formulation in the ATO literature, as discussed by Song and Zipkin (2003). An alternative one-period formulation is to maximize the profit instead of minimizing the cost. In that case, the profit is defined as

$$\hat{G}(\mathbf{y} \mid \mathbf{d}) - \mathbf{c} \cdot \mathbf{y}, \text{ where } \hat{G}(\mathbf{y} \mid \mathbf{d}) = \mathbf{p} \cdot \mathbf{x} - \mathbf{h} \cdot (\mathbf{y} - \mathbf{A}\mathbf{x}),$$

for some unit profit vector $\mathbf{p}$. It can be easily checked that the optimal solution of Equation (9.2) coincides with the profit maximization formulation, when $\mathbf{b} = \mathbf{p}$. Nevertheless, because the exact optimal solution is in general difficult to obtain, it can be fruitful to consider different formulations, which may motivate different heuristics and lead to different theoretical analyses. In addition, the effectiveness of these heuristics often depends on the actual application.

We next describe another (re)formulation for the one-period ATO model, which as we will see, naturally lead to a group of ATO inventory policies. The reformulation rewrites the second-stage problem without allocation variables $\mathbf{x}$, as $\mathbf{x}$ is simply the difference between the demand ($\mathbf{d}$) and the demand shortage ($\mathbf{z}$). Specifically, let $\tilde{\mathbf{b}} = \mathbf{b} + \mathbf{h}\mathbf{A}$, and observe that $G(\mathbf{y} \mid \mathbf{d}) = \tilde{G}(\mathbf{y} \mid \mathbf{d}) + \mathbf{h} \cdot (\mathbf{y} - \mathbf{A}\mathbf{d})$, where $\tilde{G}(\mathbf{y} \mid \mathbf{d})$ is defined as:

$$\begin{aligned} \tilde{G}(\mathbf{y} \mid \mathbf{d}) &= \min_{\mathbf{z}} \tilde{\mathbf{b}} \cdot \mathbf{z} \\ \text{s.t.} \quad &\mathbf{A}(\mathbf{d} - \mathbf{z}) \leq \mathbf{y}, \\ &\mathbf{z} \leq \mathbf{d}, \\ &\mathbf{z} \in \mathbb{Z}_+^n. \end{aligned} \quad (9.3)$$

In addition, let $\tilde{\mathbf{c}} = \mathbf{c} + \mathbf{h}$, then Equation (9.2) can be rewritten as

$$-\mathbf{h} \cdot \mathbf{A}\mathbb{E}[\mathbf{D}] + \min_{\mathbf{y}} \tilde{\mathbf{c}} \cdot \mathbf{y} + \mathbb{E}_{\mathbf{D}}[\tilde{G}(\mathbf{y} \mid \mathbf{D})], \text{ s.t. } \mathbf{y} \in \mathbb{Z}_+^m. \quad (9.4)$$

While Equation (9.4) is equivalent to Equation (9.2) (in terms of the optimal solutions), it is also simpler, as the optimization problem in Equation (9.4) contains fewer number of decision variables. The simpler structure in Equation (9.4) has proved to be helpful in the development of an approximation algorithm for the one-period ATO formulations, which we discuss in Sections 9.2.2.1 and 9.2.2.2.

The one-period ATO model is appropriate for studying systems where the leftover inventory after assembly is either salvaged or discarded. It can also be used to solve lost sales or backlogging discrete-time dynamic models without lead time, as shown by Van Mieghem and Rudi (2002). Finally, the model has served as a foundation of many heuristics for the more general dynamic ATO systems with component lead times, where the second-stage problem is viewed as a proxy to the value-to-go after the component inventory decisions are made. Examples of such

heuristics include Akçay and Xu (2004), Lu and Song (2005), and Van Jaarsveld and Scheller-Wolf (2015). In addition to numerical observations, the validity of heuristics based on one-period ATO models was theoretically demonstrated for dynamic ATO systems with long component lead times, see, e.g., Doğru et al. (2010), Reiman and Wang (2012), and Doğru et al. (2017).

9.2.1 Computational Challenges for the One-Period Formulations

Although the one-period ATO models are significantly simpler than their dynamic counterparts, they are far from easy to solve. From the computational complexity perspective, the second-stage allocation problem for arbitrary fixed inventory vector $\mathbf{y}$ is NP-hard, as it reduces to a generalization of the set-cover (DeValve et al., 2020) or the multidimensional knapsack problem (Akçay & Xu, 2004), depending whether the ATO formulation is minimizing costs or maximizing profits. In particular, the set-cover problem is not just NP-hard to solve, but also NP-hard to approximate with better than an $O(\log M)$ factor (Feige, 1998). Surprisingly, when the inventory vector $\mathbf{y}$ is not fixed, the minimization version of the ATO formulation such as Equation (9.4) yields algorithms with constant factors. Nevertheless, it is still NP-hard to be solved exactly, as shown by DeValve (2019).

It is important to note that due to the recent advances in off-the-shelf mixed-integer linear program (MILP) solvers such as CPLEX and Gurobi (see, e.g., Bixby, 2010), solving the second-stage allocation problem *given* a specific inventory $\mathbf{y}$ and demand scenario $\mathbf{d}$ has become computationally feasible for most of the practical instances. As a result, one may wonder if we can formulate Equation (9.4) as a single large-scale MILP and solve it through off-the-shelf solvers directly. More formally, as $\mathbf{D}$ has integral support, we can let Ω denote the (possibly infinite) discrete set consisting of all possible stochastic demand scenarios and π_ω be the probability demand scenario $\mathbf{d}_\omega$ occurs. Then, we can rewrite Equation (9.4) (while replacing $\tilde{\mathbf{c}}, \tilde{\mathbf{b}}$ with $\mathbf{c}, \mathbf{b}$ for notational simplicity) as:

$$\begin{aligned} & -\mathbf{h} \cdot \mathbf{A} \mathbb{E}[\mathbf{D}] + \min_{\mathbf{y}, \mathbf{z}_\omega} \mathbf{c} \mathbf{y} + \sum_{\omega \in \Omega} \pi_\omega \mathbf{b} \cdot \mathbf{z}_\omega \\ \text{s.t.} \quad & \mathbf{z}_\omega \leq \mathbf{d}_\omega, \omega \in \Omega \\ & \mathbf{y}, \mathbf{z} \geq 0, \text{integer.} \end{aligned} \tag{9.5}$$

In practice, the MILP represented in Equation (9.5) is difficult to solve even for moderate sizes of m and n , as the number of variables and constraints quickly explodes with the size of Ω . Note that we may reduce the size of Ω through some sacrifices on the solution quality using the classical sample average approximation approach (see, e.g., Birge & Louveaux, 2011; Shapiro et al., 2009). Unfortunately, if Equation (9.5) is approximated with a large number of samples, the formulation would still be challenging to solve with the state-of-the-art solvers; and if Equation (9.5) is approximated with a small number of samples, the solution obtained from the approximated formulation may not be effective for the actual problem.

9.2.2 Heuristics

The difficulty in solving the one-period ATO models exactly has motivated various heuristics. In this section, we describe heuristics from three categories, decomposition, fractional rounding and discrete convex optimization.

9.2.2.1 Decomposition

A natural method employed to deal with the complexities of solving the one-period ATO model is to consider what a “good” inventory level would be for each component on its own, assuming the availability of other components. The key idea of this approach is that assuming the availability of other components greatly simplifies the second-stage allocation problem (i.e., we only need to compare total demand for a component with its inventory), and decomposes the problem into separate newsvendor problems for each component. This approach has been widely studied due to its intuitive appeal and simplicity, and has gone under various names including “item/component-based solution” (Lu & Song, 2005), “ignoring simultaneous stock-outs” (Van Jaarsveld & Scheller-Wolf, 2015), and “newsvendor decomposition” (DeValve et al., 2020). For the sake of brevity, we adopt the term newsvendor decomposition here to refer to this type of heuristic.

To formally describe the “newsvendor decomposition”, let the random variable $\mathbf{D}^C \in \mathbb{Z}_+^m$ denote the demand for components, with $\mathbf{D}^C = \mathbf{A}\mathbf{D}$. Then a newsvendor decomposition heuristic independently solves a newsvendor problem for each component using its demand distribution, i.e., the problem for component i uses the demand distribution D_i^C . The ordering cost in component i ’s newsvendor problem is c_i , the same as the original ATO formulation, and we let q_i denote the shortage cost for component i (we will discuss various methods for choosing q_i below). Given inventory $\mathbf{y}$ and demand $\mathbf{D}^C$, let the shortage cost across all components be denoted by

$$G^{NV}(\mathbf{y} \mid \mathbf{d}^C) = \sum_i q_i (d_i^C - y_i)^+ \quad (9.6)$$

Then, the newsvendor decomposition heuristic solves

$$\min_{\mathbf{y}} \mathbf{c} \cdot \mathbf{y} + \mathbb{E}_{\mathbf{D}}[G^{NV}(\mathbf{y} \mid \mathbf{D}^C)], \text{ s.t. } \mathbf{y} \in \mathbb{Z}_+^m. \quad (9.7)$$

The advantage of Equation (9.7) is that the shortage costs in Equation (9.6) are additively separable across the components, and so Equation (9.7) decomposes into m separate newsvendor problems, one for each component. In particular, the newsvendor problem for component i is

$$\min_{y_i} c_i y_i + q_i \mathbb{E}[(D_i^C - y_i)^+]$$

Since each newsvendor problem is a one-dimensional convex problem with a well-understood solution (i.e., inventory should be set to the “critical” fractile $(q_i - c_i) / q_i$ of the demand distribution), this greatly simplifies both the computation and interpretability of the solution.

Let $\mathbf{y}^{NV} \in \mathbb{Z}_+^m$ denote a solution to the newsvendor problem in Equation (9.7). Then the newsvendor decomposition heuristic simply uses $\mathbf{y}^{NV}$ in the original one-period ATO problem in Equation (9.5). It is clear that the ordering cost $\mathbf{c} \cdot \mathbf{y}^{NV}$ is the same in the newsvendor problem (9.7) and the original problem in Equation (9.5), so the main question is how well the newsvendor shortage costs, $\mathbb{E}_{\mathbf{D}}[G^{NV}(\mathbf{y}^{NV} \mid \mathbf{D}^C)]$, approximate the true expected shortage costs, $\sum_{\omega \in \Omega} \pi_{\omega} \mathbf{b} \cdot \mathbf{z}_{\omega}(\mathbf{y}^{NV})$, where $\mathbf{z}_{\omega}(\mathbf{y}^{NV})$ denotes the optimal $\mathbf{z}_{\omega}$ decision given inventory vector $\mathbf{y}^{NV}$ in Equation (9.5),¹ and this motivates various choices of $\mathbf{q}$ proposed in the literature.

In the work of Lu and Song (2005), a few alternatives are suggested for setting q_i . Letting $\boldsymbol{\mu} = \mathbb{E}[\mathbf{D}]$ denote expected product demand, and $\boldsymbol{\mu}^C = \mathbb{E}[\mathbf{D}^C] = \mathbf{A}\boldsymbol{\mu}$ denote the expected

component demand, they note a natural approach is to set the following shortage cost for component i

$$q_i^w = \sum_j a_{ij} \frac{\mu_j b_j}{\mu_i^c}$$

The shortage cost q_i^w can be thought of as the result of first “splitting” the shortage cost of each product among all the components it uses proportional to demand, then adding this up across all products i is used in. This is an intuitive way to split up the shortage costs, but Lu and Song (2005) noted that it doesn’t always provide the best performance, and so suggest a number of other ways to set the shortage costs based on insights derived from their demand model. One such method is the simple adjustment to q_i^w that sets the following shortage cost for component i

$$q_i^u = \sum_j a_{ij} \frac{\mu_j b_j}{\mu_i^c} - c_i.$$

In a model that assumes D^c following a multivariate Poisson distribution with a certain correlation structure, Lu and Song (2005) proved that using the costs q_i^u for all i in Equation (9.7) gives a solution y^{NV} which is a component-wise upper bound on the optimal solution of the original problem in Equation (9.5). This is helpful for their implementation of a greedy procedure taking advantage of discrete convexity to compute the true optimal solution under their demand assumptions. They also show in numerical simulations that this shortage cost definition (and others they propose) provide good bounds on the optimal solution.

The work of DeValve et al. (2020) suggests another method for general demand distributions. The main idea is to choose q_i for all i so that $\sum_i a_{ij} q_i \leq b_j$ for all j . This leads to an intuitive result that the optimal value of Equation (9.7) provides a lower bound on the optimal value of Equation (9.5). A few methods for choosing such q_i are proposed, and the lower bound result allows for comparing the heuristic with an optimal solution. The main results are that the newsvendor decomposition heuristic can be a multiplicative factor $\max_j \sum_i a_{ij}$ away from the optimal solution in the worst case, a bound which is shown to be tight for this and other newsvendor decomposition heuristics.

The study of Van Jaarsveld and Scheller-Wolf (2015) suggests yet another method for newsvendor decompositions that follows a slightly generalized framework. Here a separate newsvendor problem is still solved for each component i assuming the demand distribution D_i^c , however, rather than assuming a constant unit shortage cost of q_i , the shortage costs for component i is selected based on the original shortage costs b_j used for each product j . This can also be thought of as relaxing the second-stage problem in (9.3) to allow a separate shortage decision $\mathbf{z}_i$ for each component i (where $\mathbf{A}_i$ denotes the i^{th} row of $\mathbf{A}$):

$$\begin{aligned} \tilde{G}_i^{NV}(y_i \mid \mathbf{d}) &= \min_{\mathbf{z}_i} \sum_i \tilde{\mathbf{b}} \cdot \mathbf{z}_i \\ \text{s.t. } \mathbf{A}_i(\mathbf{d} - \mathbf{z}_i) &\leq y_i, \\ \mathbf{z}_i &\leq \mathbf{d}, \\ \mathbf{z}_i &\in \mathbb{Z}_+^n, \end{aligned} \tag{9.8}$$

then the resulting newsvendor problem for component i is

$$\min_{y_i} c_i y_i + \mathbb{E}_{\mathbf{D}}[\tilde{G}_i^{NV}(y_i | \mathbf{D})], \text{ s.t. } y_i \in \mathbb{Z}_+. \quad (9.9)$$

For binary matrix $\mathbf{A}$, Equation (9.9) can be solved efficiently for each i and again the resulting vector of inventory levels can be used as a heuristic in the original problem. In Van Jaarsveld & Scheller-Wolf (2015), this heuristic is observed to perform well relative to optimal when the shortage penalty costs are large relative to holding costs, roughly corresponding to a regime with high service levels. Intuitively, it is reasonable to expect such a newsvendor approach to work well in this regime, because with high service levels, the probability of simultaneous stock-outs is low, and thus the approximation above should be good.

9.2.2.2 Fractional rounding

Aside from decomposition, solutions for the one-period ATO formulations may also be obtained via rounding schemes. Rounding is commonly applied for solving an optimization problem with integral variables. Generally speaking, it first finds a fractional solution by relaxing the integrality constraints, and then converts it into an integral solution under some appropriate rounding. For the sake of succinctness, we present all rounding schemes for the one-period ATO model under formulation (9.5).

First, we note that Equation (9.5) is significantly easier to solve when $\mathbf{y}$ and $\mathbf{z}$ are allowed to be fractional. Without the integrality constraints, Equation (9.5) reduces to a stochastic linear program, which can be effectively handled through a number of other standard stochastic programming algorithms such as stochastic subgradient descent, sample average approximation, and Benders decomposition (see, e.g., Birge & Louveaux, 2011). Therefore, we can quickly find a good fractional solution through the stochastic linear program, and then attempt to round it into a feasible integral solution. Theoretically, one can use a stochastic subgradient descent method to find a fractional solution for Equation (9.5) with an objective at most δ units higher than the optimal with probability $1 + \varepsilon$ in $O(\delta^{-2}\varepsilon^{-2})$ iterations (DeValve et al., 2020).

To introduce the rounding schemes, let $(\mathbf{y}^{LP}, \mathbf{z}_{\omega}^{LP})$ be a fractional solution to Equation (9.5). Next, we briefly describe three schemes which round $(\mathbf{y}^{LP}, \mathbf{z}_{\omega}^{LP})$ into integral solutions: In the first rounding scheme, the value of $\mathbf{y}$ is determined by rounding $\mathbf{y}^{LP}$ to the nearest integer. Then, for the demand shortage decision $\mathbf{z}_{\omega}$, the scheme ignores $\mathbf{z}_{\omega}^{LP}$ and instead fulfills demand $\mathbf{d}_{\omega}$ (hence determining the demand shortage) on a first-come-first-serve basis, where we assume demand is arriving sequentially unit by unit. For the second scheme, the value of $\mathbf{y}$ is determined by rounding down $\mathbf{y}^{LP}$, while the value of $\mathbf{z}_{\omega}$ for each ω is chosen by rounding up $\mathbf{z}_{\omega}^{LP}$. Finally, for the third scheme, $\mathbf{y}$ and $\mathbf{z}_{\omega}$ are selected as

$$y_i = \lfloor \alpha y_i^{LP} \rfloor, \quad (9.10)$$

$$z_{j\omega} = \min \left(\left\lfloor \frac{\alpha}{\alpha - 1} z_{j\omega}^{LP} \right\rfloor, d_{j\omega} \right), \quad (9.11)$$

for some $\alpha > 1$. We note that checking the feasibility of the first two rounding schemes is straightforward, but more complex for the third rounding scheme (see analysis of DeValve et al., 2020). Also, the performance of the third rounding scheme depends on

the parameter α , and the exact value of α can be optimized via a binary search for given problem instances.

The first rounding scheme is proposed by Van Jaarsveld & Scheller-Wolf (2015),² while the second and third schemes are proposed by DeValve et al. (2020). In numerical experiments, all three rounding schemes have been shown to have good performance, and generally perform better than the decomposition methods. We note in some circumstances, it is plausible that the rounding of the second-stage variables, $\mathbf{z}$, is not needed. This is because once scenario ω is realized, we only need to focus on the integer program in the second stage corresponding to ω . Such problems can be often solved via modern MILP solvers, which may offer some improvement to the rounding policies. Nevertheless, the rounding of the second-stage variables speeds up the second-stage problem significantly and offers useful structures for theoretical analysis.

The rounding schemes are also shown to have better theoretical guarantees than newsvendor decomposition methods. In DeValve et al. (2020), it is shown that the second rounding scheme is asymptotically optimal when the mean or mean absolute deviation of the component demand is scaled to infinity and is guaranteed to achieve a 2 approximation factor of the optimal objective for any problem instance.³ When the second and third rounding schemes are combined with some appropriate values of α , we can obtain a rounding policy that is guaranteed to achieve a 1.8 approximation factor.

We briefly describe the theoretical approach for deriving the approximation factor described by DeValve et al. (2020). The approximation factor is obtained through the cost of the rounded solution relative to the cost of the optimal fractional solution, which also provides a bound relative to the true optimal integer solution since the cost of the optimal fractional solution is less than the cost of the optimal integer solution. As a result, the approximation factor obtained is also known as the *integrality gap*. To establish the integrality gap (and hence approximation factor) for the second rounding scheme, DeValve et al. (2020) used the optimal dual variables of the fractional problem and their properties (e.g., dual feasibility and complementary slackness) to bound the cost of the rounded-up shortages. This in turn allows for the application of strong duality to derive that the rounded-up shortage costs are no more than the optimal primal cost of the fractional problem. In other words, the rounded second-stage solution has cost less than $\mathbf{c} \cdot \mathbf{y}^{LP} + \sum_{\omega \in \Omega} \pi_{\omega} \mathbf{b} \cdot \mathbf{z}_{\omega}^{LP}$. Meanwhile, since we rounded down the first-stage variables, their cost is less than the first-stage fractional solution, $\mathbf{c} \cdot \mathbf{y}^{LP}$. Thus, the total cost of the rounded solution is less than

$$2\mathbf{c} \cdot \mathbf{y}^{LP} + \sum_{\omega \in \Omega} \pi_{\omega} \mathbf{b} \cdot \mathbf{z}_{\omega}^{LP}, \quad (9.12)$$

which is clearly less than two times the optimal fractional cost. Note that the primal-dual analysis critically takes advantage of the fact that the first-stage cost can be used to help bound the second-stage cost, as it is NP-hard to approximate with better than an $O(\log M)$ factor the optimal solution of the second-stage problem for any fixed first-stage solution.

DeValve et al. (2020) identified another rounding scheme that improves the approximation factor to 1.8, by taking the minimum cost solution of the second and third rounding schemes. The analysis is based on the observation that the approximation factor 2 derived from the upper bound in Equation (9.12) is only multiplied by the first-stage cost, while the third rounding scheme with an appropriate α can yield a rounded solution with total cost no less than

$$\alpha \mathbf{c} \cdot \mathbf{y}^{LP} + \frac{\alpha}{\alpha - 1} \sum_{\omega \in \Omega} \pi_{\omega} \mathbf{b} \cdot \mathbf{z}_{\omega}^{LP}.$$

Finally, the minimum cost solution of the second and third rounding schemes achieves an approximation factor of 1.8 when α is selected to be 1.5.

We also note that when the number of demand scenarios is large, formulation (9.5) should be approximated via sample average approximation (SAA). In that case, it is advantageous to obtain multiple candidate solutions through multiple sample approximations, following the procedures such as the one outlined by Kleywegt et al. (2002). This idea was employed by Akçay and Xu (2004) in the ATO setting. For the SAA formulation, the rounding schemes described in this chapter can be applied the same way as the exact formulation, thus used to significantly speed up the SAA procedures when the number of products and components are large.

9.2.2.3 Discrete convex optimization

Alternatively, the one-period ATO formulations may be solved using established techniques in the discrete optimization problem literature. This approach for deriving solution methods was outlined by Zipkin (2016), who observed that for the general assembly structure, one may restrict $\mathbf{y}$ to a polyhedron to ensure that the constraints in the second-stage problem form a polymatroid, which would then imply that the objective corresponding to $\mathbf{y}$ is *cover- L^1 -convex* within the polyhedron. The notion of cover- L^1 -convexity is closely related to the concepts of L^1 -convex, a notion of discrete convexity developed in (Murota, 2003). When the objective function is L^1 -convex, the optimization problem can be solved through different variants of steepest descent or Lagrangian relaxation-based algorithms. As a first step toward establishing the validity of this approach, Zipkin (2016) conducted some numerical experiments to show that when $\mathbf{y}$ is restricted to ensure that the second-stage problem is a polymatroid optimization, there is only a small loss in the optimality of the solutions.

Some theoretical progress has been made via the discrete convexity approach for ATO systems with special structures. The two main classes of assembly structures that have been studied with this approach are the “tree families” of Zipkin (2016) and the “chained BOM” of Doğru et al. (2017). The key property shared by these classes is that the assembly structure is a *laminar set family*, i.e., sets with a hierarchical inclusion structure.⁴ The key difference is that tree families consider the sets of products used by the components, while chained BOMs consider the sets of components used by products (i.e., the sets defined by either the non-zero rows or columns of the assembly matrix $\mathbf{A}$).

The analysis of Zipkin (2016) shows that the laminar structure of tree families guarantees the second-stage allocation problem for any demand realization is a polymatroid optimization, which has a closed-form greedy solution. The properties of the closed-form second-stage solution are then used to show the objective as a function of the first-stage variables, $\mathbf{y}$, which is discrete convex. In Doğru et al. (2017), the authors directly analyze the second-stage allocation problem to show that the laminar structure allows the optimization to be partitioned into a series of greedy allocations, which they use to show that the objective is discrete convex. Thus, the analyses in Zipkin (2016) and Doğru et al. (2017) show that for these special laminar structures, known algorithms for discrete convex optimization can be applied. Finally, it has been observed by Doğru et al. (2017) that L^1 -convexity fails to hold for general assembly

structures. Thus, the application of L^h -convex optimization algorithms for general ATO problems remains an open problem.

9.2.3 Discussions on One-Period Models

In this section, we present a relatively standard one-period ATO model. There are several applications of one-period models with additional features, for which we refer interested readers to the survey of Atan et al. (2017). Because the standard one-period ATO model is already difficult to solve, the models with additional features typically focus on systems with more restrictive assembly structures, such as ATO systems with one or two end products. It would be interesting to see if the techniques developed in the standard ATO model can be extended to incorporate some of the additional features considered in the literature, and we will discuss some of these features in more detail in Section 9.4.

For the standard one-period ATO model introduced in this chapter, there are several interesting open problems for researchers to pursue. First, one may identify special ATO structures besides the laminar families in which the one-period model can be solved exactly in the future. Second, the performance guarantees for the fractional rounding schemes introduced here for the general one-period ATO models are not necessarily tight, and one may design algorithms with better approximation factors. Third, in some applications, there may be additional constraints on $\mathbf{y}$, motivated by settings where the component supplier has limited capacity, or there are already some on-hand inventories. It would be interesting to see how the existing tools can be applied to those problems. Finally, it would be interesting to study one-period models with nonlinear costs in either the first or second-stage costs. For example, an ATO system may have fixed setup costs for ordering component inventories or fixed penalty costs for not satisfying the entire demand for a product. See DeValve (2021) for a model of fixed costs and initial analysis.

We also note that the techniques for analyzing theoretical properties of the one-period ATO models (e.g., primal-dual analysis, discrete convexity) have proved to be fruitful in analyzing other dynamic inventory systems settings, (see, e.g., Levi et al., 2008; Chen & Li, 2021). Therefore, it would be interesting to see if these techniques can be leveraged to study other dynamic ATO models in the future.

9.3 DYNAMIC ATO MODELS

In this section, we discuss the dynamic ATO models. Different from single-period models, here we need to explicitly account for component replenishment lead times, which is the time it takes for the components to arrive after orders are placed. Unsatisfied product demand is either backlogged or lost. Leftover component inventories are carried over for future use.

Due to the presence of replenishment lead times, the structure of the optimal control policy for dynamic ATO models is very complex and is largely unknown until recently. Therefore, the majority of early works focus on specific classes of policies that are commonly seen in practice. In particular, for continuous-review models, most papers assume the first-come, first-served (FCFS) allocation rule, and develop tools to evaluate and optimize the independent base-stock (IBS) replenishment policies. See, for example, Song (1998, 2002), Song et al.

(1999), Song and Yao (2002), Lu et al. (2003, 2005), Lu and Song (2005), Zhao and Simchi-Levi (2006), and Van Jaarsveld and Scheller-Wolf (2015). For periodic-review models, because the demand during each period is batched and filled at the end of the period, different allocation rules have been considered. These include a fixed priority rule (Zhang, 1997), the FCFS rule (Hausman et al., 1998), and a fair-share rule Agrawal and Cohen (2001). Akçay and Xu (2004) studied a product-based allocation rule that makes optimal or near-optimal allocation decisions within each period. All these studies apply FCFS to demand between periods. In other words, these allocation rules considered in the periodic-review models, if implemented in a continuous-review environment, reduce to FCFS.

Even under these specific classes of policies, the evaluation of key performance metrics, such as the order fill rate, the average inventory, and the average backorders, is computationally challenging, because they involve high-dimensional probabilities. Therefore, the main thrust of this stream of research is to develop computationally efficient performance evaluation tools to enable firms to quantify the trade-off between inventory and service and to develop efficient exact and approximate algorithms to compute the optimal base-stock levels to minimize the long-run average cost. We refer the reader to Song and Zipkin (2003) and Atan et al. (2017) for reviews of these developments.

Our focus below is on the recent progress in understanding the form of optimal and asymptotically optimal control policies with component replenishment lead times. We consider backlogging models in the first two subsections. We then discuss lost sales models in the third subsection.

Next, we present notations for a general continuous-review ATO system with backorders. We note that most of the results also hold in periodic-review systems. For any $t \geq 0$, the demand process of product $1 \leq j \leq n$ (and component $1 \leq i \leq m$) are denoted as

$$\mathcal{D}_j(t) = \text{cumulative demand of product } j \text{ in } (0, t],$$

$$\mathcal{D}_i^C(t) = \sum_j A_{ij} \mathcal{D}_j(t) = \text{cumulative demand for component } i \text{ in } (0, t].$$

We assume the demand process $\mathcal{D}_j(t)$ for any product j is stationary and has independent increments. This implies the demand process $\mathcal{D}_i^C$ for any component i is also stationary and has independent increments. For each component i , let L_i be the replenishment lead time for component i , a positive constant.

A control policy is determined by the replenishment of components and demand fulfilled over time. For any $t \geq 0$, $1 \leq i \leq m$ and $1 \leq j \leq n$, define

$$O_i(t) = \text{cumulative component } i \text{ orders in } [0, t], \text{ with } O_i(0) = 0,$$

$$X_j(t) = \text{cumulative product } j \text{ demands satisfied in } [0, t], \text{ with } X_j(0) = 0.$$

We call $O_i(t)$, $1 \leq i \leq m$, $t \geq 0$, an *order policy/process*, and $X_j(t)$, $1 \leq j \leq n$, $t \geq 0$, an *allocation policy/process*.

For any policy $\pi = \{(O_i(t), X_j(t)), t \geq 0\}$, let $I_i(t, \pi)$ denote the on-hand inventory of component i at time t , $B_j(t, \pi)$ denote the backorders for product j at time t , and $IP_i(t, \pi)$ denote the inventory position of component i , after ordering at time t . The dynamics of I_i , B_j and IP_i are:

$$I_i(t, \pi) = I_i(0) + O_i(t - L_i) - \sum_j A_{ij} X_j(t),$$

$$B_j(t, \pi) = B_j(0) + \mathcal{D}_j(t) - X_j(t),$$

$$IP_i(t, \pi) = IP_i(0) + O_i(t) - \mathcal{D}_i^C(t),$$

where $IP_i(0)$ is the initial inventory position. We say that policy π is admissible if it is adapted to $\mathcal{F}_t$ (where $\mathcal{F}_t$ denotes the natural filtration at time t), and satisfies the following inequalities for all times $t \geq 0$:

$$\sum_j A_{ij} X_j(t) \leq I_i(0) + O_i(t - L_i),$$

$$X_j(t) \leq B_j(0) + \mathcal{D}_j(t).$$

Let h_i and b_j denote the unit holding cost rate of component i and unit backorder cost rate of product j , respectively. Then the total inventory cost incurred at time t under π is

$$C(t, \pi) = \sum_i h_i \cdot I_i(t, \pi) + \sum_j b_j \cdot B_j(t, \pi). \quad (9.13)$$

The objective function for an admissible policy π in a dynamic ATO model is typically the total long-run average expected cost, defined as

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T C(t, \pi) dt \right]; \quad (9.14)$$

or the total discounted expected cost, for a discount rate $\beta > 0$, defined as

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\int_0^T e^{-\beta t} C(t, \pi) dt \right]. \quad (9.15)$$

While the total discounted expected cost is more realistic as the present is always worth more than the future, the total long-run average expected cost has the advantage of allowing for steady-state analysis, a tool that often provides more tractability to the problem.

9.3.1 Exact Optimal Policies for N- and W-Systems

For two specially structured ATO systems – the N- and W-systems, under a cost symmetry condition, the optimal policy can be completely characterized. Even though these cases are very special, they illuminate the complexity of the ATO systems, the need to synchronize decisions, and the connections to the classic single-item base-stock policy (i.e., the IBS policy), the FCFS rule, as well as the balanced base-stock (BBS) policy for the single-product

assembly systems. They also play a vital role in devising asymptotically optimal policies and effective heuristic policies for more complex systems.

For illustration purposes, we first describe the model and results of Lu et al. (2015) for the simplest N-system. In the N-system, there are only two components, $i = 1, 2$, and two products, $j = 1, 2$. Product 1 consists of one unit of component 1 only, and product 2 is an assembly of one unit of each component. Thus, component 1 is the common component shared by both products, while component 2 is the product 2 specific component. In the context of computers, for example, component 1 may be a hard disk and component 2 may be a memory card. The hard disk may be sold directly to a customer as a portable hard disk drive (product 1) or may be assembled together with the memory card into a laptop (product 2). In the context of automobiles, component 1 may be an engine and component 2 may be the body of the car. The engine may either be sold separately as product 1 or be assembled with the car body into a car (product 2). Clearly, the N-system possesses the basic elements of an ATO system – with both assembly and distribution structures. Thus, it serves as an important building block for general multi-product, multi-component ATO systems.

Next, denote $L = \min\{L_1, L_2\}$ and $\Delta_i = L_i - L$. Let

D_i^C = a generic random variable having the same distribution as $\mathcal{D}_i^C(L)$

ΔD_i^C = a generic random variable having the same distribution as $\mathcal{D}_i^C(\Delta_i)$.

For each product j , let $D_j(t) = \mathcal{D}_j(t - L, t]$,

D_j = a generic random variable having the same distribution as $\mathcal{D}_j(L)$.

Note that for the dynamic ATO model, we use $\mathcal{D}_i^C$ (or $\mathcal{D}_j$) to represent the demand processes for component i (or product j), and D_i^C or D_j to represent the corresponding distribution of the demand over the minimum lead time. We say that an admissible policy $\pi = \{(O_i(t), X_j(t)), t \geq 0\}$ is a no-holdback (NHB) policy if for all $t \geq 0$,

$$B_1(t, \pi) \times I_1(t, \pi) = 0, \quad B_2(t, \pi) \times \min\{I_1(t, \pi), I_2(t, \pi)\} = 0.$$

When π is NHB, we also call the allocation rule $\{X_j(t), t \geq 0\}$ an NHB allocation rule. Thus, under an NHB policy, a demand is backordered if and only if there is no on-hand inventory of at least one of its components.

The NHB allocation rule was first described and analyzed by Song and Zhao (2009), followed by Dođru et al. (2010) and Lu et al. (2010). Both FCFS and NHB are commonly seen in practice; Kapuściński et al. (2004) described an example of NHB at Dell and Xu et al. (2009) described an example of FCFS at Amazon.com.

The NHB rule describes a class of allocation rules, which allocates a component to a demand only if such an allocation will result in the fulfillment of the demand. When there are multiple types of backorders waiting for the same common component, however, how to allocate that component among different backorders (i.e., the backorder clearing rule) still needs to be specified. Song and Zhao (2009) assumed the FCFS backorder clearing rule, and therefore they call the allocation policy a modified FIFO (MFIFO) rule. Lu et al. (2010) showed that, for the N-system, W-system, and their generalizations, under the IBS replenishment policy and

a symmetric cost condition specified below, the NHB rule is sample-path optimal among all possible allocation rules, regardless of the type of backorder clearing rule. The IBS replenishment assumption is relaxed in Lu et al. (2015), who showed that for any given replenishment policy, a NHB allocation rule is optimal under the cost symmetry condition. Lu et al. (2015) further show that assuming the NHB rule is followed, the optimal replenishment policy is a *coordinated base-stock* (CBS) policy.

In the following, we further describe the results of Lu et al. (2015) for the N-system. We start by stating the symmetric cost condition used in their paper.

$$(\text{Cost symmetry}) \quad b_1 = b_2 + h_2 = c^{sym}. \quad (9.16)$$

Note that the cost symmetry condition indicates that backlogging a product 1 demand incurs the same cost as backlogging a product 2 demand while holding one unit of component 2 in inventory. Therefore, when component 2 is available, there is no immediate cost difference for the allocation of a unit of component 1 among different product demands. We can either use this unit to satisfy a product 1 demand (and hence backlog a unit of product 2, incurring cost $b_2 + h_2$), or use it to satisfy a product 2 demand (and hence backlog a product 1, incurring cost b_1). This cost structure is reasonable in certain situations in practice if product 1 serves an external market that enjoys a higher profit margin or faces a higher shortage cost, while product 2 serves the internal production, as an intermediate product (or a subassembly). For instance, in the computer example, product 1 can be a high-end portable hard disk competing with several substitutes in the market, while product 2 is used to assemble a laptop in a medium price range with a flexible delivery time window. In the automobile example, the engine (component 1) may be used to satisfy an emergency supply contract for a spare part that incurs a high delay penalty, while product 2 is used to assemble a car which can tolerate a longer delivery time.

Under the symmetric cost condition Equation (9.16), Lu et al. (2015) establishes that if the common component has a shorter lead time ($L_1 < L_2$), the optimal admissible policy constitutes a CBS replenishment policy and an NHB allocation rule, and a similar type of policy is optimal when the common component has a longer lead time ($L_1 > L_2$). When the lead times are identical ($L_1 = L_2 = L$), the CBS policy reduces to an IBS policy, i.e., the optimal policy constitutes an IBS replenishment policy and an NHB allocation rule. To provide some intuition for their analysis, we describe the CBS replenishment policy in the first case with $L_1 < L_2$.

In the case where $L_1 < L_2$, define $G(y_1, y_2)$ as the expected inventory (i.e., holding and back-order) cost for any fixed time t , given the inventory position of component i at $t - L_i$ is y_i , $i = 1, 2$. Formally,

$$G(y_1, y_2) = \sum_{\tau=1,2} h_{\tau} E(y_{\tau} - D_{\tau}) + (c^{sym} + h_1) E((D_1 - y_1)^+ \vee (D_2 - y_2)^+).$$

It can be shown that $G(y_1, y_2)$ is $L^{\natural}$ -convex. That is, it is jointly convex, submodular, and its Hessian matrix is a diagonally dominant M-matrix.

The structural property of $G(y_1, y_2)$ can be leveraged to show that the optimal CBS replenishment policy has the form $(\tilde{s}_1(\cdot), \tilde{s}_2)$, where $\tilde{s}_1(\cdot)$ is a one-dimensional function and $\tilde{s}_2$ is a singular value. Specifically, component 2 uses a standard base-stock policy with base-stock

level $\tilde{s}_2$, while component 1 adopts a state-dependent base-stock policy, and the base-stock level $s_1(\cdot)$ depends only on a one-dimensional state variable, namely, the aggregate demand for product 2 during the last $L_2 - L_1$ periods. The values of $(s_1(\cdot), \tilde{s}_2)$ can be fully characterized, and we refer interested readers to the paper of Lu et al. (2015).

The CBS policy is closely related to the well-known *balanced base-stock* policy, proved to be optimal by Rosling (1989) for the single-product assembly system that consists of product 2 only. It is noteworthy that a similar policy is optimal with the addition of a second final product, i.e., product 1. The combination of the CBS replenishment policy and the NHB rule is also optimal for the W-system and its generalizations under a similar cost symmetry condition; see Dođru et al. (2010) for identical deterministic lead times and Lu et al. (2015) for general deterministic lead times. Chen et al. (2021) show that this exact characterization breaks down for the M-system, but a CBS replenishment policy combined with a periodic-review priority (PRP) allocation rule is asymptotically optimal.

9.3.2 Asymptotically Optimal Policies

For the N- and W-systems when the costs are asymmetric, Lu et al. (2015) identified an asymptotically optimal policy when the product demand rates are high. Under the asymptotic regime, the optimal replenishment policy is a CBS policy, and the optimal allocation policy is the periodic-review NHB policy with a priority-based backorder clearing (PBC) rule. The periodic-review policy allocates components at discrete review points separated by an infinitesimal interval, which is similar to that in Plambeck and Ward (2006). Chen et al. (2021) extended this result to the M-system and its generalizations. Interestingly, the form of the asymptotically optimal policy is essentially the same as that of the optimal policy in the symmetric cost case for the simple-structured systems. In addition, Dođru et al. (2010) provided extensive discussions on the effectiveness of various backorder clearing rules, including PBC and priority with reservation (PR).

The dynamic ATO model with large component order lead time is studied in a series of works by Dođru et al. (2010), Reiman and Wang (2012, 2015), Dođru et al. (2017), and Reiman et al. (2021). For the ATO models with identical order lead times, Dođru et al. (2010) proposed a SP-based policy, in which the order and allocation decisions are made based on the solution of a one-period ATO SP model. The asymptotic optimality of the SP-based policy is established in Reiman and Wang (2015), with the order lead time being scaled to infinity. The policy and asymptotic analysis is later extended to non-identical lead times with a multi-stage SP-based policy (Reiman et al., 2021). For more details on ATO models with large lead times, we refer interested readers to the recent survey of Goldberg et al. (2019), which provides a comprehensive overview of both the SP-based policy and the theoretical approaches to establish asymptotic optimality.

We also note a stream of work by Plambeck and Ward (2006, 2007) and Plambeck (2008) on high-volume demand ATO production-inventory systems. In this stream, the decision variables include the component production capacities, assembled product production sequence, and product prices. Once a component's production capacity is chosen, the production facility produces the component at full capacity, so there is no inventory decision, and the ATO production-inventory system behaves similarly to a queueing network. Recently, Wan and Wang (2015) considered a system similar to that of Plambeck and Ward (2006) and derived asymptotically optimal continuous-review component allocation policies.

9.3.3 Lost Sales Model

Aside from ATO models with backordered demand, researchers have also studied ATO models with lost sales, where unsatisfied demand will be lost. The lost sales model with general component order lead time is arguably harder than the backorder model, as the optimal policy for even the single-item inventory model is not fully understood (Zipkin, 2008). As a result, progress in understanding the optimal/near-optimal policies for the lost sales model has focused on endogenous lead times, where an order is placed to a single-server production system with exponential processing times.

Benjaafar and ElHafsi (2006) considered the special case of a *single* end-product (i.e., $n = 1$) and multiple demand classes with lost sales. They assumed the supply process for each component is a single-server queue with exponentially distributed processing times. They showed that under Markovian assumptions on demand and production, the optimal replenishment policy is a *state-dependent base-stock policy*, and the optimal allocation rule is a multi-level *state-dependent rationing policy* that depends on the inventory levels of all other components. More specifically, the state-dependent base-stock policy orders component i whenever its inventory level (denoted by I_i) falls below some base-stock level $s_i(\mathbf{I}_{-i})$, where $\mathbf{I}_{-i}$ is the inventory vector for components not labeled i . Similarly, the state-dependent rationing policy satisfies the arriving demand for product j if the inventory vector is above the rationing level $\mathbf{r}_j(\mathbf{I})$, where $\mathbf{I}$ denotes the inventory vector for all components. Interestingly, in numerical studies, Benjaafar and ElHafsi (2006) found that setting a base-stock and rationing level for every state may not be necessary, as a heuristic policy with a constant rationing level and coordinated base-stock policy achieves very similar performance. ElHafsi et al. (2008) extended the above results to nested systems.

Analysis for the lost sales ATO model with exponential processing-time single-server supply systems and multiple end products is studied by Nadar et al. (2014) while allowing for batch production of components, but assuming at most one job in the supply system at any time. In Nadar et al. (2014), the authors study the “generalized M-system”, i.e., a system with n products and $n - 1$ components, where the first $n - 1$ products each use a single component, and the final product uses all of the components. They showed that the optimal policy for the generalized M-system is a combination of a *lattice-dependent base-stock policy* and a *lattice-dependent rationing policy*. To specify the lattice-dependent policies, first consider the one-dimensional sub-lattice over $\mathbb{Z}_+^m$, defined as $L(\mathbf{x}, \Delta) = \{\mathbf{x} + k\Delta : k \in \mathbb{Z}_+\}$, for some initial vector $\mathbf{x}$ and common difference vector Δ . For any fixed common difference vector Δ , observe that $\cup_{\mathbf{x}} L(\mathbf{x}, \Delta) = \mathbb{Z}_+^m$, and for any $\mathbf{x}, \mathbf{x}'$, we either have $L(\mathbf{x}, \Delta)$ and $L(\mathbf{x}', \Delta)$ do not intersect or one is contained in another. Thus, we can partition $\mathbb{Z}_+^m$ as a set of disjoint sub-lattices in the form of $L(\mathbf{x}, \Delta)$ for any fixed Δ , and each Δ defines a sub-lattice partition. A lattice-dependent base-stock policy orders a batch of component i whenever the inventory vector (for all components) is below the base-stock level $S_i(\mathbf{y})$ on the sub-lattice $L(\mathbf{x}, \Delta^i)$, where the superscript on Δ^i indicates that for different components the base-stock level may depend on different sub-lattice partitions. Similarly, a lattice-dependent rationing policy satisfies the arriving demand for product j if the inventory vector is above the rationing level $R_j(\mathbf{y})$ on the sub-lattice $L(\mathbf{x}, \tilde{\Delta}^j)$. Note that searching for the best common difference vectors Δ^i and $\tilde{\Delta}^j$ can be extremely time-consuming. In (Nadar et al., 2016), rules-of-thumb for choosing the common difference vectors were described and shown to work very well in numerical experiments.

The lattice-dependent ATO policy has proved to be successful in solving ATO systems outside of the generalized M-systems. In Nadar et al. (2016), computational experiments

indicate that the lattice-dependent ATO policy has superior performance compared to the state-dependent policies and constant backorder/rationing policies proposed by Benjaafar and ElHafsi (2006) in terms of both objective value and computation time. Moreover, in over 22,500 instances of the general problem tested, Nadar et al. (2016) found that the lattice-dependent policy, remarkably, matches the global optimal solution in all instances. Also, Nadar et al. (2018) showed that the lattice-dependent policy can be combined with the state-aggregation idea from approximate dynamic programming. In numerical studies, the authors showed that the approach can solve ATO systems with up to 22 components.

9.4 EXTENSIONS AND RESEARCH DIRECTIONS

In this section, we present several directions of active research that extend the existing ATO models. In addition, note that the structure of the ATO system, i.e., holding components and allocating a subset of components only after demand arrivals, is not unique to manufacturing settings. Therefore, we also identify several directions with the common theme that subsets of products share common resources (corresponding to components), which may benefit from the existing research on ATO models.

9.4.1 Online Resource Allocation with Component Commonality

A significant extension to the one-period ATO model is to have the components ordered once, while allowing for dynamic (or unknown) demand arrivals. In literature thus far, models in this spirit only considered special types of ATO assembly structures (e.g., Hsu et al., 2006; Bernstein et al., 2007, 2011). For general ATO structures, this extended ATO model is at the intersection of research with component commonality and online resource allocation. A prominent example is the class of network revenue management (NRM) problems, where demand arrives over time for a set of products that consume overlapping subsets of resources (see, e.g., Talluri & Van Ryzin, 2006 for a detailed problem description). A typical assumption in NRM models is that the capacity of resources is fixed before the time horizon and thus the primary decision is how to allocate these fixed resources over time. This perspective could be combined with the ATO problem to consider both investing in resource capacity (akin to ordering component inventory in ATO), as well as resource allocation over time. In particular, intuition from the ATO problem suggests that the ability to invest in more resource capacity (at a cost) may offset some of the difficulties of allocating resources, and exploiting this trade-off may lead to better-performing policies. In addition, it has now been documented that the dynamic allocation decisions can be approximated with a one-period allocation similar to our one-period ATO model, which is known as the offline allocation in the online resource allocation literature. Specifically, recent research has shown that dynamic allocation policies can achieve constant regret in performance compared to offline allocations (Jasin & Kumar, 2012; Bumpensanti & Wang, 2020; Vera et al., 2021). However, such results are asymptotic in nature, and it would be interesting to develop non-asymptotic theoretical guarantees with the addition of resource capacity decisions.

9.4.2 E-Retailing

The recent growth of large e-retailers has led to many interesting research directions. One direction that is particularly relevant to the ATO literature is the *multi-item fulfillment problem*, where

the e-retailer receives orders for multiple items, and then decides on how to fulfill its orders using available inventories (see, e.g., Xu et al., 2009; Acimovic & Graves, 2015; Jasin & Sinha, 2015). At a high level, a multi-item order in e-retailing fulfillment is similar to a product demand arrival in an ATO problem, where each item in the order is similar to a type of component. The main difference, however, is that an e-retailer typically has some flexibility in how they choose to fulfill multi-item orders, in terms of which warehouse they will ship the order from, whether they should “split” fulfillment of different items in the order among multiple warehouses, or not fulfill some of the items in the order (Jasin & Sinha, 2015). This difference is important, as even for the deterministic problem, one needs to solve an LP with a very large number of decision variables. To deal with the issue of large numbers of variables, we refer interested readers to Jasin and Sinha (2015), and the recent works of Ma (2022) and Amil et al. (2022).

In the real world, the e-retailer not only needs to solve the multi-item fulfillment problem, but also faces the decision of how to replenish their inventories over time. Acimovic and Graves (2017) study the problem with both fulfillment and replenishment, but only through separable approximations, which reduce the multi-item problems to multiple single-item problems. Thus far, few papers have studied in detail the multi-item problem with both fulfillment and replenishment, and we identify this as a potential direction where the analysis in the ATO literature can prove fruitful. For instance, assuming the FCFS allocation rule, Van Jaarsveld and Scheller-Wolf (2015) develop an efficient algorithm to compute the optimal IBS policy for industrial scale, dynamic ATO systems. Assuming the PRP allocation policy, Chen et al. (2021) present a scalable algorithm to obtain the optimal IBS policy for general dynamic ATO systems. Their approach is to decompose the ATO system into a set of distribution subsystems, each consisting of only one component and all products that require this component. They show that each distribution subsystem can be explicitly solved in a manner similar to a newsvendor solution. In addition, we note that general models including both component commonality and fulfillment flexibility are called *newsvendor networks* in the literature, and Van Mieghem and Rudi (2002) made early progress on replenishment policies for such systems, with recent generalizations made by DeValve and Myles (2022). Optimizing such models with integrality constraints, even in the one-period setting, is computationally challenging (DeValve, 2021) and further research on effective strategies is needed.

Another important problem in e-retailing that has not received much academic attention is the item placement problem. In the item placement problem, the e-retailer needs to choose the items to carry in each of its distribution centers. This problem was studied recently by Chen and Graves (2020) and Jehl (2020), where Chen and Graves (2020) focused on developing a scalable optimization framework for solving such a problem with millions of items for large e-retailers, while Jehl (2020) considered features such as multi-item orders. Chen and Graves (2020) and Jehl (2020) both focused on numerical approaches for finding effective item placement plans. On the more theoretical side, a recent work by DeValve et al. (2021) shows that a simple greedy method for the item placement problem (with only single-item orders) achieves an approximation factor of 0.432, and it would be interesting to see whether the theoretical approximation factor can be further improved.

9.4.3 Assembly Design

Another research direction takes a step back from inventory ordering and assembly, and asks how the assembly structure of an ATO system should be designed in the first place. In particular, during the product design or manufacturing process development phase, how should

component commonality be integrated into the design of a line of products? There can be various objectives for this design problem, including minimizing costs or assembly times, or maximizing profit or product availability. The decisions may vary based on how early in the product design phase the commonality decision is considered, but typical research in the area assumes there is already a line of products to be produced and a set of components they require.

The existing research has considered whether to invest in making some components common (Van Mieghem, 2004), and how to aggregate different groups of components into sub-assemblies called “vanilla-boxes” (Swaminathan & Tayur, 1998). Interesting future system design research questions include how to add new products to an existing ATO system with new or existing components, how to update assembly structures to adapt to changing demand, and extending the insights on commonality adoption of Van Mieghem (2004) to more general settings. For instance, Song and Zhao (2009) considered a continuous-review ATO system with positive lead times and found that the value of component commonality depends strongly on component costs, lead times, and dynamic allocation rules. Under certain conditions, several previous findings based on single-period models or dynamic models with identical lead times do not hold. Another potential challenge in assembly design is that the demand for different products may change depending on the set of final products manufactured.

9.4.4 ATO Systems with Endogenous Prices

With a few exceptions, the ATO literature treats the prices for the final product as exogenous and fixed. One such exception is Plambeck and Ward (2006), who studied a dynamic ATO and component production model wherein the prices of the final products are also decisions. In Plambeck and Ward (2006), the authors derived a policy where price and replenishment decisions are determined through the solution of a one-period problem and showed it to be asymptotically optimal in the high-demand volume regime. Oh et al. (2014) derived several structural properties of the optimal joint inventory and pricing policy, and used them to propose a heuristic policy that decouples the replenishment, pricing, and allocation decisions. Despite the progress in dynamic models, an interesting question that is not yet resolved is whether we can effectively solve the one-period ATO problem with pricing decisions for large systems and integrality constraints. Another interesting and related question is solving the one-period ATO problem with *responsive pricing* studied by Chod et al. (2010), where the price of the products is a linear function of the total production.

In both Plambeck and Ward (2006) and Oh et al. (2014), the authors assumed that the demand function is given. It is interesting to investigate demand functions that arise from rational choice models in economics theory. The rational choice models for products with multiple components have been studied in the context of pricing and bundling (see, e.g., Bakos & Brynjolfsson, 1999; Alaei et al., 2019; Abdallah et al., 2021; Ma & Simchi-Levi, 2021). However, unlike the ATO models, the models considered in the bundling literature do not consider component ordering decisions. Recently, Song and Xue (2021) present a dynamic model to analyze the optimal joint replenishment, pricing, and bundling decisions over time. Instead of a fixed bill of materials in the typical ATO system, the bundling decision dynamically determines the optimal product configurations along with the product prices.

The authors showed that the optimal policy is characterized by an overstock set in each period. For items in this set, it is optimal not to order. The optimal order-up-to-levels for the rest of the items, the bundling and pricing decisions, and the bundle assembly quantities all depend on the overstock levels. Future research is needed for developing more efficient algorithms and data-driven policies.

NOTES

1. Equivalently, the true expected shortage costs can be represented by $\mathbb{E}_{\mathbf{D}}[\tilde{G}(\mathbf{y}^{NV} | \mathbf{D})]$ in the equivalent Equation (9.4).
2. In (Van Jaarsveld & Scheller-Wolf, 2015), allocation decisions are actually made dynamically as orders arrive, but we can view the dynamic allocation as a heuristic for generating a feasible static allocation decision.
3. By approximation factor, we mean that the objective of the solution achieves no higher than two times the objective of the *optimization* problem defined in Equation (9.5). We note that it does not include the term $-\mathbf{h} \cdot \mathbf{A}\mathbb{E}[\mathbf{D}]$, as it is an additive constant not included in the optimization.
4. Formally, a laminar family is a collection of sets with the property that, for any two sets in the family, either one contains the other or their intersection is empty.

REFERENCES

- Abdallah, T., Asadpour, A., & Reed, J. (2021). Large-scale bundle-size pricing: A theoretical analysis. *Operations Research*.
- Acimovic, J., & Graves, S. C. (2015). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management*, 17(1), 34–51.
- Acimovic, J., & Graves, S. C. (2017). Mitigating spillover in online retailing via replenishment. *Manufacturing & Service Operations Management*, 19(3), 419–436.
- Agrawal, N., & Cohen, M. A. (2001). Optimal material control in an assembly system with component commonality. *Naval Research Logistics (NRL)*, 48(5), 409–429.
- Akçay, Y., & Xu, S. H. (2004). Joint inventory replenishment and component allocation optimization in an assemble-to-order system. *Management Science*, 50(1), 99–116.
- Alaei, S., Makhdoui, A., & Malekian, A. (2019). Optimal subscription planning for digital goods. Available at SSRN 3476296.
- Amil, A., Makhdoui, A., & Wei, Y. (2022). Multi-item order fulfillment revisited: Lp formulation and prophet inequality. Available at SSRN 4176274.
- Atan, Z., Ahmadi, T., Stegehuis, C., de Kok, T., & Adan, I. (2017). Assemble-to-order systems: A review. *European Journal of Operational Research*, 261(3), 866–879.
- Bakos, Y., & Brynjolfsson, E. (1999). Bundling information goods: Pricing, profits, and efficiency. *Management Science*, 45(12), 1613–1630.
- Benjaafar, S., & ElHafsi, M. (2006). Production and inventory control of a single product assemble-to-order system with multiple customer classes. *Management Science*, 52(12), 1896–1912.
- Bernstein, F., DeCroix, G. A., & Wang, Y. (2007). Incentives and commonality in a decentralized multiproduct assembly system. *Operations Research*, 55(4), 630–646.
- Bernstein, F., DeCroix, G. A., & Wang, Y. (2011). The impact of demand aggregation through delayed component allocation in an assemble-to-order system. *Management Science*, 57(6), 1154–1171.
- Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.
- Bixby, R. E. (2010). Mixed-integer programming: It works better than you may think. *FERC Conference*.
- Bumpensanti, P., & Wang, H. (2020). A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science*, 66(7), 2993–3009.
- Chen, A. I., & Graves, S. C. (2020). Item aggregation and column generation for online-retail inventory placement. *Manufacturing & Service Operations Management*.
- Chen, S., Lu, L., Song, J.-S. J., & Zhang, H. (2021). Optimizing assemble-to-order systems: Decomposition heuristics and scalable algorithms. *HKUST Business School Research Paper* (2021–33).
- Chen, X., & Li, M. (2021). Discrete convex analysis and its applications in operations: A survey. *Production and Operations Management*, 30(6), 1904–1926.
- Chod, J., Pyke, D., & Rudi, N. (2010). The value of flexibility in make-to-order systems: The effect of demand correlation. *Operations Research*, 58(4-part-1), 834–848.
- DeValve, L. (2019). Practical algorithms for managing uncertain demand in complex systems. PhD thesis, Duke University.

- DeValve, L. (2021). Cost balancing for sparse newsvendor networks. Available at SSRN 3961613.
- DeValve, L., & Myles, J. (2022). Base-stock policies are close to optimal for newsvendor networks. Available at SSRN 4187297.
- DeValve, L., Pekeć, S., & Wei, Y. (2020). A primal-dual approach to analyzing ATO systems. *Management Science*, 66(11), 5389–5407.
- DeValve, L., Pekeć, S., & Wei, Y. (2021). Approximate submodularity in network design problems. *Operations Research*.
- Doğru, M. K., Reiman, M. I., & Wang, Q. (2010). A stochastic programming based inventory policy for assemble-to-order systems with application to the W model. *Operations Research*, 58(4-part-1), 849–864.
- Doğru, M. K., Reiman, M. I., & Wang, Q. (2017). Assemble-to-order inventory management via stochastic programming: Chained BOMs and the M-system. *Production and Operations Management*, 26(3), 446–468.
- Feige, U. (1998). A threshold of $\ln(n)$ for approximating set cover. *Journal of the ACM (JACM)*, 45(4), 634–652.
- Goldberg, D. A., Reiman, M. I., & Wang, Q. (2019). A survey of recent progress in the asymptotic analysis of inventory systems. *Production and Operations Management*.
- Hausman, W. H., Lee, H. L., & Zhang, A. X. (1998). Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, 109(3), 646–659.
- Hsu, V. N., Lee, C. Y., & So, K. C. (2006). Optimal component stocking policy for assemble-to-order systems with lead-time-dependent component and product pricing. *Management Science*, 52(3), 337–351.
- Jasin, S., & Kumar, S. (2012). A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research*, 37(2), 313–345.
- Jasin, S., & Sinha, A. (2015). An lp-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Operations Research*, 63(6), 1336–1351.
- Jehl, T. (2020). Data-driven decision making algorithms for internet platforms. Ph.D. thesis, University of California, Berkeley.
- Kapuściński, R., Zhang, R. Q., Carbonneau, P., Moore, R., & Reeves, B. (2004). Inventory decisions in Dell's supply chain. *Interfaces*, 34(3), 191–205.
- Kleywegt, A. J., Shapiro, A., & Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2), 479–502.
- Levi, R., Roundy, R., Shmoys, D., & Sviridenko, M. (2008). A constant approximation algorithm for the one-warehouse multiretailer problem. *Management Science*, 54(4), 763–776.
- Lu, L., Song, J.-S., & Zhang, H. (2015). Optimal and asymptotically optimal policies for assemble-to-order N- and W-systems. *Naval Research Logistics*, 62(8), 617–645.
- Lu, Y., & Song, J.-S. (2005). Order-based cost optimization in assemble-to-order systems. *Operations Research*, 53(1), 151–169.
- Lu, Y., Song, J.-S., & Yao, D. D. (2003). Order fill rate, leadtime variability, and advance demand information in an assemble-to-order system. *Operations Research*, 51(2), 292–308.
- Lu, Y., Song, J.-S., & Yao, D. D. (2005). Backorder minimization in multiproduct assemble-to-order systems. *IEEE Transactions*, 37(8), 763–774.
- Lu, Y., Song, J.-S., & Zhao, Y. (2010). No-holdback allocation rules for continuous-time assemble-to-order systems. *Operations Research*, 58(3), 691–705.
- Ma, W. (2022). Simple and order-optimal correlated rounding schemes for multi-item e-commerce order fulfillment. arXiv preprint arXiv:2207.04774.
- Ma, W., & Simchi-Levi, D. (2021). Reaping the benefits of bundling under high production costs. *International Conference on Artificial Intelligence and Statistics*. PMLR, 1342–1350.
- Murota, K. (2003). *Discrete convex analysis, vol. 10*. SIAM.
- Nadar, E., Akan, M., & Scheller-Wolf, A. (2014). Optimal structural results for assemble-to-order generalized M-systems. *Operations Research*, 62(3), 571–579.
- Nadar, E., Akan, M., & Scheller-Wolf, A. (2016). Experimental results indicating lattice-dependent policies may be optimal for general assemble-to-order systems. *Production and Operations Management*, 25(4), 647–661.
- Nadar, E., Akcay, A., Akan, M., & Scheller-Wolf, A. (2018). The benefits of state aggregation with extreme-point weighting for assemble-to-order systems. *Operations Research*, 66(4), 1040–1057.

- Oh, S., Sourirajan, K., & Ettl, M. (2014). Joint pricing and production decisions in an assemble-to-order system. *Manufacturing & Service Operations Management*, 16(4), 529–543.
- Plambeck, E. L. (2008). Asymptotically optimal control for an assemble-to-order system with capacitated component production and fixed transport costs. *Operations Research*, 56(5), 1158–1171.
- Plambeck, E. L., & Ward, A. R. (2006). Optimal control of a high-volume assemble-to-order system. *Mathematics of Operations Research*, 31(3), 453–477.
- Plambeck, E. L., & Ward, A. R. (2007). Note: A separation principle for a class of assemble-to-order systems with expediting. *Operations Research*, 55(3), 603–609.
- Reiman, M. I., Wan, H., & Wang, Q. (2023). Asymptotically optimal inventory control for assemble-to-order systems. *Stochastic Systems*, 13(1), 128–180.
- Reiman, M. I., & Wang, Q. (2012). A stochastic program based lower bound for assemble-to-order inventory systems. *Operations Research Letters*, 40(2), 89–95.
- Reiman, M. I., & Wang, Q. (2015). Asymptotically optimal inventory control for assemble-to-order systems with identical lead times. *Operations Research*, 63(3), 716–732.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). *Lectures on stochastic programming: Modeling and theory*. SIAM.
- Song, J.-S. (1998). On the order fill rate in a multi-item, base-stock inventory system. *Operations Research*, 46(6), 831–845.
- Song, J.-S. (2002). Order-based backorders and their implications in multi-item inventory systems. *Management Science*, 48(4), 499–516.
- Song, J.-S., Xu, S. H., & Liu, B. (1999). Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes. *Operations Research*, 47(1), 131–149.
- Song, J. S., & Xue, Z. (2021). Demand shaping through bundling and product configuration: A dynamic multiproduct inventory-pricing model. *Operations Research*, 69(2), 525–544.
- Song, J.-S., & Yao, D. D. (2002). Performance analysis and optimization of assemble-to-order systems with random lead times. *Operations Research*, 50(5), 889–903.
- Song, J.-S., & Zhao, Y. (2009). The value of component commonality in a dynamic inventory system with lead times. *Manufacturing & Service Operations Management*, 11(3), 493–508.
- Song, J.-S., & Zipkin, P. (2003). Supply chain operations: Assemble-to-order systems. In S.C. Graves, and A.G. de Kok (Eds.), *Supply chain management: Design, coordination and operation, handbooks in operations research and management science* (Vol. 11, pp. 561–596). Elsevier.
- Swaminathan, J. M., & Tayur, S. R. (1998). Managing broader product lines through delayed differentiation using vanilla boxes. *Management Science*, 44(12-part-2), S161–S172.
- Talluri, K. T., & Van Ryzin, G. J. (2006). *The theory and practice of revenue management*, vol. 68. Springer Science & Business Media.
- Van Jaarsveld, W., & Scheller-Wolf, A. (2015). Optimization of industrial-scale assemble-to-order systems. *INFORMS Journal on Computing*, 27(3), 544–560.
- Van Mieghem, J. A. (2004). Commonality strategies: Value drivers and equivalence with flexible capacity and inventory substitution. *Management Science*, 50(3), 419–424.
- Van Mieghem, J. A., & Rudi, N. (2002). Newsvendor networks: Inventory management and capacity investment with discretionary activities. *Manufacturing & Service Operations Management*, 4(4), 313–335.
- Vera, A., Banerjee, S., & Gurvich, I. (2021). Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research*.
- Wan, H., & Wang, Q. (2015). Asymptotically-optimal component allocation for assemble-to-order production–inventory systems. *Operations Research Letters*, 43(3), 304–310.
- Xu, P. J., Allgor, R., & Graves, S. C. (2009). Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management*, 11(2), 340–355.
- Zhang, A. X. (1997). Demand fulfillment rates in an assemble-to-order system with multiple products and dependent demands. *Production and Operations Management*, 6(3), 309–324.
- Zhao, Y., & Simchi-Levi, D. (2006). Performance analysis and evaluation of assemble-to-order systems with stochastic sequential lead times. *Operations Research*, 54(4), 706–724.
- Zipkin, P. (2008). On the structure of lost-sales inventory models. *Operations Research*, 56(4), 937–944.
- Zipkin, P. (2016). Some specially structured assemble-to-order systems. *Operations Research Letters*, 44(1), 136–142.

10. Inventory models with returns and remanufacturing

Xiting Gong and Sean X. Zhou

10.1 INTRODUCTION

Remanufacturing is an industrial process that restores end-of-life goods to their original working condition (USITC, 2012). Due to its high economic and environmental values, remanufacturing has been widely adopted in many industries, e.g., electrical appliances, machinery, and car parts, in recent years. USITC (2012) reported that remanufacturing was an important and growing activity in many industrial sectors in the United States and worldwide, and between 2009 and 2011, the value of US remanufactured production grew by 15% to at least \$43 billion, supporting 180,000 full-time US jobs. Moreover, remanufacturing is a critical sustainable manufacturing strategy, promoted and supported by many governments worldwide. This also motivates manufacturers to engage in collecting and remanufacturing product returns, and then reselling/reusing them.

In response to the industry needs, researchers have been studying inventory models with product returns and remanufacturing since the 1970s. Different from conventional inventory models, those models consider remanufacturing of returned items into serviceable products for fulfilling customer demands. Those models can be either deterministic or stochastic (depending on whether there is randomness in product demands and/or returns), or either continuous-review or periodic-review (depending on decision epochs). In the past decades, many such models have been developed and analyzed, with a major research goal of deriving optimal or effective heuristic policies. We refer to Fleischmann et al. (1997) for a review of early models and Ilgin and Gupta (2012) and Souza (2013) for reviews of more recent models.

In this chapter, we provide a detailed review of recent studies on periodic-review stochastic inventory models with returns and remanufacturing. It starts with a classic repairable inventory model studied by Simpson (1978), and then covers its various extensions, including single-stage models with a single type of returns (or single-return for short), single-stage models with multiple types of returns (or multi-return for short), multi-echelon models, and models with differentiated remanufactured and new products. For each of those studies, we provide a detailed model description, formulation, and summarize its main results. Most of those studies focus on characterizing the optimal policy via dynamic programming. The rest of them focus on developing efficient heuristic policies when the optimal policy is intractable computationally due to the curse of dimensionality. We will conclude the chapter with some directions for future research.

We note that there is a stream of literature on inventory models with returns but *without remanufacturing*, where returns directly go to stock for serving customer demand rather than kept in a separate buffer until they are remanufactured or disposed of. They include, for example, series models with returns (DeCroix et al., 2005), assembly models with returns (DeCroix

& Zipkin, 2005), and assemble-to-order models with returns (DeCroix et al., 2009). Since those models do not consider remanufacturing, they are beyond the scope of this chapter.

The rest of this chapter is organized as follows. In Section 10.2, we review single-stage single-return models. In Section 10.3, we review multi-return or multi-echelon inventory models. In Section 10.4, we review models with differentiated remanufactured and new products. We conclude the chapter in Section 10.5 with directions for future research. Throughout this chapter, we use “increasing” and “decreasing” in a non-strict sense, that is, they represent “non-decreasing” and “non-increasing,” respectively. For real numbers x and y , we denote $x^+ = \max\{x, 0\}$, $x \vee y = \max\{x, y\}$, and $x \wedge y = \min\{x, y\}$.

10.2 SINGLE-STAGE, SINGLE-RETURN MODELS

In this section, we review several single-stage, single-return inventory models. In Section 10.2.1, we review a classic inventory model with remanufacturing. After that, we review models with finite capacities in Section 10.2.2, a model with core acquisition and pricing in Section 10.2.3, models with non-identical lead times in Section 10.2.4, and a model with dependent demands and returns in Section 10.2.5.

10.2.1 A Classic Inventory Model with Remanufacturing

In this subsection, we review a classic inventory model with remanufacturing studied by Simpson (1978). As Simpson (1978) uses “repairing,” here we may use “repairing” and “remanufacturing” interchangeably.

Consider a firm managing an inventory system of serviceable and repairable products over a planning horizon of T periods, indexed forward by $t = 1, \dots, T$. A repairable product can be remanufactured into a serviceable product. At the beginning of each period t , the firm reviews its serviceable inventory level $x_{t,0}$ and the total inventory level $x_{t,1}$ (including both serviceable and repairable products). Next, it decides on purchasing z_t units of serviceable products from a supplier at a unit price p , repairing w_t units available in its repairable inventory with a unit cost r , and disposing u_t units at zero salvage value or disposal cost. The decisions are constrained by $z_t \geq 0$, $w_t \geq 0$, $u_t \geq 0$, and $w_t + u_t \leq x_{t,1} - x_{t,0}$. The purchasing and repairing decisions have zero lead times (i.e., the purchased and repaired units are available to use immediately). Thus, the serviceable and total inventory levels become $x_{t,0} + z_t + w_t$ and $x_{t,1} + z_t - u_t$ after all the decisions. After that, random demand D_t and random returns R_t are realized. Here, D_t and R_t can be correlated in period t , but are independent in different periods. The random demand D_t is satisfied by the serviceable inventory as much as possible; and unmet demand is backlogged to the next period, incurring a unit backlogging cost b . The leftover serviceable and repairable inventories are carried to the next period, incurring unit holding costs h and s , respectively. All the cost parameters are non-negative. Besides, $p > r$, since otherwise repairing would not be economical. There is a one-period discount factor $0 \leq \alpha \leq 1$. The firm’s objective is to minimize the expected total discounted cost over the T -period planning horizon.

The firm’s optimization problem in the above model can be formulated as a dynamic program as follows. The state variables in each period t are $x_{t,0}$ and $x_{t,1}$, with the state space $\mathbb{S} := \{(x_0, x_1) \in \mathfrak{R}^2 \mid x_0 \leq x_1\}$. The system dynamics from period t to period $t+1$ are given by

$$x_{t+1,0} = x_{t,0} + z_t + w_t - D_t, \quad x_{t+1,1} = x_{t,1} + z_t - u_t - D_t + R_t.$$

Denote $V_t(x_0, x_1)$ as the firm's minimum expected total discounted cost from period t to T , given the system state (x_0, x_1) in period t . Then, we can write the firm's optimality equations as follows (for notational convenience, we suppress the subscript t unless confusion would otherwise arise): for $t = 1, \dots, T$,

$$V_t(x_0, x_1) = \min_{\substack{z \geq 0, w \geq 0, u \geq 0 \\ w+u \leq x_1 - x_0}} \{pz + rw + s(x_1 - x_0 - w - u) + G_t(x_0 + z + w) \\ + \alpha \mathbb{E}[V_{t+1}(x_0 + z + w - D_t, x_1 + z_t - u_t - D_t + R_t)]\}, \quad (10.1)$$

where $G_t(x) := \mathbb{E}[h(x - D_t)^+ + b(D_t - x)^+]$. The boundary condition is given by $V_{T+1}(x_0, x_1) \equiv 0$. Denote $(z_t^*(x_0, x_1), w_t^*(x_0, x_1), u_t^*(x_0, x_1))$ as an optimal solution of (z, w, u) to Equation (10.1).

By analyzing the structural properties of the value function $V_t(x_0, x_1)$, Simpson (1978) proves that the firm's optimal policy has the following simple structure.

Theorem 10.1 (Simpson, 1978) *For $t = 1, \dots, T$, there exist non-negative constants $\theta_t, \delta_t, \xi_t$, with $\theta_t \geq \delta_t$, such that $w_t^*(x_0, x_1) = (x_1 - x_0) \wedge (\theta_t - x_0)^+$, $z_t^*(x_0, x_1) = (\delta_t - x_1)^+$, and $u_t^*(x_0, x_1) = (x_1 - x_0) \wedge (x_1 - \theta_t - \xi_t)^+$.*

In words, the optimal policy for period t operates as follows. If the total inventory level x_1 is less than δ_t , then it is optimal to remanufacture all repairable products and raise the total inventory level to δ_t by further purchasing. If the total inventory level x_1 is more than δ_t but the serviceable inventory level x_0 is less than θ_t , then it is optimal to raise the latter to θ_t by remanufacturing only, subject to the availability of repairable product inventory. Finally, if the total inventory level x_1 is greater than $\theta_t + \xi_t$, then it is optimal to bring it down to $\theta_t + \xi_t$ by disposing of some repairable product inventory, again subject to its availability.

There are two salient properties of Simpson's model. First, since repairing is cheaper than purchasing (i.e., $r < p$), the optimal policy always gives higher priority to repairing over purchasing (i.e., the purchase option is resorted to only after all repairable units are repaired). This property continues to hold under many extensions of Simpson's model (e.g., the lost-sales model where unmet demand is lost rather than backlogged, models with finite repairing capacity in Section 10.2.2, a model with core acquisition and pricing in Section 10.2.3, and a model with multiple types of returns Section 10.3.1). Second, the value function $V_t(x_0, x_1)$ for each period t is *additively convex* in (x_0, x_1) on $\mathbb{S}$. That is, there exist univariate convex functions $f_{t,0}(x)$ and $f_{t,1}(x)$ such that $V_t(x_0, x_1) = f_{t,0}(x_0) + f_{t,1}(x_1)$ for all $(x_0, x_1) \in \mathbb{S}$. This property is crucial for the simple structure of the firm's optimal policy described in Theorem 10.1. This property continues to hold under several extensions of Simpson's model (e.g., a model with core acquisition (but with exogenous selling prices) in Section 10.2.3, a model with multiple types of returns in Section 10.3.1, and multi-echelon models in Section 10.3.3).

When unmet demand is backlogged, the model with positive, identical purchasing and repairing lead times can be transformed into Simpson's (1978) model with zero lead times. To this end, we should define the state variables $x_{t,0}$ and $x_{t,1}$ in period t as the serviceable inventory position (i.e., the inventory level of serviceable product plus all pipeline inventories, i.e., remanufactured/ordered) and the total inventory position (i.e., $x_{t,0}$ plus repairable inventory), respectively, and define the value function $V_t(x_0, x_1)$ properly. We refer interested readers to

Inderfurth (1997) for the detailed formulation and results. In the remainder of this chapter, whenever possible, we will stick to the notation used in this subsection.

10.2.2 Capacity Constraints

In this subsection, we review capacitated inventory models with remanufacturing studied by Gong and Chao (2013). Those models incorporate one or more of the following three types of capacities into Simpson's (1978) model: manufacturing (or purchase) capacity K_m , remanufacturing (or repairing) capacity K_r , and the total manufacturing/remanufacturing capacity K , with $K \geq \max\{K_r, K_m\}$. Meanwhile, to follow the term used in most of the other papers we will review on remanufacturing, we will refer to repairable products as returned products in the rest of this chapter. For simplicity, those models do not allow for the disposal of returned products, and thus only involve manufacturing and remanufacturing decisions in each period.

Same as Simpson (1978), Gong and Chao (2013) formulate the firm's optimization problem as a dynamic program. The state variables in period t are the returned inventory level x_t^0 and the total inventory level x_t^1 (note that they are different from those used in Section 10.2.1 and this new definition of state facilitates the analysis of the models with capacity constraints), with the state space $\mathfrak{R}_+ \times \mathfrak{R}$. The decision variables in period t are the returned inventory level y_t^0 and the total inventory level y_t^1 after the manufacturing and remanufacturing decisions are made. Besides the regular decision constraints $0 \leq y_t^0 \leq x_t^0$ and $y_t^1 \geq x_t^1$, due to capacities, there are one or more of the following constraints:

$$\begin{aligned} x_t^0 - y_t^0 &\leq K_r, \\ y_t^1 - x_t^1 &\leq K_m, \\ (x_t^0 - y_t^0) + (y_t^1 - x_t^1) &\leq K. \end{aligned}$$

The system dynamics from period t to period $t+1$ are given by

$$x_{t+1}^0 = y_t^0 + R_t, \quad x_{t+1}^1 = y_t^1 - D_t + R_t.$$

Denote $V_t(x^0, x^1)$ as the firm's minimum expected total discounted cost from period t onwards, given the system state (x^0, x^1) in period t . In the presence of any capacity constraint (regardless of its type), the value function $V_t(x^0, x^1)$ is not decomposable into univariate functions, and the firm's optimal inventory policy does not have a simple structure as that in Simpson's (1978) model. Gong and Chao (2013) prove that the value function satisfies the following key structural property.

Proposition 10.1 (Gong & Chao, 2013) *For the general model with capacities K_r , K_m , and K , the value function $V_t(x_t^0, x_t^1)$ is $L^{\natural}$ -convex, $t = 1, \dots, T$.*

A function $f : V \rightarrow \mathfrak{R}$ is $L^{\natural}$ -convex if $\psi(v, \zeta) = f(v - \zeta \mathbf{e})$ is submodular on $\{(v, \zeta) \mid v - \zeta \mathbf{e} \in \mathbb{V}\}$, where $\mathbb{V} \subset \mathfrak{R}^n$ is a lattice and $\mathbf{e}$ is an n -dimensional vector of all ones. Proposition 10.1 implies that $V_t(x_t^0, x_t^1)$ is jointly convex, submodular, and its Hessian matrix is diagonal dominant (if twice differentiable). This structural property plays a key role in characterizing the firm's optimal policy. We summarize below the structures of the optimal policies for different capacitated

models derived from Gong and Chao (2013). For the model with capacities K_r and either K or K_m , the optimal remanufacturing and manufacturing policies for period t are a modified remanufacture-down-to policy and a modified total-up-to policy, respectively. Specifically, it is optimal to remanufacture the returned inventory down to a remanufacture-down-to level as much as possible, subject to capacity K_r , and raise the total inventory level via manufacturing to a total-up-to level as much as possible, subject to capacity K_m or the remaining total capacity after remanufacturing. Further, the remanufacture-down-to level is a partly constant increasing function of the total inventory level x_t^1 with slopes at most one; and the total-up-to level is a partly constant increasing function of the returned inventory level x_t^0 with slopes at most one.

As discussed in Section 10.2.1, the optimal policy gives higher priority to remanufacturing over manufacturing when there is no capacity constraint. This property continues to hold when there are a remanufacturing capacity K_r and/or a total capacity K , but it does not always hold when there is a manufacturing capacity K_m . The reason for the latter is as follows. The manufacturing capacity K_m , if not used in any period, is wasted and cannot be carried to future periods. By contrast, the returned products, if not remanufactured, can be carried to future periods. Thus, it can be more cost-effective to manufacture some products (for utilizing the manufacturing capacity) while keeping some returned products for future use. As a result, the optimal policy does not always give higher priority to remanufacturing when there is a manufacturing capacity.

Recently, Gong and Liu (2021) extend Gong and Chao's (2013) model to a more general one with positively dependent random manufacturing and remanufacturing capacities. For this more general model, they prove that the value function $V_t(x^0, x^1)$ remains L^3 -convex. Utilizing this property, they partially characterize the structure of the optimal policy for each period by two increasing control functions with slopes at most one. For the models with deterministic manufacturing or remanufacturing capacity, they completely characterize the optimal policy by those control functions. For the special case with unlimited manufacturing capacity, they further characterize the optimal policy and obtain additional insights. For example, while the optimal policy always gives higher priority to remanufacturing when it has a deterministic capacity, this property fails to hold when remanufacturing has a random capacity. Moreover, in that case, the optimal policy may never use remanufacturing even if it is less costly than manufacturing.

10.2.3 Returned Product Acquisition and Pricing

In this subsection, we review an inventory model with returned product (or called core) acquisition and pricing studied by Zhou and Yu (2011). This model incorporates two additional features into Simpson's (1978) model. First, random customer demand in period t , denoted by $D_t(p_t)$, depends on the selling price p_t of the serviceable product in the following linear form:

$$D_t(p_t) = a - bp_t + \varepsilon_t^d, \quad \forall p_t \in [p_l, p_u],$$

where $\{\varepsilon_t^d : 1 \leq t \leq T\}$ are *i.i.d.* continuous random variables with mean zero. Second, random product returns in period t , denoted by $R_t(e_t)$, depends on the acquisition effort, e_t , exerted by the firm, and has the following additive form:

$$R_t(e_t) = f(e_t) + \varepsilon_t^r, \quad \forall e_t \geq 0,$$

where $\{\varepsilon_t^r : 1 \leq t \leq T\}$ are *i.i.d.* non-negative, continuous random variables, and $f(e) \geq 0$ is a strictly increasing, concave function. The random variables ε_t^d and ε_t^r can be correlated in each period, but are independent in different periods. In each period t , besides the manufacturing, remanufacturing, and disposal decisions in Simpson's (1978) model, the firm needs to decide on the selling price p_t and the acquisition effort e_t . In the presence of the pricing decisions, the objective of the firm is to maximize the expected total discounted profit over a finite planning horizon.

Zhou and Yu (2011) also formulate the firm's optimization problem as a dynamic program, with two state variables $x_{t,0}$ and $x_{t,1}$ for period t defined in Section 10.2.1. Denote $V_t(x_0, x_1)$ as the firm's maximum expected total discounted profit from period t to T , given the system state (x_0, x_1) in period t . When the selling prices are exogenous, Zhou and Yu (2011) show that the value function $V_t(x_0, x_1)$ in each period t is additively concave. In other words, the decomposition result in Simpson's (1978) model continues to hold after incorporating the firm's acquisition efforts. As a result, the firm's optimal inventory policy has the same structure as that in Theorem 10.1 for Simpson's (1978) model. In addition, the optimal acquisition effort in period t is decreasing in the total inventory level x_1 .

For the general model with endogenous selling prices, the value function $V_t(x_0, x_1)$ is no longer decomposable (while the priority of remanufacturing over manufacturing remains). Consequently, the optimal policy for each period becomes much more involved. Zhou and Yu (2011) prove that it can be characterized by a set of constants and control functions. We refer interested readers to Theorem 8 in Zhou and Yu (2011) for details. Besides, they prove that the optimal selling price in period t is decreasing in the post-production serviceable and total inventory levels, while the optimal acquisition effort in period t is increasing in the post-production serviceable level and decreasing in the post-production total inventory level.

10.2.4 Non-Identical Manufacturing and Remanufacturing Lead Times

In this subsection, we review inventory models with non-identical manufacturing and remanufacturing lead times. Since the optimal policies for the models are very complex and in general intractable computationally due to the state tracking pipeline inventories, almost all the existing studies focus on developing effective heuristic policies. To our knowledge, Inderfurth (1997) is the only study which considers the optimal policy for such models. However, this study only characterizes the optimal policy structure when the returned products are *not* stocked over periods and the manufacturing lead time exceeds the remanufacturing lead time for one period. In what follows, we review the models studied by Kiesmüller (2003) and Xin (2021) as well as their proposed heuristic policies. The inventory models with non-identical manufacturing and remanufacturing lead times are related to dual-sourcing inventory models with different lead times reviewed in Chapter 9 in this book, with the added complexity of remanufacturing quantity in each period constrained by the available returned inventory.

Kiesmüller (2003) considers an infinite-horizon remanufacturing inventory model with non-identical lead times. The manufacturing lead time L_p can be either larger or smaller than the remanufacturing lead time L_r . The demands in different periods are *i.i.d.* random variables, and so are product returns in different periods. Unmet demands are backlogged with unit cost h_b . The unit holding costs for serviceable and returned products are h_s and h_r , respectively, with

$h_s > h_r$. Kiesmüller (2003) does not consider disposal of returned products and assumes zero manufacturing and remanufacturing costs. The objective is to minimize the long-run average holding and backlogging cost. Since the optimal policy is expected to be very complex, Kiesmüller (2003) proposes a dual-index (S, M) policy based on two different inventory positions for remanufacturing and manufacturing decisions, respectively. Depending on whether L_r is larger than L_p , the two inventory positions are defined as follows. In period t , the remanufacturing inventory position $X_u(t)$ is defined as

$$X_u(t) := I_s(t) + \sum_{i=0}^{L_r \wedge L_p} p(t - (L_p - L_r)^+ - i) + \sum_{i=1}^{L_r} u(t - i)$$

and the manufacturing inventory position $X_p(t)$ is defined as

$$X_p(t) := I_s(t) + \sum_{i=1}^{L_p} p(t - i) + \begin{cases} \sum_{i=0}^{L_p} u(t - (L_r - L_p + i)), & \text{if } L_r > L_p; \\ I_R(t) + \sum_{i=1}^{L_r} u(t - i), & \text{if } L_r < L_p. \end{cases}$$

where $I_s(t)$ is the serviceable inventory level in period t , $I_R(t)$ is the returned inventory level in period t , and $p(t')$ and $u(t')$ are the manufacturing and remanufacturing quantities in period t' , respectively. Then, in each period t , a dual-index (S, M) policy raises the inventory position $X_p(t)$ to the base-stock level S via manufacturing and the inventory position $X_u(t)$ to the base-stock level M via remanufacturing, subject to available returned inventory.

Kiesmüller (2003) compares the numerical performances of the best dual-index (S, M) policy and the best single-index (S, M) policy. Here, a single-index (S, M) policy uses one inventory position defined as the serviceable inventory level plus all pipeline inventories, and applies the base-stock levels S and M to manufacturing and remanufacturing decisions, respectively. The numerical results show that the best dual-index (S, M) policy performs significantly better than the best single-index (S, M) policy, especially when the lead time difference is large.

Recently, Xin (2021) considers an infinite-horizon remanufacturing inventory model with manufacturing lead time larger than remanufacturing lead time. Different from Kiesmüller (2003), Xin (2021) considers positive unit manufacturing and remanufacturing costs, and allows disposal of returned products. For his model, Xin (2021) proposes a constant-order (L, U) threshold policy, where $L \leq U$. This policy manufactures a constant amount of new product in each period, brings the total inventory level down to U by disposing of returned product, and raises the serviceable inventory level up to L by remanufacturing. Xin (2021) proves that this class of policies is asymptotically optimal as the manufacturing lead time grows large.

Zhou et al. (2011) and Tao and Zhou (2014) both propose heuristic policies for finite-horizon models with non-identical lead times. For non-identical lead times, Zhou et al. (2011) propose a heuristic policy for a model with multiple types of returns, based on the optimal policy for the model with identical lead times. By extending their proposed heuristic policy

for the model with identical lead times and dependent demands and returns, Tao and Zhou (2014) propose a heuristic policy for a model with non-identical lead times. We refer interested readers to these studies for their proposed heuristic policies and numerical performances for the models with non-identical lead times. In the next subsection and Section 10.3.1, we will review the identical-lead-time models studied by Tao and Zhou (2014) and Zhou et al. (2011), respectively.

10.2.5 Dependent Demands and Returns

In many situations, product returns in a period depend on demand/sales that occurred in earlier periods, i.e., returns and demands are dependent over time. However, directly incorporating such dependence into the dynamic program will result in a high-dimensional state space (as the firm needs to keep track of the earlier sales/demand) and the optimal policy too complicated to be implemented in practice. In this subsection, we review an inventory model with dependent demands and returns by Tao and Zhou (2014) and their proposed heuristic policy for this model.

Tao and Zhou (2014) consider an inventory model similar to the one studied by Simpson (1978), except that it does not allow disposal of returned products and that it considers a general stochastic process for demands and returns in different periods. Specifically, at the beginning of each period t , the firm observes an information set f_t , which contains all the relevant information available at that time. For example, it can include the realized past demands $(d_1, \dots, d_{t-1})$ and product returns $(r_1, \dots, r_{t-1})$, and possibly some additional information denoted by $(e_1, \dots, e_t)$. The information set f_t is a specific realization in the set of all possible realizations, denoted by $\mathcal{F}_t$, of the random vector $F_t = (D_1, R_1, E_1, \dots, D_{t-1}, R_{t-1}, E_{t-1}, E_t)$. Given the information set f_t , the conditional joint distribution of the future demands and returns $(D_t, R_t, \dots, D_T, R_T)$ is known. The evolution of F_t is exogenous as it does not depend on the firm's decisions. This demand and return model is rather general. We refer to Tao and Zhou (2014) for two concrete examples of this model used in the literature.

For this model, the system state in period t consists of not only $(x_{t,0}, x_{t,1})$ defined in Section 10.2.1 (i.e., serviceable and total inventory levels, respectively), but also the information set f_t . Accordingly, one can readily formulate the firm's problem into a dynamic program, which is similar to Equation (10.1) except that there is no-disposal decision u_t , the value functions $V_t(\cdot)$ and $V_{t+1}(\cdot)$ should be modified as $V_t(\cdot, f_t)$, $V_{t+1}(\cdot, F_{t+1})$, respectively, and the expectations over (D_t, R_t, F_{t+1}) should be modified as conditional expectations given f_t . Following similar analysis in Simpson (1978), the firm's optimal policy can be shown to have a structure similar to Theorem 10.1 except that the thresholds in each period t now depend on f_t .

Although the optimal policy has a simple structure, computing the state-dependent parameters is intractable when f_t is high-dimensional. In view of this, Tao and Zhou (2014) develop a heuristic policy, called *approximation balancing policy* β , to manage this model. This policy is closely related to approximation algorithms for stochastic inventory systems reviewed in Chapter 12 in this book. We next construct this policy.

The first step of constructing the heuristic policy is to transform the original problem with cost parameters (p, r, h, s, b) defined in Section 10.2.1 to a problem with unit manufacturing cost p_t for period t , zero unit remanufacturing cost, unit period- t holding cost h_t for serviceable

products, zero unit holding cost for returned products, and unit backlogging cost b_t for period t . Here, for $t = 1, \dots, T$, the cost parameters p_t , h_t , and b_t are defined by

$$\begin{aligned} p_t &= p - r + s_t, \\ h_t &= h + r - s_t - \alpha(r - s_{t+1}), \\ b_t &= b - (r - s_t) + \alpha(r - s_{t+1}), \end{aligned}$$

where $s_t = \sum_{j=t}^T \alpha^{j-t} s$ for $t = 1, \dots, T$ and $s_{T+1} = 0$. Tao and Zhou (2014) show that for any feasible policy π , the expected total discounted cost C^π for the original problem equals the sum of the expected total discounted cost $\tilde{C}^\pi$ for the transformed problem under the same policy π and a policy-independent cost C_0 . That is, $C^\pi = \tilde{C}^\pi + C_0$ for any feasible policy π . Denote C^* and $\tilde{C}^*$ as the optimal expected total discounted costs for the original and the controllable part of the transformed problems, respectively.

The second step is to define the marginal holding cost and the backlogging cost associated with the firm's decisions. For each period t , for any serviceable inventory level $x_{t,0}$ and manufacturing quantity z_t (≥ 0), define the marginal holding cost as

$$H_t(x_{t,0}, z_t) = \sum_{j=t}^T h_j \left(z_t - \left(\sum_{n=t}^j D_n - x_{t,0} \right)^+ \right),$$

and the backlogging cost as

$$B_t(x_{t,0} + z_t) = b_t(x_{t,0} + z_t - D_t)^-.$$

We are now ready to present the approximation balancing policy β . For each period t , after reviewing the system state $(x_{t,0}, x_{t,1}, f_t)$, let w_t^β be the smallest solution of w_t to the following equation on $[0, \infty)$:

$$\mathbb{E}[H_t(x_{t,0}, w_t) | f_t] = \mathbb{E}[B_t(x_{t,0} + w_t) | f_t]. \quad (10.2)$$

The left-hand side of Equation (10.2) equals zero when $w_t = 0$, is convex increasing in w_t , and approaches infinity when $w_t \rightarrow \infty$. The right-hand side of Equation (10.2) is non-negative when $w_t = 0$, is convex decreasing in w_t , and approaches zero when $w_t \rightarrow \infty$. Therefore, w_t^β always exists. In addition, when $w_t^\beta > x_{t,1} - x_{t,0}$, let z_t^β be the smallest solution of z_t to the following equation on $[0, \infty)$:

$$\mathbb{E}[H_t(x_{t,0}, x_{t,1} - x_{t,0} + z_t) | f_t] + p_t z_t = \mathbb{E}[B_t(x_{t,1} + z_t) | f_t]. \quad (10.3)$$

One can easily verify that z_t^β always exists when $w_t^\beta > x_{t,1} - x_{t,0}$. Finally, the remanufacturing balancing policy β is defined as follows: For $t = 1, \dots, T$, if $w_t^\beta \leq x_{t,1} - x_{t,0}$, then remanufacture w_t^β ; otherwise remanufacture all the returned products and manufacture z_t^β units.

Since computing w_t^β and z_t^β involves solving only two equations with convex functions, the remanufacturing balancing policy β can be computed efficiently. Same as the optimal policy,

this policy also always remanufactures the returned product first. Tao and Zhou (2014) prove the following result on its theoretical performance.

Theorem 10.2 (Tao & Zhou, 2014) $\tilde{C}^\beta \leq 2\tilde{C}^*$. Thus, $C^\beta \leq 2C^*$ if $C_0 \geq 0$. That is, when $C_0 \geq 0$, the expected total discounted cost under the remanufacturing balancing policy β is at most twice the optimal cost.

The numerical results in Tao and Zhou (2014) show that the relative cost errors of the approximation balancing policy β are in general much smaller than the theoretical bound, with an average error below 10% and a maximum error of 24.26% across the 576 instances tested. They use a similar balancing idea to develop policies for models with non-identical lead times. Although these policies do not have provable performance bounds, their overall numerical performances are similar to the case with zero lead times.

Finally, they consider two extensions: a model with multiple types of returns (see Section 10.3.1) and a model with a total manufacturing and remanufacturing capacity K (see Section 10.2.2). For each extension, they develop a remanufacturing balancing policy and prove that Theorem 10.2 still holds for this policy.

Recently, Tao et al. (2020) develop heuristic policies for remanufacturing inventory systems with dependent demands and returns and non-identical manufacturing and remanufacturing lead times. To minimize the expected long-run average cost, they propose a class of manufacturing/remanufacturing policies called *forecast-adjusted base-stock policies*, and provide an exact procedure to evaluate its performance. In addition, they apply heavy traffic approximations to derive closed-form expressions for near-optimal policy control parameters and demonstrate the effectiveness of the heuristics and the approximations through numerical studies.

10.3 MULTI-RETURN OR MULTI-ECHELON MODELS

In remanufacturing systems, it is often observed that products are returned in different physical conditions and so incur different remanufacturing costs. Moreover, yield of the remanufacturing process is random due to various quality problems with returned products. In this section, we review an inventory model with multiple types of returns in Section 10.3.1, a multi-return model with random yield in Section 10.3.2, and multi-echelon inventory models with remanufacturing in Section 10.3.3.

10.3.1 Multiple Types of Returns

In this subsection, we review an inventory model with multiple types of returns studied by Zhou et al. (2011). On the basis of Simpson's (1978) model, this model needs the following additional notation and assumptions. In each period t , the firm receives K types of random returns with varying physical conditions, denoted by $R_{t,1}, \dots, R_{t,K}$. They can be correlated within the same period, but random demands and returns in different periods are independent. For $k = 1, \dots, K$, the unit remanufacturing cost of a type- k return is r_k , with $r_1 \leq r_2 \leq \dots \leq r_K < p$. There are unit holding cost s and unit disposal cost u for returns, regardless of their types. It is assumed that $p + u > r_K$, since otherwise the firm will never remanufacture type- K returns. The firm's objective is to find the optimal manufacturing, remanufacturing, and disposal policy that minimizes its expected total discounted cost over a finite planning horizon.

When there are K types of returns, the system state in each period t needs to be expanded to a $(K+1)$ -dimensional vector $\mathbf{x}_t = (x_{t,0}, \dots, x_{t,K})$, where $x_{t,0}$ is the serviceable inventory level and $x_{t,k}$ is the aggregate inventory level of serviceable product and type-1 to type- k returned products, $k = 1, \dots, K$. The state space is given by $\mathbb{V} := \{\mathbf{x} \in \mathfrak{R}^{K+1} \mid x_0 \leq x_1 \leq \dots \leq x_K\}$.

For $t = 1, \dots, T$, denote $V_t(\mathbf{x}_t)$ as the optimal expected total discounted cost from period t to T , given the system state $\mathbf{x}_t$ in period t . For $t = 1, \dots, T$, the optimality equation can be written as

$$V_t(\mathbf{x}_t) = \min_{\mathbf{w}_t, \mathbf{y}_t} \left\{ \sum_{k=1}^K (r_k w_{t,k} + s(y_{t,k} - y_{t,k-1})) + p(y_{t,0} - x_{t,0} - \sum_{k=1}^K w_{t,k}) \right. \\ \left. + \sum_{k=1}^K u(x_{t,k} - x_{t,k-1} - (y_{t,k} - y_{t,k-1}) - w_{t,k}) + G_t(y_{t,0}) \right. \\ \left. + \alpha \mathbb{E}[V_{t+1}(y_{t,0} - D_t, y_{t,1} - D_t + R_{t,1}, \dots, y_{t,K} + \sum_{k=1}^K R_{t,k} - D_t)] \right\},$$

subject to the constraints

$$0 \leq w_{t,k} \leq x_{t,k} - x_{t,k-1} - (y_{t,k} - y_{t,k-1}), \quad k = 1, \dots, K, \\ y_{t,0} \leq y_{t,1} \leq \dots \leq y_{t,K}, \\ 0 \leq y_{t,0} - x_{t,0} - (w_{t,1} + \dots + w_{t,K}).$$

The boundary condition is given by $V_{T+1}(\cdot) \equiv 0$.

In the above equation, the decision variables are $\mathbf{w}_t = (w_{t,1}, \dots, w_{t,K})$ and $\mathbf{y}_t = (y_{t,0}, \dots, y_{t,K})$, where $w_{t,k}$ is the remanufacturing quantity of type- k cores, $y_{t,0}$ is the inventory level of the serviceable product after manufacturing and remanufacturing decisions but before the demand D_t is realized, and $y_{t,k}$ is the aggregate inventory level of the serviceable product and type-1 to type- k cores after all decisions are made but before demand and returns occur. Then, the manufacturing quantity is $y_{t,0} - x_{t,0} - \sum_{k=1}^K w_{t,k}$, and the disposal quantity of type- k returned product is $x_{t,k} - x_{t,k-1} - (y_{t,k} - y_{t,k-1}) - w_{t,k}$, $k = 1, \dots, K$. The above constraints ensure that the manufacturing, remanufacturing and disposal quantities are all non-negative; and the ending inventory levels of all types of returned product are non-negative.

This inventory model has $K+1$ state variables and $2K+1$ decision variables (including one manufacturing decision, K remanufacturing decisions, and K disposal decisions) in each period. Zhou et al. (2011) establish the following structural properties of its value function:

Proposition 10.2 (Zhou et al., 2011) For $t = 1, \dots, T$,

- (a) $V_t(\mathbf{x}_t)$ is additively convex in $\mathbf{x}_t$ on $\mathbb{V}$;
- (b) $-(r_{k+1} - r_k) \leq \partial V_t(\mathbf{x}_t) / \partial x_{t,k} \leq 0$, for $k = 1, \dots, K-1$, and $\partial V_t(\mathbf{x}_t) / \partial x_{t,K} \geq -(p - r_K)$.

Based on these properties, they further show that the optimal policy has the following simple structure.

Theorem 10.3 (Zhou et al., 2011) For $t = 1, \dots, T$, the optimal manufacturing, remanufacturing, and disposal policies for period t are determined by two sets of parameters, $\{\xi_{t,k}, k = 0, \dots, K\}$ and $\{\eta_{t,k}, k = 1, \dots, K\}$, satisfying $\xi_{t,K} \leq \dots \leq \xi_{t,0}$, $\eta_{t,K} \leq \dots \leq \eta_{t,1}$, and $\xi_{t,k} \leq \eta_{t,k+1}$, $k = 0, \dots, K-1$, with $\xi_{t,-1} = \eta_{t,0} = \infty$, in the following manner: let $\mathbf{x}$ be the system state at the beginning of period t ,

- (i) there exists a unique $m \in \{0, 1, \dots, K\}$ such that, if $\xi_{t,m} \leq x_m < \xi_{t,m-1}$, then $y_{t,0}^* = \dots = y_{t,m}^* = x_m$, or if $x_m < \xi_{t,m} \leq x_{m+1}$, then $y_{t,0}^* = \dots = y_{t,m}^* = \xi_{t,m}$;
- (ii) there exists a unique $l \in \{m+1, \dots, K\}$ such that, if $\eta_{t,l} \leq x_{l-1} < \eta_{t,l-1}$, then $y_{t,l}^* = \dots = y_{t,K}^* = x_{l-1}$, or if $x_{l-1} < \eta_{t,l} \leq x_l$, then $y_{t,l}^* = \dots = y_{t,K}^* = \eta_{t,l}$; and
- (iii) for all $k \in \{m+1, \dots, l-1\}$, $y_{t,k}^* = x_k$.

We refer to Zhou et al. (2011) for detailed interpretations of the above optimal manufacturing, remanufacturing, and disposal policies. Proposition 10.2 and Theorem 10.3 extend the results for the single-return model by Simpson (1978) to a more general model with multiple types of returns. Further, Zhou et al. (2011) show that these results hold after relaxing the assumption of identical holding costs and identical disposal costs across different types of returned product to the following condition:

$$-(r_{k+1} - r_k) \leq \frac{s_k - s_{k+1}}{1 - \alpha} \leq u_k - u_{k+1}, \quad \forall k = 1, \dots, K-1, \quad (10.4)$$

where s_k and u_k are unit holding and disposal costs of type- k returned product, respectively, $k = 1, \dots, K$. The key observations for establishing Proposition 10.2 and Theorem 10.3 are as follows. Since $r_1 < r_2 < \dots < r_K < p$ and under the condition in Equation (10.4), the optimal policy gives decreasing priorities to remanufacturing type-1 to type- K types of returned product and manufacturing, and increasing priorities to disposing of type-1 to type- K types of returned product. With these priority properties of the optimal policy, we can transform the remanufacturing inventory system with multiple types of returns to serial inventory systems and then establish the desired results. We refer to Gong and Wang (2021) for a simple proof of Proposition 10.2 and detailed discussions on the intuitions.

10.3.2 Random Yield

In this subsection, we review an inventory model with multiple types of returns and random yield studied by Tao et al. (2012). This model is similar to the one studied by Zhou et al. (2011), except that the remanufacturing processes are subject to random yields and disposal of returned product is not allowed. Specifically, the remanufacturing process of type- k returned product is subject to a stochastically proportional yield δ_k , which is a random variable with support on $[0, 1]$ and mean $\bar{\delta}_k$, $k = 1, \dots, K$. Hence, if the firm remanufactures w_k units of type- k returned product, it would generate $\delta_k w_k$ serviceable products. In each period t , the firm needs to decide on an ordering quantity (or manufacturing quantity with perfect yield) z_t and a remanufacturing quantity $w_{t,k}$ for type- k returned product, $k = 1, \dots, K$.

For this model with random yield, the value function is no longer decomposable, and the optimal policy is very complicated. Tao et al. (2012) show that the value function is jointly convex and satisfies some additional properties. Based on these properties, they prove the following properties of the optimal policy.

Theorem 10.4 (Tao et al., 2012) For $t = 1, \dots, T$, suppose the system state is $(I_t, J_{t,1}, \dots, J_{t,K})$, where I_t denotes the serviceable inventory level, and $J_{t,k}$ denotes the inventory level of type- k returned product, $k = 1, \dots, K$, the optimal policy satisfies the following properties:

- (i) the optimal ordering quantity z_t^* is decreasing in I_t ;
- (ii) for any $J_{t,k} > 0$, the optimal remanufacturing quantity w_t^* satisfies $J_{t,k} \geq w_t^* > 0$ when $z_t^* > 0$;
- (iii) for any k with deterministic δ_k , $w_{t,k}^* = J_{t,k}$ when $z_t^* > 0$;
- (iv) for any k_1 with $J_{t,k_1} > 0$ and k_2 with deterministic δ_{k_2} satisfying $(1-\alpha)r_{k_1} - s_{k_1}/\bar{\delta}_{k_1} < (1-\alpha)r_{k_2} - s_{k_2}/\bar{\delta}_{k_2}$, if $w_{t,k_2}^* > 0$, then $J_{t,k_1} \geq w_{t,k_1}^* > 0$, and furthermore, $w_{t,k_1}^* = J_{t,k_1}$ if δ_{k_1} is deterministic.

These properties provide insights into the behaviors of the optimal policy. For example, properties (i) and (ii) in Theorem 10.4 show that the firm should manufacture less with a higher serviceable inventory level, and it should start remanufacturing the core of each type before it manufactures, which appears intuitive. On the other hand, Tao et al. (2012) report several counter-intuitive observations on the behaviors of the optimal policy. First, they observe that the optimal policy is not a state-independent trigger policy. That is, the optimal decision on whether or not to manufacture or remanufacture is not solely determined by a state-independent constant. This observation is in contrast to the conventional production problems with random yield, for which the optimal policy is shown to be a state-independent trigger policy (see, e.g., Henig & Gerchak, 1990). Second, they observe that the optimal remanufacturing quantities may not be monotone in their respective returned product inventories. By contrast, when yields are deterministic, Zhou et al. (2011) prove that they are non-decreasing in the returned product inventories. Third, they observe that the optimal remanufacturing quantities may not be monotone in the serviceable inventory level. By contrast, the optimal manufacturing quantity is proven to be decreasing in the serviceable inventory level. We refer to Tao et al. (2012) for detailed explanations of these counter-intuitive behaviors of the optimal policy.

Finally, Tao et al. (2012) develop three simple heuristic policies for this model with random yield. The first heuristic is called the *deterministic heuristic*. This heuristic is developed by treating the random yield rate δ_k as deterministic and replacing it with $\bar{\delta}_k$ for all $k = 1, \dots, K$. Recall that Zhou et al. (2011) have studied the model with deterministic yield and proved that its value function is decomposable and the optimal policy can be efficiently computed (see Proposition 10.2 and Theorem 10.3 in Section 10.3.1). The second heuristic is called the *myopic heuristic*, as it ignores the future and always treats the next period as the end of planning horizon. The third heuristic is called the *hybrid heuristic*. This heuristic approximates the value function in the next period with the value function for the model with deterministic yield. Tao et al. (2012) prove that the hybrid heuristic possesses similar properties to the optimal policy presented in Theorem 10.4. Through a numerical study, they show that the hybrid heuristic outperforms the other two heuristics.

10.3.3 Multi-Echelon Models

In this subsection, we review multi-echelon (or series) inventory models with remanufacturing studied by Decroix (2006). Without remanufacturing, the seminal paper by Clark and Scarf (1960) shows that echelon base-stock policy is optimal for series inventory systems. To

simplify notation, following Decroix (2006), we review the models with two stages. Since the recovery facility receiving returned products can be at either the upstream or downstream stage, there are two separate models corresponding to these two scenarios. In what follows, we first review the simpler model with upstream product remanufacturing and then the more complex model with downstream product remanufacturing.

10.3.3.1 Upstream product remanufacturing

Consider a two-stage series inventory system, where customer demands occur at stage 1, stage 1 orders from stage 2 with positive lead time L_1 , and stage 2 orders from an outside supplier with positive lead time L_2 . In addition, random product returns arrive at the remanufacturing facility at stage 2, which remanufactures returned products into serviceable products with lead time L_2 . Similar to the other models reviewed in this section, the remanufactured products are indistinguishable from the products ordered from the outside supplier in stage 2. In each period, the firm needs to decide the quantities shipped from upstream to stages 1 and 2, and the remanufacturing and disposal quantities at the remanufacturing facility. The system cost includes unit ordering costs at both stages, unit remanufacturing and disposal costs at the remanufacturing facility, unit holding and backlogging costs at stage 1, unit holding cost for items at stage 2 or in transit to stage 1, and unit holding cost for items at the remanufacturing facility or in transit to stage 2. The firm's objective is to minimize the expected total discounted cost over the T -period planning horizon.

Decroix (2006) formulates the firm's optimization problem as a dynamic program. The state variables in each period t consist of the following:

- $x_{t,1}$ = the inventory position at stage 1 (which equals its inventory level and the pipeline inventories from stage 2 to stage 1);
- $\hat{x}_{t,2}$ = echelon net inventory at stage 2 prior to arrival of shipment;
- $\mathbf{s}_{t,2} = (s_{t,2}^1, \dots, s_{t,2}^{L_2})$ = vector of shipments in transit to stage 2;
- $x_{t,u}$ = returned product inventory level at the remanufacturing facility.

Define $C_t(x_1, \hat{x}_2, x_u, \mathbf{s}_2)$ as the minimum expected total discounted cost over a $(T - t + 1)$ -period planning horizon when starting at state $(x_1, \hat{x}_2, x_u, \mathbf{s}_2)$. In addition, define $\tilde{C}_t(x_1)$ as the minimum expected discounted cost at stage 1 over an $(T - t + 1)$ -period horizon when starting at state x_1 , considering only stage 1 costs and ignoring potential upstream shortages that could prevent stage 1 from ordering its desired quantity. For this latter problem, the optimal policy with $T - t + 1$ periods remaining is known to be a base-stock policy, with the base-stock level denoted by $\tilde{x}_{1t}^*$. Decroix (2006) proves the following decomposition result that facilitates the characterization of optimal policy.

Proposition 10.3 (DeCroix, 2006) *For $T - t + 1 \geq L_1$, there exist functions $q_t(\hat{x}_2, x_u, \mathbf{s}_2)$ such that $C_t(x_1, \hat{x}_2, x_u, \mathbf{s}) = \tilde{C}_t(x_1) + q_t(\hat{x}_2, x_u, \mathbf{s}_2)$.*

For $t = T - L_1 + 1, \dots, T$, stage 1 shall not order as its order would not arrive before the end of the planning horizon. Following Proposition 10.3, the optimal ordering policy at stage 1 with $T - t + 1$ periods remaining is to bring the inventory position x_1 up to the base-stock level $\tilde{x}_{1t}^*$ subject to the available inventory $\hat{x}_2 + s_{t,2}^{L_2}$ at stage 2. To analyze the optimal ordering/remanufacturing/disposal policy at stage 2, Decroix (2006) formulates the recursive equations on $q_t(\hat{x}_2, x_u, \mathbf{s}_2)$. Further, when $T - t + 1 \geq L_1 + L_2$, we can aggregate the state variables $\hat{x}_2$ and

s_2 into the inventory position $x_2 := \hat{x}_2 + s_2^1 + \dots + s_2^{L_2}$, and reduce the function $q_t(\hat{x}_2, x_u, s_2)$ to a bivariate function $\tilde{q}_t(x_2, x_u)$ for any $T - t + 1 \geq L_1 + L_2$. Again, for $t = T - L_1 - L_2 + 1, \dots, T$, stage 2 should neither remanufacture nor order because when those arrive at stage 2, stage 1 has stopped ordering. Finally, by noting that the recursive equations on $\tilde{q}_t(x_2, x_u)$ are identical to those studied by Simpson (1978), Decroix (2006) characterizes the optimal policy at stage 2 and the remanufacturing facility as follows.

Proposition 10.4 (DeCroix, 2006) *For $T - t + 1 \geq L_1 + L_2$, the optimal policy at stage 2 and the remanufacturing facility has the same form as that in Theorem 10.1 for the single-stage system.*

10.3.3.2 Downstream product remanufacturing

Suppose now that the product remanufacturing occurs at stage 1 with lead time L_1 . In this case, Decroix (2006) assumes that the disposal of returned products is not allowed in order to obtain a simple structure of optimal policy. The other settings are almost the same as those described in Section 10.3.3.1 for the case of upstream product remanufacturing, except that the state variable $\hat{x}_{t,2}$ needs to include the returned product inventory $x_{t,u}$ at the remanufacturing facility. With slight abuse of notations, we still use $C_t(x_1, \hat{x}_2, x_u, s_2)$ to denote the minimum expected total discounted cost of the current problem over an $(T - t + 1)$ -period horizon when starting at state $(x_1, \hat{x}_2, x_u, s_2)$. In addition, define $\tilde{C}_t(x_1, x_u)$ as the minimum expected discounted cost at stage 1 and the remanufacturing facility when starting at state (x_1, x_u) , ignoring potential upstream shortages that could prevent stage 1 from ordering its desired quantity. In this setting, the optimal policy is a modification of that in Theorem 10.1 for the single-stage system (recall the current problem does not allow disposal of cores). Similar to Proposition 10.3 for the case of upstream product remanufacturing, Decroix (2006) proves the following important decomposition result for the case of downstream product remanufacturing.

Proposition 10.5 (DeCroix, 2006) *For $T - t + 1 \geq L_1$, there exist functions $q_t(\hat{x}_2, s_2)$ such that $C_t(x_1, \hat{x}_2, x_u, s) = \tilde{C}_t(x_1, x_u) + q_t(\hat{x}_2, s_2)$.*

Following this result, the optimal policy at stage 1 and the remanufacturing facility is a modified version of that in Theorem 10.1 by incorporating the no-disposal assumption and the available inventory $\hat{x}_2 + s_2^{L_2}$ at stage 2. To analyze the optimal ordering policy at stage 2, Decroix (2006) formulates the recursive equations on $q_t(\hat{x}_2, s_2)$ and aggregates the state variables $\hat{x}_2$ and s_2 into the inventory position $x_2 := \hat{x}_2 + s_2^1 + \dots + s_2^{L_2}$ when $T - t + 1 \geq L_1 + L_2$. After that, Decroix (2006) shows that the optimal ordering policy at stage 2 in each period is an echelon base-stock policy.

To summarize, when the remanufacturing facility is located at stage 2 (or the most upstream for a general N -stage system), the optimal policy for managing the system is a simple combination of the optimal policies for managing a traditional series inventory system without remanufacturing and a single-stage system with remanufacturing. By contrast, when the remanufacturing facility is located at stage 1 (or any downstream stage of an N -stage system), the optimal policy still has a simple structure when disposal of cores is not allowed. When the remanufacturing facility is located at a downstream stage and disposal of returned products is allowed, however, the decomposition result in Proposition 10.5 no longer holds. In this case, the structure of the optimal policy remains unknown and is expected to be very complex.

10.4 MODELS WITH DIFFERENTIATED REMANUFACTURED AND NEW PRODUCTS

So far, all the models we have reviewed in Sections 10.2 and 10.3 assume that customers do not differentiate between remanufactured and manufactured/ordered products. This is applicable to the scenarios such as single-use containers, parts for replacement services, and ordered/manufactured products are also remanufactured ones while using new components instead of those harvested from returned products. In this section, we review two models for which remanufactured and manufactured products are treated differently (we refer to the latter as new product): a joint inventory and pricing optimization model in Section 10.4.1, and a perishable inventory model with remanufacturing in Section 10.4.2.

10.4.1 Joint Inventory and Pricing Optimization with Differentiated Remanufactured and New Products

In this subsection, we review a joint inventory and pricing model with differentiated remanufactured and new products studied by Yan et al. (2017).

Consider a firm selling both new and remanufactured products over a finite planning horizon with T periods. At the beginning of each period t , the firm reviews the inventory levels of both products. Second, the firm sets the selling price p_1 for the new product and p_2 for the remanufactured product. The value of a new product is modeled by a random variable v with distribution $F(\cdot)$, and that of a remanufactured product is $\eta(v)$, with $0 \leq \eta(v) \leq v$ (since customers typically value remanufactured products less than new products). It is also assumed that $\eta(v)$ and $v - \eta(v)$ are strictly increasing in v . Thus, a customer's utility of buying a new (respectively, remanufactured) product is $v - p_1$ (respectively, $\eta(v) - p_2$). A customer will choose to buy the product that gives her a higher non-negative utility. Hence, the probabilities for a customer to buy new and remanufactured products, denoted by $\lambda_1(p_1, p_2)$ and $\lambda_2(p_1, p_2)$, can be computed as

$$\lambda_1(p_1, p_2) = \Pr(v - \eta(v) \geq p_1 - p_2, v \geq p_1), \quad (10.5)$$

$$\lambda_2(p_1, p_2) = \Pr(v - \eta(v) < p_1 - p_2, \eta(v) \geq p_2). \quad (10.6)$$

The demands in period t for new and remanufactured products are random and depend on both prices, and they are modeled by

$$D_{it}(p_1, p_2) = \lambda_i(p_1, p_2)d_t + \varepsilon_{it}, i = 1, 2,$$

where $\varepsilon_{it} \in [\underline{\varepsilon}_i, \bar{\varepsilon}_i]$ is a random noise with mean zero, $i = 1, 2$, d_t is a deterministic number representing the potential total demand for the firm's products in period t , and $\lambda_1(p_1, p_2)$ and $\lambda_2(p_1, p_2)$ are, respectively, given in Equations (10.5) and (10.6). The random noises ε_{1t} and ε_{2t} can be correlated, but they are independent of those of other periods.

For ease of analysis, Yan et al. (2017) convert the pricing decisions p_1 and p_2 to the equivalent decisions λ_1 and λ_2 (i.e., the fractions of customers who purchase new and remanufactured

products). To this end, they express the prices p_1 and p_2 as functions of λ_1 and λ_2 through Equations (10.5) and (10.6), and set the feasible set Ω of (λ_1, λ_2) as

$$\Omega = \{(\lambda_1, \lambda_2) : 0 \leq \lambda_1 \leq 1, 0 \leq \lambda_2 \leq 1, 0 \leq \lambda_1 + \lambda_2 \leq 1\}.$$

Besides, the firm needs to decide the manufacturing quantity for the new product, with unit cost c_1 . The timing of this decision depends on the firm's production strategy for the new product, and it is before (respectively, after) demands are realized under the make-to-stock (respectively, make-to-order) strategy. After the firm's pricing (and manufacturing, if applicable) decisions, random demands for both products in period t are realized and satisfied as much as possible. Unsatisfied demands are backlogged, incurring unit shortage costs π_0 and π for new and remanufactured products, respectively. Excess inventories of new and remanufactured products are carried to the next period, incurring unit holding costs h_0 and h , respectively. Random return R_t is realized at the end of period t , and remanufactured immediately with a unit cost c_2 , with $c_2 \leq c_1$. The manufacturing and remanufacturing lead times are both zero. The firm's objective is to maximize its expected total discounted profit over the planning horizon of T periods.

In what follows, we separately present the results for models under make-to-order and make-to-stock strategies for the new product.

10.4.1.1 Make-to-order model

Under the make-to-order (MTO) strategy, the firm manufactures new products after demands are realized in each period. Since manufacturing lead time is zero, demand for the new product is always satisfied, and there is no need to hold new product inventory. As a consequence, the system state consists of a single variable, i.e., the inventory level of remanufactured product. By analyzing the firm's optimality equations, Yan et al. (2017) establish the following structural properties of the optimal pricing policy.

Theorem 10.5 (Yan et al., 2017) *For $t = 1, \dots, T$, suppose the starting inventory level of remanufactured product in period t is x_0 , the optimal fractions of customers $(\lambda_{1t}^*(x_0), \lambda_{2t}^*(x_0))$ and the optimal prices $(p_{1t}^*(x_0), p_{2t}^*(x_0))$ satisfy the following properties:*

- (i) $\lambda_{1t}^*(x_0)$ is decreasing in x_0 , while $\lambda_{1t}^*(x_0) + \lambda_{2t}^*(x_0)$ and $x_0 - \lambda_{2t}^*(x_0)d_t$ are increasing in x_0 ;
- (ii) $p_{2t}^*(x_0)$ is decreasing in x_0 , while $p_{1t}^*(x_0) - p_{2t}^*(x_0)$ is increasing in x_0 .

Yan et al. (2017) also analyze how the optimal customer choice probabilities λ_{it}^* depend on the inventory holding cost h_0 and the backlogging cost π_0 of the remanufactured product, and obtain several monotonic properties. For example, they prove that $\lambda_{1t}^*(x_0)$ is decreasing in h_0 and increasing in π_0 , while conversely, $\lambda_{2t}^*(x_0)$ is increasing in h_0 and decreasing in π_0 .

10.4.1.2 Make-to-stock model

Under the make-to-stock (MTS) strategy, the firm manufactures new product before demands are realized in each period. In this case, the firm also needs to hold inventory of new product and decides the manufacturing quantity (or equivalently, the post-production inventory level z for the new product). The system state consists of two variables: the inventory level x_0 of remanufactured product, and the inventory level u of new product. Yan et al. (2017) establish the following structural properties of the optimal production policy for the new product.

Theorem 10.6 (Yan et al., 2017) For $t = 1, \dots, T$, given system state (x_0, u) in period t , the optimal production policy for new product is a base-stock policy with base-stock level $z_t^0(x_0)$, i.e., if $u < z_t^0(x_0)$, then manufacture to raise the inventory level of new product to $z_t^0(x_0)$; otherwise, manufacture nothing. In addition, $z_t^0(x_0)$ is decreasing in x_0 .

Yan et al. (2017) also study the optimal pricing policy under MTS and offer insights into why the monotonic properties in Theorem 10.5 for the MTO system break down for the MTS system. The monotonic properties of the optimal policy on h_0 and π_0 for the MTO system cannot be extended to the MTS system either. Besides, they compare the optimal profits under the MTO and MTS systems, and establish lower and upper bounds for the difference.

Furthermore, Yan et al. (2017) extend the above base models to include general production lead times, returned product acquisition, and sales-dependent product returns. Due to space limitations of this chapter, we omit reviewing these extensions.

10.4.2 Perishable Inventory

In this subsection, we review a perishable inventory model with remanufacturing studied by Fu et al. (2019), which is motivated by an emerging practice in the cut flower industry.

Consider a firm selling a perishable product with a lifetime of two periods over a planning horizon of T periods. At the beginning of each period t , the firm reviews the inventory level $x_{t,1}$ of the serviceable product with one-period remaining lifetime (hereafter referred to as product 1) and the inventory level w_t of the returned product (also with one-period remaining lifetime). Second, the firm decides on the remanufacturing quantity $q_{t,1}$ from the returned product and the manufacturing quantity $q_{t,2}$, subject to constraints $0 \leq q_{t,1} \leq w_t$ and $q_{t,2} \geq 0$. The remanufactured products are indistinguishable from product 1, but they are distinguishable from the manufactured products with a two-period remaining lifetime (hereafter referred to as product 2). The manufacturing and remanufacturing lead times are both zero. Third, random demands $D_{t,1}$ and $D_{t,2}$ for products 1 and 2 are realized and satisfied by the firm's on-hand inventories. If product i is stock-out, then a deterministic proportion $\gamma_{i,3-i} \in [0, 1]$ of the unsatisfied demand for product i will purchase product $3-i$ (if available) as a substitute, $i = 1, 2$. The eventual unsatisfied demands are lost. Fourth, the leftover returned product is disposed of, the unsold product 1 is salvaged with unit value s , and the unsold product 2 is carried to the next period with unit holding cost h . At the end of period t , the random return R_t , modeled as a random proportion ε_t of the sold product 2 in period t , is realized and carried to the next period. Here, $\varepsilon_1, \dots, \varepsilon_T$ are independent random variables with support $[0, 1]$. With a discount factor $\alpha \in (0, 1]$, the firm's objective is to maximize the expected total discounted profit over the planning horizon of T periods.

Additional parameters of the model are summarized as follows:

- c_1 = unit remanufacturing cost;
- c_2 = unit manufacturing cost;
- p_i = selling price of product i , $i = 1, 2$;
- h = unit holding cost of unsold product 2.

It is natural to assume $0 \leq s < c_1 < p_1$, $0 \leq c_2 < p_2$, and $p_1 < p_2$. Besides, Fu et al. (2019) impose the following two inequalities on the cost parameters:

$$p_1 - s \geq \gamma_{12}(p_2 + h), \quad \text{and} \quad p_2 + h \geq \gamma_{21}(p_1 - s) + \alpha p_1.$$

As explained in detail by Fu et al. (2019), these inequalities ensure that the firm has no incentive to substitute one product with the other product when the former is still in stock, and they are crucial in establishing the concavity of the firm's value function.

The firm's optimization problem in this model can be formulated as a dynamic program as follows. The state variables in each period t are x_t and w_t , with the state space $\mathfrak{R}_+^2$. The system dynamics from period t to period $t+1$ are given by

$$\begin{aligned} x_{t+1} &= (q_{t,2} - D_{t,2} - \gamma_{12}(D_{t,1} - (x_t + q_{t,1}))^+)^+, \\ w_{t+1} &= R_t = \min\{q_{t,2}, D_{t,2} + \gamma_{12}(D_{t,1} - (x_t + q_{t,1}))^+\} \varepsilon_t. \end{aligned}$$

Denote $V_t(x, w)$ as the firm's maximum expected total discounted profit from period t to T , given the system state (x, w) in period t . Fu et al. (2019) establish several structural properties of the value function $V_t(x, w)$, including its concavity. Based on the concavity result, they then characterize the firm's optimal manufacturing and remanufacturing policy as follows.

Theorem 10.7 (Fu et al., 2019) *For $t = 1, \dots, T$, suppose the system state is (x, w) in period t . The optimal remanufacturing policy is a modified base-stock policy with the base-stock level y_t^* , i.e., $q_{t,1}^* = \max\{x, \min\{x + w, y_t^*\}\}$, and the optimal manufacturing policy is given by*

$$q_{t,2}^*(x, w) = \Phi_t^*(\max\{x, \min\{x + w, y_t^*\}\}),$$

where the function $\Phi_t^*(x)$ decreases in x and converges to a constant $q_{t,2}^L$ when $x \rightarrow \infty$.

Therefore, the optimal manufacturing quantity $q_{t,2}^*(x, w)$ decreases in the inventory level of product 1 after remanufacturing, and converges to the constant $q_{t,2}^L$ when the inventory level of product 1 is sufficiently large. Fu et al. (2019) provide upper and lower bounds on $q_{t,2}^L$, both of which are newsvendor fractile solutions. Furthermore, they conduct a numerical study to investigate the operational effects of remanufacturing and demand substitution, and the economic and environmental values of remanufacturing. Finally, they extend some of the results to a model with a joint production capacity and a model with a product lifetime longer than two periods.

10.5 SUMMARY AND FUTURE DIRECTIONS

We have reviewed some recent studies on periodic-review inventory models with product returns and remanufacturing, including single-stage single-return models, single-stage multi-return models, multi-echelon models, and models with differentiated remanufactured and new products. We end this chapter by identifying three directions for future research.

First, when considering product returns and remanufacturing, lost-sales models seem not yet been well studied. Jia et al. (2016) consider a lost-sales model and argue that lost sales better fit the remanufacturing practices that motivate their research. Lost sales introduce additional challenges in the analysis, in particular, the decomposition results presented in the earlier sections for the backlogging models would no longer hold. Moreover, if allowing positive manufacturing/

remanufacturing lead times, lost-sales models result in high-dimensional stochastic optimization problems, calling for the development of effective heuristic inventory policies.

Second, most of the existing models assume that product returns are independent of past demands or sales. This assumption provides tractability, but it may be violated in many practical scenarios in which returns are highly dependent on past demands or sales. However, the studies on models with dependent demands and returns are still limited. See the models reviewed in Sections 10.2.5 and 10.4.2 for examples. We refer to Chou et al. (2020) for a detailed review of the studies of such models. When returns depend on past demands or sales, the optimal control policy likely becomes very complex and computationally intractable because of the high dimensionality of the resulting dynamic program. Hence, it is expected that the focus shall be more on the development of effective heuristic policies.

Finally, as remanufacturing mostly is conducted by manufacturers of electronics, machinery, engines, etc., it is likely that the system involves multiple products/components in an assemble-to-order (ATO) type of setting. The study of ATO systems with product returns and remanufacturing is relatively scarce. This can be a fruitful research direction provided the literature on much more well-studied ATO systems without remanufacturing.

REFERENCES

- Chou, M. C., Sim, C.-K., & Yuan, X.-M. (2020). Policies for inventory models with product returns forecast from past demands and past sales. *Annals of Operations Research*, 288, 137–180.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- DeCroix, G. (2006). Optimal policy for a multiechelon inventory system with remanufacturing. *Operations Research*, 54(3), 532–543.
- DeCroix, G., Song, J.-S., & Zipkin, P. (2005). A series system with returns: Stationary analysis. *Operations Research*, 53(2), 350–362.
- DeCroix, G., Song, J.-S., & Zipkin, P. (2009). Managing an assemble-to-order system with returns. *Manufacturing & Service Operations Management*, 11(1), 144–159.
- DeCroix, G., & Zipkin, P. (2005). Inventory management for an assembly system with product or component returns. *Management Science*, 51(8), 1250–1265.
- Fleischmann, M., Bloemhof-Ruwaard, J. M., Dekker, R., van der Laan, E. A., Van Nunen, J. A. E. E., & Van Wassenhove, L. (1997). Quantitative models for reverse logistics: A review. *European Journal of Operational Research*, 103(1), 1–17.
- Fu, K., Gong, X., & Liang, G. (2019). Managing perishable inventory systems with product returns and remanufacturing. *Production and Operations Management*, 28(6), 1366–1386.
- Gong, X., & Chao, X. (2013). Optimal control policy for capacitated inventory systems with remanufacturing. *Operations Research*, 61(3), 603–611.
- Gong, X., & Liu, S. (2021). Managing hybrid manufacturing/remanufacturing inventory systems with random production capacities. Available at SSRN <https://ssrn.com/abstract=3967151>.
- Gong, X., & Wang, T. (2021). Preservation of additive convexity and its applications in stochastic optimization problems. *Operations Research*, 69(4), 1015–1024.
- Henig, M., & Gerchak, Y. (1990). The structure of periodic review policies in the presence of random yield. *Operations Research*, 38(4), 634–643.
- Ilgin, M. A., & Gupta, S. M. (2012). *Remanufacturing modeling and analysis*. CRC Press.
- Inderfurth, K. (1997). Simple optimal replenishment and disposal policies for a product recovery system with leadtimes. *OR Spectrum*, 19(2), 111–122.
- Jia, J., Xu, S. H., & Guide Jr, V. D. R. (2016). Addressing supply-demand imbalance: designing efficient remanufacturing strategies. *Production and Operations Management*, 25(11), 1958–1967.

- Kiesmüller, G. P. (2003). A new approach for controlling a hybrid stochastic manufacturing/remanufacturing system with inventories and different leadtimes. *European Journal of Operational Research*, 147(1), 62–71.
- Simpson, V. (1978). Optimum solution structure for a repairable inventory system. *Operations Research*, 26(2), 270–281.
- Souza, G. C. (2013). Closed-loop supply chains: A critical review, and future research. *Decision Sciences*, 44(1), 7–38.
- Tao, Z., Gao, X., & Zhou, S. X. (2020). Stationary analysis for inventory management when product return depends on past demand. Working paper.
- Tao, Z., & Zhou, S. X. (2014). Approximation balancing policies for inventory systems with remanufacturing. *Mathematics of Operations Research*, 39(4), 1179–1197.
- Tao, Z., Zhou, S. X., & Tang, C. S. (2012). Managing a remanufacturing system with random yield: properties, observations, and heuristics. *Production and Operations Management*, 21(5), 797–813.
- USITC. (2012). Remanufactured goods: An overview of the U.S. and global industries, markets, and trade. Tech. rep., USITC Publication, Washington, DC.
- Xin, L. (2021). *Asymptotic analysis of a remanufacturing system with non-identical lead times*. Available at SSRN. URL <https://ssrn.com/abstract=3760906>.
- Yan, X., Chao, X., Lu, Y., & Zhou, S. X. (2017). Optimal policies for selling new and remanufactured products. *Production and Operations Management*, 26(9), 1746–1759.
- Zhou, S. X., Tao, Z., & Chao, X. (2011). Optimal control of inventory systems with multiple types of remanufacturing products, manufacturing & service oper. *Management*, 13(1), 20–34.
- Zhou, S. X., & Yu, Y. (2011). Optimal product acquisition, pricing, and inventory management for systems with remanufacturing. *Operations Research*, 59(2), 514–521.

11. Approximation algorithms for stochastic inventory systems

Cong Shi

11.1 INTRODUCTION

Most (if not all) core problems studied in inventory management fall into the category of multistage stochastic optimization models. Particularly, one has to make multiple, typically dependent decisions over time to optimize a certain objective function under the uncertainty of how the system will evolve in the future time horizon. Unfortunately, it is usually computationally intractable to find exact optimal solutions for these fundamental models, and even finding good solutions is very challenging, both in theory and practice. Thus, in many scenarios, there is a challenging gap in leveraging the respective models being used in effective and practical policies.

In this book chapter, we will survey several recent algorithmic and performance analysis techniques that can be applied to core stochastic inventory models to construct provably near-optimal algorithms, particularly algorithms that admit worst-case performance guarantees (see Levi (2014) and Shi (2014)). The notion of worst-case performance guarantees has been used extensively in computer science in the analysis of approximation algorithms for combinatorial NP-hard problems (see Vazirani (2001) and Williamson and Shmoys (2010)).

Definition 11.1 An α -approximation algorithm for an optimization problem is a polynomial-time algorithm that for all instances of the problem produces a solution whose value is within a factor of α of the value of an optimal solution.

For an α -approximation algorithm, we call α the *worst-case performance guarantee*, *approximation ratio*, or *approximation factor* of the algorithm. Note that $\alpha > 1$ for minimization problems, while $\alpha < 1$ for maximization problems.

Intuitively, the notion of a worst-case guarantee could be viewed in the context of a “game” between the algorithm designer and an adversary opponent. The algorithm designer proposes an algorithm, and the adversary attempts to generate the worst instance, for which the relative gap between the cost of the proposed algorithm and the optimal cost is maximized. The worst-case guarantee is simply a statement to what extent the adversary can be successful with respect to a given algorithm.

11.1.1 State-of-the-Art Overview

The concept of approximation algorithms has been applied to several deterministic problems in inventory management (e.g., Silver and Meal (1973), Roundy (1993), Levi et al. (2006, 2008b, 2008c), Shen et al. (2009), Cheung et al. (2016), Nagarajan and Shi (2016), DeValve et al. (2020)). This book chapter, however, focuses on *stochastic* inventory systems.

The recent stream of research on designing approximation algorithms for the multiperiod stochastic inventory control problems was initiated by Levi et al. (2007a) who proposed a 2-approximation algorithm for the basic uncapacitated backlogged model. Levi et al. (2007a) introduced a novel marginal cost accounting scheme that is used to decompose the total cost by decisions (rather than by periods) and devised a cost-balancing rule to balance between the marginal holding and backlogging costs. Truong (2014) provided a 2-approximation algorithm for the same problem via a look-ahead (myopic) optimization approach. Leveraging a new idea termed *forced backlogging cost*, Levi et al. (2008d) proposed a 2-approximation algorithm for the capacitated backlogged model. Levi et al. (2008a) designed a 2-approximation algorithm for the lost-sales inventory model with positive lead times. This was clearly a breakthrough result as the problem has a high dimension due to the need to keep track of the pipeline inventory. The analysis is much more involved (compared to the backlogging counterpart), which relies crucially on the notion of truncated inventory positions. Levi and Shi (2013) gave a 3-approximation algorithm for the uncapacitated backlogged model with fixed costs (also known as the stochastic lot-sizing problem). Note the fixed cost component introduces nonlinearity into the objective function, and therefore a novel randomized cost-balancing rule is applied to resolve this difficulty. Subsequently, Shi et al. (2014) gave a 4-approximation algorithm for the capacitated backlogged problem with fixed costs.

There has also been a series of studies devoted to perishable inventory systems. Perishable inventory systems (with a general positive product lifetime $m \geq 1$) are notoriously hard to solve, due to the need to keep track of the age information of the on-hand inventory which leads to an m -dimensional state vector. Chao et al. (2015) proposed an approximation algorithm with a worst-case guarantee between 2 and 3 for perishable inventory systems. They developed a nested marginal cost accounting scheme for perishable inventory systems. This scheme is similar in spirit to that developed by Levi et al. (2007a), but has a more complex and nested structure due to the multi-dimensional inventory state representing the age distribution of on-hand inventory. The worst-case performance analysis departs from previous studies, which rely on the existence of a one-to-one matching between the supply and demand units when the inventory units are consumed in a first-in-first-out manner. That is, when analyzing the performances of the approximation algorithms, all the previous studies “*geometrically*” match product units in a one-to-one manner for the systems operating under two different policies; and the costs for each pair of matched units can be readily compared. However, the perishability of products destroys this matching mechanism. To overcome this difficulty, they introduced a key new concept, called the trimmed on-hand inventory level, defined as the part of on-hand inventory units ordered before a particular time. This key concept allows for cost comparison between two different policies. Subsequently, Chao et al. (2018) and Zhang et al. (2016) studied perishable inventory systems with capacity and with setup costs, respectively. Zhang et al. (2019) also tackled the perishable inventory control problem and obtained a 2-approximation for several special cases of the model.

There has been a blossom of theories and methods developed for more complex stochastic inventory systems. Tao and Zhou (2014) designed a 2-approximation algorithm for stochastic inventory systems with remanufacturing. Leveraging a new concept called the delayed forced holding and production cost, Jiang et al. (2019) designed a 2-approximation algorithm for stochastic inventory systems under α or β service level constraints. Levi et al. (2017) gave a 2-approximation algorithm for multi-echelon (serial) inventory systems. Chu and Shen (2010) gave a 1.26-approximation algorithm for the one-warehouse-multi-retailer (OWMR) system

under service level constraints. Xin (2021) gave a 1.79-approximation algorithm for continuous review lost-sales inventory systems under Poisson demands. The main idea is to establish a lower bound (using the backlogging system) and an upper bound (that is the minimum between constant-order and base-stock policies). It is shown that beyond a given threshold α^* , the constant-order stays the same but OPT continues to grow. So it suggests that the worst-case ratio can only be within $[0, \alpha^*]$. Furthermore, the author establishes that the ratio of upper bound to lower bound is monotone in α within $[0, \alpha^*]$, and hence claims the worst-case ratio is achieved at α^* , resulting in 1.79. Truong and Roundy (2011) designed a 2-approximation algorithm for the multimachine, multiproduct lost-sales system with capacity expansion decisions. For the distribution-free or the black-box model, Levi et al. (2007b) also proposed an approximation scheme based on a sample average approximation approach for multiperiod stochastic inventory problems.

Besides constant-factor approximation algorithms, there has also been a stream of recent and growing studies focusing on asymptotic analysis (with parameter scaling) of stochastic inventory systems (especially high-dimensional ones). We refer interested readers to a comprehensive survey by Goldberg et al. (2019).

We tabulate all the existing key (constant) approximation ratios in Table 11.1.

It is also important to mention that in computational experiments (e.g., Hurley et al. (2007)), the approximation policies perform significantly better than the worst-case guarantees, in many cases within a couple of percentages of optimum. Moreover, in many important settings, these policies outperform simple myopic policies.

Table 11.1 Summary of current approximation results for stochastic inventory control systems

Stochastic Inventory Control Systems	Approx. Ratio	References
Backlogged	2	Levi et al. (2007a), Truong (2014)
Lost-Sales	2	Levi et al. (2008a)
Backlogged, Capacity	2	Levi et al. (2008d)
Backlogged, Fixed Cost	3	Levi and Shi (2013)
Backlogged, Fixed Cost, Capacity	4	Shi et al. (2014)
Service Level	2	Jiang et al. (2019)
Backlogged, Perishable	2 to 3	Chao et al. (2015), Zhang et al. (2019)
Backlogged, Perishable, Capacity, Lead Time	3	Chao et al. (2018)
Backlogged, Perishable, Fixed Cost	3 to 4	Zhang et al. (2016)
Backlogged, Remanufacturing	2	Tao and Zhou (2014)
Backlogged, Serial System	2	Levi et al. (2017)
Service Level, One-Warehouse-Multi-Retailer	1.26	Chu and Shen (2010)
Lost-Sales, Poisson Demand, Continuous Review	1.79	Xin (2021)
Lost-Sales, Capacity Expansion, Multiproduct	2	Truong and Roundy (2011)

11.2 STOCHASTIC INVENTORY SYSTEMS: BASIC MODEL

11.2.1 Model Formulation

We consider a finite planning horizon of T periods indexed $t = 1, \dots, T$. The demands over these periods are random variables, denoted by $D_1, \dots, D_T$, and the goal is to coordinate a sequence of orders over the planning horizon to satisfy these demands with minimum expected cost. As a general convention, from now on we will refer to a random variable and its realization using capital and lower-case letters.

In each period $t = 1, \dots, T$, four types of costs are incurred, a per-unit ordering cost c_t for ordering any number of units at the beginning of period t , a per-unit holding cost h_t for holding excess inventory from period t to $t+1$, a per-unit backlogging penalty b_t that is incurred for each unsatisfied unit of demand at the end of period t . Unsatisfied units of demand are usually called *backorders*. Each unit of unsatisfied demand incurs a per-unit backlogging penalty cost b_t in each period t until it is satisfied. In addition, we consider a model with a lead time of L periods between the time an order is placed and the time at which it actually arrives. We assume that the lead time is a known integer L . We assume without loss of generality that the discount factor is equal to 1, and that $c_t = 0$ and $h_t, b_t \geq 0$, for each t . (If $c_t > 0$, one can readily derive an equivalent system with $c_t = 0$ and modified h_t and b_t .)

At the beginning of each period s , we observe what is called an *information set* denoted by f_s . The information set f_s contains all of the information that is available at the beginning of time period s . More specifically, the information set f_s consists of the realized demands $d_1, \dots, d_{s-1}$ over the interval $(1, s)$, and possibly some exogenous information. The information set f_s in period s is one specific realization in the set of all possible realizations of the random vector $F_s = (D_1, \dots, D_{s-1})$. The set of all possible realizations is denoted by $\mathcal{F}_s$. The observed information set f_s induces a given conditional joint distribution of the future demands $(D_s, \dots, D_T)$. For ease of notation, D_t will always denote the random demand in period t according to the conditional joint distribution in some period $s \leq t$, where it will be clear from the context to which period s it refers. The index t will be used to denote a general time period, and s will always refer to the current period. The only assumption on the demands is that for each $s = 1, \dots, T$, and each $f_s \in \mathcal{F}_s$, the conditional expectation $\mathbb{E}[D_t | f_s]$ is well defined and finite for each period $t \geq s$. In particular, we allow for non-stationary and correlation between the demands in different periods.

The traditional approach to studying these models has been dynamic programming. Using the dynamic programming approach, it can be shown that *state-dependent base-stock policies are optimal* (see, e.g., Zipkin (2000)). However, the computational complexity of the resulting dynamic programs is very sensitive to the dimension of the sets $\mathcal{F}_s$. In particular, in many practical scenarios, these sets are of high dimension, which leads to dynamic programming formulations that are computationally intractable. In fact, it has been shown by Halman et al. (2009) that this model is #P-hard, even for the special case of independent discrete, finite support demands.

11.2.2 Dual-Balancing Policy

We shall describe the dual-balancing policy following Levi et al. (2007a). This policy is based on two major ideas:

Marginal cost accounting scheme. The standard dynamic programming approach directly assigns to the decision of how many units to order in each period only the expected holding and backlogging costs incurred in that period, although this decision might affect the costs in future periods. Instead, the marginal cost accounting scheme assigns to the decision in each period all the expected costs that, once this decision is made, become unaffected by any decision made in future periods. These costs may still depend on future demands.

Cost balancing. The idea of cost balancing was used in the past to construct heuristics with constant performance guarantees for deterministic inventory problems (e.g., Silver and Meal (1973)). The key observation in the above model is that any policy in any period incurs potential expected costs due to over-ordering (namely, expected holding costs of carrying excess inventory) and under-ordering (namely, expected backlogging costs incurred when demand is not met on time). The dual-balancing policy described below simply aims to repeatedly balance the expected (marginal) holding cost against the expected (marginal) backlogging cost. The cost balancing depends on the marginal cost accounting scheme.

11.2.2.1 Marginal cost accounting

We first introduce a *marginal holding cost accounting approach*. Without loss of generality, assume that the ordered supply units are consumed on a *first-ordered, first-consumed basis*. The key observation under this assumption is that once an order is placed in some period, then the expected holding cost that the units just ordered will incur over the rest of the planning horizon is a function only of the realized demands over the rest of the horizon, not of any future orders. Hence, within each period, we can associate the overall expected holding cost that is incurred by the units ordered in this period over the entire horizon. We note that similar ideas of holding cost accounting were used previously in the context of models with continuous time, infinite horizon, and stationary (Poisson distributed) demand (see, e.g., the work of Axsäter and Lundell (1984) and Axsäter (1990)). More specifically, let x_s be the *inventory position* at the beginning of period s that captures the total sum of the physical *on-hand inventory* and the *outstanding orders* (placed in past periods, but still on the way) minus the pending backlogged demand. Say now that q_s units were ordered in period s , and consider a future period $t \geq s+L$. Then the holding cost incurred by the q_s units ordered in period s at the end of period t is $h_t(q_s - (D_{[s,t]} - x_s)^+)^+$, where $x^+ = \max(x, 0)$ and $D_{[s,t]} = \sum_{j=s}^t D_j$ is the cumulative demand over the interval $[s, t]$. Observe that if $D_{[s,t]} \leq x_s$, then none of the q_s units has been yet consumed. When $D_{[s,t]}$ exceeds x_s , the q_s units are used to satisfy the demand until all of them are consumed. It follows that the total holding cost incurred by the q_s units ordered in period s over the entire horizon is equal to

$$H_s = H_s(Q_s) \triangleq \sum_{t=s+L}^T h_t(Q_s - (D_{[s,t]} - X_s)^+)^+. \quad (11.1)$$

Because X_s and Q_s are realized at the beginning of period s (whereas, x_s and q_s are the realizations of X_s and Q_s , respectively), then, as seen from the beginning of period s , this quantity depends only on future demands and not on any future decisions.

In addition, in an uncapacitated model the decision of how many units to order in each period affects the expected backlogging cost in only a single future period, namely, a lead time

ahead. Now let Π_s be the backlogging cost incurred in period $s+L$, for each $s = 1-L, \dots, T-L$. In particular, it is straightforward to verify that

$$\Pi_s \triangleq b_{s+L}(D_{[s,s+L]} - (X_s + Q_s))^+, \quad (11.2)$$

where $D_j \triangleq 0$ with probability 1 for each $j \leq 0$. (Observe that the supply units captured by $X_s + Q_s$ will become available by time period $s+L$, and that no order placed after period s will arrive by time period $s+L$.)

Now let $\mathcal{C}(P)$ be the cost of a feasible policy P and use the superscript P to relate the respective quantities to that policy. Clearly,

$$\mathcal{C}(P) \triangleq \sum_{t=1-L}^0 \Pi_t^P + H_{(-\infty, 0]} + \sum_{t=1}^{T-L} (H_t^P + \Pi_t^P), \quad (11.3)$$

where $H_{(-\infty, 0]}$ denotes the total holding cost incurred by units ordered before period 1 (given as an input). We note that the first two expressions $\sum_{t=1-L}^0 \Pi_t^P$ and $H_{(-\infty, 0]}$ are the same for any feasible policy and each realization of demand, and therefore we will omit them. Because they are non-negative, this will not affect our approximation results. Also observe that, without loss of generality, it can be assumed that $Q_t^P = H_t^P = 0$ for any policy P and each period $t = T-L+1, \dots, T$, because nothing that is ordered in these periods can be used within the given planning horizon. We now can re-define

$$\mathcal{C}(P) \triangleq \sum_{t=1}^{T-L} (H_t^P + \Pi_t^P). \quad (11.4)$$

The cost accounting scheme in Equation (11.4) above is marginal; i.e., in each period we account for all the expected costs that become unaffected by any future decision.

11.2.2.2 Policy description

The dual-balancing policy (denoted by DB) is conceptually simple. The decision in each period, how much to order, is based on the following balancing quantity. Given the observed information set f_t in each period t , compute the *balancing quantity* q_t^{DB} which balances the expected marginal holding cost incurred by the units ordered against the expected backlogging cost in period $t+L$. That is, q_t^{DB} uniquely solves

$$\mathbb{E}[H_t^{DB}(q_t^{DB}) | f_t] = \mathbb{E}[\Pi_t^{DB}(q_t^{DB}) | f_t]. \quad (11.5)$$

Note that the left-hand side (LHS) of Equation (11.5) is monotone increasing (starting from 0 to ∞) and the right-hand side (RHS) is monotone decreasing (starting from a finite value to 0), and therefore this balancing quantity is unique and can be efficiently computed via bisection search.

11.2.2.3 Worst-case analysis

We provide an abbreviated argument that this dual-balancing policy gives a 2-approximation. By the policy construction, we know that, with probability 1, $\mathbb{E}[H_t^{DB} | F_t] = \mathbb{E}[\Pi_t^{DB} | F_t] \triangleq Z_t^{DB}$

for each period $t = 1, \dots, T - L$. This implies that, for each period t , $\mathbb{E}[H_t^{DB} + \Pi_t^{DB} | F_t] = 2Z_t^{DB}$, which further implies that the total cost incurred by the dual-balancing policy is given by

$$\mathbb{E}[\mathcal{C}(DB)] = \sum_{t=1}^{T-L} \mathbb{E}[H_t^{DB} + \Pi_t^{DB}] = 2 \cdot \sum_{t=1}^{T-L} \mathbb{E}[Z_t^{DB}]. \quad (11.6)$$

To complete the worst-case analysis, we would like to show that the expected cost of an optimal policy denoted by OPT is at least $\sum_{t=1}^{T-L} \mathbb{E}[Z_t^{DB}]$, i.e., at least half of the cost of DB . This will be done by amortizing the cost of OPT against the cost of DB . For each period $t = 1, \dots, T - L$, let Y_t^P be the inventory position of a policy P in period t after ordering, i.e., $Y_t^P = X_t^P + Q_t^P$. In the subsequent analysis, we will use a random partition of periods $t = \{1, 2, \dots, T - L\}$ to the following two sets:

- The set $\mathcal{T}_H \triangleq \{t : Y_t^{OPT} > Y_t^{DB}\}$ consists of periods in which OPT had a higher inventory position than DB after ordering.
- The set $\mathcal{T}_\Pi \triangleq \{t : Y_t^{OPT} \leq Y_t^{DB}\}$ consists of periods in which the inventory position of OPT does not exceed that of DB after ordering.

It is clear that the above two sets are mutually exclusive and exhaustive.

Next, we will argue that the total holding cost incurred by OPT is higher than the marginal holding cost incurred by DB in periods that belong to $\mathcal{T}_H$, and that the total backlogging cost incurred by OPT is higher than the backlogging cost incurred by DB associated with periods within $\mathcal{T}_\Pi$.

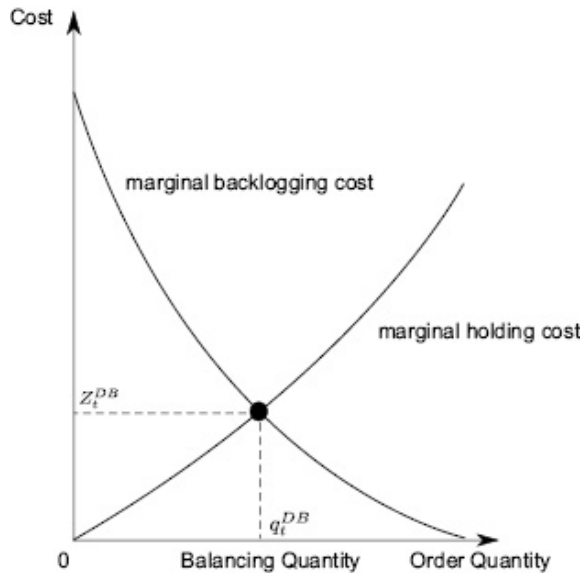


Figure 11.1 *Illustration of the DB policy*

Lemma 11.1 (Levi et al. (2007a)) *The holding cost and backlogging cost incurred by OPT are denoted by H^{OPT} and Π^{OPT} , respectively. Then, with probability 1,*

$$H^{OPT} \geq \sum_t H_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_H), \quad \Pi^{OPT} \geq \sum_t \Pi_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_\Pi).$$

The high-level idea of Lemma 11.1 is as follows. In each period $t \in \mathcal{T}_H$, the inequality $Y_t^{DB} < Y_t^{OPT}$ holds and implies that the q_t^{DB} units ordered by the DB policy in period t have been ordered by OPT either in period t or even earlier (if we match the supply units against the demand units using a FIFO rule). Thus, the marginal holding cost these q_t^{DB} units incur under OPT is higher than that under DB. It is important to note that this argument only works for marginal holding costs as they decompose the total holding costs into each decision (independent of other decisions). Traditional per-period-based holding cost accounting will not work.

On the other hand, in each period $t \in \mathcal{T}_\Pi$, the inequality $Y_t^{DB} \geq Y_t^{OPT}$ holds and implies that the backlogging incurred by OPT at the end of period $t+L$ will be higher than that of the DB in that period. Note that without ordering capacities, the marginal backlogging cost accounting is the same as the traditional per-period-based backlogging cost accounting.

Theorem 11.1 (Levi et al. (2007a)) *For each instance of the stochastic lot-sizing problem, the expected cost of the dual-balancing policy DB is at most twice the expected cost of an optimal policy OPT, i.e.,*

$$\mathbb{E}[\mathcal{C}(DB)] \leq 2 \cdot \mathbb{E}[\mathcal{C}(OPT)]. \quad (11.7)$$

The proof is almost immediate. We have shown that

$$\mathbb{E}[\mathcal{C}(DB)] = \sum_{t=1}^{T-L} \mathbb{E}[H_t^{DB} + \Pi_t^{DB}] = 2 \cdot \sum_{t=1}^{T-L} \mathbb{E}[Z_t^{DB}]. \quad (11.8)$$

Lemma 11.1 implies that

$$\begin{aligned} \mathbb{E}[\mathcal{C}(OPT)] &\geq \mathbb{E}\left[\sum_{t \in \mathcal{T}_H} H_t^{DB} + \sum_{t \in \mathcal{T}_\Pi} \Pi_t^{DB}\right] \\ &= \sum_{t=1}^{T-L} \mathbb{E}\left[H_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_H) + \Pi_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_\Pi)\right] \\ &= \sum_{t=1}^{T-L} \mathbb{E}\left[\mathbb{E}\left[H_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_H) + \Pi_t^{DB} \cdot \mathbf{1}(t \in \mathcal{T}_\Pi) \mid F_t\right]\right] \\ &= \sum_{t=1}^{T-L} \mathbb{E}\left[\left(\mathbf{1}(t \in \mathcal{T}_H) + \mathbf{1}(t \in \mathcal{T}_\Pi)\right) Z_t^{DB}\right] \\ &= \sum_{t=1}^{T-L} \mathbb{E}\left[Z_t^{DB}\right] \\ &= \frac{1}{2} \cdot \mathbb{E}[\mathcal{C}(DB)]. \end{aligned}$$

We note that the above worst-case analysis is tight in the sense that there exists a set of instances on which the ratio between the expected cost of the policy and the optimal expected cost converges to 2 (see Levi et al. (2007a)). On the other hand, in computational experiments (Hurley et al. (2007)), the dual-balancing policy performs significantly better than the worst-case guarantee, in many cases within a couple of percentages of optimum. Moreover, in many important settings the dual-balancing policies outperform simple myopic policies.

11.3 STOCHASTIC INVENTORY SYSTEMS WITH FIXED COSTS

We now add a fixed cost component to the basic model described in Section 11.2.1. More precisely, a fixed ordering cost K is incurred in each period with a strictly positive ordering quantity. (It should be noted that our analysis remains valid for non-stationary K_t satisfying $K_{t+1} \leq K_t$, which is commonly assumed in the literature.) Adding the fixed cost component to (11.4), we can re-write

$$\mathcal{C}(P) \triangleq \sum_{t=1}^{T-L} (K \cdot \mathbf{1}(Q_t^P > 0) + H_t^P + \Pi_t^P). \quad (11.9)$$

To address the nonlinearity induced by the fixed costs, a randomized decision rule is employed to balance the expected fixed ordering costs, holding costs and backlogging costs, in each period. In particular, the order quantity in each period is decided based on a carefully designed randomized rule that chooses among various possible order quantities with carefully chosen probabilities.

11.3.1 Randomized Cost-Balancing Policy

To describe the policy, we modify the definition of the information set f_t to also include the randomized decisions of the randomized balancing policy up to period $t-1$. Thus, given the information set f_t , the inventory position at the beginning of period t is known. However, the order quantity in period t is still *unknown* because the policy randomizes among various order quantities. We denote the randomized cost-balancing policy by RB . The decision in each period, whether to order and how much to order, is based on the following quantities.

- Compute the *balancing quantity* $\hat{q}_t$ which balances the expected marginal holding cost incurred by the units ordered against the expected backlogging cost in period $t+L$. That is, $\hat{q}_t$ uniquely solves

$$\mathbb{E}[H_t^{RB}(\hat{q}_t) | f_t] = \mathbb{E}[\Pi_t^{RB}(\hat{q}_t) | f_t] \triangleq \theta_t. \quad (11.10)$$

- Compute the *holding-cost-K quantity* $\tilde{q}_t$ that solves $\mathbb{E}[H_t^{RB}(\tilde{q}_t) | f_t] = K$, i.e., $\tilde{q}_t$ is the order quantity that brings the expected marginal holding cost to K .
- Compute $\mathbb{E}[\Pi_t^{RB}(\tilde{q}_t) | f_t]$, i.e., the expected backlogging cost if one orders $\tilde{q}_t$ units in period t .
- Compute $\mathbb{E}[\Pi_t^{RB}(0) | f_t]$, i.e., the expected backlogging cost resulting from not ordering in period t .

Based on the above quantities computed, the following randomized rule is used in each period t . Let P_t denote our ordering probability which is *a priori* random. With the observed information set f_t , the ordering probability $p_t = P_t | f_t$ in period t is defined differently in the two cases below.

Case (I)

If the balancing cost exceeds K , i.e., $\theta_t \geq K$, the *RB* policy orders the balancing quantity $q_t^{RB} = \hat{q}_t$ with probability $p_t = 1$. The intuition is that when $\theta_t \geq K$, the fixed ordering cost K is less dominant compared to marginal holding and backlogging costs. Moreover, if the *RB* policy does not place an order, the conditional expected backlogging cost is potentially large. Thus, it is worthwhile to order the balancing quantity $q_t^{RB} = \hat{q}_t$ with probability $p_t = 1$.

Case (II)

If the balancing cost is less than K , i.e., $\theta_t < K$, the *RB* policy orders the holding-cost- K quantity (i.e., $q_t^{RB} = \tilde{q}_t$) with probability p_t and nothing with probability $1 - p_t$. That is,

$$q_t^{RB} = \begin{cases} \tilde{q}_t, & \text{with probability } p_t \\ 0, & \text{with probability } 1 - p_t \end{cases} \quad (11.11)$$

The probability p_t is computed by solving the following equation

$$p_t K = p_t \cdot \mathbb{E}[\Pi_t^{RB}(\tilde{q}_t) | f_t] + (1 - p_t) \cdot \mathbb{E}[\Pi_t^{RB}(0) | f_t]. \quad (11.12)$$

The underlying reason behind the choice of this particular randomization in Equation (11.12) is that the policy perfectly balances the three types of costs, namely, the marginal holding cost, the marginal backlogging cost, and the fixed ordering cost associated with the period t . In particular, since we order the holding-cost- K quantity with probability p_t and nothing with probability $1 - p_t$, the conditional expected marginal holding cost in this case is

$$\mathbb{E}[H_t^{RB}(q_t^{RB}) | f_t] = p_t \mathbb{E}[H_t^{RB}(\tilde{q}_t) | f_t] + (1 - p_t) \mathbb{E}[H_t^{RB}(0) | f_t] = p_t K. \quad (11.13)$$

By the construction of p_t in Equation (11.12), the conditional expected backlogging cost is

$$\mathbb{E}[\Pi_t^{RB}(q_t^{RB}) | f_t] = p_t \mathbb{E}[\Pi_t^{RB}(\tilde{q}_t) | f_t] + (1 - p_t) \mathbb{E}[\Pi_t^{RB}(0) | f_t] = p_t K. \quad (11.14)$$

Since p_t is the ordering probability in Case (II), the expected fixed ordering cost is $p_t K$. It can be shown that Equation (11.12) has the following solution,

$$0 \leq p_t = \frac{\mathbb{E}[\Pi_t^{RB}(0) | f_t]}{K - \mathbb{E}[\Pi_t^{RB}(\tilde{q}_t) | f_t] + \mathbb{E}[\Pi_t^{RB}(0) | f_t]} < 1. \quad (11.15)$$

The inequalities in Equation (11.15) follow from the fact that $\theta_t < K$ and $\tilde{q}_t > \hat{q}_t$, which implies that $\mathbb{E}[\Pi_t^{RB}(\tilde{q}_t) | f_t] < \mathbb{E}[\Pi_t^{RB}(\hat{q}_t) | f_t] = \theta_t < K$.

11.3.2 Worst-Case Analysis

To obtain a 3-approximation algorithm, one wishes to show that on expectation the cost of an optimal policy can “pay” for at least one-third of the expected cost of the randomized cost-balancing policy. The periods are decomposed into subsets which we will define explicitly. For certain well-behaved subsets, we want to show that the holding and backloging costs incurred by an optimal policy can “pay” for one-third of the cost incurred by the RB policy. The difficulty arises in analyzing the remaining subset of *problematic periods*, for which it is not *a priori* clear how to “pay” for their cost. These problematic periods are further partitioned into intervals defined by each pair of two consecutive orders placed by the optimal policy. It can be shown that the total expected cost incurred by the RB policy in problematic periods within each interval, does not exceed $3K$. This implies that the fixed ordering cost incurred by an optimal policy can “pay” on the expectation of one-third of the cost incurred by the randomized cost-balancing policy in problematic periods. Let Z_t^{RB} be a random variable defined as

$$Z_t^{RB} \triangleq \mathbb{E}[H_t^{RB}(Q_t^{RB}) | F_t] = \mathbb{E}[\Pi_t^{RB}(Q_t^{RB}) | F_t]. \quad (11.16)$$

Note that Z_t^{RB} is a random variable that is realized with the information set in period t . Observe that by the construction of the RB policy, the random variable Z_t^{RB} is well defined since the expected marginal holding costs and the expected marginal backloging costs are always balanced. That is, the conditional expected marginal holding cost is always equal to the conditional expected marginal backloging cost. In addition, the expected fixed ordering cost in period t is also Z_t^{RB} by the construction of the algorithm, and therefore we have the following lemma.

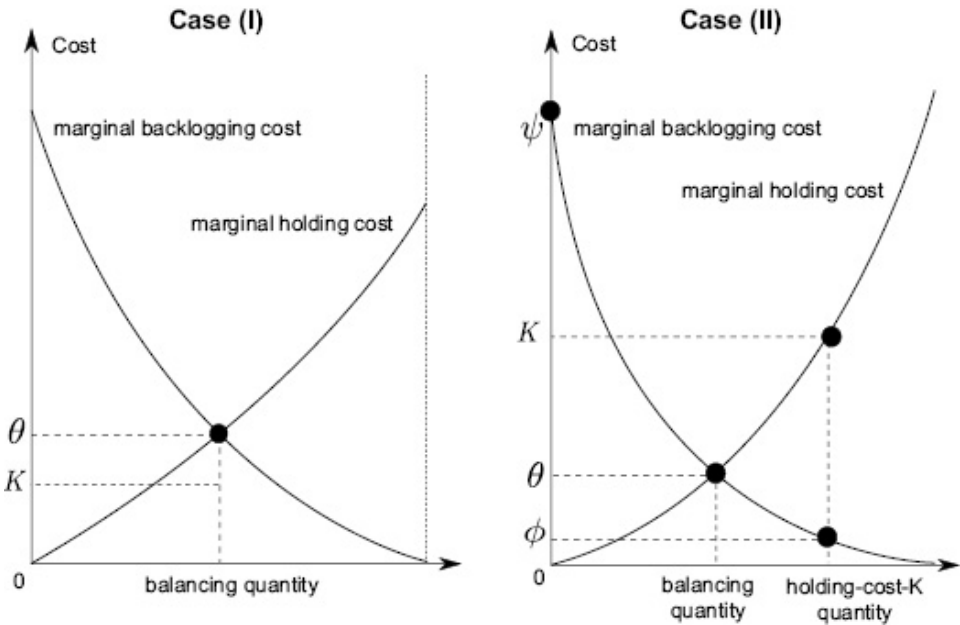


Figure 11.2 Illustration of the RB policy

Lemma 11.2 (Levi and Shi (2013)) *Let $\mathcal{C}(RB)$ be the total cost incurred by the RB policy. Then we have,*

$$\mathbb{E}[\mathcal{C}(RB)] \leq 3 \cdot \sum_{t=1}^{T-L} \mathbb{E}[Z_t^{RB}]. \quad (11.17)$$

To complete the worst-case analysis, we would like to show that the expected cost of an optimal policy denoted by OPT is at least $\sum_{t=1}^{T-L} \mathbb{E}[Z_t^{RB}]$. This will be done by amortizing the cost of OPT against the cost of the RB policy. In particular, we shall show that on expectation OPT pays for a large fraction of the cost of the RB policy. In the subsequent analysis, we will use a random partition of periods $t = \{1, 2, \dots, T-L\}$ for the following sets.

- The set $\mathcal{T}_{1H} \triangleq \{t : \Theta_t \geq K \text{ and } Y_t^{OPT} > Y_t^{RB}\}$ consists of periods in which the balancing cost Θ_t exceeds K and the optimal policy had a higher inventory position than that of the RB policy after ordering (recall that if $\Theta_t \geq K$ then the RB policy orders the balancing quantity with probability 1 and the value Y_t^{RB} is known deterministically (i.e., realized) with F).
- The set $\mathcal{T}_{1\pi} \triangleq \{t : \Theta_t \geq K \text{ and } Y_t^{OPT} \leq Y_t^{RB}\}$ consists of periods in which the balancing cost exceeds K and the inventory position of the optimal policy does not exceed that of the RB policy after ordering.
- The set $\mathcal{T}_{2H} \triangleq \{t : \Theta_t < K \text{ and } Y_t^{OPT} \geq X_t^{RB} + \tilde{Q}_t^{RB}\}$ consists of periods in which the balancing cost is less than K and, in such periods, the inventory position of the RB policy after ordering would be either X_t^{RB} if no order was placed, or $X_t^{RB} + \tilde{Q}_t^{RB}$ if the *holding-cost- K quantity* is ordered, depending on the randomized decision of the RB policy. However, the inventory position of OPT after ordering exceeds even $X_t^{RB} + \tilde{Q}_t^{RB}$. (Note again that the quantity $\tilde{Q}_t^{RB}$ is known deterministically (i.e., realized) with F .)
- Analogous to $\mathcal{T}_{2H}$, the set $\mathcal{T}_{2\pi} \triangleq \{t : \Theta_t < K \text{ and } X_t^{RB} \geq Y_t^{OPT}\}$ consists of periods in which the inventory position of OPT after ordering is below X_t^{RB} .
- The set $\mathcal{T}_{2M} \triangleq \{t : \Theta_t < K \text{ and } X_t^{RB} < Y_t^{OPT} < X_t^{RB} + \tilde{Q}_t^{RB}\}$ consists of periods in which the balancing cost is less than K and the inventory position of OPT after ordering is within $(X_t^{RB}, X_t^{RB} + \tilde{Q}_t^{RB})$. Thus, whether the RB policy or OPT has more inventory depends on whether the RB policy placed an order.

Note that the sets $(\mathcal{T}_{1H} - \mathcal{T}_{2M})$ are disjointed and the union makes a complete set. Conditioning on f_t , it is already known which part of the partition period t belongs.

Next, we will show that the total holding cost incurred by OPT is higher than the marginal holding cost incurred by the RB policy in periods that belong to $\mathcal{T}_{1H} \cup \mathcal{T}_{2H}$, and that the total backlogging cost incurred by OPT is higher than the backlogging cost incurred by the OPT policy associated with periods within $\mathcal{T}_{1\pi} \cup \mathcal{T}_{2\pi}$.

Lemma 11.3 (Levi and Shi (2013)) *The holding cost and backlogging cost incurred by OPT are denoted by H^{OPT} and Π^{OPT} , respectively. Then, with probability 1,*

$$H^{OPT} \geq \sum_t H_t^{RB} \cdot \mathbf{1}(t \in \mathcal{T}_{1H} \cup \mathcal{T}_{2H}), \Pi^{OPT} \geq \sum_t \Pi_t^{RB} \cdot \mathbf{1}(t \in \mathcal{T}_{1\pi} \cup \mathcal{T}_{2\pi}).$$

The high-level idea of Lemma 11.3 is as follows. In each period $t \in \mathcal{T}_{1H} \cup \mathcal{T}_{2H}$, the inequality $Y_t^{RB} < Y_t^{OPT}$ holds and implies that the Q_t^{RB} units ordered by the *OPT* policy in period t have been ordered by *OPT* either in period t or even earlier. Thus, the holding cost they incur under *OPT* is higher than those incurred under the *RB* policy. On the other hand, in each period $t \in \mathcal{T}_{1H} \cup \mathcal{T}_{2H}$, the inequality $Y_t^{RB} \geq Y_t^{OPT}$ holds and implies that the backlogging incurred by *OPT* at the end of period $t+L$ will be higher than that of the *RB* in that period. We are still left with the problematic set $\mathcal{T}_{2M}$. Note that in this particular set, whether the *RB* policy or *OPT* has more inventory depends on whether the *RB* policy placed an order. Fortunately, Lemma 11.4 shows that the fixed ordering costs incurred by *OPT* can cover the randomized balancing costs in $\mathcal{T}_{2M}$ (via a clever binary-tree type argument).

Lemma 11.4 (Levi and Shi (2013)) *The expected randomized cost in set $\mathcal{T}_{2M}$ is less than the total expected fixed ordering cost incurred by *OPT*, i.e.,*

$$\mathbb{E} \left[\sum_t Z_t^{RB} \cdot \mathbf{1}(t \in \mathcal{T}_{2M}) \right] \leq \mathbb{E} \left[\sum_{t=1}^{T-L} K \cdot \mathbf{1}(Q_t^{OPT} > 0) \right].$$

As an immediate consequence of Lemmas 11.3 and 11.4, we obtain the following result. Let $\mathcal{C}(OPT)$ be the total cost incurred by the cost-balancing policy *RB*. Following the same conditional expectation argument as in the basic model, we have

$$\mathbb{E}[\mathcal{C}(OPT)] \geq \sum_{t=1}^{T-L} \mathbb{E}[Z_t^{RB}]. \quad (11.18)$$

Theorem 11.2 (Levi and Shi (2013)) *For each instance of the stochastic lot-sizing problem, the expected cost of the randomized cost-balancing policy *RB* is at most three times the expected cost of an optimal policy *OPT*, i.e.,*

$$\mathbb{E}[\mathcal{C}(RB)] \leq 3 \cdot \mathbb{E}[\mathcal{C}(OPT)]. \quad (11.19)$$

11.4 PERISHABLE INVENTORY SYSTEMS WITH M SHELF LIFE

We now focus on a stochastic periodic-review perishable inventory system over a planning horizon of T (possibly infinite) periods, indexed by $t = 1, \dots, T$. The lifetime of the product is m periods, i.e., a product perishes after staying m periods in stock. Our model allows for a non-stationary and generally correlated demand process. We assume that the order lead time is zero, i.e., an order placed at the beginning of a period can be used in the same period. This is a common assumption in the perishable inventory literature (see Karaesmen et al. (2011)).

In each period t , four types of costs may occur: a unit ordering cost $\hat{c}$, a unit holding cost $\hat{h}$ for leftover inventory, a unit backlogging cost $\hat{b}$ for unsatisfied demand, and a unit outdating cost $\hat{\theta}$ for expired products. There is also a one-period discount factor α , with $0 < \alpha \leq 1$ when $T < \infty$ and $0 < \alpha < 1$ when $T = \infty$. We assume that $\hat{b} > (1 - \alpha)\hat{c}$ and $\hat{\theta} + \alpha\hat{c} \geq 0$. Thus, $\hat{\theta}$ can be

negative, and in this case it can be interpreted as a unit salvage value. Following Nahmias (1975) we assume that any remaining inventory at the end of the planning horizon can be salvaged with a return of $\hat{c}$ per-unit and unsatisfied demand can be satisfied by an emergency order at a cost of $\hat{c}$ per-unit. We note that our analysis can be extended to the case with a unit salvage value $\hat{v}$ for any on-hand inventory and a unit penalty cost $\hat{p}$ for any unsatisfied demand at the end of the planning horizon, as long as $\hat{v} \leq \hat{c}$ and $\hat{b} + \alpha\hat{p} > \hat{c}$.

For each period t , the sequence of events is as follows.

- First, at the beginning of period t , the information set $f_t \in \mathcal{F}_t$ and the inventory vector

$$\mathbf{x}_t = (x_{t,1}, \dots, x_{t,m-1}) \quad (11.20)$$

are observed, where $x_{t,i}$ is the quantity of on-hand products whose remaining lifetime is i periods, $i = 1, \dots, m-2$, and $x_{t,m-1}$ is the quantity of on-hand products whose remaining lifetime is $m-1$ periods minus the quantity of backlogged demands (if any). Thus, $x_{t,1}, \dots, x_{t,m-2}$ are always non-negative; while $x_{t,m-1}$ can be positive or negative. For simplicity, we assume that the inventory system is initially empty at the beginning of period 1, i.e., $x_{1,i} = 0$, for all $i = 1, \dots, m-1$; but our analysis and results can be extended to the case with an arbitrary initial state.

- Second, an order with quantity q_t is placed, incurring an ordering cost $\hat{c}q_t$. Denote y_t as the total inventory level after receiving the order in period t . Then, $y_t = \sum_{i=1}^{m-1} x_{t,i} + q_t$.
- Third, the demand in period t is realized and satisfied as much as possible by the on-hand inventory using the FIFO issuing policy, i.e., the oldest inventory is consumed first when demand arrives. At the end of period t , if $y_t - D_t \geq 0$, then the excess inventory incurs a holding cost $\hat{h}(y_t - D_t)$. Following Nahmias (1975), we assume that all excess inventory (including the inventory which perishes at the end of this period) incurs a holding cost. On the other hand, if $y_t - D_t < 0$, then the system incurs a backlogging cost $\hat{b}(D_t - y_t)$. Furthermore, if the inventory with one period remaining life $x_{t,1} > D_t$, then $e_t := (x_{t,1} - D_t)^+$ units perish and incur an outdated cost $\hat{\theta}e_t$.
- Finally, the system proceeds to the subsequent period $t+1$. By the definition of the inventory vector $\mathbf{x}_t$ and the FIFO issuing policy, we obtain the following state transition from $\mathbf{x}_t$ to $\mathbf{x}_{t+1}$:

$$\begin{aligned} x_{t+1,j} &= \left(x_{t,j+1} - \left(D_t - \sum_{i=1}^j x_{t,i} \right)^+ \right)^+, \text{ for } 1 \leq j \leq m-2, \\ x_{t+1,m-1} &= q_t - \left(D_t - \sum_{i=1}^{m-1} x_{t,i} \right)^+. \end{aligned} \quad (11.21)$$

We remark that in defining the inventory state $\mathbf{x}_t$ in Equation (11.20), it is convenient and natural to combine the inventory having $m-1$ periods of remaining life with the number of backlogs in $x_{t,m-1}$. This is because when demand arrives, by the FIFO issuing policy, it is first met by $x_{t,1}$, and when $x_{t,1}$ is consumed then the remaining demand is met by $x_{t,2}$. This process continues and when (and if) $x_{t,m-2}$ also depletes to 0, the remaining demand will be satisfied

by $x_{t,m-1}$. Clearly, when the demand is large, this last number will continue to go down after reaching 0, representing the backlog level. We also note that inventory only outdates through the first dimension, $x_{t,1}$, of vector $\mathbf{x}_t$, while backlogs always stay in the last dimension, $x_{t,m-1}$ (hence backlogs will *not* disappear after m periods). Moreover, if in period t there are backlogs (thus $x_{t,m-1}$ is negative and $x_{t,j} = 0$ for $j = 1, \dots, m-2$), then with Equation (11.21) in the next period $x_{t+1,j}$ will be equal to 0 for all $j = 1, \dots, m-2$, but $x_{t+1,m-1}$ can be positive or negative, depending on whether q_t is greater or less than $D_t - x_{t,m-1}$.

Then the expected total discounted cost incurred under a given policy P that orders q_t in period t can be written as

$$\mathcal{C}(P) = \mathbb{E} \left[\sum_{t=1}^T \alpha^{t-1} \left(\hat{c}q_t + \hat{h}(Y_t - D_t)^+ + \hat{b}(D_t - Y_t)^+ + \hat{\theta}e_t \right) - \alpha^T \hat{c} \sum_{i=1}^{m-1} X_{T+1,i} \right].$$

Next, we carry out a cost transformation to obtain an equivalent model with the unit ordering cost equal to 0. This will enable us to assume, without loss of generality, that the unit ordering cost is 0 in the subsequent analysis.

Proposition 11.1 (Chao et al. (2015)) *For every perishable inventory system with cost parameters $\hat{c}$, $\hat{h}$, $\hat{b}$, and $\hat{\theta}$, there is an equivalent system with non-negative cost parameters $c = 0$, $h = \hat{h} + (1 - \alpha)\hat{c}$, $b = \hat{b} - (1 - \alpha)\hat{c}$, and $\theta = \hat{\theta} + \alpha\hat{c}$. And the expected total discounted cost can be rewritten as*

$$\mathcal{C}(P) = \mathbb{E} \left[\sum_{t=1}^T \alpha^{t-1} \left(h(Y_t - D_t)^+ + b(D_t - Y_t)^+ + \theta e_t \right) \right] + \sum_{t=1}^T \alpha^{t-1} \hat{c} \mathbb{E}[D_t].$$

11.4.1 Nested Marginal Cost Accounting Scheme

We develop a new marginal cost accounting scheme for perishable inventory systems, similar in spirit to that of Levi et al. (2007a). Our marginal cost accounting scheme exhibits a nested structure due to the multi-dimensionality of the system state. The main idea underlying this approach is to decompose the total cost in terms of the marginal costs of individual decisions. That is, we associate the decision in period t with its affiliated cost contributions to the system. These marginal costs may include costs (associated with the decision) incurred in both the current and subsequent periods.

Given the inventory vector $\mathbf{x}_t = (x_{t,1}, \dots, x_{t,m-1})$ at the beginning of period t , and that a policy P orders q_t , we aim to compute the marginal cost contributions to the system by these q_t units on the holding, outdating, and backlogging costs. To this end, for $i = 1, \dots, m-1$, we let $B_t(\mathbf{x}_t, i)$ denote the number of outdated units in periods $[t, t+i-1]$ given that the inventory vector at the beginning of period t is $\mathbf{x}_t$, with the convention that $B_t(\mathbf{x}_t, 0) \equiv 0$. Then, for $1 \leq i \leq m-1$, we have

$$B_t(\mathbf{x}_t, i) = \max \left\{ \sum_{j=1}^i x_{t,j} - D_{[t, t+i-1]}, B_t(\mathbf{x}_t, i-1) \right\}. \quad (11.22)$$

To see why this is true, note that $\sum_{j=1}^i x_{t,j} - B_t(\mathbf{x}_t, i-1)$ is the number of non-expired units in $x_{t,1}, \dots, x_{t,i}$ that would meet demands in periods $[t, t+i-1]$. These units, if not consumed, will expire at the end of period $t+i-1$. Thus $\left(\sum_{j=1}^i x_{t,j} - B_t(\mathbf{x}_t, i-1) - D_{[t,t+i-1]}\right)^+$, if positive, would be the number of units that will expire at the end of period $t+i-1$. Adding $B_t(\mathbf{x}_t, i-1)$ to it gives the total number of expired units in $[t, t+i-1]$, which is Equation (11.22).

The nested structure in the auxiliary function $B_t(\cdot, \cdot)$ follows from the fact that some inventory units reach their expiration date before meeting the demand, and have to be discarded from the on-hand inventory. Using this auxiliary function, the number of outdated units in period $t+i-1$, for $1 \leq i \leq m-1$, is given as

$$e_{t+i-1} = \left(\sum_{j=1}^i x_{t,j} - B_t(\mathbf{x}_t, i-1) - D_{[t,t+i-1]} \right)^+,$$

and the number of outdated units in period $t+m-1$ is

$$e_{t+m-1} = \left(q_t + \sum_{j=1}^{m-1} x_{t,j} - B_t(\mathbf{x}_t, m-1) - D_{[t,t+m-1]} \right)^+.$$

11.4.1.1 Nested marginal holding cost accounting

We first focus on the marginal holding cost accounting of a given policy P . The holding cost for the q_t units ordered in period t may be incurred in any period from t to $t+m-1$ (after which the remaining ones will perish), or T , whichever is smaller. Let $H_t^P(q_t)$ be the discounted marginal holding cost (to period 1) incurred by these q_t units. Then it follows from the FIFO issuing policy that

$$H_t^P(q_t) := h \sum_{i=t}^{(t+m-1) \wedge T} \alpha^{i-1} \left(q_t - (D_{[t,i]} + B_t(\mathbf{x}_t, i-t) - \sum_{j=1}^{m-1} x_{t,j})^+ \right)^+, \quad (11.23)$$

where the auxiliary function $B_t(\mathbf{x}_t, i)$ is given recursively via Equation (11.22). To see why Equation (11.23) is valid, note that the total number of units in $\mathbf{x}_t$ that do not expire until $t+i$ is $\sum_{j=1}^{m-1} x_{t,j} - B_t(\mathbf{x}_t, i)$, thus the net demand after consuming the units in $\mathbf{x}_t$ is $\left(D_{[t,t+i]} - (\sum_{j=1}^{m-1} x_{t,j} - B_t(\mathbf{x}_t, i))\right)^+$. Hence, the number of unconsumed units from q_t at the end of period $t+i$ is $(q_t - (D_{[t,t+i]} + B_t(\mathbf{x}_t, i) - \sum_{j=1}^{m-1} x_{t,j})^+)^+$.

Because the marginal holding cost is computed based on the nested structure of the auxiliary function $B_t(\cdot, \cdot)$, we call it *the nested marginal holding cost accounting*. Note that the marginal holding cost associated with the q_t units ordered in period t is only affected by future demands but not by future decisions.

11.4.1.2 Nested marginal outdated cost accounting

Similarly, we can compute the marginal outdated cost associated with the q_t units ordered by policy P in period t using the following nested scheme. For $t = 1, \dots, T-m+1$,

$$\begin{aligned}\Theta_t^P(q_t) &:= \alpha^{t+m-2} \theta e_{t+m-1} \\ &= \alpha^{t+m-2} \theta \left(q_t + \sum_{j=1}^{m-1} x_{t,j} - B_t(\mathbf{x}_t, m-1) - D_{[t, t+m-1]} \right)^+, \end{aligned} \quad (11.24)$$

where $B_t(\cdot, \cdot)$ is defined in (11.22); and for $t = T - m + 2, \dots, T$, we have $\Theta_t^P \equiv 0$ since the ordered units do not expire within the planning horizon.

11.4.1.3 Marginal backlogging cost accounting

For each period $t = 1, \dots, T$, the discounted (to period 1) marginal backlogging cost of the q_t units ordered in period t by policy P can be expressed as

$$\Pi_t^P(q_t) := \alpha^{t-1} b \left(D_t - \sum_{i=1}^{m-1} x_{t,i} - q_t \right)^+, \quad (11.25)$$

which is exactly the same as the traditional backlogging cost using the period-by-period accounting scheme. The intuition is that any negative consequence of under-ordering can be corrected by placing an order in the next period; thus it suffices to only consider the backlogging cost incurred in the current period.

11.4.1.4 Total cost of a given policy

Note that the marginal costs defined above, $H_t^P(q_t)$, $\Theta_t^P(q_t)$, and $\Pi_t^P(q_t)$, are random as they depend on future demands. Since the system is initially empty, the expected total system cost $\mathcal{C}(P)$ of policy P can be obtained by summing Equations (11.23), (11.24), and (11.25) over t from 1 to T , and then taking expectations. Thus, through Equation (11.1) we have

$$\mathcal{C}(P) = \mathbb{E} \left[\sum_{t=1}^T \left(H_t^P(q_t) + \Pi_t^P(q_t) + \Theta_t^P(q_t) \right) \right] + \sum_{t=1}^T \alpha^{t-1} \hat{c} \mathbb{E}[D_t]. \quad (11.26)$$

Ignoring the constant terms, we can write the effective cost of a policy P as

$$C(P) = \mathbb{E} \left[\sum_{t=1}^T \left(H_t^P(q_t) + \Pi_t^P(q_t) + \Theta_t^P(q_t) \right) \right]. \quad (11.27)$$

11.4.2 Proportional-Balancing (PB) Policy

For each period $t = 1, \dots, T$, with an observed information set $f_t \in \mathcal{F}_t$, the PB policy orders $q_t^{PB} = q_t$ that balances a proportion of the expected marginal holding and outdating costs with the expected backlogging cost as follows:

$$\frac{mh+\theta}{2(m-1)h+\theta} \mathbb{E}[H_t^{PB}(q_t) + \Theta_t^{PB}(q_t) | f_t] = \mathbb{E}[\Pi_t^{PB}(q_t) | f_t]. \quad (11.28)$$

It can be verified that the LHS of Equation (11.28) is an increasing convex function of the order quantity q_i , which equals 0 when $q_i=0$ and approaches infinity when q_i tends to infinity. On the other hand, the RHS of Equation (11.28) is a decreasing convex function of the order quantity q_i , which equals a non-negative number when $q_i=0$ and tends to 0 when q_i goes to infinity. Since q_i can take any non-negative real value and both functions are continuous, q_i in (11.28) is well defined. Furthermore, since LHS minus RHS of Equation (11.28) is increasing in q_i , q_i^{PB} can be very efficiently computed using bisection methods. It should be noted that q_i^{PB} is a function of f_i and $\mathbf{x}_i$, but for simplicity we make this dependency implicit.

Theorem 11.3 (Chao et al. (2015)) *The proportional-balancing policy for the perishable inventory system with $m \geq 2$ periods of product lifetime has a worst-case performance guarantee of $\left(2 + \frac{(m-2)h}{mh + \theta}\right)$, i.e., for any instance of the problem,*

$$C(PB) \leq \left(2 + \frac{(m-2)h}{mh + \theta}\right) C(OPT).$$

Theorem 11.3 shows that, when the product lifetime $m=2$, the *PB* policy has a worst-case performance guarantee of 2; while for a general lifetime m , the *PB* policy has a worst-case performance guarantee between 2 and 3.

11.4.3 Worst-Case Analysis

The arguments used in the literature on proving worst-case performance guarantees for approximation algorithms utilize a “unit-matching” approach (see, e.g., Levi et al. (2007a, 2008a, 2008d, 2017), Levi and Shi (2013)). In a sense, the approach is geometric, and it relies on the correspondence of units in the systems operating under different policies throughout the planning horizon, and then it compares the costs incurred by the matched units in different systems. However, the unit-matching approach fails to work for perishable inventory systems because the inventory units can perish and the number of outdated units differs in systems operating under different policies.

To overcome this difficulty, we develop an algebraic approach for comparing different systems. A key concept in our approach is the *trimmed on-hand inventory level*, which is defined as the part of on-hand inventory units ordered before any given particular time. These trimmed inventory levels serve as a generalization of the traditional inventory level, as they provide critical (partial) information on the ages of the products on-hand. Due to the nature of perishable systems, it is impossible to quantify the effect of the decision made in the current period t on future costs only through the traditional total inventory level Y_t . The trimmed inventory levels provide a tractable way to analyze this effect, and also provide the *right* framework for coupling the marginal holding and outdated costs in different systems. More technically, the difference between the trimmed inventory levels of our policy and the optimal policy *OPT* can be bounded by the difference between the outdated units of the two policies. An essential part of this worst-case analysis presented below is based on this new concept.

We now compare the *PB* policy with the optimal policy *OPT*. To this end, we make the dependency of the relevant quantities on the policy, *PB* or *OPT*, explicit. For each realization of demands $D_1, \dots, D_T$ and the exogenous information $W_1, \dots, W_T$, we compare and analyze the inventory processes of the systems operating under these two policies.

Given a realization $f_T \in \mathcal{F}_T$, let $\mathcal{T}_H$ ($\mathcal{T}_\Pi$) be the set of periods in which the optimal policy has more (less) total inventory level than the *PB* policy does.

$$\mathcal{T}_H = \{t \in [1, T] : Y_t^{OPT} \geq Y_t^{PB}\}, \quad \mathcal{T}_\Pi = \{t \in [1, T] : Y_t^{OPT} < Y_t^{PB}\}.$$

Lemma 11.5 (Chao et al. (2015)) *For each realization $f_T \in \mathcal{F}_T$, we have*

$$\sum_{t \in \mathcal{T}_H} H_t^{PB} \leq \sum_{t=1}^T H_t^{OPT} + (m-2) \frac{h}{\theta} \sum_{t=1}^T \Theta_t^{OPT}.$$

Proof. Lemma 11.5 is one of the key technical results of this chapter. Its complete proof is complicated and lengthy. To illustrate the main ideas of our argument, we provide below the proof for the following, weaker, result:

$$\sum_{t \in \mathcal{T}_H} H_t^{PB} \leq \sum_{t=1}^T H_t^{OPT} + (m-1) \frac{h}{\theta} \sum_{t=1}^T \Theta_t^{OPT}. \quad (11.29)$$

For ease of illustration, we also assume that $\alpha = 1$.

As said earlier, an important concept in proving our main technical result is what we refer to as the *trimmed on-hand inventory level*, denoted by $Y_{t,s}$ for any $s \geq t \geq 1$, which is defined as *the part of on-hand inventory at the beginning of period s which is ordered in period t or earlier*. From the definition of $Y_{t,s}$, it holds that $Y_{t,s} = 0$ when $s \geq t + m$, and

$$Y_{t,s} = (Y_t - D_{[t,s)} - e_{[t,s)})^+, \quad s = t, t+1, \dots, t+m-1. \quad (11.30)$$

For any period $t = 1, \dots, T$, we define the notation $R(t)$ as follows: if the set $\{s \in \mathcal{T}_H : s > t\}$ is nonempty, then $R(t) := \min\{s \in \mathcal{T}_H : s > t\}$; otherwise, $R(t) := T + 1$. In addition, for any $s \geq 1$, denote $\tilde{H}_s$ as the part of the holding cost incurred in period s associated with the products ordered in periods $\{t : t \in \mathcal{T}_H, t \leq s\}$. Since the lifetime of the products is m , all products ordered in period t or earlier will leave the system by the end of period $t+m-1$. Then, it follows that for any $t \in \mathcal{T}_H$, $\tilde{H}_s = 0$ when $t+m \leq s \leq R(t)-1$. Consequently, by the definitions of H_t , $\tilde{H}_s$, and $R(t)$, we have

$$\sum_{t \in \mathcal{T}_H} H_t = \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{R(t)-1} \tilde{H}_s = \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} \tilde{H}_s.$$

For each period $s \in [t, R(t)-1]$, from its definition, $\tilde{H}_s$ is clearly no greater than the part of the holding cost incurred in period s associated with the products ordered in periods $1, \dots, t$, which can be expressed as $h(Y_{t,s} - D_s)^+$. Since $\sum_{t=1}^T H_t = h \sum_{s=1}^T (Y_{s,s} - D_s)^+$ and $Y_{t,s} \geq Y_{s,s}$ for any $t \leq s$, the following inequalities hold for any policy:

$$\sum_{t \in \mathcal{T}_H} H_t \leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} (Y_{t,s} - D_s)^+ \leq \sum_{t=1}^T H_t. \quad (11.31)$$

In particular, we have

$$\sum_{t \in \mathcal{T}_H} H_t^{PB} \leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} (Y_{t,s}^{PB} - D_s)^+; \quad (11.32)$$

$$\sum_{t=1}^T H_t^{OPT} \geq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} (Y_{t,s}^{OPT} - D_s)^+. \quad (11.33)$$

Subtracting Equation (11.33) from Equation (11.32), we obtain

$$\begin{aligned} \sum_{t \in \mathcal{T}_H} H_t^{PB} - \sum_{t=1}^T H_t^{OPT} &\leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} \left((Y_{t,s}^{PB} - D_s)^+ - (Y_{t,s}^{OPT} - D_s)^+ \right) \\ &\leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} (Y_t^{PB} - Y_t^{OPT} - e_{[t,s]}^{PB} + e_{[t,s]}^{OPT})^+ \\ &\leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t}^{(t+m) \wedge R(t)-1} (e_{[t,s]}^{OPT} - e_{[t,s]}^{PB})^+, \end{aligned} \quad (11.34)$$

where the second inequality follows from Equation (11.30) and $a^+ - b^+ \leq (a - b)^+$ for any real numbers a and b , and the last one holds because $Y_t^{PB} \leq Y_t^{OPT}$ when $t \in \mathcal{T}_H$.

Thus, it follows from $(e_t^{OPT} - e_t^{PB})^+ \leq e_t^{OPT}$ and $e_{[t,t]}^{OPT} = 0$ that

$$\begin{aligned} \sum_{t \in \mathcal{T}_H} H_t^{PB} - \sum_{t=1}^T H_t^{OPT} &\leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t+1}^{(t+m) \wedge R(t)-1} e_{[t,s]}^{OPT} \leq h \sum_{t \in \mathcal{T}_H} \sum_{s=t+1}^{(t+m) \wedge R(t)-1} e_{[t,R(t)-1]}^{OPT} \\ &\leq (m-1)h \sum_{t \in \mathcal{T}_H} \sum_{\ell=t}^{R(t)-2} e_{\ell}^{OPT} \leq (m-1)h \sum_{t=1}^T e_t^{OPT} \\ &= (m-1)h \sum_{t=m}^T e_t^{OPT} = (m-1) \frac{h}{\theta} \sum_{t=1}^T \Theta_t^{OPT}, \end{aligned}$$

where the last inequality follows from $e_t^{OPT} = 0$ for $t \leq m-1$ as the system is initially empty, $\Theta_t^{OPT} = \theta e_{t+m-1}^{OPT}$ for $1 \leq t \leq T-m+1$, and $\Theta_t^{OPT} = 0$ when $T-m+1 < t \leq T$. This proves a weaker form of Lemma 2, i.e., the result in Equation (11.29).

Lemma 11.6 (Chao et al. (2015)) For each realization $f_T \in \mathcal{F}_T$, we have $\sum_{t \in \mathcal{T}_H} \Theta_t^{PB} \leq \sum_{t=1}^T \Theta_t^{OPT}$

Proof For brevity, we only prove below the result for the special case when $\alpha=1$. Since any products ordered after period $T-m+1$ do not perish within the planning horizon, we only need to consider periods $t=1, \dots, T-m+1$. For each realization f_T and the resulting $\mathcal{T}_H$, we partition

the periods $\{1, \dots, T - m + 1\}$ as follows: First, start in period $T - m + 1$ and search backward for the latest period $t \in \mathcal{T}_H$ such that $\Theta_t^{PB} > \Theta_t^{OPT}$. If no such period exists, then we terminate the partition process. Otherwise, let t' be that period and mark the periods $t', t' - 1, \dots, (t' - m)^+ + 1$. Next, repeat the above procedure over periods $1, \dots, (t' - m)^+$ until the remaining set of periods is empty. As a result, this procedure partitions the periods $\{1, \dots, T - m + 1\}$ into marked and unmarked periods. Let $\mathcal{T}_M$ denote the set of marked periods.

We first consider any period $t \in \mathcal{T}_H \setminus \mathcal{T}_M$. Then, it follows from the definition of $\mathcal{T}_M$ that $\Theta_t^{PB} \leq \Theta_t^{OPT}$. Consequently

$$\sum_{t \in \mathcal{T}_H \setminus \mathcal{T}_M} \Theta_t^{PB} \leq \sum_{t \in \mathcal{T}_H \setminus \mathcal{T}_M} \Theta_t^{OPT}. \quad (11.35)$$

Since the set $\mathcal{T}_M$ is made up of disjoint intervals, we consider a representative interval with its largest period being t , i.e., this interval consists of periods $(t - m)^+ + 1, \dots, t - 1, t$. Then, by the construction of $\mathcal{T}_M$, $t \in \mathcal{T}_H$ and $\Theta_t^{PB} > \Theta_t^{OPT}$. Since $\Theta_t = \theta e_{t+m-1}$, we have $e_{t+m-1}^{PB} > e_{t+m-1}^{OPT} \geq 0$. Note that e_t is the number of perished products in period t and it satisfies the following identity for any feasible policy:

$$e_{t+m-1} = (Y_t - D_{[t, t+m-1]} - e_{[t, t+m-1]})^+. \quad (11.36)$$

Thus it follows from Equation (11.36) and $e_{t+m-1}^{PB} > 0$ that

$$\begin{aligned} e_{[t, t+m-1]}^{PB} &= (Y_t^{PB} - D_{[t, t+m-1]} - e_{[t, t+m-1]}^{PB})^+ + e_{[t, t+m-1]}^{PB} \\ &= Y_t^{PB} - D_{[t, t+m-1]}. \end{aligned} \quad (11.37)$$

On the other hand, for the *OPT* policy we have

$$\begin{aligned} e_{[t, t+m-1]}^{OPT} &= (Y_t^{OPT} - D_{[t, t+m-1]} - e_{[t, t+m-1]}^{OPT})^+ + e_{[t, t+m-1]}^{OPT} \\ &\geq Y_t^{OPT} - D_{[t, t+m-1]}. \end{aligned} \quad (11.38)$$

Subtracting Equation (11.38) from Equation (11.37) yields

$$e_{[t \vee m, t+m-1]}^{PB} - e_{[t \vee m, t+m-1]}^{OPT} = e_{[t, t+m-1]}^{PB} - e_{[t, t+m-1]}^{OPT} \leq Y_t^{PB} - Y_t^{OPT} \leq 0,$$

where the equality holds since $e_t^{PB} = e_t^{OPT}$ for $1 \leq t \leq m - 1$ and the last inequality follows from $t \in \mathcal{T}_H$. This proves, by $\Theta_s = \theta e_{s+m-1}$ for any period s , that for all t ,

$$\Theta_{[(t-m+1) \vee 1, t]}^{PB} \leq \Theta_{[(t-m+1) \vee 1, t]}^{OPT}.$$

As the above result holds for any of the disjoint intervals of $\mathcal{T}_M$, adding them up yields

$$\sum_{t \in \mathcal{T}_M} \Theta_t^{PB} \leq \sum_{t \in \mathcal{T}_M} \Theta_t^{OPT}. \quad (11.39)$$

Finally, since

$$\mathcal{T}_H \subset (\mathcal{T}_H \setminus \mathcal{T}_M) \cup \mathcal{T}_M \subset \{1, 2, \dots, T\},$$

we obtain, using Equation (11.35) and Equation (11.39), that

$$\sum_{t \in \mathcal{T}_H} \Theta_t^{PB} \leq \sum_{t \in \mathcal{T}_H \setminus \mathcal{T}_M} \Theta_t^{PB} + \sum_{t \in \mathcal{T}_M} \Theta_t^{PB} \leq \sum_{t \in \mathcal{T}_H \setminus \mathcal{T}_M} \Theta_t^{OPT} + \sum_{t \in \mathcal{T}_M} \Theta_t^{OPT} \leq \sum_{t=1}^T \Theta_t^{OPT}.$$

This completes the proof of Lemma 11.6 when $\alpha = 1$.

Note that for each perished unit ordered in periods $1, \dots, T$, it must stay in the system for exactly m periods. Thus, for any policy, we have the following inequality

$$mh \sum_{t=1}^T \Theta_t \leq \theta \sum_{t=1}^T H_t. \quad (11.40)$$

Combining this inequality with Lemmas 11.5 and 11.6 leads to the following result.

Corollary 11.1 (Chao et al. (2015)) *For each realization $f_T \in \mathcal{F}_T$, we have*

$$\sum_{t \in \mathcal{T}_H} (H_t^{PB} + \Theta_t^{PB}) \leq \left(1 + \frac{(m-2)h}{mh+\theta}\right) \sum_{t=1}^T (H_t^{OPT} + \Theta_t^{OPT}). \quad (11.41)$$

Proof We apply Lemmas 11.5 and 11.6, and Equation (11.40) to obtain

$$\begin{aligned} \sum_{t \in \mathcal{T}_H} (H_t^{PB} + \Theta_t^{PB}) &\leq \sum_{t=1}^T H_t^{OPT} + \frac{(m-2)h}{\theta} \sum_{t=1}^T \Theta_t^{OPT} + \sum_{t=1}^T \Theta_t^{OPT} \\ &= \sum_{t=1}^T (H_t^{OPT} + \Theta_t^{OPT}) + \frac{(m-2)h}{mh+\theta} \left(1 + \frac{mh}{\theta}\right) \sum_{t=1}^T \Theta_t^{OPT} \\ &\leq \left(1 + \frac{(m-2)h}{mh+\theta}\right) \sum_{t=1}^T (H_t^{OPT} + \Theta_t^{OPT}), \end{aligned}$$

thereby proving the corollary.

Lemma 11.7 (Chao et al. (2015)) *For each realization $f_T \in \mathcal{F}_T$, we have $\sum_{t \in \mathcal{T}_\Pi} \Pi_t^{PB} \leq \sum_{t=1}^T \Pi_t^{OPT}$.*

Proof From the definition of Π_t and $\mathcal{T}_\Pi$, we have

$$\sum_{t \in \mathcal{T}_\Pi} \Pi_t^{PB} = b \sum_{t \in \mathcal{T}_\Pi} \alpha^{t-1} (D_t - Y_t^{PB})^+ \leq b \sum_{t \in \mathcal{T}_\Pi} \alpha^{t-1} (D_t - Y_t^{OPT})^+ \leq \sum_{t=1}^T \Pi_t^{OPT},$$

where the first inequality holds since $Y_t^{OPT} < Y_t^{PB}$ when $t \in \mathcal{T}_\Pi$.

With the preparations above, we put together our main result, i.e., Theorem 11.3.

Proof For each period $t = 1, \dots, T$, denote Z_t^{PB} as the conditional expected *balanced cost* by the *PB* policy in period t . That is,

$$Z_t^{PB} = \frac{mh+\theta}{2(m-1)h+\theta} \mathbb{E}[H_t^{PB} + \Theta_t^{PB} \mid F_t] = \mathbb{E}[\Pi_t^{PB} \mid F_t].$$

Note that Z_t^{PB} is a random variable before period t ; and in period t , $F_t = f_t$ is realized and its value is the expected balanced cost conditional on the observed information set f_t . Using the marginal cost accounting scheme and a standard argument of conditional expectations, we have

$$\begin{aligned} C(PB) &= \sum_{t=1}^T \mathbb{E}[H_t^{PB} + \Theta_t^{PB} + \Pi_t^{PB}] = \sum_{t=1}^T \mathbb{E}[\mathbb{E}[H_t^{PB} + \Theta_t^{PB} + \Pi_t^{PB} \mid F_t]] \\ &= \left(2 + \frac{(m-2)h}{mh+\theta}\right) \sum_{t=1}^T \mathbb{E}[Z_t^{PB}]. \end{aligned} \quad (11.42)$$

Applying Corollary 11.1, Lemma 11.7, and the fact that $\{t \in \mathcal{T}_H\}$ and $\{t \in \mathcal{T}_\Pi\}$ are completely determined by F_t , we obtain

$$\begin{aligned} C(OPT) &= \mathbb{E}\left[\sum_{t=1}^T (H_t^{OPT} + \Theta_t^{OPT}) + \sum_{t=1}^T \Pi_t^{OPT}\right] \\ &\geq \mathbb{E}\left[\frac{1}{1 + \frac{(m-2)h}{mh+\theta}} \sum_{t \in \mathcal{T}_H} (H_t^{PB} + \Theta_t^{PB}) + \sum_{t \in \mathcal{T}_\Pi} \Pi_t^{PB}\right] \\ &= \sum_{t=1}^T \mathbb{E}\left[\frac{1}{1 + \frac{(m-2)h}{mh+\theta}} \mathbf{1}(t \in \mathcal{T}_H) (H_t^{PB} + \Theta_t^{PB}) + \mathbf{1}(t \in \mathcal{T}_\Pi) \Pi_t^{PB}\right] \\ &= \sum_{t=1}^T \mathbb{E}\left[\mathbb{E}\left[\frac{1}{1 + \frac{(m-2)h}{mh+\theta}} \mathbf{1}(t \in \mathcal{T}_H) (H_t^{PB} + \Theta_t^{PB}) + \mathbf{1}(t \in \mathcal{T}_\Pi) \Pi_t^{PB} \mid F_t\right]\right] \\ &= \sum_{t=1}^T \mathbb{E}\left[\frac{1}{1 + \frac{(m-2)h}{mh+\theta}} \mathbf{1}(t \in \mathcal{T}_H) \mathbb{E}[H_t^{PB} + \Theta_t^{PB} \mid F_t] + \mathbf{1}(t \in \mathcal{T}_\Pi) \mathbb{E}[\Pi_t^{PB} \mid F_t]\right] \\ &= \sum_{t=1}^T \mathbb{E}[(\mathbf{1}(t \in \mathcal{T}_H) + \mathbf{1}(t \in \mathcal{T}_\Pi)) Z_t^{PB}] = \sum_{t=1}^T \mathbb{E}[Z_t^{PB}]. \end{aligned}$$

Thus, it follows from Equation (11.42) that $C(PB) \leq \left(2 + \frac{(m-2)h}{mh+\theta}\right) C(OPT)$.

11.5 FUTURE RESEARCH DIRECTIONS

We have discussed several core stochastic inventory models in greater depth in this book chapter. Also, we have provided the summary of current approximation ratios for various stochastic inventory systems in Table 11.1. It is clear that the journey has just started, and there are always more questions than answers. Here we would like to provide a list of important *open questions*.

- Joint pricing and inventory control problems
- Stochastic dual-sourcing inventory control problems
- Stochastic inventory control problems with random yield
- Stochastic perishable inventory systems with depletion decisions
- Stochastic joint-replenishment problems (JRP)
- Stochastic one-warehouse multi-retailer problems
- Stochastic assemble-to-order systems (ATO)

Solving any of the above problems with constant (or parametric) performance guarantees would readily yield a high-quality Ph.D. thesis or a top-tier publication in OR/MS. We hope this book chapter can inspire students and researchers to delve into these hard problems (which can often lead to novel ideas and techniques).

Equally importantly, applying these approximation techniques in practical inventory control problems would be vital in demonstrating their efficacy. Conducting any thorough empirical studies (with real data) on the developed approximation algorithms would be of great interest and importance to our community.

REFERENCES

- Axsäter, S. (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research*, 38(1), 64–69.
- Axsäter, S., & Lundell, P. (1984). In-process safety stock. In *Proceedings of the 23rd IEEE conference on decision and control* (pp. 839–842). IEEE Control Systems Society.
- Chao, X., Gong, X., Shi, C., Yang, C., Zhang, H., & Zhou, S. X. (2018). Approximation algorithms for capacitated perishable inventory systems with positive lead times. *Management Science*, 64(11), 5038–5061.
- Chao, X., Gong, X., Shi, C., & Zhang, H. (2015). Approximation algorithms for perishable inventory systems. *Operations Research*, 63(3), 585–601.
- Cheung, M., Elmachoub, A. N., Levi, R., & Shmoys, D. B. (2016). The submodular joint replenishment problem. *Mathematical Programming*, 158(1), 207–233.
- Chu, Y. L., & Shen, Z. M. (2010). A power-of-two ordering policy for one-warehouse multiretailer systems with stochastic demand. *Operations Research*, 58(2), 492–502.
- DeValve, L., Pekeć, S., & Wei, Y. (2020). A primal-dual approach to analyzing ATO systems. *Management Science*, 66(11), 5389–5407.
- Goldberg, D. A., Reiman, M. I., & Wang, Q. (2019). A survey of recent progress in the asymptotic analysis of inventory systems. *Forthcoming in Production and Operations Management*.
- Halman, N., Klabjan, D., Mostagir, M., Orlin, J., & Simichi-Levi, D. (2009). A fully polynomial time approximation scheme for single-item stochastic lot-sizing problems with discrete demand. *Mathematics of Operations Research*, 34(3), 674–685.
- Hurley, G., Jackson, P., Levi, R., Roundy, R. O., & Shmoys, D. B. (2007). New policies for stochastic inventory control models—theoretical and computational results [Working paper]. MIT.

- Jiang, Y., Shi, C., & Shen, S. (2019). Service Level constrained inventory systems. *Production and Operations Management*, 28(9), 2365–2389.
- Karaesmen, I. Z., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. *International Series in Operations Research and Management Science*, 151, 393–436.
- Levi, R. (2014). *Provably near-optimal approximation algorithms for operations management models*, chap. 8. INFORMS Tutorial Series (pp. 179–192).
- Levi, R., Janakiraman, G., & Nagarajan, M. (2008a). A 2-approximation algorithm for stochastic inventory control models with lost-sales. *Mathematics of Operations Research*, 33(2), 351–374.
- Levi, R., Lodi, A., & Sviridenko, M. (2008b). Approximation algorithms for the multi-item capacitated lot-sizing problem via flow-cover inequalities. *Mathematics of Operations Research*, 33(2), 461–474.
- Levi, R., Pál, M., Roundy, R. O., & Shmoys, D. B. (2007a). Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 821–838.
- Levi, R., Roundy, R., Truong, V. A., & Wang, X. (2017). Provably near-optimal balancing policies for multi-echelon stochastic inventory control models. *Mathematics of Operations Research*, 42(1), 256–276.
- Levi, R., Roundy, R. O., & Shmoys, D. B. (2006). Primal-dual algorithms for deterministic inventory problems. *Mathematics of Operations Research*, 31(2), 267–284.
- Levi, R., Roundy, R. O., & Shmoys, D. B. (2007b). Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 821–839.
- Levi, R., Roundy, R. O., Shmoys, D. B., & Sviridenko, M. (2008c). A constant approximation algorithm for the one-warehouse multi-retailer problem. *Management Science*, 54(4), 763–776.
- Levi, R., Roundy, R. O., Shmoys, D. B., & Truong, V. A. (2008d). Approximation algorithms for capacitated stochastic inventory models. *Operations Research*, 56(5), 1184–1199.
- Levi, R., & Shi, C. (2013). Approximation algorithms for the stochastic lot-sizing problem with order lead times. *Operations Research*, 61(3), 593–602.
- Nagarajan, V., & Shi, C. (2016). Approximation algorithms for inventory problems with submodular or routing costs. *Mathematical Programming*, 160(1–2), 225–244.
- Nahmias, S. (1975). Optimal ordering policies for perishable inventory-II. *Operational Research*, 23(4), 735–749.
- Roundy, R. O. (1993). Efficient, effective lot-sizing for multi-product, multi-stage production systems. *Operations Research*, 41(2), 371–386.
- Shen, Z. M., Shu, J., Simchi-Levi, D., Teo, C. P., & Zhang, J. (2009). Approximation algorithms for general one-warehouse multi-retailer systems. *Naval Research Logistics*, 56(7), 642–658.
- Shi, C. (2014). Approximation algorithms for stochastic optimization problems in operations management. In J. J. Cochran (Ed.) *Wiley encyclopedia of operations research and management science*, (pp. 1–20). John Wiley & Sons.
- Shi, C., Zhang, H., Chao, X., & Levi, R. (2014). Approximation algorithms for capacitated stochastic inventory systems with setup costs. *Naval Research Logistics*, 61(4), 304–319.
- Silver, E. A., & Meal, H. C. (1973). A heuristic selecting lot-size requirements for the case of a deterministic time varying demand rate and discrete opportunities for replenishment. *Production and Inventory Management*, 14, 64–74.
- Tao, Z., & Zhou, S. X. (2014). Approximation balancing policies for inventory systems with remanufacturing. *Mathematics of Operations Research*, 39(4), 1179–1197.
- Truong, V.-A. (2014). Approximation algorithm for the stochastic multiperiod inventory problem via a look-ahead optimization approach. *Mathematics of Operations Research*, 39(4), 1039–1056.
- Truong, V.-A., & Roundy, R. O. (2011). Multidimensional approximation algorithms for capacity-expansion problems. *Operations Research*, 59(2), 313–327.
- Vazirani, V. J. (2001). *Approximation algorithms*. Springer.
- Williamson, D. P., & Shmoys, D. B. (2010). *The design of approximation algorithms*. Cambridge University Press.

- Xin, L. (2021). A 1.79-approximation algorithm for a continuous review lost-sales inventory model. *Forthcoming in Operations Research*.
- Zhang, C., Ayer, T., & White, C. C. (2019). 2-approximation policies for perishable inventory systems when FIFO is an optimal issuing policy [Working paper]. Duke University.
- Zhang, H., Shi, C., & Chao, X. (2016). Approximation algorithms for perishable inventory systems with setup costs. *Operations Research*, 64(2), 432–440.
- Zipkin, P. H. (2000). *Foundations of inventory management*. The McGraw Hill Companies.

PART II

INTERFACES

12. Information and incentives in inventory management

Bharadwaj Kadiyala, Hau Lee, and Özalp Özer

12.1 INTRODUCTION

The theory and practice of inventory management have gone hand-in-hand for the past several decades. What started out as the study of centralized multi-echelon inventory management has evolved to capture the reality of today's global supply chains involving interactions among multiple firms, with their own institutional objectives and information, adhering to terms of complex supply chain agreements, and catering to a host of different and innovative business models. In these rapidly evolving business environments, oftentimes the firm-level incentives do not align with that of the supply chain, and hence the supply chain as a whole may not provide the best value to its end-customer. We shall explore, in this chapter, how inventory may be used as a lever to mediate conflicting incentives in supply chains and avoid, as observed in some cases, a market failure.

The issue of misaligned incentives and misreported private information has been widely observed in practice. For example, long-term collaborative practices such as *vendor-managed inventory* (VMI) have proven difficult to maintain over multiple planning horizons, resulting in companies terminating such agreements (e.g., Kouvelis et al., 2006; Brinkhoff et al., 2015). One frequently cited reason for such failed relationships has been the declining of trust (both in terms of credibility and capability) among firms implementing such partnerships. In the case of VMI, point-of-sale (POS) data is usually transferred to the supplier by the retailer on a regular basis to enable the supplier to perform the inventory management task. However, POS data suffers from two limitations: they are censored demand realizations (which complicate statistical inference) and they do not convey private demand information (e.g., in-store promotions) that the retailer may possess. In the classical case, Hammond (2006) describes the stern opposition to VMI practice from Barilla's distributors. The article points out the difficulty Barilla had in incorporating promotional data, which is separate from the usual electronic data interchange (EDI) information, into their forecasting process. Giorgio Maggiali, then Director of Logistics at Barilla, noted,

We're grappling with how to treat these promotions in our operations planning processes, including forecasting, manufacturing, and logistics.

Ineffectively managing inventory led to the disappointment of some distributors over VMI implementation, and eventually falling out of the relationship.¹

The above issues (in a slightly different form) are also applicable to contemporary online marketplace platforms such as Amazon, Tabao, and eBay. These platforms which serve as an intermediary connect manufacturers/sellers with potential consumers. The e-commerce ecosystem on these platforms allows them to access and analyze sales data of the individual sellers

and in turn provide valuable market insights to the sellers. However, individual sellers may not always be willing to share their proprietary information related to their pricing strategy (e.g., upcoming sales promotions) potentially lowering the value created in the marketplace platform. For example, Amazon regularly replenishes inventory in their distribution centers in anticipation of consumer demand. Without accurate pricing and promotion information, Amazon may fail to accurately anticipate consumer demand, leading to either excess overage or underage costs (which can otherwise be avoided through credible communication).

Establishing a supply chain is a long-term commitment which requires firms to put in place processes that govern information and material flow. Long-term partnerships introduce a host of new (theoretical and implementation) challenges and opportunities that one may not encounter in a one-time interaction. The focus of this chapter is to highlight precisely some of these challenges and how to address them in a dynamic inventory management setting. One popular solution approach considered in the literature (and to be discussed in this chapter) to tackle the problem of incentives and information in a dynamic setting is through the use of *dynamic contracts*. The developments in economic theory in the area of dynamic mechanism design have been particularly valuable in providing suitable frameworks to analyze the dynamic settings. The landmark result of the *revelation principle* (Myerson, 1979), which forms the bedrock of the analysis in the static setting can be extended to the dynamic setting (Myerson, 1986). While there is no general methodology to solve the dynamic contracting problem, recent advances provide some guidance (e.g., Eső & Szentes, 2007; Oh & Özer, 2013; Pavan et al., 2014).

We anchor our discussions in the chapter to the role of demand-side and supply side information in dynamic inventory management problems. In Section 12.2, we discuss inventory models with a focus on different information settings (pertaining to customer demand and on-hand inventory levels) with the underlying assumption that the supply chain functions in a centralized fashion without incentive-related concerns. One may think of the results and insights in this section as the basis to quantify the *value of information* due to technological developments (EDI, Radio Frequency Identification, or superior forecasting technologies). In Section 12.3, we consider supply chain settings in which one of the firms is better informed (about demand-related or inventory-related information) than the other firm. Furthermore, the objectives of firms in the supply chain are such that without proper monetary incentives in place credible information sharing cannot take place, often leading to a lose-lose outcome. This stream of literature while still in its infancy is rich enough to provide a flavor for the modeling challenges encountered in studying the dynamic settings. We note that Section 12.3 is the main focus of this chapter. Therefore, to keep our discussions streamlined, we only briefly discuss emerging developments in the related topics and we refer the reader to prior review studies where necessary.

12.2 INFORMATION IN INVENTORY MODELS

Demand and on-hand inventory information are perhaps the two most essential inputs to effectively implement and monitor an inventory management policy. The classical inventory models, as in Clark and Scarf (1960), impose strong assumptions about the inventory manager's access to demand and on-hand inventory information to compute the optimal policy. A number of studies have since relaxed these assumptions to characterize optimal inventory policies

under various settings. In this section, we review the issues that surface when the assumptions about demand-related information (Section 12.2.1) and inventory-related information (Section 12.2.2) are relaxed, and how they can be addressed, with emphasis on modeling approaches and the insights derived.

12.2.1 Demand Information

Demand information is perhaps one of the most fundamental inputs to inventory management problems. Here, we briefly discuss literature that relaxes the demand information requirements in classical inventory management problems (Clark & Scarf, 1960; Veinott, 1966). In most supply-chain contexts, demand in each period is not independent and identically distributed (iid) according to a *known* probability distribution. On the contrary, demand (forecast) models are periodically updated based on new information that becomes available in each period. There are several statistical approaches proposed in the literature to model and forecast demand. Broadly, these statistical approaches can be classified as Bayesian, time-series, and Martingale models of forecast evolution. We refer the reader to several excellent reviews of demand models in dynamic inventory management: Gallego and Özer (2002), Özer (2011), Chen and Mersereau (2015), Chen and Lee (2017), and Kurtuluş (2017), to list a few. Research in this field is growing to accommodate other demand learning models, which are based on data-driven methods, see, for example, Ban and Rudin (2019) and Ban (2020).

Building on the earlier works of joint demand (forecast) modeling and inventory management within a single firm, researchers have since explored the *value* of demand information in supply chains involving multiple firms. In particular, as the supply chain networks have become globalized, collaborative practices for inventory management have grown in prominence. For example, VMI was popularized by the Walmart and P&G partnership in the 1980s; *collaborative planning, forecasting, and replenishment* (CPFR) in the 1990s (Aviv, 2001). One of the important drivers for this supply chain revolution in partnerships may be attributed to the rapid evolution of (information) technology around the same time (e.g., EDI, *retail link* by Walmart, and, more recently, B2C e-commerce).

At the same time, the growing complexity of the supply chain networks also brought forward several challenges. Among these is the *bullwhip effect*, i.e., the variability of the order process is higher than the variability of the demand process at each level of the supply chain. Essentially, the bullwhip effect is an inefficiency that results from the distortion of information flows in the supply chain. In one of the most celebrated papers in management science, Lee et al. (1997) identify demand signal processing, i.e., updating order quantities based on past demand observations, as one of the drivers of the bullwhip phenomenon. Nevertheless, they also suggest a natural remedy to overcome the bullwhip—demand *information sharing* between supply chain members.

Following Lee et al. (1997), the value of information sharing to increase supply chain efficiency was extensively investigated by several scholars. Broadly speaking, the goals in the follow-up studies were to (i) quantify the value of information sharing (to better negotiate agreements within a partnership program) and (ii) explore different sources (demand model, demand forecast, order policy) of information sharing leading to varying benefits in lowering supply chain costs; see, for example, Lee et al. (2000). Gavirneni et al. (1999) investigate different degrees of information sharing ranging from *no-sharing* benchmark to *partial* to *complete* information sharing. In the partial information-sharing model, the supplier knows

the retailer's end-consumer demand distribution and the retailer also shares the details of its ordering policy (value of the parameters s, S of a (s, S) policy). Based on a simulation study, Gavirneni et al. (1999) investigate how: production capacity at the upstream supplier, cost structure (ratio of penalty to holding costs), and market conditions (demand variability) impact the value of (*partial* or *complete*) information sharing.

Aviv (2001, 2002, 2007) investigates these questions under different demand models. Gallego and Özer (2001) show that demand information sharing is valuable when the upstream supplier *cannot* infer current demand from the retailer's order history. Chen and Lee (2009) consider a most general demand model based on the Martingale Model of Forecast Evolutions (MMFE) (Heath & Jackson, 1994) to quantify the value of information sharing and the bullwhip effect. They relax the assumption that the supplier knows the retailer's demand model and order policy (within the scope of information sharing). The authors find that by having information about the retailer's *projected* future orders (which are *suitably* revised), the supplier achieves identical cost savings (compared to the case with complete information about the demand model and order policy).

A relevant and important question in the above context is whether a decentralized supply chain can be coordinated using only *local* (e.g., site inventory as opposed to echelon inventory) information. In a decentralized supply chain setting with the possibility of some oversight within the supply chain (e.g., but not limited to, the headquarters overseeing different departments within an organization), Lee and Whang (1999) propose a performance measurement scheme that has the desirable properties of being incentive compatible, conserves cost (e.g., the scheme is self-supporting without payments from the headquarters) and requires only local (and *not* echelon inventory) information to manage inventory in the supply chain. Kapuściński and Parker (2021) consider a similar setting under capacity limits and illustrate how the performance scheme proposed by Lee and Whang (1999) can be suitably updated to achieve coordination.

Much of the follow-up literature on this topic has been devoted to empirically testing whether and under what conditions the bullwhip effect is actually observed (see, e.g., Cachon et al., 2007; Bray & Mendelson, 2012). Significant progress has been made in this dimension and it is outside the scope of this chapter to discuss this literature. We refer to Chen and Lee (2017) for an extensive discussion of the empirical challenges and observations in measuring the bullwhip effect using field data.

12.2.2 Inventory Information

On-hand inventory (also like demand) information is inherently dynamic and is impacted by various processes that may or may not be under the control of an inventory manager. Not surprisingly, the accuracy of inventory information has been a topic of much interest in operations management. Industry estimates for the inaccuracy are staggering. In one of the earlier studies, Raman et al. (2001) estimate that nearly 65% of inventory records were inaccurate at a store-SKU level. Following that, a number of studies have empirically investigated the extent of inventory record inaccuracy in the retail industry (Ton & Raman, 2010). We also refer to the extensive review of this topic by Chen and Mersereau (2015).

An important step to countering inaccurate inventory information is to enrich inventory management models to account for inventory inaccuracy. In particular, Atali et al. (2009) identify and explicitly model three possible drivers of inventory inaccuracy: misplacement,

shrinkage, and transaction errors. In addition, to adjusting optimal ordering policy based on the sources of inaccuracy (Atali et al., 2009), firms may also optimally decide when to audit or inspect their inventory, see K  k and Shang (2007), DeHoratius et al. (2008), Bassamboo et al. (2020), and Chen (2021). These analytical models also help practitioners quantify the value of inventory tracking technologies, such as Radio Frequency Identification (RFID). In theory, RFID technology allows retailers to digitize and simplify access to inventory information (with little human intervention). By comparing models with and without accurate inventory information, the above papers quantify the value of RFID technology. For a detailed overview of inventory models with RFID, we refer to Lee and   zer (2007). Based on field experiments at several stores of a retail chain, Hardgrave et al. (2013) empirically find that RFID helps with reducing inventory inaccuracy by about 26%.

In terms of future research and opportunities, a growing stream of literature explores the strategic role of inventory information on the demand side of operations. How should the retailer communicate its on-hand inventory information with its consumers? This question has been answered in a few different retail contexts which we discuss next. One of the issues faced by retail platforms is that the quantity of products they receive from their suppliers cannot be contracted upon precisely. In these cases, inventory is pushed to retail platforms depending on what is available with the supplier and, therefore, the on-hand inventory information is random from the platform's point of view. In a setting with two vertically differentiated products, Cui and Shin (2018) consider a model where the starting inventory of two variants of a product and the total inventory are random variables. Its realization is available only to the retail platform who *ex-post* (i.e., after observing the inventory realization) decides whether to provide disaggregate, aggregate or no-information about its inventory to its consumers. Assuming truthful disclosure, they show that *ex-post* the retailer benefits from always providing aggregate inventory information to consumers (i.e., the total inventory of the two variants of the product) rather than disaggregate inventory information. The equilibrium results are driven by the retailer's desire to lower supply-demand mismatch.

The above assumption of truthful inventory communication can be relaxed by noting that perhaps the retailer may have an interest in manipulating consumer beliefs about its on-hand inventory. For example, all major online retailers/intermediaries display some form of inventory information to their consumers: some examples include: "in stock" on Amazon.com; detailed SKU-level inventory information by IKEA and Target; "almost gone!" by Sierra Trading Post. What can consumers learn from such communication (when the retailer need not be truthful)? In one of the earlier studies, Allon and Bassamboo (2011) model the communication game between the retailer and the consumer as a cheap-talk game, i.e., the inventory-related messages do not directly affect consumer payoff, are unverifiable (by the consumer) and non-binding. They show that the messages (e.g., "buy now") shared by the retailer cannot credibly communicate any inventory-related information with the consumers.

In contrast to the above disclosure schemes, a recent emerging body of literature considers the question of how a retailer should communicate inventory information by *committing* to a signaling mechanism *ex-ante*. In particular, the retailer reveals the conditional probability distribution (based on the inventory realization) that will be used to generate the communication message, *prior* to the realization of the inventory level. Drakopoulos et al. (2021) consider such a setting in the context of personalized inventory information, i.e., there is a possibility to communicate real-time inventory information on a one-on-one basis with consumers. In particular, the paper explores the joint problem of pricing and designing a signaling mechanism,

prior to observing their on-hand inventory using a *Bayesian persuasion* framework (Kamenica & Gentzkow, 2011). The signal (about on-hand inventory) that is communicated to consumers after the realization of on-hand inventory, impacts the consumer's belief about product availability and their decision to buy the product in one of two periods (if at all). They find that personalizing the communication, i.e., sending different messages—"buy now" vs. "wait"—depending on consumer valuation, can significantly increase the retailer's profit. Küçükgül et al. (2021) consider an online retail platform's information provision problem in the context of "time-locked sales" using a *dynamic* Bayesian persuasion framework (Kremer et al., 2014). In particular, an online retail platform dynamically decides what information to provide each arriving customer to maximize its revenues. The information provided by the platform—which can potentially be any function of past sales data—fuels social learning among consumers about the valuation of the product. In such a setting, they show it is sufficient for the platform to consider providing only one of three messages: "neutral", "positive", or "negative" to impact a consumer's purchase decision.

12.3 INFORMATION AND INCENTIVES IN INVENTORY MODELS

In this section, we append the discussions in Section 12.2 by investigating how information and incentives (of sharing information) evolve in a dynamic decentralized inventory management problem. In particular, we consider the arguably more realistic scenario where one of the supply-chain firms has superior information than other firms in the supply chain, i.e., there is *information asymmetry* in the supply chain. Furthermore, we relax the assumption that firms do not strategically act on their private information. If the incentives of all firms are aligned with that of the entire supply chain then the assumption is non-binding, and hence, it can be ignored. Anecdotal evidence suggests otherwise (see Section 12.1 for examples).

Our discussion in this section focuses on two aspects of the problems that arise due to information asymmetry in dynamic inventory management problems. *First*, as in the previous section, we focus on the different *sources* of information asymmetry, pertaining to: demand information (Section 12.3.1), on-hand inventory information (Section 12.3.2), and cost information (Section 12.3.3). Furthermore, in Section 12.3.4 we consider demand information asymmetry in capacity planning models. *Second*, through our discussion of the various information asymmetry settings we wish to illustrate the breadth and depth of the *modeling approaches* employed to tackle information asymmetry in dynamic inventory management problems.

Related to the *first* point above, our goal is to highlight how inventory dynamics and information asymmetry interact in various supply chain settings. In Section 12.3.1, we consider inventory management problems where one of the firms, typically, the downstream retailer, has superior demand information than the upstream supplier in a multi-period setting. In Section 12.3.2 we consider a setting where the downstream retailer has superior information about on-hand inventory and may use that information strategically in its interaction with an upstream supplier. In Section 12.3.3, we discuss inventory management problems with information asymmetry due to cost information (shortage cost and production cost).

Related to the *second* point above, our goal is to illustrate the different solution approaches in the literature used to tackle problems of information asymmetry in dynamic inventory models. Generally speaking, the most common solution approach is to design contracts (i.e.,

order quantity–payment plan) that facilitate information sharing, improve supply chain performance, and create a win–win outcome.² The complexity in determining terms of a contract in dynamic settings, however, arises due to the fact that one needs to account for and prescribe action (and payment) for all possible contingencies that may arise in future interactions. That is, the contract dictates the terms of trade in the period the contract is signed (as in a single-period setting), and also for all future periods in a dynamic setting.

Contracts in dynamic multi-period settings can be broadly classified based on how the terms of trade in the contract are determined (*static* vs. *dynamic*) and on the duration of the contract (*short-* vs. *long-term*). If contract terms depend on the realization of some randomness in the setting, then such a contract is a dynamic contract, i.e., terms are contingent on the realization of the randomness in the environment. In contrast, the terms of a static contract can be pinned down completely at the start of the planning horizon. If the contract binds the firms for a single time period (in a multi-period relationship), then such a contract is a short-term contract. In contrast, if the contract binds the firms for the entire duration of their relationship, then such a contract is referred to as a long-term contract. For example, in a short-term dynamic contract, a firm chooses whether to offer a contract on a period-to-period basis. Whereas once a long-term dynamic contract is set in motion, it lasts until the end of the planning horizon. Note that in both cases the contract terms can be dynamically determined. In a single-shot interaction without the possibility of evolution of information or incentives, it makes sense to consider only static contracts. Of course, there is no possibility to offer long-term contracts in these settings. In a multi-period setting where information and incentives may evolve, however, all four possibilities emerge: $\{\text{static, dynamic}\} \times \{\text{short-term, long-term}\}$. Given our interest in studying settings with information evolution, in this section, we devote our discussions to dynamic contracts, which can be short- or long-term.

12.3.1 Demand Information Asymmetry

In this section, we consider inventory management problems where a downstream retailer has superior demand information than an upstream supplier. In Section 12.3.1.1, we discuss Kadiyala et al. (2020) in which the upstream supplier (statistically) learns about demand information through sales data, which in turn affects the timing (i.e., the period in which to offer the contracts) and the design of contracts. In Section 12.3.1.2, we discuss Lobel and Xiao (2017), who design contracts to be offered at the start of the planning horizon, which facilitates information sharing in an environment where the retailer's private information is non-stationary. These two papers also illustrate complementary approaches to modeling demand information asymmetry—*persistent* vs. *non-persistent*. In both approaches, a relevant parameter of the demand distribution is privately known to the retailer. In Kadiyala et al. (2020), the parameter of demand distribution remains constant over the planning horizon, whereas in the non-persistent setting, as in Lobel and Xiao (2017), the parameter evolves over time, i.e., the parameter is non-stationary.

12.3.1.1 Sharing stationary demand information

Kadiyala et al. (2020) consider an inventory management problem faced by an upstream supplier that is in a collaborative agreement, such as VMI, with a retailer. A VMI partnership provides the supplier with an opportunity to manage inventory for the supply chain in exchange for POS and inventory-level information from the retailer. However, retailers typically possess

superior local market information beyond POS data. This information is useful to the supplier for inventory planning purposes, however, it is often difficult to communicate this information even in long-term agreements such as VMI, resulting in firms terminating such agreements (see Section 12.1 for examples). Kadiyala et al. (2020) investigates how a supplier should manage inventory and update an ongoing VMI agreement to maximize profit by facilitating credible demand information sharing.

In a typical VMI agreement, at the start of each review period, the supplier decides the amount of product to produce (at unit cost c) and delivers it to the retailer. The retailer satisfies the demand to the extent possible (unmet demand is lost) and shares the POS data with the supplier. The supplier is liable for any leftover inventory, incurring unit holding cost of h per period. If the retailer stocks out, then neither the supplier nor the retailer observes the true demand realization. The retailer earns a per unit revenue r and pays the supplier a unit wholesale price w , which are all stipulated in the *ongoing* VMI agreement. The demand in each period is iid with cdf $G(\cdot)$ and probability density function (pdf) $g(\cdot)$. The downstream retailer has superior information about local demand conditions, which is modeled in a parametric fashion. There is a parameter of the demand distribution denoted by ξ which is known to the retailer but not the supplier.³ The supplier, however, has probabilistic information about the parameter ξ , i.e., prior distribution π on the set of values taken by ξ .

In the above supply chain setting, the supplier has two channels to acquire information about the unknown parameter of the demand distribution. First, based on the periodic POS data, the supplier can dynamically update his⁴ belief π_t in each period to make better inventory decisions over time (*learn* approach). Alternately, as in a static contracting problem, the supplier may seek to credibly elicit this information from the retailer at the start of the planning horizon by offering an appropriately designed menu of screening contracts (*screen* approach) within the purview of the ongoing VMI agreement. This paper explores a *learn-and-screen* approach which dynamically considers the tradeoff between choosing either information acquisition channels in each time period. Below we focus on the interplay between the learn-and-screen approaches in a dynamic inventory problem.

Suppose that the supplier offers a menu of contracts $\{S(\cdot), P(\cdot)\}$ in the first period. If the retailer with private demand information ξ chooses a contract $S(\tilde{\xi}), P(\tilde{\xi})$ from the menu, then the supplier maintains a base-stock level $S(\tilde{\xi})$ in each of the following periods in exchange for a one-time payment⁵ $P(\tilde{\xi})$ from the retailer in period one. The retailer's and the supplier's profit, denoted by Π^r and Π^s , respectively, for the remaining time horizon is given by

$$\begin{aligned}\Pi^r(S(\tilde{\xi}), P(\tilde{\xi}), \xi) &= \mathbb{E} \left[\sum_{t=1}^{\infty} \alpha^{t-1} ((r - w) \min\{S(\tilde{\xi}), D_t\} - P(\tilde{\xi})) \right] \\ \Pi^s(x_1, S(\tilde{\xi}), P(\tilde{\xi})) &= \mathbb{E} \left[\sum_{t=1}^{\infty} \alpha^{t-1} (w \min\{S(\tilde{\xi}), D_t\} - c(S(\tilde{\xi}) - x_t) - h(S(\tilde{\xi}) - D_t)^+ + P(\tilde{\xi})) \right],\end{aligned}$$

where α is the discounting factor and x_t is the starting inventory level in period t . Due to the *revelation principle*, it is sufficient for the supplier to restrict attention to contracts that facilitate truth-telling (Myerson, 1979). That is, we impose the following set of constraints

$$\Pi^r(S(\tilde{\xi}), P(\tilde{\xi}), \xi) \geq \Pi^r(S(\tilde{\xi}), P(\tilde{\xi}), \xi), \quad \forall \tilde{\xi} \neq \xi. \quad (12.1)$$

Furthermore, the menu of contracts should improve the retailer's profit over her reservation profit, which in this case is the profit obtained from the ongoing VMI agreement. However, the profit under the ongoing VMI agreement is quite complex since the supplier's inventory management problem does not admit a closed-form solution; see also Chen (2010); Bisi et al. (2011). Thus, to ensure participation we also need to have:

$$\Pi^r(S(\xi), P(\xi), \xi) \geq (r - w) \sum_{t=1}^{\infty} \alpha^{t-1} \mathbb{E}[\min\{y_t^o, D_t(\xi)\}], \quad \forall \xi, \quad (12.2)$$

where $\Pi_{\min}^r(x, \mathbf{y}^o, \xi)$ is the type- ξ retailer's reservation profit when the on-hand inventory level is x and the supplier maintains post-order inventory levels $\mathbf{y}^o = (y_1^o, y_2^o, \dots)$ if the retailer rejects the menu of contracts offered. The supplier's incentive problem can be summarized as follows:

$$\tilde{\Pi}^{sr}(x_1, \pi_1) := \max_{S(\cdot), P(\cdot)} \mathbb{E}_{\xi}[\Pi^s(x_1, S(\xi), P(\xi))]; \text{ subject to } S(\cdot) \geq x_1, (12.1), \text{ and } (12.2). \quad (12.3)$$

Kadiyala et al. (2020) solve the above contract design by first determining a closed-form upper bound $\bar{\Pi}_{\min}^r(x_1, \mathbf{y}^o, \xi)$ on the retailer's reservation $\Pi_{\min}^r(x_1, \mathbf{y}^o, \xi)$. Replacing the original reservation profit with an upper bound provides a feasible solution to the contract design problem.

In the learn-and-screen approach, the supplier also optimally decides which period to offer the screening contracts. Prior to that period, the supplier makes inventory decisions to meet demand *and* also learn about the underlying demand (in a Bayesian fashion). Thus, the learn-and-screen approach gives rise to a unique Bayesian inventory-optimal stopping problem. The corresponding value function is given by

$$\begin{aligned} & \tilde{V}(x_1, \pi_1) \\ &:= \sup_{(\mathbf{y}, \tau)} \mathbb{E} \left[\sum_{n=1}^{\tau-1} \alpha^{n-1} (cx_n + (w-c)y_n - (w+h) \int_0^{y_n} Q_n(z) dz) + \alpha^{\tau-1} \tilde{\Pi}^{sr}(x_{\tau}, \pi_{\tau}) \right], \end{aligned} \quad (12.4)$$

where $\mathbf{y} := (y_1, y_2, \dots, y_{\tau-1})$ are the post-order inventory levels prior to offering the screening contracts; $\tau \in \{1, 2, \dots\} \cup \{+\infty\}$ is a contract offering time, and Q_t is the posterior predictive distribution given by $\int_0^z g(z | \xi) \pi_t(\xi) d\xi$. The optimal policy under the learn-and-screen strategy can be obtained using a dynamic programming approach. Note that the screening contracts in the learn-and-screen approach are dynamic in that the exact terms of the contract depend on the time period in which they are offered, which in turn depends on the entire history of (random) sales and inventory decisions. Furthermore, the contracts are long-term since it is binding for the remaining time horizon.

The value function associated with the learn-and-screen approach in Equation (12.4) brings to the fore the interplay between the two sources of information acquisition. The supplier's production/inventory decisions prior to offering screening contracts determine the evolution

of the supplier's belief process. The supplier updated belief in the screening period π_t in turn determines the retailer's information rent (and also the retailer's reservation profit).

Underlying the optimal contract, there are two (sometimes, opposing) forces based on the learning dynamic and the incentive necessary for credible communication of demand information, which together determines the optimal menu of base-stock levels. To illustrate how these forces interact, consider two arbitrary but consecutive time periods. In the first period, the POS can reveal either a censored or an uncensored demand realization. A censored demand observation in the first period suggests to the supplier that the average market size must be larger than what was expected previously. Counter-intuitively, however, the optimal menu of base-stock levels offered in the following period are smaller. The increased confidence in a larger market size implies that the retailer makes greater expected profit than the previous period. As a result, the supplier lowers the menu of base-stock levels (and hence, the incentive) offered to the retailer, while still facilitating credible communication.

Likewise, an uncensored demand observation in the first period suggests to the supplier that the average market size may be smaller *or* larger than what was previously expected. The direction of this ordering depends on the magnitude of the sales observation. Following a small demand realization, the optimal menu of base-stock levels becomes larger to increase the incentive for the retailer to share the demand information in the following period. However, as the magnitude of the demand observation increases the supplier becomes more confident that the underlying average market size is large. In the event of a *large* uncensored demand observation, the supplier mimics his actions following a censored demand observation in the first period, i.e., resorts to lowering the menu of base-stock levels.

In summary, Kadiyala et al. (2020) propose and characterize a dynamic learn-and-screen approach, which suitably augments an ongoing VMI agreement to facilitate credible communication of demand information. Notably, the proposed learn-and-screen approach can be easily incorporated into an ongoing VMI agreement for the following reasons. First, the learn-and-screen approach does not disturb the terms of the ongoing VMI agreement (ownership and control of inventory, wholesale price) between the firms. Second, the form of the contract (base-stock policy) is optimal because the supplier faces the classical periodic-review inventory control problem with lost sales after demand information is (and can be) credibly shared. Third, monitoring the contract terms, after they are accepted, requires minimal effort. The supplier collects a one-time payment from the retailer, and the retailer periodically monitors the base-stock inventory level maintained by the supplier. In fact, current VMI frameworks, such as PeopleSoft Enterprise Inventory and Fulfillment Management by Oracle, already implements this feature.

12.3.1.2 Sharing non-stationary demand information

Lobel and Xiao (2017) consider a two-level supply chain consisting of an upstream supplier and a downstream retailer. The retailer owing to her proximity to consumer demand is equipped with better demand forecast information compared to the supplier. The supply chain setting considered is decentralized in that the periodic demand and inventory at the downstream retailer is not shared with the supplier.

Lobel and Xiao's (2017) paper considers an infinite horizon periodic-review inventory control problem. In each period t , the retailer first obtains a (private) demand forecast $\mu_t \in [\underline{\mu}, \bar{\mu}]$ with a cumulative distribution function (cdf) $F(\cdot)$. The retailer then places an order with the supplier for q_t units of product raising the on-hand inventory level from x_t to $x_t + q_t$. The

supplier operates in a make-to-order setting with zero lead time to produce and deliver the quantity ordered by the retailer. Furthermore, the marginal production cost is c , the retail price is r , and unit holding and backlogging costs are h, b per period, respectively. All these parameters are public information. The actual demand realized in period t is given by $\mu_t + \varepsilon_t$, where $\varepsilon_t \in [\underline{\varepsilon}, \bar{\varepsilon}]$ with cdf $G(\cdot)$, is a zero-mean random variable, capturing the error in demand forecast. Importantly, ε_t is realized after the retailer places the order for period t with the manufacturer. Both μ_t, ε_t in each period are the retailer's private information but distribution functions F, G are public information.

Given this model setup, Lobel and Xiao (2017) use a principal-agent framework to formulate and solve the supplier's (supply) contract design problem that maximizes his profit under backlogging and lost-sales settings. We highlight some of the unique aspects of the modeling framework. First, the private information in this setting is non-persistent. That is, the retailer's private demand forecast information is different in each period, drawn iid from $F(\cdot)$. Second, the focus of their paper is on long-term (dynamic) contracts. The contracts are long-term in that the contract specifies the terms of the trade for the entire time horizon. In addition, these contracts are dynamic in that, the terms of the contract in period t are dependent on all the information available until period t , i.e., they are *history-dependent*. The contract needs to be signed by the retailer in period one, although the precise realization of the future terms of trade are not realized by then. This feature of *long-term* contracts also highlights the difficulty in incentivizing the retailer to sign such a contract. Further, the supplier needs to have sufficient commitment power to convince the retailer that he would not deviate from the contract terms in the future.

To put the key results in perspective, we first note that the retailer's optimal inventory policy *without* information asymmetry is a simple base-stock policy (under both backlogging and lost-sales settings) whereby the average order size is equal to μ_t and the safety-stock is used to counter the uncertainty due to the forecasting error ε_t . With information asymmetry, the contract design problem is significantly more complex due to the space of the possible contract forms: it consists of a combination of order quantity and payment. Under both backlogging and lost-sales settings, the authors show that the optimal contract consists of a base-stock policy. Note, however, that the base-stock level in the case without information asymmetry may be different than in the case with information asymmetry. The optimal long-term contract is designed carefully to put in place the incentive necessary for the retailer to choose the appropriate base-stock levels. The optimal payment (to the supplier) structure is a combination of a fixed-fee and a wholesale price agreement.

To highlight the key modeling features, we first introduce the notion of a long-term contract more formally.⁶ We define the history of realized and forecasted demand by $h_t = \{\mu_1, \varepsilon_1, \mu_2, \varepsilon_2, \dots, \mu_t\}$. In other words, h_t is all the information available to the retailer prior to making the order decision in period t . The long-term contract offered by the supplier would elicit the retailer's private information in each period. We denote the history of reported information by the retailer as $\hat{h}_t = \{\hat{\mu}_1, \hat{\varepsilon}_1, \hat{\mu}_2, \hat{\varepsilon}_2, \dots, \hat{\mu}_t\}$. As such, $h_t \neq \hat{h}_t$.

As in the static contracting problems, the difficulty in solving the dynamic contract design problem arises due to the generality of the retailer's response $\hat{h}_t$, which makes the set of possible contracts to consider extremely large and intractable. The *revelation principle* in static contract design problems states that any feasible contract can be equivalently implemented by a contract that enables truth-telling (Myerson, 1979). This result significantly reduces the set of feasible contracts to consider, making the problem more tractable. The revelation principle

can be extended to the case of dynamic contracting under commitment (Myerson, 1986). In this case, it is sufficient to consider dynamic contracts that ensure truth-telling $h_t = \hat{h}_t$, in each period t . Even with this simplification, the dynamic contracting problem is still challenging due to the fact that truth-telling has to hold in every period.

Consider any arbitrary long-term contract $\{q_t(\hat{h}_t), T_t(\hat{h}_t)\}_{t \geq 1}$ which induces the retailer to report $\hat{h}_t$ to the supplier. Under this contract, the retailer chooses quantity $q_t(\hat{h}_t)$ and makes a payment $T_t(\hat{h}_t)$ to the supplier in period t . The retailer's profit from period t onward denoted by $\Pi_t(h_t, \hat{h}_t)$, for any reported history $\hat{h}_t$ and realized history h_t , is given by:

$$\begin{aligned} \Pi_t(h_t, \hat{h}_t) = & p\mu_t - \mathbb{E}[h(y_t - \varepsilon_t)^+ + b(\varepsilon_t - y_t)^+] - T_t(\hat{h}_t) \\ & + \delta \mathbb{E}[\max_{\hat{h}_{t+1}} \Pi_{t+1}(h_{t+1}, \hat{h}_{t+1})]. \end{aligned} \quad (12.5)$$

The above profit function consists of the retailer's revenue minus holding and backlogging costs, and the payment to the supplier in period t . Further, $y_t = q_t - \mu_t$ is the safety stock over the mean demand (μ_t) ordered to satisfy demand.

The supplier's contract design problem can be formulated as

$$\max_{\{q_t, T_t\}} \mathbb{E} \left[\sum_{t=1}^{\infty} \delta^{t-1} (T_t(h_t) - cq_t(h_t)) \right], \quad (12.6)$$

$$\text{s.t. } \Pi_t(h_t, h_t) \geq \Pi_t(h_t, \hat{h}_t), \forall t \text{ and} \quad (12.7)$$

$$\Pi_1(h_1, h_1) \geq 0. \quad (12.8)$$

Under the incentive-compatibility constraint Equation (12.7), the equilibrium response of the retailer is to truthfully share $h_t = \{\varepsilon_{t-1}, \mu_t\}$ in each period. The supplier's objective function in Equation (12.6) as a result depends on h_t . The supplier has to ensure that the retailer signs the contract in period one, when it goes into effect for the remaining time horizon. To that end, Equation (12.8) ensures that the retailer obtains at least their reservation profit, which is normalized to zero above.⁷ The manufacturer's contract design problem under the backlogging and lost-sales models is similar with the exception of the inventory dynamics. In the case of backlogging, the dynamics are linear whereas in the case of lost sales they are non-linear.

Lobel and Xiao (2017) characterize the optimal solution to the contract design problem in Equations (12.6)–(12.8) based on a relaxation (of the space of feasible contracts) approach (Eső & Szentes, 2007). Under the backlogging model, the optimal contract is a combination of the wholesale price $w(\mu_1)$ and a fixed fee $T(\mu_1)$ (only for the first period), which together induce the retailer to order according to a base-stock policy. The induced base-stock level in each period is determined by the average demand in the period μ_t and a safety-stock level y_t , which remains constant throughout the time horizon and is equal to $y_1(\mu_1)$. In other words, it is sufficient to incentivize the retailer to credibly reveal private information $h_1 = \mu_1$ in the first period.

The optimal long-term contract in the case of the lost-sales model exhibits a similar structure, i.e., wholesale price $w(\mu_1)$ and a fixed fee $T(\mu_1)$ (albeit different values), to ensure truthful communication of the retailer's private information h_t . However, due to the lost-sales inventory dynamics, the first-period demand forecast μ_1 no longer impacts the ensuing inventory problem after the first stock-out (as long as the stock-out event is credibly communicated with the supplier). In fact, the inventory problem after the stock-out event is the same regardless of the initial demand forecast μ_1 . Therefore, under the optimal long-term contract, after the first stock-out event, the supplier lowers the wholesale price to equal the production cost c . As a result, the optimal ordering policy after the stock-out event is a base-stock policy, with the base-stock level that maximizes the total supply chain profit. Nevertheless, to facilitate credible communication of stock-out event requires another payment from the retailer, which is exercised when the first stock-out happens.

In summary, Lobel and Xiao's (2017) paper justifies the use of simple wholesale price and two-part tariff contracts in dynamic inventory problems when the supplier has the commitment power to execute a long-term dynamic contract. Further, contrasting the optimal contract under backlogging and lost-sales settings reveals that the supply chain can be coordinated under the lost-sales setting after the first stock-out event.

12.3.2 Inventory Information Asymmetry

In this section, we discuss Zhang et al. (2010) who consider information asymmetry pertaining to on-hand inventory level in dynamic inventory management problems. The solution approach prescribed in the paper is based on dynamic short-term contracts (in contrast to the dynamic long-term contracts discussed in Section 12.3.1).

Zhang et al. (2010) consider a dynamic inventory problem (with lost sales) in a two-level supply chain setting where the downstream firm (the retailer) has more information about local inventory compared to the upstream firm (the supplier). There are two noteworthy aspects of the supply chain setting considered in the paper. First, periodic demand/sales observed by the retailer are not shared with the supplier. This assumption captures the plight of missing credible communication channels in real-world supply chains. Second, the authors consider a short-term contracting framework to analyze and solve the dynamic inventory problem. That is, the supplier can only commit to procurement contracts that are time-bound. There are a few reasons that motivate the short-term contracts. For one, it simplifies execution and monitoring contracts in that the terms of a short-term contract are valid only for a single period. Furthermore, the supplier may just not have the history/reputation to credibly offer and execute long-term contracts. In this sense, the supply chain considered in the paper is not fully matured to implement long-term partnership programs such as vendor-managed inventory or consignment-type arrangements.

The retailer's inventory level at the start of the time horizon x_1 is her *private* information. At the start of each period $t \geq 1$, the retailer places an order for q_t units of product with the supplier, raising her on-hand inventory level to $y_t := x_t + q_t$. The supplier has zero lead time to produce and ship the retailer's order quantity in each period. The demand in period t denoted by the random variable D_t is iid with cdf $F(\cdot)$ (and pdf $f(\cdot)$). The retailer's single-period revenue function is denoted by $v(y_t) := r\mathbb{E}[\min\{y_t, D_t\}]$, where r denotes the retail price. The retailer carries any leftover inventory in period t after satisfying the period's demand, i.e.,

$x_{t+1} := \min\{y_t, D_t\} - D_t$ at a unit holding cost h . The distribution $F(\cdot)$ and the parameters r, h are assumed to be public information.

The supplier operates in a make-to-order setting and incurs a marginal production cost c to satisfy the retailer's order. The supplier's knowledge of the retailer's initial inventory is modeled in a probabilistic fashion with cdf G_1 (and pdf g_1) with support over $[0, y_0]$. Note that G_1 may have a jump at $x_1=0$, which happens when the retailer stocks out of inventory and this discontinuity of the distribution has important consequences for the optimal solution. Only the retailer observes demand realization in each period, and as a result the retailer's inventory level in all subsequent periods is also her private information.

We briefly discuss the single-period problem (aka *newsvendor* problem) since the solution in the static setting reveals an important property of the optimal solution which also later applies to the dynamic problem. According to the *revelation principle*, it is sufficient for the supplier to restrict attention to contracts that induce *truth-telling*, i.e., the supplier offers a menu of contracts $\{s(x_1), q(x_1)\}$, such that the retailer with on-hand inventory x_1 chooses the order quantity $q(x_1)$ in exchange of the payment $s(x_1)$ to the supplier. In choosing so, the retailer credibly reveals her on-hand inventory level to be x_1 . Truth-telling is achieved by imposing the *incentive compatibility* (IC) constraint(s):

$$v_1(x_1 + q_1(x_1)) - s_1(x_1) \geq v_1(x_1 + q_1(\hat{x}_1)) - s_1(\hat{x}_1), \quad x_1, \hat{x}_1 \in [0, y_0] \quad (12.9)$$

where $\hat{x}_1$ is any arbitrary inventory level reported by the retailer when her true inventory level is x_1 . In addition, the optimal menu of contracts $\{s(x_1), q(x_1)\}$ should provide the retailer at least as much *reservation profit* as the retailer would obtain from her outside option, specified in the following *individual rationality* (IR) constraint.

$$v_1(x_1 + q_1(x_1)) - s_1(x_1) \geq v_1(x_1), \quad x_1 \in [0, y_0] \quad (12.10)$$

The above constraint features a type-dependent reservation profit $v_1(x_1)$. If the retailer rejects the supplier's menu of contracts, then she makes a profit by satisfying demand (to the extent possible) with existing inventory (which is her private information). The supplier's contract design problem is given by

$$\max_{\{s_1(x_1), q_1(x_1)\}} \int_0^{y_0} (s_1(x_1) - cq_1(x_1))g(x_1)dx_1; \quad \text{subject to (12.9), (12.10).} \quad (12.11)$$

We focus our attention on the structure of the optimal contract in the static setting, when the demand distribution $F(\cdot)$ is exponential with mean $\frac{1}{\lambda}$. Furthermore, we assume that the initial inventory level $x_1 = [y_0 - D_0]^+$, where y_0 is a known constant and D_0 is iid according to distribution F .⁸ The optimal solution in this setting exhibits a special structure:

$$q_1(x_1) = \begin{cases} \frac{1}{\lambda} \log\left(\frac{r}{c}\right), & x_1 = 0; \\ 0, & x_1 \in (0, y_0], \end{cases} \quad (12.12)$$

$$s_1(x_1) = \begin{cases} \frac{r-c}{\lambda}, & x_1 = 0; \\ 0, & x_1 \in (0, y_0]. \end{cases} \quad (12.13)$$

There are a couple of points worth highlighting here: First, the supplier transacts *only* with the retailer who has zero on-hand inventory, i.e., $x_1 = 0$. Second, the retailer with zero on-hand inventory makes *zero* profit (by accepting the supplier's contract) which is also their reservation profit since she has no inventory of her own to satisfy demand. Fundamentally, the retailer's value for additional inventory from the supplier reduces as her on-hand inventory increases. The above contract form, called the *batch-order contract* (BOC), is a consequence of this economic force taken to the extreme. What is even more remarkable is that BOC with appropriately designed terms, is optimal even in the dynamic setting under some conditions!

From a modeling perspective, the focus of the paper is on designing *short-term* contracts: A short-term contract lasts one time period and it specifies the terms of the trade, i.e., the quantity–payment schedule for *that* period. Therefore, in any period, the retailer may choose to participate in the mechanism by choosing a contract, or alternately, choose not to participate in the mechanism in that period. In the latter case, the retailer may satisfy demand in that period using inventory carried over from previous periods (if at all). In contrast to the static problem, the sequence of contract offers and responses repeats in each period. To summarize, the supplier designs and offers a menu of contract $\{s_t(x_t), q_t(x_t)\}$ in each period t to credibly elicit the retailer's private on-hand inventory information x_t .

With repeated interactions between the supplier and the retailer truthfully revealing on-hand inventory level in the first period may work against the retailer in the long term. The supplier can learn the information in the first period and may then exploit the retailer in the future. Realizing this issue, the retailer may not reveal her private information credibly in the first period. As a result, optimal contracting may result in *pooling* and *separating* equilibrium. In fact, the optimal BOC in Equations (12.12) and (12.13) in the static setting only screens the retailer with zero on-hand inventory while the other retailers with positive on-hand inventory pool, albeit in an extreme fashion.

The challenge in solving a dynamic short-term contracting problem (in general) arises from the fact that the supplier's contract offer in a period should take into account the retailer's objective in that period *and* the retailer's expectation of the supplier's contract offering in the next period. The contract offered in the next period, however, depends on the supplier's belief at the start of the next period (which in turn depends on the retailer's order quantity decision in the current period). This dynamic gives rise to the possibility that the retailer may manipulate her ordering decisions in the current period not only for immediate gains⁹ but also to impact the supplier's future (belief, and hence) contract offer. A suitable equilibrium concept for this dynamic game of incomplete information is the *perfect Bayesian equilibrium* (Fudenberg & Tirole, 1991). This equilibrium concept puts reasonable restrictions on not just actions (quantity/payment decisions) but also on the evolution of the supplier's belief about the retailer's on-hand inventory.

Now consider the dynamic version of the problem. Consider the supplier's problem in period t . The IC constraint¹⁰ is given by:

$$\begin{aligned} v(x_t + q_t(x_t)) - s_t(x_t) + \delta U_{t+1}(x_t + q_t(x_t) | x + q_t(x_t)) \\ \geq v(x_t + q_t(\hat{x}_t)) - s_t(\hat{x}_t) + \delta U_{t+1}(x_t + q_t(\hat{x}_t) | \hat{x}_t + q_t(\hat{x}_t)), \end{aligned} \quad (12.14)$$

where $U_{t+1}(\cdot)$ represents the retailer's expected value-to-go from period $t+1$ onwards. Consider the left-hand side (LHS) of the above inequality and a retailer with on-hand inventory level x_t in period t . The first part $v(x_t + q_t(x_t)) - s_t(x_t)$ is similar to LHS of Equation (12.9). The new term $\delta U_{t+1}(x_t + q_t(x_t) | x_t + q_t(x_t))$ is the retailer's expected value-to-go (discounted by a factor δ) when she truthfully reports her inventory level to be x_t by choosing a contract $q_t(x_t)$. This function encapsulates the retailer's (rational) expectation of the contract offering in the next period. Consider now the right-hand side (RHS): the term $\delta U_{t+1}(x_t + q_t(\hat{x}_t) | \hat{x}_t + q_t(\hat{x}_t))$, which represents the expected value-to-go if the retailer with on-hand inventory level x_t chooses to report, instead, $\hat{x}_t$, by choosing order quantity $q_t(\hat{x}_t)$. In this case, the supplier believes the retailer's inventory level after replenishment is $\hat{x}_t + q_t(\hat{x}_t)$ and updates his belief process accordingly. The individual rationality constraints are also updated based on the multi-period setting as follows:

$$v(x_t + q_t(x_t)) - s_t(x_t) + \delta U_{t+1}(x_t + q_t(x_t) | x_t + q_t(x_t)) \geq v(x_t) + \delta \underline{U}_{t+1}(x_t), \quad (12.15)$$

where $\underline{U}_{t+1}$ is the retailer's reservation profit obtained by not ordering from the supplier from period $t+1$ onward.

The supplier designs the menu in period t , $\{s_t(x_t), q_t(x_t)\}$ to maximize his total profit from transactions in period t and his value-to-go:

$$\max_{\{s_t(x_t), q_t(x_t)\}} \int \frac{\{s_t(x_t) - c q_t(x_t) + \delta \Pi_{t+1}(x_t + q_t(x_t))\}}{J_t(x_t + q_t(x_t) | x_t)} g_t(x_t) dx \quad (12.16)$$

where Π_{t+1} is the supplier's value-to-go from period $t+1$ onward. The standard procedure¹¹ of solving for the optimal contract to offer in period t is to maximize $J_t(x_t + q_t(x_t) | x_t)$ point-wise, which in turn depends on the structure of the first derivative of J_t .

The optimal contract in period t (other than the terminal period when it is BOC) can be quite complicated to derive in general. Considering an infinite horizon simplifies the problem to some extent due to the stationarity of the value-to-go functions and the optimal contract. In fact, an optimal short-term contract in the infinite horizon problem turns out to be astonishingly simple—a BOC with fixed (b^*, s^*) under some conditions on the cost parameters. That is, in any period, the supplier offers to supply b^* units of the product in exchange for a payment of s^* from the retailer. Like in the static case, only the retailer with zero on-hand inventory in that period would choose to participate in the mechanism. We note that the inventory dynamics under BOC (b^*, s^*) is similar to that under a (s, S) policy with $s=0$ and $S=b^*$. This resemblance is particularly counter-intuitive since there are no fixed costs in the model.

The complexity of this problem arises from keeping track of the supplier's belief process $\{G_t\}_{t \geq 1}$. However, under exponential demand distribution (as in the static case) significant simplification can be achieved. Suppose that the starting inventory in period t is given by $x_t + q_t(x_t)$ and demand is exponentially distributed. Then the starting inventory in period $t+1$, $[x_t + q_t(x_t) - D_t]^+$, with distribution G_{t+1} , is *weakly reverse exponential* (WRE), i.e.,

$$\frac{G_{t+1}(x_{t+1})}{g_{t+1}(x_{t+1})} \geq \frac{1}{\lambda}, \quad \text{for any belief distribution } G_t. \quad (12.17)$$

What is interesting about WRE property is that under exponential demand distribution, the supplier's belief in any period $t \geq 2$ satisfies WRE.¹² More remarkably, the WRE property of the G_t along with high holding cost and production cost relative to retail price ensures that $\frac{\delta J_t(x_t + q_t(x_t))}{\delta q_t} < 0$ for any $q_t > 0$ when $x_t > 0$ and a BOC is being used from period $t + 1$ onward.

Referring back to Equation (12.16), this implies that the optimal order in period t is $q_t = 0$ when $x_t > 0$, i.e., the optimal contract in period t is also BOC type.

The optimal batch-order contract can be simply characterized using first-order conditions. In addition to the analytical tractability, batch-order contracts are simpler to execute and monitor. Nevertheless, the batch-order contract and the inventory dynamics under the batch-order contract are unlike that under the base-stock policy, which is optimal for the supply chain with symmetric information (and zero fixed costs). However, with asymmetric information, the batch-order contract allows the supplier to nullify the retailer's informational advantage by only contracting with the retailer with zero on-hand inventory. As a result, the supplier is able to extract most of the channel profit leaving only the reservation profit for the retailer.

We conclude this section by noting that the general theme of the literature on dynamic contracting in inventory management is to identify *simple* contracts (and the underlying optimal inventory control) that are optimal or in some cases, near-optimal. This is especially important in the dynamic context since the optimal contract mechanism need not be unique and in general can be quite complex. The findings in this section, therefore, help us reconcile theory with practice: some of the dynamic contract forms typically observed in practice are in fact linear-price, quantity-discount along with base-stock and (s, S) -type inventory policies.

12.3.3 Cost Information Asymmetry

In this section, we consider inventory management problems with incomplete information pertaining to supply chain cost structures. In Section 12.3.3.1, we discuss Lutze and Özer (2008) who consider a supply chain setting in which the downstream retailer has superior information about a stationary shortage (penalty) cost. In particular, the authors propose and characterize long-term contracts to be offered by the supplier at the start of the planning horizon that facilitate information sharing and optimize the supplier's inventory/production costs. Section 12.3.3.2 we discuss Gao (2015) where the upstream supplier has private information about the non-stationary production cost structure. The focus of this paper is on long-term dynamic contracts which are designed and offered by the retailer to the supplier.

12.3.3.1 Inventory shortage cost information asymmetry

Lutze and Özer (2008) is one of the first papers to study dynamic inventory management under information asymmetry. They consider a two-level supply chain where the downstream retailer has private information about *shortage cost*, i.e., the unit cost of not satisfying customer demand. Both the retailer and the upstream supplier manage inventory to minimize operational costs at their respective locations. Motivated by practice, the paper considers a *promised lead-time* contract as a vehicle for sharing risk-associated demand uncertainty between the two supply chain firms.

The promised lead-time contract, designed by the supplier, specifies two parameters: the lead time τ promised by the supplier and a payment K made by the retailer to the supplier. This contract stipulates that the supplier would ship the retailer's order in its entirety, τ periods after

it is placed by the retailer. Once the order is shipped by the supplier, there is also a lead time ℓ for delivery. If the supplier does not have enough inventory on hand by the promised lead time, then he procures inventory from an alternative source to meet the retailer's order quantity. For a given promised lead time τ , the optimal base-stock levels for the retailer and the supplier, which balances the shortage and overage costs, are given as follows:

$$Y_s^*(\tau) := F_{L+1-\tau}^{-1} \left(\frac{p_s - (1-\alpha)c_s}{h_s + p_s} \right) \quad (12.18)$$

$$Y_r^*(p_r, \tau) := F_{\ell+1+\tau}^{-1} \left(\frac{p_r - (1-\alpha)c_r}{h_r + p_r} \right) \quad (12.19)$$

where p_j, c_j, h_j represent the unit penalty (shortage) cost, marginal procurement cost, and unit holding cost per period, respectively, for the supplier ($j=s$) and the retailer ($j=r$). Furthermore, L represents the supplier's procurement lead time and F is the cdf associated with end-customer demand. In Equation (12.18), the supplier's optimal base-stock level takes into account uncertainty in demand over the effective planning horizon $L+1-\tau$. Likewise, in Equation (12.19), the retailer's optimal base-stock level takes into account uncertainty in demand over the effective planning horizon $\ell+1+\tau$. As expected, it follows from Equation (12.19) that the retailer's optimal base-stock level Y_r^* increases in p_r .

The promised lead-time contract accomplishes two things: first, the retailer is guaranteed shipment of her entire order, thus eliminating the uncertainty on her supply side. Second, the supplier has a longer planning horizon (thereby, greater flexibility) to better plan and meet the retailer's order quantity. While both the firms have potential benefits from the arrangement, tension arises since the retailer prefers a smaller promised lead time τ , whereas the supplier would prefer to offer a larger τ . In an environment where the supplier knows the retailer's shortage cost p_r , the supplier can optimally design the promised lead time so that the retailer finds it in her best interest to accept the contract.

In the asymmetric information setting, which is the primary focus of Lutze and Özer (2008), the supplier does not know p_r . In this case, the above-mentioned tension renders any communication between the retailer and the supplier as uninformative cheap talk; see Proposition 3 of Lutze and Özer (2008). In particular, the supplier knows that the retailer's shortage cost is one of N values: $p_1, \dots, p_N$ with a prior distribution given by $\lambda_1, \dots, \lambda_N$, where $\lambda_i \geq 0$ for all i and $\sum_i \lambda_i = 1$. The supplier's problem is to design a menu of promised lead-time contracts which are offered at the start of the planning horizon. Due to the *revelation principle*, the supplier can restrict the search for the optimal menu contracts to the menu of contracts that facilitates truth-telling (Myerson, 1979). That is, the search for the optimal menu of contracts can be restricted to a menu of N promised lead-time contracts, such that the retailer chooses a contract from the menu that communicates her private truthfully and also maximizes her profit. The supplier's optimal contract design problem can be formulated as:

$$\min_{(\tau_i, K_i)_{i=1, \dots, N}} \sum_{i=1}^N \lambda_i (G_s^*(\tau_i) - K_i) \quad (12.20)$$

$$\text{s.t. } G_r^*(p_i, \tau_i) + K_i \leq U_r^{\max}, \forall i = 1, \dots, N \quad (12.21)$$

$$G_r^*(p_i, \tau_i) + K_i \leq G_r^*(p_i, \tau_j) + K_j, \forall j \neq i, \quad (12.22)$$

where $\tau_i \in \{0, \dots, L+1\}$ for all $i \in \{1, \dots, N\}$. The supplier's objective function in Equation (12.20) consists of the supplier's optimal inventory cost, $(G_s^*(\tau_i) - K_i)$, associated with promised lead-time contract (τ_i, K_i) offered to the retailer with shortage cost p_i . Equations (12.21) and (12.22) feature the retailer's participation and incentive-compatibility constraints under the truth-telling contract mechanism, respectively. In particular, $G_r^*(p_i, \tau_i) + K_i$ is the optimal inventory cost incurred by the retailer with shortage cost p_i if she accepts a promised lead-time contract (τ_i, K_i) from the menu offered by the supplier. We also highlight here that the analysis of mechanism design problem with finitely many retailer types is markedly different compared to the problem with a continuum of types; we refer the reader to Lovejoy (2006) for details.

Next, we discuss some of the important insights that can be gleaned from the optimal menu of contracts characterized in the paper. As noted above, a retailer with a higher shortage cost is offered a shorter promised lead time in return for a higher payment to the supplier. In fact, all the retailers (except the one with the *smallest* shortage cost) are offered a promised lead time that is *shorter* than their first-best promised lead time (i.e., the optimal promised lead time when the supplier knows the retailer's shortage cost). As a result, the supplier bears more risk associated with demand uncertainty when dealing with a retailer with a high shortage cost.

In addition to the contract design problem, the paper also considers the question of when should the supplier forgo working with a retailer? The supplier may want to keep inventory cost below a threshold, which may be due to a limited operational budget or an outside option of working with a different retailer (who presents a lower risk). From our above discussions note that, as the retailer's shortage cost increases, the supplier ends up bearing more of the inventory risk. By extending this intuition, Lutze and Özer (2008) characterize a *cutoff-type* policy, whereby the supplier only transacts with retailers whose shortage costs do not exceed a certain cutoff level.

12.3.3.2 Production cost information asymmetry

Gao (2015) considers a supply chain setting with product cost information asymmetry, by abstracting away from the specific supply side factors contributing to the information asymmetry. The upstream supplier has dynamic private information about the *supply state* which directly impacts his production cost structure. The downstream buyer (a traditional retailer) makes periodic inventory decisions and carries leftover inventory to meet demand. The supply chain setting considered in the paper assumes that the retailer has greater bargaining power compared to the supplier, and is therefore, modeled as the *principal* who designs and offers long-term dynamic contracts to the supplier (the *agent*). As is standard in related literature, these contracts are offered on a *take-it-or-leave-it* basis.

The supplier's supply state in period $t \geq 1$ denoted by z_t is drawn iid from a publicly known cdf $G(\cdot)$ and pdf $g(\cdot)$. The supply state information z_t is the supplier's private information. Thus, the supplier's private information is non-persistent in this setting. The supplier's marginal production cost $c(z_t)$ is convex decreasing in the supply state. The buyer places an order with the supplier and carries inventory in each period to satisfy end-consumer demand. Any unmet demand is lost. The retailer incurs a unit holding cost h per period and the unit revenue is given by r .¹³

At the start of time horizon, the retailer designs and offers a long-term contract $\{T_t(\hat{h}_t), q_t(\hat{h}_t)\}_{t \geq 1}$ to the supplier, where $\hat{h}_t$ denotes the supplier's report of the supply state until period t , i.e., $\hat{h}_t = (\hat{z}_1, \dots, \hat{z}_t)$. The terms of the contract evolve based on all the information available from prior periods. In each period t , the retailer orders $q_t(x_t, \hat{h}_t)$ raising the on-hand inventory level to $y_t = q_t + x_t$, where x_t is the starting inventory in period t and makes a payment $T_t(x_t, \hat{h}_t)$ to the supplier.¹⁴ As noted in our discussion in Section 12.3.1.2, the search for the optimal *dynamic* contracts can be restricted to contracts that induce truth-telling (Myerson, 1986). The caveat, however, is that truth-telling needs to be induced in each period t .

The supplier's cost function from period t onward $u_t(x_t, h_t, \hat{h}_t)$ depends on her report process $\hat{h}_t$ and her expected cost from period $t+1$ onward based on the report in period t . This cost function satisfies a dynamic programming recursion:

$$u_t(x_t, h_t, \hat{h}_t) = c(z_t)q_t(x_t, \hat{h}_t) - T_t(x_t, \hat{h}_t) + \gamma \mathbb{E}[u_{t+1}(x_{t+1}, h_{t+1})], \quad (12.23)$$

where $x_{t+1} = (y_t - D_t)^+$, demand $D_t, t \geq 1$ is iid drawn from pdf f and γ is discounting factor. The dynamic IC constraints can be represented as

$$u_t(x_t, h_t, h_t) \leq u_t(x_t, h_t, \hat{h}_t), \quad \forall t. \quad (12.24)$$

The supplier's total production cost under the optimal dynamic contract should be lower than what it would be without signing the contract. Thus, the IR constraints can be formulated as

$$u_t(x_t, h_t, h_t) \leq 0, \quad \forall t. \quad (12.25)$$

The IR constraints need to be satisfied in each period in a dynamic contract, else, the supplier may not persist with the dynamic contract for the entire time horizon. Given the above, the retailer's dynamic contracting problem can be formulated as follows:

$$\min_{(T_t, q_t)} \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t (h[y_t - D_t]^+ - r \min\{y_t, D_t\}) \right], \quad \text{subject to (12.24), (12.25).}$$

The solution approach is based on first finding the optimal solution to a relaxed problem in which the IC and IR constraints need to be satisfied only in the first period, i.e., a static contract design problem. In this relaxation, the supply state in all remaining periods is assumed to be publicly known. Gao (2015) shows that the optimal solution for the relaxed problem can be suitably modified to satisfy the above dynamic IC and IR constraints. In particular, the payment structure T_t can be displaced over time while still ensuring the IC and IR constraints are satisfied in each period. The buyer's cost under this modified dynamic contract is equal to that under the dynamic contract for the relaxed problem, thereby establishing optimality of the modified dynamic contract for the problem with dynamic IC and IR constraints.

The optimal dynamic contract has some of the features of the optimal (dynamic) long-term contract proposed in Lobel and Xiao (2017) for the lost-sales setting. The optimal policy is a state-dependent base-stock policy and the base-stock level that minimizes total supply chain

cost can be implemented after a finite number of periods. In Gao (2015), it can be implemented from the second period onward whereas it was the *first* stock-out event that triggered the switch to channel-efficient base-stock levels. It is, therefore, reassuring to know the robustness of base-stock policy in asymmetric information settings.

We conclude this section by noting that there is also related research that considers cost information asymmetry in a static setting; see, for example, Ha (2001) and Corbett and Tang (1999) for an extensive review.

12.3.4 Demand Information Asymmetry in Capacity Management

In this section, we consider capacity management problems with asymmetric demand forecast information, which evolves in a dynamic fashion. The problem of capacity management shares some similar features to inventory management, with the important difference that unused capacity cannot be carried over to future periods.

Ren et al. (2010) investigate if truthful demand forecast information sharing can emerge in a long-term relationship between an upstream supplier and a downstream retailer without using pricing levers. They investigate this question in the context of the supplier's multi-period capacity planning problem (similar to a multi-period newsvendor model). Demand in each period is modeled using a multiplicative form $D_t = \theta_t \cdot X$, where X is a non-negative normal random variable with a given mean and standard deviation, which are publicly known. The scaling factor $\theta_t \in \{\theta^l, \theta^h\}$ determines if demand for that period is forecasted to be low or high, with probability α and $1 - \alpha$, respectively.

The retailer *privately* observes forecast θ_t and decides what messages $m_t \in \{l, h\}$, indicating low or high demand forecast, to communicate to the supplier. Based on this message, the supplier decides to build capacity for the period K_t at a unit cost of c . Following that, the retailer *privately* observes demand realization D_t and places an order q_t with the supplier. The supplier produces and ships $\min\{q_t, K_t\}$ to the retailer. The retailer pays the supplier unit price r and sells in the market for unit price p . The supplier incurs a unit cost h for any leftover capacity in the period and the retailer incurs a unit cost g for any unmet demand. The single-period profit functions for the supplier and the retailer, respectively, are:

$$u_t = r \min\{K_t, q_t\} - h(K - q_t)^+ - cK_t$$

$$v_t = (p - r) \min\{K_t, q_t\} - g(D_t - q_t)^+.$$

The total profit expected profit of the supplier is $\mathbb{E}[\sum_{t \geq 1} \delta^{t-1} u_t]$ and that of the retailer is $\mathbb{E}[\sum_{t \geq 1} \delta^{t-1} v_t]$. In a single-period setting, the authors show that forecast information cannot be truthfully shared.

For a multi-period setting, the authors identify a *review strategy* that facilitates credible communication of the demand forecasts. The review strategy consists of a review phase of length R periods and a credibility assessment threshold. In each review phase, the supplier maintains a *score* for the retailer which is updated at end of each period based on the information provided (forecasts and order sizes) by the retailer. During the review period, the supplier trusts the retailer's messages $\{L, H\}$ to be true. The review strategy consists of statistical hypothesis tests (depending on whether the reported forecast is low or high) to check whether

the information shared by the retailer matches the long-run averages given by the primitives (μ, σ, α which are publicly known). That is, demand forecast should be high on an average $1 - \alpha$ fraction of the periods and the average demand corresponding to high demand forecast periods should be $\theta_h \mu$. If the retailer passes the statistical tests during the review phase (i.e., their score exceeds a credibility assessment threshold), then another review period commences. If the retailer fails statistical tests during the review phase, then the supplier punishes the retailer by discarding the messages for a certain number of periods, i.e., resorts to the single-period equilibrium of the game. Ren et al. (2010) show that for a sufficiently large: discount rate, review period length, and credibility assessment threshold; truth-telling can be sustained in equilibrium under the review strategy. The equilibrium emerges due to the threat of the uninformed firm “punishing” the informed firm in the future for inaccurate forecast reports.

The findings in this paper illustrate how a “scorecard” based system may also be used to elicit private information which may improve capacity/production planning in the supply chain (in addition to, for example, evaluating supplier effectiveness).

We conclude this section by highlighting related research that has considered incentive problems in multi-period capacity planning problems. Oh and Özer (2013) propose a general framework to model multiple evolutions of forecasts generated by multiple firms. Using this framework, they introduce the Martingale Model of Asymmetric Forecast Evolutions (MMAFE) and propose a mechanism for an upstream supplier to elicit a downstream retailer’s information credibly before making a single-shot capacity decision. There are two unique aspects to the mechanism design problem considered in this paper. First, due to dynamically evolving forecasts, they investigate *when* is the right time for the supplier to offer the mechanism (screening contracts) to the retailer. Second, in contrast to the other static/dynamic mechanism design problems considered above, the supplier builds capacity *even* if the retailer rejects the menu of contracts.¹⁵ Thus, the retailer’s decision to accept/reject the mechanism is explicitly handled in their paper. The resulting mechanism design problem not does admit standard solution methodology and we refer to Oh and Özer (2013) for further details.

Feng et al. (2015) study a model of dynamic interactions, in particular, a dynamic bargaining game between a buyer (with private demand information) and a seller that ensues prior to a one-time demand realization. In each round of negotiation, one of the firms moves to offer a contract, i.e., the informed firm offers a contract to signal type or the uninformed firm may offer contracts to screen the other firm. The negotiation continues until an agreement on quantity and payment for the trade of a product is reached. In the process, the contract offers are updated by each party, based on the outcomes of the previous negotiation stages.

Liu et al. (2019) consider an innovative multi-period agreement for sharing capacity (built by a single supplier firm) among multiple downstream manufacturing firms. The supplier makes a one-time capacity decision at the start of the time horizon, and in each period, the capacity has to be allocated among all the firms. All firms (the supplier and manufacturers) have private information about their demand. In addition, the supplier also has access to a spot market where it can choose to sell its capacity. The authors propose a unique multi-firm, multi-period agreement that is *budget-balanced*, i.e., the total payment received is equal to the total payment made in *each* period as a part of the agreement. In other words, the partnership is financially self-supporting. The proposed agreement not only ensures truthful sharing of private demand information but also that the supplier builds (*ex-ante*) efficient capacity and the capacity is allocated efficiently (*ex-post*) among all the firms.

12.4 CONCLUSION AND FUTURE DIRECTIONS

In this chapter, we have reviewed inventory management literature with special focus on information and incentives in dynamic settings. This literature, while still in its infancy, already illustrates a variety of problems that can be handled equipped with the machinery of dynamic mechanism design. There are several promising research directions at the confluence of information and incentives both in a traditional inventory management and emerging contexts. Below we highlight a few:

- *Information design approach.* In contrast to monetary transfer as a part of the contract mechanisms discussed in the chapter, non-price levers such as information may also be used by the informed firm to communicate its private information. This approach has been explored in the context of demand-side of operations but a similar approach may also be used in managing inventory within a supply chain.
- *Social and environmental considerations.* Recent research has shown the value and impact of incentives to motivate socially- and environmentally responsible operations from upstream firms; see, for example, Porteous et al. (2015); Kraft et al. (2020). Supply chain settings such as these are fraught with incomplete information (e.g., about sourcing methods) and unobservable upstream actions (e.g., labor practices). In these supply chain contexts, besides knowing the quantity of inventory, firms also need to know where the inventory is being sourced from. For example, the US government recently had announced a ban on imports of cotton and tomatoes from the Xinjiang area of China, including products made with those materials, due to human rights violations (Swanson, 2021). In that sense, it is not just information about inventory but also about the sourcing of materials that go into the product that matters. An important question in this context is what role does/can inventory management (under incomplete information or moral hazard) play in improving the social and environmental performance of the supply chain?
- *Empirical research.* The theoretical contract forms discussed in this chapter need to be reconciled with what is observed in practice. In particular, examining the performance of long-term contracts based on field data may validate the theoretical insights as well as provide opportunities for future research.

NOTES

1. Other examples of failed VMI partnerships include Spartan Stores, Inc., a grocery chain, halted its VMI programs after 12 months of operations, blaming part of the failure on the fact that the VMI vendors had not taken promotions data into account (Mathews, 1995). Furthermore, K-Mart cut VMI contracts from more than 300 to about 50, blaming poor performance (partially) on the fact that the suppliers did not have adequate forecasting skills (Fisher, 1997).
2. Given the focus of the chapter is on dynamic inventory models, we refer the reader to Cachon (2003) and Chen (2003) for an extensive treatment of static incomplete information settings.
3. Further, larger ξ represents larger average demand, i.e., $G(\cdot | \xi_1) \geq G(\cdot | \xi_2)$ if $\xi_1 \leq \xi_2$.
4. Throughout the chapter, we will refer to the downstream firm, who is typically the retailer, as “she” and the upstream supplier/manufacturer as “he”.
5. Equivalently, a payment can be charged on a per-period basis.
6. We also refer the reader to Zhang and Zenios (2008) for a related long-term contracting model.

7. We refer the reader to Kadiyala et al. (2021) for an extensive discussion of the impact of the retailer's outside option (due to an alternative sourcing option) on the contract design problem in a single-period setting.
8. This assumption implies that the supplier knows the retailer's inventory level prior to designing the current menu of contracts. Technically, this assumption ensures that the $\frac{G_1(x)}{g_1(x)} = \lambda$, which ensures that the first derivative of the objective function is always negative for $x > 0$. Lacking this assumption, the structure of the contract has an additional segment wherein the retailer truthfully reveals her private information and the optimal quantity is based on their type.
9. Anand et al. (2008) show how the retailer may distort her actions (order quantity) in a period to obtain better wholesale price in the future periods.
10. With a slight abuse of notation we use $v(\cdot)$ to the single-period profit function which includes the holding cost.
11. IC constraints are rewritten in terms of local conditions using the *envelope theorem* and substituted back in to the objective function to obtain J_r .
12. The distribution G_1 also has WRE property if, for example, the starting inventory x_1 is carried over from an earlier period $[y_0 - D_0]^+$, with D_0 also exponentially distributed.
13. The paper also discusses the case in which the retailer incurs a fixed cost in each period for placing an order with the supplier.
14. It is to be noted that the buyer's inventory level x_t is also observable to the supplier.
15. There is no possibility of credible information sharing in the event the retailer rejects the menu of contracts. Nevertheless, the supplier can still make (coarse) inference about the retailer's type.

REFERENCES

- Allon, G., & Bassamboo, A. (2011). Buying from the babbling retailer? The impact of availability information on customer behavior. *Management Science*, 57(4), 713–726.
- Anand, K., Anupindi, R., & Bassok, Y. (2008). Strategic inventories in vertical contracts. *Management Science*, 54(10), 1792–1804.
- Atali, A., Lee, H., & Özer, Ö. (2009). If the inventory manager knew: Value of visibility and RFID under imperfect inventory information. Available on SSRN, The University of Texas at Dallas.
- Aviv, Y. (2001). The effect of collaborative forecasting on supply chain performance. *Management Science*, 47(10), 1326–1343.
- Aviv, Y. (2002). Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand. *Manufacturing and Service Operations Management*, 4(1), 55–74.
- Aviv, Y. (2007). On the benefits of collaborative forecasting partnerships between retailers and manufacturers. *Management Science*, 53(5), 777–794.
- Ban, G.-Y. (2020). Confidence intervals for data-driven inventory policies with demand censoring. *Operations Research*, 68(2), 309–326.
- Ban, G.-Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1), 90–108.
- Bassamboo, A., Moreno, A., & Stamatopoulos, I. (2020). Inventory auditing and replenishment using point-of-sales data. *Production and Operations Management*, 29(5), 1219–1231.
- Bisi, A., Dada, M., & Tokdar, S. (2011). A censored-data multiperiod inventory problem with newsvendor demand distributions. *Manufacturing and Service Operations Management*, 13(4), 525–533.
- Bray, R. L., & Mendelson, H. (2012). Information transmission and the bullwhip effect: An empirical investigation. *Management Science*, 58(5), 860–875.
- Brinkhoff, A., Özer, Ö., & Sargut, G. (2015). All you need is trust? An examination of inter-organizational supply chain projects. *Production and Operations Management*, 24(2), 181–200.
- Cachon, G. P. (2003). Supply chain coordination with contracts. In S. Graves & T. de Kok (Eds.), *Supply Chain Management: Design, Coordination and Operation*. Elsevier.

- Cachon, G. P., Randall, T., & Schmidt, G. M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9(4), 457–479.
- Chen, F. (2003). Information sharing and supply chain coordination. In S. Graves & T. de Kok (Eds.), *Supply Chain Management: Design, Coordination and Operation*. Elsevier.
- Chen, L. (2010). Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Operations Research*, 58(2), 396–413.
- Chen, L. (2021). Fixing phantom stockouts: Optimal data-driven shelf inspection policies. *Production and Operations Management* [Forthcoming].
- Chen, L., & Lee, H. L. (2009). Information sharing and order variability control under a generalized demand model. *Management Science*, 55(5), 781–797.
- Chen, L., & Lee, H. L. (2017). *Modeling and measuring the bullwhip effect* (pp. 3–25). Springer International Publishing.
- Chen, L., & Mersereau, A. J. (2015). *Analytics for operational visibility in the retail store: The cases of censored demand and inventory record inaccuracy* (pp. 79–112). Springer.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- Corbett, C. J., & Tang, C. S. (1999). *Designing supply contracts: Contract type and information asymmetry* (pp. 269–297). Springer.
- Cui, R., & Shin, H. (2018). Sharing aggregate inventory information with customers: Strategic cross-selling and shortage reduction. *Management Science*, 64(1), 381–400.
- DeHoratius, N., Mersereau, A. J., & Schrage, L. (2008). Retail inventory management when records are inaccurate. *Manufacturing and Service Operations Management*, 10(2), 257–277.
- Drakopoulos, K., Jain, S., & Randhawa, R. (2021). Persuading customers to buy early: The value of personalized information provisioning. *Management Science*, 67(2), 828–853.
- Eső, P., & Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. *Review of Economic Studies*, 74(3), 705–731.
- Feng, Q., Lai, G., & Lu, L. X. (2015). Dynamic bargaining in a supply chain with asymmetric demand information. *Management Science*, 61(2), 301–315.
- Fisher, M. L. (1997). What is the right supply chain for your product? *Harvard Business Review*, 75(March–April).
- Fudenberg, D., & Tirole, J. (1991). Perfect Bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53(2), 236–260.
- Gallego, G., & Özer, Ö. (2001). Optimal use of demand information in supply chain management. In J. Song & D. Yano (Eds.), *Supply chain structures*, chap. 5. (pp. 119–160) Boston: Springer.
- Gallego, G., & Özer, Ö. (2002). *Optimal use of demand information in supply chain management* (pp. 119–160). Springer.
- Gao, L. (2015). Long-term contracting: The role of private information in dynamic supply risk management. *Production and Operations Management*, 24(10), 1570–1579.
- Gavirneni, S., Kapuściński, R., & Tayur, S. (1999). Value of information in capacitated supply chains. *Management Science*, 45(1), 16–24.
- Ha, A. Y. (2001). Supplier-buyer contracting: Asymmetric cost information and cutoff level policy for buyer participation. *Naval Research Logistics (NRL)*, 48(1), 41–64.
- Hammond, J. H. (2006). Barilla SpA (D). Harvard business school case 9-695-066. Harvard University.
- Hardgrave, B. C., Aloysius, J. A., & Goyal, S. (2013). Rfid-enabled visibility and retail inventory record inaccuracy: Experiments in the field. *Production and Operations Management*, 22(4), 843–856.
- Heath, D. C., & Jackson, P. L. (1994). Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. *IIE Transactions*, 26(3), 17–30.
- Kadiyala, B., Özer, Ö., & Bensoussan, A. (2020). A mechanism design approach to vendor managed inventory. *Management Science*, 66(6), 2628–2652.
- Kadiyala, B., Özer, Ö., & Oh, S. (2021). Sourcing and information sharing under disintermediation [Working paper]. University of Texas at Dallas.
- Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6), 2590–2615.

- Kapuściński, R., & Parker, R. (2021). Conveying demand information in serial supply chains with capacity limits [Working paper]. University of Michigan.
- Kök, A. G., & Shang, K. H. (2007). Inspection and replenishment policies for systems with inventory record inaccuracy. *Manufacturing and Service Operations Management*, 9(2), 185–205.
- Kouvelis, P., Chambers, C., & Wang, H. (2006). Supply chain management research and production and operations management review, trends, and opportunities. *Production and Operations Management*, 15(3), 449–469.
- Kraft, T., Valdés, L., & Zheng, Y. (2020). Motivating supplier social responsibility under incomplete visibility. *Manufacturing and Service Operations Management*, 22(6), 1268–1286.
- Kremer, I., Mansour, Y., & Perry, M. (2014). Implementing the “wisdom of the crowd”. *Journal of Political Economy*, 122(5), 988–1012.
- Küçükgül, C., Özer, Ö., & Wang, S. (2021). Engineering social learning: Information design of time-locked sales campaigns for online platforms. *Management Science* [Forthcoming].
- Kurtuluş, M. (2017). *Collaborative forecasting in retail supply chains* (pp. 39–61). Springer International Publishing.
- Lee, H., & Özer, Ö. (2007). Unlocking the value of RFID. *Production and Operations Management*, 16(1), 40–64.
- Lee, H., & Whang, S. (1999). Decentralized multi-echelon supply chains: Incentives and information. *Management Science*, 45(5), 633–640.
- Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The Bullwhip effect. *Management Science*, 43(4), 546–558.
- Lee, H. L., So, K. C., & Tang, C. S. (2000). The value of information sharing in a two-level supply chain. *Management Science*, 46(5), 626–643.
- Liu, F., Lewis, T. R., Song, J.-S., & Kuribko, N. (2019). Long-term partnership for achieving efficient capacity allocation. *Operations Research*, 67(4), 984–1001.
- Lobel, I., & Xiao, W. (2017). Technical note—Optimal long-term supply contracts with asymmetric demand information. *Operations Research*, 65(5), 1275–1284.
- Lovejoy, W. S. (2006). Optimal mechanisms with finite agent types. *Management Science*, 52(5), 788–803.
- Lutze, H., & Özer, Ö. (2008). Promised lead-time contracts under asymmetric information. *Operations Research*, 56(4), 898–915.
- Mathews, R. (1995). Spartan pulls the plug on VMI. *Progressive Grocer*, 74(11), 64–64.
- Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica: Journal of the Econometric Society*, 61–73.
- Myerson, R. B. (1986). Multistage games with communication. *Econometrica*, 54(2), 323–358.
- Oh, S., & Özer, Ö. (2013). Mechanism design for capacity planning under dynamic evolutions of asymmetric demand forecasts. *Management Science*, 59(4), 987–1007.
- Özer, Ö. (2011). *Inventory management: Information, coordination, and rationality* (pp. 321–365). Springer.
- Pavan, A., Segal, I., & Toikka, J. (2014). Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2), 601–653.
- Porteous, A. H., Rammohan, S. V., & Lee, H. L. (2015). Carrots or sticks? Improving social and environmental compliance at suppliers through incentives and penalties. *Production and Operations Management*, 24(9), 1402–1413.
- Raman, A., DeHoratius, N., & Ton, Z. (2001). Execution: The missing link in retail operations. *California Management Review*, 43(3), 136–152.
- Ren, Z. J., Cohen, M. A., Ho, T. H., & Terwiesch, C. (2010). Information sharing in a long-term supply chain relationship: The role of customer review strategy. *Operations Research*, 58(1), 81–93.
- Swanson, A. (2021). U.S. Bans all cotton and tomatoes from Xinjiang region of China. *The New York Times*. <https://www.nytimes.com/2021/01/13/business/economy/xinjiang-cotton-tomato-ban.html>
- Ton, Z., & Raman, A. (2010). The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, 19(5), 546–560.

- Veinott, A. F. (1966). The status of mathematical inventory theory. *Management Science*, 12(11), 745–777.
- Zhang, H., Nagarajan, M., & Sošić, G. (2010). Dynamic supplier contracts under asymmetric inventory information. *Operations Research*, 58(5), 1380–1397.
- Zhang, H., & Zenios, S. (2008). A dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation. *Operations Research*, 56(3), 681–696.

13. Joint pricing and inventory decisions

Xin Chen, Peng Hu, and Zhenyu Hu

13.1 INTRODUCTION

Matching supply and demand in supply chains is at the core of our economy. Many enterprises ranging from large companies such as Amazon and Walmart to small brick-and-mortar stores rely on efficient supply and demand matching for their survival and success. Recognizing the importance, companies spend billions of dollars on talents and technologies to improve its efficiency. Despite their effort, it remains a daunting challenge in practice. On the one hand, supply processes can be quite complicated: complex supply chain structures; supply disruptions and uncertainties; long lead times; economy of scale and scope; etc. On the other hand, demand can be highly variable and uncertain.

Facing these challenges, academic research has a long history of developing effective inventory management strategies. For example, the optimality of base-stock policies or (s, S) policies is now well understood in various stochastic inventory models (see, for instance, Chapter 9 in Simchi-Levi et al., 2014; Zipkin, 2000). More recently, pricing has become another important and effective lever to shape demand in response to real-time supply processes. In a basic multi-period model with linear ordering cost, Federgruen and Heching (1999) demonstrate the significant benefit of making joint inventory and pricing decisions and prove that a base-stock list-price policy is optimal. When a fixed ordering cost is present, a (s, S) policy is shown to be optimal for replenishment under certain conditions and price is a function of the inventory position (see, for instance, Chen & Simchi-Levi, 2004a, 2004b). More recent research has developed and analyzed models with more complex supply processes and demand processes, which we set to survey in this chapter.

There are several notable prior survey papers on quantitative models of joint pricing and inventory decisions. Among these papers, Eliashberg and Steinberg (1993) review the literature up to the year 1991 on the interface of operations and marketing with an emphasis on integrated inventory and pricing models; Elmaghraby and Keskinocak (2003) focus on a few key papers on dynamic pricing in the presence of inventory considerations; Chan et al. (2004) provide a comprehensive review of coordinated pricing and inventory models including markdown and clearance pricing; Yano and Gilbert (2003) focus on papers in which inventory replenishment is critical and comprehensively survey EOQ-type models and deterministic models emphasizing demand smoothing (as a result of convex production cost). The three papers, Elmaghraby and Keskinocak (2003), Chan et al. (2004), and Yano and Gilbert (2003), cover the related literature up to the year 2004. Chen and Simchi-Levi (2012) provide a survey covering many papers which appear after the publication of Elmaghraby and Keskinocak (2003), Chan et al. (2004), and Yano and Gilbert (2003) and up to the year 2012, and include strategic models on supply chain competition, coordination and cooperation built upon operational and tactical inventory and pricing models. Chen and Chen (2015) survey dynamic pricing models in which inventory is not considered or inventory is given or ordered only at the beginning of the planning horizon.

The purpose of this chapter is to provide an up-to-date survey of the academic research on the single-product joint pricing and inventory dynamic models assuming known uncertainty distributions. We primarily cover relevant papers appearing after Chen and Simchi-Levi (2012). Similar to Chen and Simchi-Levi (2012), our intention is not to provide a comprehensive collection of all relevant papers but to focus on a few selected papers to highlight key up-to-date developments. There is active research dealing with joint pricing and inventory models in which the distributions of relevant uncertainties are unknown and have to be learned, for which we refer readers to Chapter 12 in this book.

The survey is divided into three parts. The first part, Section 13.2, introduces the basic multi-period inventory and pricing model along the line of Federgruen and Heching (1999) and reviews some recent papers that generalize the model to incorporate more general demand or supply functions. A major focus is on deriving conditions under which simple base-stock type policies are optimal. The second part, Section 13.3, presents models incorporating consumer behaviors to capture inter-temporal effects on demand. We cover models with reference price effects and models with strategic consumers. For these models, simple control policies such as a reference price-dependent base-stock list-price policy and a bang-bang policy are shown to be optimal. The third part, Section 13.4, deals with models with inter-temporal effects on supply, and covers models with replenishment lead times and models with perishable products with finite lifetimes. The literature has established certain monotone properties of the optimal policy with limited sensitivities using L^1 -convexity, a key concept from discrete convex analysis, and developed several heuristics to deal with replenishment lead times. Finally, some concluding remarks and thoughts on future research are provided in Section 13.5.

13.2 BASIC MODELS

Consider a firm that makes pricing and inventory replenishment decisions over a planning horizon of T periods. At the beginning of each period t , the firm reviews its current inventory level, denoted as x , and decides the price $p \in [\underline{p}, \bar{p}]$ and the ordering quantity q that replenishes the inventory to the level y (the ordering quantity then satisfies $q = y - x$). In the basic model here, order lead time is assumed to be zero. Hence, the order placed arrives immediately and incurs an ordering cost $C(y - x)$. The cost function $C(\cdot)$ can be general, however, for the most part of this chapter we will assume a linear ordering cost, i.e., $C(q) = cq$ with c being the unit ordering cost.

The demand $D(p, \xi)$ in period t is assumed to be a function of both the price p and a random noise ξ , with the random noise being independent and identically distributed (i.i.d.) across periods and realized after the pricing and inventory decisions are made. One commonly used functional form of the demand is

$$D(p, \xi) = \zeta d(p) + \varepsilon, \quad (13.1)$$

where $\xi := (\zeta, \varepsilon)$ and $\mathbb{E}[\zeta] = 1, \mathbb{E}[\varepsilon] = 0$. The above form is referred to as *additive demand* when $\zeta \equiv 1$ and *multiplicative demand* when $\varepsilon \equiv 0$. The mean demand function $d(p)$ can be derived from the reservation-price model in consumer choice theory. In the reservation-price model, a representative customer has a private valuation (or reservation price) v for the firm's product; the firm knows only that v follows a distribution $F(\cdot)$. The customer would

purchase the product only if the net surplus $v - p \geq 0$, which leads to a purchase probability of $\bar{F}(p) := 1 - F(p)$.¹ Assuming a total market size of N , then the mean demand can be computed as $d(p) = N\bar{F}(p)$.² For example, when v is uniformly distributed on $[0, \bar{v}]$, we have the linear demand: $d(p) = N(1 - p/\bar{v})$; when v is exponentially distributed with mean μ_v , we have the exponential demand: $d(p) = Ne^{-p/\mu_v}$; when v follows logistic distribution with mean μ_v and unit scale parameter, we have the logit demand: $d(p) = N \frac{e^{\mu_v - p}}{1 + e^{\mu_v - p}}$. We refer readers

to Chapter 7 in Talluri and van Ryzin (2004) for a more comprehensive account of consumer choice theory and aggregate demand models.

By assuming $F(\cdot)$ to be strictly increasing on its support, one can ensure the mean demand function $d(p)$ to be strictly decreasing in price p and define the inverse demand function $p(d)$ on $[\underline{d}, \bar{d}]$, where $\underline{d} = d(\bar{p})$, $\bar{d} = d(\underline{p})$. Sometimes it is more convenient to view the expected demand level d instead of the price p as the decision variable so that the function form in Equation (13.1) becomes linear in d . Correspondingly, given the expected demand d and inventory level y , we can define the expected revenue $\Pi(d, y)$ collected from fulfilling the random demand. When the unfulfilled demand—paying the current price $p(d)$ —can be backlogged, the expected revenue is simply $\Pi(d, y) = p(d)d$. On the other hand, when the unfilled demand is lost, the expected revenue is $\Pi(d, y) = p(d)\mathbb{E}[D(p(d), \xi) \wedge y]$. Here, $x \wedge y := \min\{x, y\}$. After the fulfillment of the demand, a holding or shortage penalty cost

$$H(y - D(p(d), \xi)) := h(y - D(p(d), \xi))^+ + b(D(p(d), \xi) - y)^+$$

is incurred, where $x^+ := \max\{x, 0\}$, and h and b are the unit holding and shortage penalty costs respectively. The inventory level carried over to the next period, denoted as $\tilde{x}$, is then $\tilde{x} = y - D(p(d), \xi)$ in the backlogging case and $\tilde{x} = (y - D(p(d), \xi))^+$ in the lost-sales case.

Let $V_t(x)$ be the profit-to-go function when the inventory level is x at the beginning of period t . We can write the Bellman equation for $t = 1, \dots, T$ as

$$V_t(x) = \max_{y \geq x, d \in [\underline{d}, \bar{d}]} \left\{ \Pi(d, y) - C(y - x) - \mathbb{E}[H(y - D(p(d), \xi))] + \gamma \mathbb{E}[V_{t+1}(\tilde{x})] \right\}, \quad (13.2)$$

with the terminal condition $V_{T+1}(x) = 0$.³ In Equation (13.2), $\gamma \in [0, 1]$ is the discount factor, and $\Pi(d, y)$ and $\tilde{x}$ are defined above depending on whether a backlogging model or a lost-sales model is assumed. We let $y^*(x)$ and $d^*(x)$ denote the optimal order-up-to level and expected demand-level decisions.

The characterization of the optimal solution to Equation (13.2) critically depends on the concavity of $\Pi(d, y)$. Consider the backlogging case with linear ordering cost and demand following the form of Equation (13.1). In this case, $\Pi(d, y) = p(d)d$ is independent of the inventory level y , and its concavity in d can be guaranteed by assuming valuation v has an increasing hazard rate, i.e., $f(p)/\bar{F}(p)$ is increasing. This assumption covers a wide range of valuation distributions including the uniform, exponential and logistic distributions mentioned above.⁴ It follows by induction that $V_t(x)$ is also concave for $t = 1, \dots, T$. To see this, note that because under the induction hypothesis that $V_{t+1}(\cdot)$ is concave, $\mathbb{E}[V_{t+1}(y - \zeta d - \varepsilon)]$ is then jointly concave in y and d . With linear ordering cost, both the ordering cost $c(y - x)$ and the

inventory-related cost $\mathbb{E}[H(y - \zeta d - \varepsilon)]$ are jointly convex in (x, y, d) . Hence, Equation (13.2) has a concave objective with linear constraints and the concavity is preserved under maximization, i.e., $V_t(x)$ is also concave. In addition, by ignoring the constraint $y \geq x$ in Equation (13.2), we can obtain an optimal solution to the relaxed problem, denoted as (s, d_s) , that is independent of x . The concavity of the objective function then ensures that

$$y^*(x) = \begin{cases} s, & \text{if } x \leq s, \\ x, & \text{if } x > s, \end{cases} \text{ and } d^*(x) = \begin{cases} d_s, & \text{if } x \leq s, \\ \tilde{d}(x), & \text{if } x > s, \end{cases}$$

where

$$\tilde{d}(x) = \operatorname{argmax}_{d \in [\underline{d}, \bar{d}]} \{ \Pi(d, x) - \mathbb{E}[H(x - \zeta d - \varepsilon)] + \gamma \mathbb{E}[V_{t+1}(x - \zeta d - \varepsilon)] \}.$$

Furthermore, one can establish that the objective function in Equation (13.2) is supermodular in (x, y, d) with the feasible region being a lattice. Therefore, both $y^*(x)$ and $d^*(x)$ are increasing in x (see, for instance, Theorem 2.2.8 in Simchi-Levi et al., 2014). The constant s (which could be different from period to period) is referred to as the base-stock level. When the inventory is below the base-stock level, it is optimal to order up to the base-stock level and charge a “regular” price $p(d_s)$. When the inventory is above the base-stock level, it is optimal not to order anything and charge a discounted price $p(\tilde{d}(x))$ —the higher x is, the deeper the discount. This structure of the optimal policy is known as the *base-stock list-price* policy.

A significant amount of literature on joint pricing and inventory management is devoted to generalizing the above flow of arguments to more complicated models. A major class of models studies a non-linear ordering cost function $C(\cdot)$, in which case the base-stock list-price policy may fail to be optimal. In particular, when $C(\cdot)$ involves a fixed cost and a linear variable cost, then an (s, S, p) policy, in which inventory is raised to the order-up-to level S only if it drops below the reorder point s , can be established to be optimal under certain conditions (see Chen and Simchi-Levi, 2004a, 2004b). We refer readers to Chen and Simchi-Levi (2012) for a comprehensive survey of many of these papers.

Another stream considers the lost-sales model, where the challenge mainly comes from the fact that the expected single-period revenue $\mathbb{E}[(pD(p, \xi)) \wedge (py)]$ and the expected future profit $\mathbb{E}[V_{t+1}((y - D(p, \xi))^+)]$ are possibly non-concave, even if $pD(p, \xi)$ and $V_{t+1}(y)$ are concave. A common approach in the literature is to identify sufficient conditions to ensure the tractability of such a model. Kocabiykoglu and Popescu (2011) and Lu and Simchi-Levi (2013) focus on the expected single-period revenue. For dynamic models, Huh and Janakiraman (2008) consider a stationary system with a fixed ordering cost, and identify sufficient conditions for the optimality of (s, S) -type policies by using sample-path arguments. Feng et al. (2020), on the other hand, notice that the term py is jointly concave on $\mathcal{A}$ if any two vectors (p_1, y_1) and (p_2, y_2) in $\mathcal{A}$ are not strictly unordered, i.e., $(p_1 - p_2)(y_1 - y_2) \leq 0$. Thus, by restricting the attention to pricing policies that are only decreasing in the inventory level y , Feng et al. (2020) establish concavity of the expected single-period revenue for linear ordering costs. In addition, they apply the property of *stochastic concavity in midpoint* to ensure the objective function is concave along these decreasing price paths. Consequently, they show that the base-stock ordering policy is optimal among the class of policies with the price decreasing in the order-up-to inventory level.

Before we dive into the more specific topics on modeling the inter-temporal effects on either the demand or supply side, we briefly mention some of the recent progress—appearing after Chen and Simchi-Levi (2012)—that seeks to either generalize the demand or supply side of the problem in Equation (13.2) without any inter-temporal consideration. Feng et al. (2014) use the functional form $D(p, \varepsilon) = d(p) + \sigma(p)\varepsilon$ as opposed to Equation (13.1) to model the stochastic demand, where ε has zero mean and unit variance. One advantage of this model is that it allows more flexibility in modeling of how price affects the mean and variance of the demand. They identify conditions that generalize the one used in Federgruen and Heching (1999) to ensure the optimality of base-stock list-price policy. Restricting to additive demand and certain noise distributions, Chen and Zhang (2014) manage to characterize the optimal policy by imposing the weaker property of quasi-concavity of the expected revenue $pd(p)$ (and also the holding and shortage penalty cost function $H(\cdot)$) when fixed cost is present. Shen et al. (2018) consider the general functional form $D(p, \xi)$ and present conditions based on stochastic order and monotone property in ξ that guarantee the optimality of base-stock list-price policy while Bensoussan et al. (2019) propose an alternative approach based on proving concavity of the price-optimized profit function. Besides generalizing the functional form of the demand, inventory-dependent demand is considered by Yang and Zhang (2014). In addition to price, a firm can withhold or dispose of its on-hand inventory to shape demand, and they characterize the optimal ordering, pricing and reallocation/disposal policy. Lu et al. (2014) study a quantity-based pricing problem where a customer is assumed to purchase either a unit or a bundle of multiple units at different prices. They show that the optimal ordering policy remains a base-stock policy and characterize conditions under which a price discrimination based on quantity is deployed.

For the supply side, Li and Zheng (2006) and Feng (2010) incorporate supply uncertainty by considering random yield and random capacity, respectively. Gong et al. (2014) further study joint pricing and inventory control in the context of dual sourcing with both suppliers subject to random disruption. To address the non-convexity induced by the random capacity, Chen et al. (2018) and Chen and Gao (2019) develop a transformation technique that converts the nonconvex minimization problem to an equivalent convex minimization problem (note that while neither Chen et al. (2018) nor Chen and Gao (2019) analyze joint inventory and pricing problems, the developed methodologies can be applicable).

Feng and Shanthikumar (2018) provide a general perspective on modeling random demand as well as possibly random supply in Equation (13.2). Instead of parameterizing the random demand via the random noise ξ , they model demand as a stochastic function $\{\tilde{D}(p), p \in [\underline{p}, \bar{p}]\}$ of the price, where $\tilde{D}(p)$ is a random variable whose distribution is parameterized by p . Similarly they model the random supply as a stochastic function of order quantity $\{\tilde{S}(q), q \geq 0\}$, which includes random capacity and random yield models as special cases. They establish conditions directly on the stochastic functions $\tilde{D}(p)$ and $\tilde{S}(q)$ using the notion of *stochastic linearity in midpoint* to establish the concavity of objective in Equation (13.2) for the backlogging and linear ordering cost case.

13.3 CONSUMER-BEHAVIOR MODELS

From the reservation-price model introduced in Section 13.2, one can see that consumers' purchasing decisions (and hence the mean demand) in one period depend only on the price

in that period and is independent of the price(s) in all the other periods. In many practical scenarios—especially when consumers have repeated interactions with the product, however, consumers may either take into account prices seen in the past or strategize on timing the purchase in anticipation of a future discount. Such behaviors are evidenced by both our own anecdotal shopping experiences as well as empirical studies. We introduce in this section two different consumer-behavior models that result in inter-temporal dependence on demands and discuss their impact on optimal pricing and inventory decisions.

13.3.1 Reference Price Effect

The concept of a reference price stems from the principle of reference dependence in the well-known prospect theory (Kahneman & Tversky, 1979). It argues that consumers form price expectations from historically observed prices, which serves as a reference point against which a selling price is evaluated. A purchasing instance is then classified as gain or loss depending on whether the selling price is below or above the reference price. According to prospect theory, consumers react more strongly to losses than to gains of the same magnitude—a phenomenon known as *loss aversion*.

To model a reference price effect, we build on the reservation-price model introduced in Section 13.2. A consumer with valuation v who faces a selling price p acquires a direct *economic surplus*: $v - p$ from purchasing or consuming the product. In addition, a consumer with reference price r obtains a *gain/loss surplus* $\lambda(r - p)$, where $\lambda(z) = \lambda_+ z^+ - \lambda_- (-z)^+$. Here, $\lambda_+, \lambda_- \geq 0$ and loss aversion postulated by prospect theory is captured by $\lambda_- > \lambda_+$.⁵ The consumer's overall utility from purchasing the product is then $v - p + \lambda(r - p)$. Similar to the reservation-price model, consumers purchase the product if the utility from purchasing is non-negative and hence giving rise to a purchase probability $\bar{F}(p - \lambda(r - p))$ and mean demand function $d(r, p) = N\bar{F}(p - \lambda(r - p))$. When v is uniformly distributed on $[0, v]$, for instance, we have the piece-wise linear demand: $d(r, p) = N(1 - p/v + \lambda(r - p)/v)$ that is widely used in both empirical (Greenleaf, 1995; Natter et al., 2007) and analytical works (Kopalle et al., 1996; Fibich et al., 2003; Chen et al., 2016, 2017).

The *reference price effect* on demand is defined as $R(z, p) := d(p + z, p) - d(p, p) = N(F(p) - F(p - m(z)))$, which measures the change in demand at price p due to discount or surcharge z . The demand is said to be *more responsive to losses (gains)* if $R(z, p) < -R(-z, p)$ ($R(z, p) > -R(-z, p)$) for all $z > 0$. As we shall see, such asymmetry in demand response is critical in determining the structure of the optimal dynamic pricing strategy. Contrary to what one might expect, there is no general relationship between the demand's responsiveness to losses and the consumer's loss aversion. Hu and Nasiry (2018) find that the valuation distribution $F(\cdot)$ plays a decisive role in determining the direction of demand response: the demand tends to be more responsive to losses (gains) if $F(\cdot)$ is convex (concave). In particular, if one seeks to approximate the demand function via the following piece-wise linear function

$$d(r, p) = b - ap + \eta_+(r - p)^+ - \eta_-(p - r)^+, \quad (13.3)$$

where $a, b, \eta_+, \eta_- \geq 0$, then it is possible that $\eta_+ > \eta_-$, i.e., demand being more responsive to gains, even though consumers may be loss averse.⁶ Such an observation is also validated by the empirical result of Greenleaf (1995) who argues that “this market level result [demand being

more responsive to gains] need not contradict previous results that the reverse relationship holds for purchases at the household level” (p. 90).

As in Chen et al. (2016), we assume demand can be backlogged and follows additive form, i.e., $D(r, p, \varepsilon) = d(r, p) + \varepsilon$. The ordering cost is assumed to be linear and for simplicity is normalized to zero: $c=0$ (see footnote 3). The reference price r_t in period t is generated from an exponentially smoothed adaptive expectations process (see, e.g., Mazumdar et al., 2005):

$$r_t = \alpha r_{t-1} + (1 - \alpha)p_{t-1}$$

with $\alpha \in [0, 1)$ and $t = 2, \dots, T$. In other words, a consumer is assumed to form the reference price by taking a weighted average of the selling price and reference price encountered in the last period.⁷ The parameter α is referred to as the *memory factor*, and as α increases a consumer adapts more slowly to new price information.

Let $V_t(x, r)$ be the profit-to-go function when the inventory level is x and the reference price is r at the beginning of period t . The Bellman equation in this case is given by

$$\begin{aligned} V_t(x, r) = \max_{y \geq x, p \in [\underline{p}, \bar{p}]} \{ & pd(r, p) - \mathbb{E}[H(y - d(r, p) - \varepsilon)] \\ & + \gamma \mathbb{E}[V_{t+1}(y - d(r, p) - \varepsilon, \alpha r + (1 - \alpha)p)] \}, \end{aligned} \quad (13.4)$$

with $V_{T+1}(\cdot, \cdot) = 0$. Note that compared to Equation (13.2), the reference price effect results in an additional state variable r . When $R(r - p, p) \equiv 0$, $V_t(x, r)$ is independent of r and hence Equation (13.4) reduces to Equation (13.2). Another important special case of Equation (13.2) is when $\varepsilon \equiv 0$ and the initial inventory level is zero. In this case, $V_t(x, r)$ is independent of x and we have a pure pricing problem:

$$\tilde{V}_t(r) = \max_{p \in [\underline{p}, \bar{p}]} \{ pd(r, p) + \gamma \tilde{V}_{t+1}(\alpha r + (1 - \alpha)p) \}, \quad (13.5)$$

with $\tilde{V}_{T+1}(\cdot) = 0$. It can be shown inductively that $\tilde{V}_t(r) - \min_{\tilde{y}} \mathbb{E}[H(\tilde{y} - \varepsilon)] \geq V_t(x, r)$ for any $t = 1, \dots, T$ and state (x, r) , i.e., the problem in Equation (13.5) serves as an upper bound to the problem in Equation (13.4).

In what follows, we analyze the solution structure to Equation (13.4) by focusing on the piece-wise linear demand in Equation (13.3) that is more responsive to losses, i.e., $\eta_- \geq \eta_+$ —a setting considered in Chen et al. (2016). We then discuss the challenges involved in the gain-sensitive demand, i.e., $\eta_+ > \eta_-$, and the more general non-linear demand.

As we have seen in Section 13.2, the key argument in establishing the base-stock list-price policy is to show the joint concavity of the per-period profit function in state and decision as well as the concavity of the value-to-go function. However, even in the simpler special case of $\eta_- = \eta_+$, the revenue $pd(r, p)$ is not jointly concave in (r, p) , and $V_t(x, r)$ may not be concave in r . To resolve this issue, Chen et al. (2016) introduce the transformation $W_t(x, r) = V_t(x, r) - \Lambda r^2$ for some $\Lambda > 0$ and the Bellman equation in Equation (13.4) can be equivalently rewritten as

$$\begin{aligned} W_t(x, r) = \max_{y \geq x, p \in [\underline{p}, \bar{p}]} \{ & \Pi(r, p) - \mathbb{E}[H(y - d(r, p) - \varepsilon)] \\ & + \gamma \mathbb{E}[W_{t+1}(y - d(r, p) - \varepsilon, \alpha r + (1 - \alpha)p)] \}, \end{aligned}$$

where $\Pi(r, p) = pd(r, p) - \Lambda r^2 + \gamma \Lambda (\alpha r + (1 - \alpha)p)^2$. Chen et al. (2016) show that under the technical condition $\eta_- - \eta_+ \leq 2a(1 - \alpha)$, there always exists a positive constant Λ such that $\Pi(r, p)$ is jointly concave in (r, p) . One can then establish the joint concavity of $W_t(x, r)$ and the following structural result on the optimal solution $y^*(x, r)$ and $p^*(x, r)$.

Proposition 13.1 (Theorems 1 and 2 in Chen et al., 2016)

Let $(s(r), p_s(r))$ be an optimal solution to

$$\max_{y, p \in [\underline{p}, \bar{p}]} \left\{ \Pi(r, p) - \mathbb{E}[H(y - d(r, p) - \varepsilon)] + \gamma \mathbb{E}[W_{t+1}(y - d(r, p) - \varepsilon, \alpha r + (1 - \alpha)p)] \right\}.$$

If $x \leq s(r)$, then $y^*(x, r) = s(r)$ and $p^*(x, r) = p_s(r)$; otherwise, no order is placed, i.e., $y^*(x, r) = x$ and $p^*(x, r)$ is decreasing in x .

The optimal policy characterized in Proposition 13.1 can be viewed as a reference price-dependent base-stock list-price policy. When the inventory level is below the reference price-dependent base-stock $s(r)$, it is optimal to order up to $s(r)$ and charge a price $p_s(r)$ that only depends on the current reference price; otherwise, no order is placed and a discount is applied. Chen et al. (2016) further show that $\alpha r + (1 - \alpha)p_s(r)$ is increasing in r , and in the special case of $\eta_- = \eta_+$, $s(r)$ is also increasing in r .

In the infinite horizon setting, i.e., $T = \infty$, Chen et al. (2016) establish that under the optimal policy, whenever an order takes place, the optimal price as well as reference price trajectory will exactly follow that of the pure pricing problem in Equation (13.5)—a setting studied in Popescu and Wu (2007). In particular, it is shown in Popescu and Wu (2007) that there exists an interval of steady-state reference prices $[\underline{r}, \bar{r}]$ such that whenever the initial reference price $r \in [\underline{r}, \bar{r}]$, then $p_t^* = r_t^* = r$ for any $t \geq 1$, and when $r > \bar{r}$ ($r < \underline{r}$) then $\lim_{t \rightarrow \infty} p_t^* = \lim_{t \rightarrow \infty} r_t^* = \bar{r}$ ($\lim_{t \rightarrow \infty} p_t^* = \lim_{t \rightarrow \infty} r_t^* = \underline{r}$). In other words, at a steady-state reference price $r \in [\underline{r}, \bar{r}]$, a constant-pricing strategy is optimal. Correspondingly, it is optimal to order up to a steady-state base-stock level $s(r)$ and charge the constant price r in every period.

Both the concavity property in Proposition 13.1 and the long-run optimality of the constant-pricing strategy crucially depend on the assumption $\eta_- \geq \eta_+$. When the demand is more responsive to gains ($\eta_+ > \eta_-$), then $pd(r, p)$ is not even concave in p (see Figure 13.1 for a comparison and illustration). Hence, even for the pure pricing problem in Equation (13.5), the objective function is non-concave in the decision. It is shown by Popescu and Wu (2007) that no steady state exists in this case. Hu et al. (2016b) further demonstrate that due to the non-concavity, the optimal pricing strategy in general can be highly discontinuous in the state, and the resulting cyclic patterns of optimal prices can be very complex. In the special case of $\eta_- = 0$ and $\alpha = 0$, Hu et al. (2016b) prove that a *cyclic skimming pricing* strategy is optimal. In this strategy, the firm charges a regular price when the consumer's reference price is below some threshold; the firm then gradually applies a deeper and deeper discount over time until a consumer's reference price falls below the threshold again, and the firm repeats the cycle. Similar patterns are also observed in Hu and Nasiry (2018) for exponential demand, i.e., when $F(\cdot)$ is exponentially distributed, in which case the demand tends to be more responsive to gains as well.

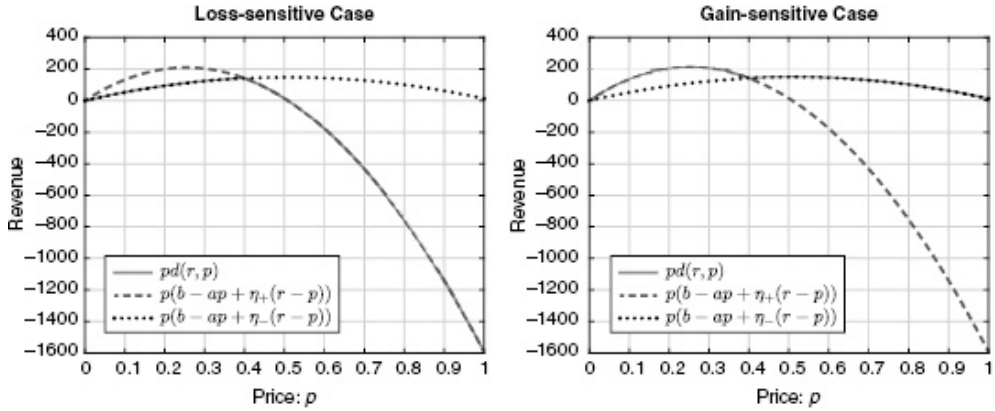


Figure 13.1 Comparison of the per-period revenue function $pd(r, p)$

Characterizing the optimal joint pricing and inventory policy of Equation (13.4) for the case $\eta_+ > \eta_-$ or general non-linear demand remains largely open. One recent attempt is made by Wang et al. (2019) who consider the case when $F(\cdot)$ follows a logistic distribution. However, they impose the strong assumption that average purchase utility $\mu_v - p + \lambda(r - p)$ is always nonnegative to ensure the concavity of $pd(r, p)$ in p and the applicability of the transformation technique in Chen et al. (2016). In general, the problem still suffers from the non-concavity issue and the base-stock policy may no longer be optimal. Given the complexity of characterizing the optimal policy, providing good computational heuristics is a promising future research direction. Chen et al. (2017) develop efficient algorithms to compute the optimal prices in Equation (13.5) when demand is piece-wise linear. Yet, effective heuristics for the general Equation (13.4) are still lacking.

13.3.2 Strategic Consumers

Several works build demand models on the behavior of strategic consumers who choose the time of purchasing so as to maximize their utilities. As an example, here we introduce in detail the model of Hu et al. (2016a), who incorporates strategic customer behavior in the context of a perishable product with a lifetime of only one period. More general models of managing perishable inventory (but without strategic consumer behavior) are discussed in Section 13.4, and we briefly review other models on strategic consumer behavior at the end of this section.

The model is motivated by daily operations of firms that sell perishable goods such as fresh seafood and bakery products, where each period consists of one regular-sales phase (e.g., the day) and one clearance phase (e.g., the evening). The firm sells the product at an exogenous price p during the regular sales, and any inventory leftover from the phase can be sold at a given discount price $p_0 \leq p$ in the clearance phase. Hu et al. (2016a) observe that any leftover inventory from this clearance phase has to be disposed of because it is not fresh enough for selling in the future; however, items sold during the clearance sales are still suitable for consumption in the following regular-sales phase (e.g., the next day), which means that markdown may cannibalize future sales at regular price.

In particular, Hu et al. (2016a) label the time periods such that each period t starts with a clearance phase, which is followed by a regular phase (in the next day). As a result, inventory is carried over from the regular phase in the previous period to the clearance phase in the current period (which then perishes) while inter-temporal substitution of demand occurs between the two phases within a period. To model the inter-temporal substitution behavior, suppose M consumers appear in the market in period t , where M are independent and identically distributed across time periods. It is assumed that $1-\alpha$ portion of consumers only buy in the regular-sales phase; among the remaining α portion of consumers, who try to buy in the clearance phase, a fraction ρ of such consumers can afford the regular price p but are attracted by the markdown. Hence, given z units to sell in the clearance phase, $z \wedge (\alpha M)$ consumers are satisfied in this phase, and $\rho(\alpha M - z)^+$ unsatisfied consumers will return for a second attempt in the following phase. That is, the total demand in the following regular-sales phase is given by

$$D(M, z) = \rho(\alpha M - z)^+ + (1 - \alpha)M.$$

Notice that the aggregate demand model given above can be formulated from a micro individual consumer choice model by constructing a consumer utility function and comparing expected utilities from purchasing in either phase (see Section 5.1 in Hu et al., 2016a).

The firm faces a trade-off between product spoilage and inter-temporal demand substitution: on one hand, it generates extra revenue for the otherwise disposed of product in the clearance phase by attracting these consumers who cannot afford the full price; on the other hand, it leads to a potential loss because some consumers who can afford the full price may forward buy the product during clearance sales and stock up for future consumption. In other words, markdown sales allow the firm to sell to consumers who cannot afford the regular price p at the expense of a lower revenue from consumers who can afford p but choose to buy early at the discount price p_0 .

The firm's objective is to maximize the total expected discounted profit by dynamically adjusting its inventory and markdown decisions periodically. Specifically, given x units of unsold items at the beginning of a clearance phase, the firm needs to decide $z \in [0, x]$ units of leftover to markdown, and y units to order (which is also the order-up-to level due to perishability) for the next regular-sales phase. Hence, the expected per-period profit obtained in the two phases is

$$\Pi(y, z) = p_0 \mathbb{E}[z \wedge (\alpha M)] - cy + p \mathbb{E}[y \wedge D(M, z)],$$

where c is the marginal ordering cost. Since $(y - D(M, z))^+$ units of unsold items are carried over to the next clearance phase, the Bellman equation for the firm's problem is

$$V_t(x) = \max_{y \geq 0, z \in [0, x]} \left\{ \Pi(y, z) + \gamma \mathbb{E}[V_{t+1}((y - D(M, z))^+)] \right\}.$$

Recall that demand can only be shifted from a clearance phase to the next regular-sales phase, and the firm carries inventory only from a regular-sales phase to the next clearance phase. Such a feature creates complicated dynamics for inventory carried over between two consecutive periods, and also results in the per-period profit function $\Pi(y, z)$ to be neither convex

nor concave. Hence, traditional approaches introduced in Section 13.2 in solving dynamic inventory models cannot be applied directly to characterize the associated optimal policy $[y^*(x), z^*(x)]$.

Nevertheless, Hu et al. (2016a) notice that the single-period profit $\Pi(y, z)$ is quasi-concave in y with changeover $\tau = \operatorname{argmax}\{\mathbb{E}[y \wedge (\alpha M)] - \frac{c}{\alpha} y\}$, and it is quasi-convex in z . Furthermore, the authors manage to show the preservation of P quasi-convexity in terms of z in the above multi-period maximization problem under some conditions, and using the quasi-convexity property they further establish for the infinite horizon problem that the optimal markdown decision $z^*(x)$ has the so-called *bang-bang* structure. Specifically, one of the main results of Hu et al. (2016a) is stated below:

Proposition 13.2 (Theorem 3 and Proposition 4 in Hu et al., 2016a)

If M follows a two-point distribution, then $z^*(x) = x$ when $x \geq s$ and $z^*(x) = 0$ when $x < s$ for some s decreasing in p_0 .

Proposition 13.2 implies that the firm should either put all of the inventory on sale or dispose of all of it at the beginning of each clearance phase, where the choice depends on whether the inventory level x falls above or below a certain cutoff level s . Furthermore, as the discounted price p_0 increases, the threshold s decreases, implying that the firm is more likely to sell in the clearance phase. This simple bang-bang optimal policy allows the firm to develop clear guidelines to manage discount sales for leftover products.

Hu et al. (2016a) also conduct extensive numerical explorations to examine the optimal policy under general distribution of M , the sensitivity of optimal policy to mean and standard deviation of demand, the loss of efficiency under some static policies in comparison with the optimal dynamic policy, etc. They also extend the base model by allowing the firm to endogenously determine the markdown price p_0 and they numerically confirm that the bang-bang structure of the markdown quantity decision remains valid in this case.

We remark that compared to the joint pricing and inventory problems we reviewed in Sections 13.2 and 13.3.1, the prices p and p_0 are given exogenously in the model of Hu et al. (2016a), and they model instead the markdown decision via the capacity control variable z . This is similar to the quantity-based models (versus price-based models) in the revenue management literature (see Section 5.1.1 in Talluri & van Ryzin, 2004 for a discussion on this modeling issue). There are also papers that consider price-based models with strategic consumers. Using a similar two-stage (full-price stage and markdown stage) model, Wu et al. (2015) adopt the concept of reference price introduced in Sections 13.3.1 to model consumers' belief about future markdown price; the markdown price in turn affects consumer's reference price in the next period according to the exponential smoothing process. In comparison to the two-stage model in Hu et al. (2016a) and Wu et al. (2015) that restrict the inter-temporal demand substitution to occur only between the two stages within the same period, Chen and Shi (2019) allow consumers to strategically time their purchase at any time after their arrival in an EOQ-type deterministic model. They use a quite different mechanism design approach to derive the optimal ordering and pricing policy.

13.4 MODELS WITH LEAD/LIFETIME

While Section 13.3 studies the inter-temporal effect in demand, this section can be viewed as exploring the inter-temporal effect in supply. Two commonly encountered temporal attributes

of a product in practice are lead time and lifetime. When orders are not delivered immediately, one needs to monitor the pipeline stock before deciding the price and order quantity for the current period. Similarly, when the product is perishable, the inventory should be sorted by its remaining lifetime which in turn affects what the right price is and how much more to order. In this section, we first present the general model by Chen et al. (2014) that considers the lead time and lifetime of a product simultaneously and discuss some structural properties of the optimal policy. The general model includes a non-perishable product with positive lead time and a perishable product with zero lead time—two frequently studied settings—as special cases. We then briefly review some available heuristics in the literature for these special cases.

Consider a product whose lifetime lasts l periods, and upon being ordered and manufactured it takes k periods to be delivered to the firm. Clearly, we require $k < l$ so that the product does not perish in transit. Chen et al. (2014) use an $(l-1)$ -dimensional vector $\mathbf{s} = (s_1, \dots, s_{l-1})$ to record the system state at the beginning of a period t . When $s_i \geq 0$, s_i represents the amount of inventory—either on-hand or in transit—with a residual lifetime of no more than i periods. Note that an item is in transit if and only if it has a residual lifetime of more than $l-k$ periods. In particular, s_{l-k} is the total on-hand inventory level and s_{l-1} is the inventory position of the system. When $s_i < 0$, the value $|s_i|$ is interpreted as the additional units that *should* have been ordered $l-i$ periods ago to make s_i zero. For $i \leq l-k-1$, $s_i < 0$ simply implies that there is no on-hand inventory with a residual lifetime of no more than i periods, however, its exact value, resulting from the accounting scheme used in the state transition below, is less meaningful. For $i = l-k$, $s_i < 0$ indicates there is a backlog in the system and its value records the amount of backlogging demand. For $i > l-k$, s_i is then the net inventory position of items with a residual lifetime of no more than i periods after subtracting away backlogged demand. By our definition, we have $s_1 \leq s_2 \leq \dots \leq s_{l-1}$.

A more natural state representation commonly used in the perishable inventory literature (see, Nahmias, 1982) is $\mathbf{x} = (x_1, \dots, x_{l-1})$, where x_i represents the amount of inventory, on-hand or in transit, that has a residual lifetime of i periods and whenever there is a backlog, $x_{l-k} < 0$ records its information. Using $\mathbf{s}$, we can represent $\mathbf{x}$ as $x_1 = s_1^+$, $x_i = s_i^+ - s_{i-1}^+$ for $i \leq l-k-1$, $x_{l-k} = s_{l-k} - s_{l-k-1}^+$ and $x_i = s_i - s_{i-1}$ for $i > l-k$. As we will see below, the station transition under $\mathbf{x}$ is much more complicated compared to that under $\mathbf{s}$.

Based on the system state, the firm decides the order-up-to level $s_l \geq s_{l-1}$ by placing an order $q = s_l - s_{l-1}$ and the price p that is uniformly charged for inventories of all ages. We assume an additive demand form, i.e., $D(p, \varepsilon) = d(p) + \varepsilon$ or $D(p(d), \varepsilon) = d + \varepsilon$. Demand is fulfilled as much as possible using on-hand inventory and unfulfilled demand is backlogged (we refer readers to Chen et al., 2014 for the lost-sales case). In contrast to the model in Section 13.2, the inventory is now vertically differentiated by the residual lifetime, and it becomes important to specify how the demand is fulfilled using inventories of different ages. Here, we assume that the firm has full control over how the inventory is issued and can also decide how much inventory to carry over to the next period by intentionally disposing of the on-hand stocks on top of those that are due to expire.⁸ It can then be argued that it is always optimal to deplete the oldest on-hand inventory first. That is, the inventory is used to fulfill the demand (or be disposed of) on a first-in-first-out (FIFO) basis. Without the ability to dispose of inventory, however, Chen et al. (2014) provide an example that shows a FIFO issuing policy may no longer be optimal. When the realized demand is smaller than the on-hand inventory level, i.e., $d + \varepsilon < s_{l-k}$, we let w denote the inventory depletion decision, which should satisfy $s_1 \vee (d + \varepsilon) \leq w \leq s_{l-k}$. Note that if $d + \varepsilon < s_1$, then on top of fulfilling demand, one is still required to dispose of all the remaining expiring inventory and hence $w \geq s_1$. On the other hand, when the realized demand is higher than the on-hand inventory, i.e., $d + \varepsilon \geq s_{l-k}$, all on-hand inventory (if any) will be

depleted and we let $w = d + \varepsilon$ so that $s_{l-k} - w$ records the back-order information. The two cases can be combined by requiring⁹

$$s_1 \vee (d + \varepsilon) \leq w \leq s_{l-k} \vee (d + \varepsilon).$$

Given the above decisions, the firm collects revenue $p(d)d$ and holding or shortage penalty cost $H(s_{l-k} - w) = h(s_{l-k} - w)^+ + b(w - s_{l-k})^+$ (again, for simplicity, we normalize ordering cost to zero). In addition, inventory being disposed of incurs a per-unit disposal cost θ —resulting in a total disposal cost $\theta(w - d - \varepsilon)$. Under the FIFO depletion policy, the state then evolves according to¹⁰

$$\tilde{\mathbf{s}} = (s_2 - w, \dots, s_{l-k} - w, s_{l-k+1} - w, \dots, s_l - w).$$

To illustrate the state transition, consider the following example. Let $l = 5, k = 2$ and $\mathbf{s} = (2, 5, 6, 11)$. We equivalently have $\mathbf{x} = (2, 3, 1, 5)$ with the first three components being inventory on-hand and the last component being inventory in transit. Let $s_l = 13$ (or $q = 2$) and $w = d + \varepsilon = 7$. We then have $\tilde{\mathbf{s}} = (-2, -1, 4, 6)$. Correspondingly, one can compute $\tilde{\mathbf{x}} = (0, 0, 4, 2)$. Notice that the arriving order of five units is used first to immediately satisfy one unit of back-order at the beginning of the next period—leaving four units of on-hand stock. Also, note that the exact values of the first two components in $\tilde{\mathbf{s}}$ —as long as being negative—are irrelevant. For instance, if the current state is $\mathbf{s}' = (1, 3, 6, 11)$, i.e., $\mathbf{x}' = (1, 2, 3, 5)$, then $\tilde{\mathbf{s}}' = (-4, -1, 4, 6)$ contains the same information as $\tilde{\mathbf{s}}$.

Let $V_t(\mathbf{s})$ be the profit-to-go function when the system state is $\mathbf{s}$ at the beginning of period t . We can write the Bellman equation for $t = 1, \dots, T$ as:

$$V_t(\mathbf{s}) = \max_{s_l \geq s_{l-1}, d \in [\underline{d}, \bar{d}]} \{p(d)d + \mathbb{E}[g_t(\mathbf{s}, s_l, d \mid \varepsilon)]\}, \quad (13.6)$$

where

$$g_t(\mathbf{s}, s_l, d \mid \varepsilon) = \max_{s_1 \vee (d + \varepsilon) \leq w \leq s_{l-k} \vee (d + \varepsilon)} \{-\theta(w - d - \varepsilon) - H(s_{l-k} - w) + \gamma V_{t+1}(\tilde{\mathbf{s}})\},$$

and $V_{T+1}(\cdot) = 0$. With zero lead time ($k = 0$), Equation (13.6) requires no modification except $s_1, \dots, s_{l-1}$ now all represent on-hand inventory positions.

On the other hand, for the special case of a non-perishable product ($l = +\infty$), distinguishing the residual lifetime of on-hand inventory becomes unnecessary and, with a slight abuse of notation, it is sufficient to use a k -dimensional vector $\mathbf{s} = (s_0, s_1, \dots, s_{k-1})$ to represent the state. Here, s_i is the net inventory position of items that takes no more than i periods to arrive. In particular, s_0 is the net on-hand inventory level and s_{k-1} is the net inventory position of the system. It is no longer necessary to dispose of any inventory and hence $w = d + \varepsilon$, and Equation (13.6) can be simplified to

$$\tilde{V}_t(\mathbf{s}) = \max_{s_k \geq s_{k-1}, d \in [\underline{d}, \bar{d}]} \{p(d)d - \mathbb{E}[H(s_0 - d - \varepsilon)] + \gamma \mathbb{E}[\tilde{V}_{t+1}(\tilde{\mathbf{s}})]\}, \quad (13.7)$$

where $\tilde{s}_i = s_{i+1} - (d + \varepsilon)$ for $i = 0, \dots, k-1$.

By employing the concept of L^1 -concavity (see Chapter 2 in Simchi-Levi et al., 2014 for an introduction and Chen & Li, 2020 for a survey of its applications to the operations model), Chen et al. (2014) establish the following structural properties for the optimal solution to Equation (13.6).

Proposition 13.3 (Theorems 1 in Chen et al., 2014)

Let $(s_l(\mathbf{s}), d(\mathbf{s}))$ be the optimal solution to the Equation (13.6). Then $s_l(\mathbf{s})$ and $d(\mathbf{s})$ are increasing in $\mathbf{s}$, and for any $\delta \geq 0$,

$$s_l(\mathbf{s} + \delta \mathbf{e}) \leq s_l(\mathbf{s}) + \delta, d(\mathbf{s} + \delta \mathbf{e}) \leq d(\mathbf{s}) + \delta,$$

where $\mathbf{e}$ is a vector whose components are all ones.

Unlike the optimal policy for the basic model in Equation (13.2) where one can find a state-independent base-stock level, the optimal order-up-to level here in general depends on the whole pipeline inventory in a complicated way. However, Proposition 13.3 shows that as the pipeline inventory positions become larger, one should order up to a higher level of inventory position and in the meantime lower the price. In addition, the inequalities in Proposition 13.3 establish a limited sensitivity result: a unit increase in the pipeline inventory position results in at most one unit increase in the optimal order-up-to level and optimal demand level.

The structural properties of $(s_l(\mathbf{s}), d(\mathbf{s}))$ established in Proposition 13.3 can be translated to that with respect to $\mathbf{x}$. Let $\hat{q}(\mathbf{x})$ and $\hat{d}(\mathbf{x})$ be the optimal order quantity and demand level when the system state is represented by $\mathbf{x}$. Then one can show that

$$\begin{aligned} -\delta &\leq \hat{q}(\mathbf{x} + \delta \mathbf{e}_{l-1}) - \hat{q}(\mathbf{x}) \leq \dots \leq \hat{q}(\mathbf{x} + \delta \mathbf{e}_1) - \hat{q}(\mathbf{x}) \leq 0, \\ 0 &\leq \hat{d}(\mathbf{x} + \delta \mathbf{e}_{l-1}) - \hat{d}(\mathbf{x}) \leq \dots \leq \hat{d}(\mathbf{x} + \delta \mathbf{e}_1) - \hat{d}(\mathbf{x}) \leq \delta, \end{aligned}$$

where $\mathbf{e}_i$ denotes the i -th unit vector for $i = 1, \dots, l-1$. In other words, on the one hand, the optimal order quantity $\hat{q}(\mathbf{x})$ is decreasing in the inventory level of each age, and it is more sensitive to younger inventory as opposed to the inventory that is about to perish. The optimal demand level, on the other hand, is increasing in the inventory level of each age, and it is most sensitive to the inventory that is close to expiring.

In the case of a non-perishable product, Proposition 13.3 directly carries over to Equation (13.7), where $\mathbf{s} = (s_0, s_1, \dots, s_{k-1})$ and the optimal order-up-to level is denoted by $s_k(\mathbf{s})$. One can similarly arrive at the conclusion that the optimal order quantity (demand level) is decreasing (increasing) in each outstanding order, and it is more sensitive to orders placed recently (earlier)—recovering the main insight from Pang et al. (2012).

Although Proposition 13.3 provides valuable qualitative insights into the optimal solution to the Equation (13.6), it is in general computationally intractable to solve Equation (13.6) due to the well-known “curse of dimensionality.” The literature has developed heuristics for the special cases of zero lead time ($k=0$) and a non-perishable product ($l = +\infty$) separately. For the zero lead time case, both Li et al. (2009) and Chen et al. (2014) use the one-dimensional state: s_{l-1} , the current inventory position, as an approximation of the state information. The

expected disposal cost associated with the current order is then upper bounded by a quantity that only depends on the order-up-to level s_l and the current demand level d . By solving the corresponding one-dimensional dynamic program, one then obtains a base-stock list-price heuristic. Compared to the optimal policy, the heuristic ignores the lifetime information of the current inventory and it takes into account the perishability via the approximated future disposal cost. Although Chen et al. (2014) and Li et al. (2009) differ in their approximations of the expected disposal cost, the numerical results of Chen et al. (2014) show that the performance of the two heuristics are very close and both are near-optimal in the infinite horizon setting with the long-run average profit criterion.

For the case of a non-perishable product, Bernstein et al. (2016) use the myopic demand level (price) $d^M(s_0)$ that solves

$$\max_{d \in [\underline{d}, \bar{d}]} \{p(d)d - \mathbb{E}[H(s_0 - d - \varepsilon)]\}$$

to first approximate the optimal demand-level decision in Equation (13.7). To solve for the ordering decision, one still faces a k -dimensional program, i.e., a Bellman equation of the form in Equation (13.7) with d substituted by $d^M(s_0)$. By further approximating $d^M(s_0)$ as a linear function of the current net inventory level s_0 , Bernstein et al. (2016) reduce the state space to a one-dimensional state which they call “price-deflated inventory position.” The heuristic ordering policy solved from the resulting one-dimensional dynamic program is again a base-stock policy. In comparison, Chen et al. (2019) propose a different heuristic that first approximates the ordering decisions via a constant-order policy. That is, in each period, a constant amount of new inventory q is ordered and as a result the pipeline inventory levels are also q . For a given q , the pricing decisions can then be solved by the following one-dimensional dynamic program

$$v_t(s_0; q) = \max_{d \in [\underline{d}, \bar{d}]} \{p(d)d - \mathbb{E}[H(s_0 - d - \varepsilon)] + \gamma \mathbb{E}[v_{t+1}(s_0 + q - d - \varepsilon; q)]\}.$$

The best constant-order quantity can then be found via a bi-section search. Chen et al. (2019) further prove that the above constant-order dynamic pricing heuristic is asymptotically optimal as the lead time k goes to infinity in the infinite horizon setting with the long-run average profit criterion.

13.5 CONCLUSION

Our above discussion highlights several notable advancements regarding integrated inventory and pricing models, including: (a) more general conditions on the demand or supply side to guarantee the simple structure of the optimal policies in the basic multi-period model; (b) new models that take into account consumers' behavior either through reference price models or explicitly modeling strategic consumers; (c) new models and methodologies dealing with multi-dimensional Markov decision processes arising from settings with replenishment lead times and perishable products with limited lifetimes.

There are many interesting open research problems. First, there are very limited results on multi-product models. We refer the reader to Chen et al. (2013) for a two-product case, and Song and Xue (2021), Song et al. (2021) and the references therein for recent progress. Admittedly, the optimal policies of inventory models with multiple products are usually hard

to characterize and not very insightful. Looking at different asymptotic regimes such as a long lead time or large demand size and using the insights derived there to help design effective heuristics might be a fruitful direction to pursue.

Second, incorporating consumer-behavior models is only a start. There are many different consumer-behavior models in the literature. For instance, for reference price models, there are different reference price evolution models other than the one touched upon here, and modeling the consumers' heterogeneity may better capture the underlying demand. In multi-product settings, we would expect much richer consumers' behaviors. For example, when evaluating a product, consumers may form a reference price based on both the product's historical prices and the shelf prices of other competing brands; see Chapters 2 and 5 in Hu (2015) for a discussion on reference price models in multi-product settings and the related empirical literature. It would be interesting to integrate those behavioral models into dynamic inventory decision frameworks. Of course, the challenge is how to balance model accuracy and tractability. As opposed to considering consumer behavior, there is also a stream of literature that seeks to understand the managerial behavior in inventory and pricing decisions, for which we refer readers to Ramachandran et al. (2018) and Chapter 17 of this book.

Third, developing data-driven inventory and pricing models would be valuable. We refer to Chapter 12 of this book for a survey on learning in inventory management, Chen and Chen (2015) for a survey of dynamic pricing based on robust optimization approaches and demand learning approaches, and Qin et al. (2019) which develops sampling complexity of the sample approximation approach to a single-product inventory and pricing model.

Fourth, there is very limited activity on explicitly modeling inventory and pricing competition in multi-period settings and developing methodologies to analyze and solve competition models. We believe this is primarily due to the complexity involved in modeling and analyzing dynamic games. Nevertheless, as online retailing and omnichannel retailing drive fierce competition, we would expect multi-period competition models could provide valuable managerial insights and suggest implementable strategies. We refer to the earlier survey by Chen and Simchi-Levi (2012) for relevant papers.

Finally, developing compelling case studies on the deployment of integrated inventory and pricing models and algorithms built upon industrial practice and data will greatly enhance our understanding of the theory and practice and provide useful benchmarks to test our models and algorithms. We hope future research will make great strides in these directions.

ACKNOWLEDGMENT

We would like to thank Professor Jeannette Song, Professor Andrew Davis and Professor Jordan Tong for their valuable suggestions.

NOTES

1. Alternatively, one can arrive at the same formula by using the random utility model commonly adopted in the multi-product setting. In random utility model, consumer's utility from purchasing is $\mu_v - p + \varepsilon_v$, where $\mu_v - p$ is the deterministic component and ε_v is the random component, and the utility from no purchasing is zero. Then with $F(\cdot)$ being the distribution of $v := \mu_v + \varepsilon_v$, the purchase probability is $\bar{F}(p)$.

2. Customers are implicitly assumed to be infinitesimal here. In a finite population model where there are N customers with customers' valuation being i.i.d. random variables from $F(\cdot)$, one can directly model the stochastic demand as $\tilde{D}(p) = B(N, F(p))$, where $B(N, F(p))$ is the binomial random variable.
3. In many cases when demand backlogging and linear ordering cost are assumed, an alternative terminal condition $V_{T+1}(x) = cx$ is often assumed for analytical convenience so that the ordering cost can be absorbed into the cost function $H(\cdot)$ and can then be normalized to zero.
4. In fact, $R(d, y) = p(d)d$ in the reservation-price model is concave if and only if the *virtual valuation*: $p - \bar{F}(p)/f(p)$ is increasing (see pp. 315–316 in Talluri and van Ryzin, 2004, for a derivation). The condition is also commonly assumed in the auction literature (for instance, Myerson, 1981) and is satisfied when v has increasing hazard rate.
5. The piece-wise linear form here is widely applied in the literature (see, for instance, Kőszegi and Rabin, 2006). One is referred to Hu and Nasiry (2018) for some discussions on the nonlinear functional form.
6. See Example 3 in Hu and Nasiry (2018). Also note that Equation (13.3) is exact if v is uniformly distributed on $[0, \bar{v}]$ with $b = N, a = N / \bar{v}, \eta_+ = N\lambda_+ / \bar{v}, \eta_- = N\lambda_- / \bar{v}$. In this case, $\eta_- > \eta_+$ if and only if $\lambda_- > \lambda_+$, i.e., demand response inherits consumers' psychological bias.
7. Readers are referred to Nasiry and Popescu (2011) for an alternative model based on minimum price encountered in the history and Briesch et al. (1997) for a discussion of reference price formation in a multi-product setting.
8. Such full control in how the inventory is issued can be observed, for instance, in blood banks. In certain cases, it can also be indirectly achieved by the firm's inventory display strategy. Of course, another commonly observed case is when the firm has no control in how the inventory is depleted. In such case, due to a uniform price on all items, consumers are expected to buy the freshest items first resulting in a last-in-first-out (LIFO) sequence.
9. If intentional disposal is not allowed, then we have $w = s_1 \vee (d + \varepsilon)$
10. If one uses $\mathbf{x}$ as the state, then the state $\tilde{\mathbf{x}}$ in the next period is characterized by $\tilde{x}_i = (x_{i+1} - (w - \sum_{j=1}^i x_j)^+)^+$ for $i = 1, \dots, l - k - 1, \tilde{x}_{l-k} = x_{l-k+1} - (w - \sum_{j=1}^{l-k} x_j)^+ \tilde{x}_i = (x_{i+1} - (w - \sum_{j=1}^i x_j)^+)^+, \tilde{x}_i = x_{i+1}$ for $i = l - k + 1, \dots, l - 2$ and $\tilde{x}_{l-1} = q$.

REFERENCES

- Bensoussan, A., Xie, Y., & Yan, H. (2019). Joint inventory-pricing optimization with general demands: An alternative approach for concavity preservation. *Production and Operations Management*, 28(9), 2390–2404.
- Bernstein, F., Li, Y., & Shang, K. (2016). A simple heuristic for joint inventory and pricing models with lead time and backorders. *Management Science*, 62(8), 2358–2373.
- Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research*, 24(2), 202–214.
- Chan, L., Shen, Z., Simchi-Levi, D., & Swann, J. (2004). *Coordination of pricing and inventory decisions: A survey and classification*. Kluwer.
- Chen, H., & Zhang, Z. (2014). Joint inventory and pricing control with general additive demand. *Operations Research*, 62(6), 1335–1343.
- Chen, M., & Chen, Z.-L. (2015). Recent developments in dynamic pricing research: Multiple products, competition, and limited demand information. *Production and Operations Management*, 24(5), 704–731.
- Chen, X., & Gao, X. (2019). Stochastic optimization with decisions truncated by positively dependent random variables. *Operations Research*, 67(5), 1321–1327.
- Chen, X., Gao, X., & Pang, Z. (2018). Preservation of structural properties in optimization with decisions truncated by random variables and its applications. *Operations Research*, 66(2), 340–357.
- Chen, X., Hu, P., & He, S. (2013). Preservation of supermodularity in parametric optimization problems with nonlattice structures. *Operations Research*, 61(5), 1166–1173.

- Chen, X., Hu, P., & Hu, Z. (2017). Efficient algorithms for the dynamic pricing problem with reference price effect. *Management Science*, 63(12), 4389–4408.
- Chen, X., Hu, P., Shum, S., & Zhang, Y. (2016). Dynamic stochastic inventory management with reference price effects. *Operations Research*, 64(6), 1529–1536.
- Chen, X., & Li, M. (2020). Discrete convex analysis and its applications in operations: A survey. *Production and Operations Management*.
- Chen, X., Pang, Z., & Pan, L. (2014). Coordinating inventory control and pricing strategies for perishable products. *Operations Research*, 62(2), 284–300.
- Chen, X., & Simchi-Levi, D. (2004a). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, 52(6), 887–896.
- Chen, X., & Simchi-Levi, D. (2004b). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, 29(3), 698–723.
- Chen, X., & Simchi-Levi, D. (2012). Pricing and inventory management. In Ö. Özer & R. Phillips (Eds.), *The Oxford handbook of pricing management*. Oxford University Press.
- Chen, X., Stolyar, A., & Xin, L. (2019). Asymptotic optimality of constant-order policies in joint pricing and inventory control models. Available at: SSRN 3375203.
- Chen, Y., & Shi, C. (2019). Joint pricing and inventory management with strategic customers. *Operations Research*, 67(6), 1610–1627.
- Eliashberg, J., & Steinberg, R. (1993). *Marketing-production joint decision making*. 5 (pp. 829–877). North Holland.
- Elmaghraby, W., & Keskinocak, P. (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10), 1287–1309.
- Federgruen, A., & Heching, A. (1999). Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3), 454–475.
- Feng, Q. (2010). Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Science*, 56(12), 2154–2172.
- Feng, Q., Luo, S., & Shanthikumar, J. G. (2020). Integrating dynamic pricing with inventory decisions under lost sales. *Management Science*, 66(5), 1783–2290.
- Feng, Q., Luo, S., & Zhang, D. (2014). Dynamic inventory–pricing control under backorder: Demand estimation and policy optimization. *Manufacturing and Service Operations Management*, 16(1), 149–160.
- Feng, Q., & Shanthikumar, J. G. (2018). Supply and demand functions in inventory models. *Operations Research*, 66(1), 77–91.
- Fibich, G., Gavious, A., & Lowengart, O. (2003). Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research*, 51(5), 721–734.
- Gong, X., Chao, X., & Zheng, S. (2014). Dynamic pricing and inventory management with dual suppliers of different lead times and disruption risks. *Production and Operations Management*, 23(12), 2058–2074.
- Greenleaf, E. (1995). The impact of reference price effects on the profitability of price promotions. *Marketing Science*, 14(1), 82–104.
- Hu, P., Shum, S., & Yu, M. (2016a). Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Operations Research*, 64(1), 118–134.
- Hu, Z. (2015). Dynamic pricing with reference price effects [Ph.D. Thesis]. University of Illinois at Urbana-Champaign.
- Hu, Z., Chen, X., & Hu, P. (2016b). Dynamic pricing with gain-seeking reference price effects. *Operations Research*, 64(1), 150–157.
- Hu, Z., & Nasiry, J. (2018). Are markets with loss-averse consumers more sensitive to losses? *Management Science*, 64(3), 1384–1395.
- Huh, W. T., & Janakiraman, G. (2008). (s, S) optimality in joint inventory-pricing control: An alternate approach. *Operations Research*, 56(3), 783–790.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kocabiyykoglu, A., & Popescu, I. (2011). An elasticity perspective on the newsvendor with price sensitive demand. *Operations Research*, 59(2), 301–312.

- Kopalle, P., Rao, A., & Assuncao, J. (1996). Asymmetric reference price effects and dynamic pricing policies. *Marketing Science*, 15(1), 60–85.
- Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4), 1133–1165.
- Li, Q., & Zheng, S. (2006). Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research*, 54(4), 696–705.
- Li, Y., Lim, A., & Rodrigues, B. (2009). Note—Pricing and inventory control for a perishable product. *Manufacturing and Service Operations Management*, 11(3), 538–542.
- Lu, Y., Chen, Y., Song, M., & Yan, X. (2014). Optimal pricing and inventory control policy with quantity-based price differentiation. *Operations Research*, 62(3), 512–523.
- Lu, Y., & Simchi-Levi, D. (2013). On the unimodality of the profit function of the pricing newsvendor. *Production and Operations Management*, 22(3), 615–625.
- Mazumdar, T., Raj, S., & Sinha, I. (2005). Reference price research: Review and propositions. *Journal of Marketing*, 69(4), 84–102.
- Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1), 58–73.
- Nahmias, S. (1982). Perishable inventory theory: A review. *Operations Research*, 30(4), 680–708.
- Nasiry, J., & Popescu, I. (2011). Dynamic pricing with loss-averse consumers and peak-end anchoring. *Operations Research*, 59(6), 1361–1368.
- Natter, M., Reutterer, T., Mild, A., & Taudes, A. (2007). Practice prize report? An assortmentwide decision-support system for dynamic pricing and promotion planning in DIY retailing. *Marketing Science*, 26(4), 576–583.
- Pang, Z., Chen, F. Y., & Feng, Y. (2012). A note on the structure of joint inventory-pricing control with leadtimes. *Operations Research*, 60(3), 581–587.
- Popescu, I., & Wu, Y. (2007). Dynamic pricing strategies with reference effects. *Operations Research*, 55(3), 413–429.
- Qin, H., Simchi-Levi, D., & Wang, L. (2019). Data-driven approximation schemes for joint pricing and inventory control models. Available at: SSRN 3354358.
- Ramachandran, K., Terseyaglu, N., & Xia, Y. (2018). Multidimensional decision making in operations: An experimental investigation of joint pricing and quantity decisions. *Management Science*, 64(12), 5461–5959.
- Shen, X., Bao, L., & Yu, Y. (2018). Coordinating inventory and pricing decisions with general price-dependent demands. *Production and Operations Management*, 27(7), 1355–1367.
- Simchi-Levi, D., Chen, X., & Bramel, J. (2014). *The logic of logistics: Theory, algorithms, and applications for logistics management* (3rd ed.). Springer.
- Song, J.-S. J., Song, Z. X., & Shen, X. (2021). Demand management and inventory control for substitutable products. Available at: SSRN 3866775.
- Song, J.-S., & Xue, Z. (2021). Demand shaping through bundling and product configuration: A dynamic multiproduct inventory-pricing model. *Operations Research*, 69(2), 525–544.
- Talluri, K., & van Ryzin, G. (2004). *The theory and practice of revenue management*. Springer.
- Wang, Q., Zhao, N., Wu, J., & Zhu, Q. (2019). Optimal pricing and inventory policies with reference price effect and loss-averse customers. *Omega*, 99, 102174.
- Wu, S., Liu, Q., & Zhang, R. Q. (2015). The reference effects on a retailer's dynamic pricing and inventory strategies with strategic consumers. *Operations Research*, 63(6), 1320–1335.
- Yang, N., & Zhang, R. (2014). Dynamic pricing and inventory management under inventory-dependent demand. *Operations Research*, 62(5), 1077–1094.
- Yano, C. A., & Gilbert, S. M. (2003). *Coordinated pricing and production/procurement decisions: A review*. Kluwer.
- Zipkin, P. (2000). *Foundations of inventory management*. McGraw Hill.

14. Statistical learning in inventory management

Wang Chi Cheung and David Simchi-Levi

14.1 SINGLE-PERIOD SETTING: NEWSVENDOR MODEL

The newsvendor model, first studied by Arrow et al. (1960) and Scarf (1960), is one of the most fundamental models in the field of inventory management. In the newsvendor model, a decision-maker (DM) aims to fulfill a stochastic demand, denoted D , for a product over a single period. At the start of the period, the DM has no on-hand inventory. The dynamics during the period follow the sequence of events:

1. The DM decides to order y units of inventory.
2. The DM observes the realization of stochastic demand D .
3. The DM accounts for the operational costs.
 - If $y \geq D$, the DM incurs an overage cost of $h \times (y - D)$, where $h > 0$ is the per-unit overage cost.
 - Otherwise, the DM incurs an underage cost of $b \times (D - y)$, where $b > 0$ is the per-unit underage cost.

The goal of the DM is to minimize the sum of the expected underage and overage costs. Importantly, ordering decision y has to be made *before* observing the realization of stochastic demand D . Denote $(z)^+ = \max\{z, 0\}$. The DM's goal can be formulated as the stochastic optimization problem

$$\min_{y \geq 0} C(y), \text{ where } C(y) = \mathbb{E}[h(y - D)^+ + b(D - y)^+]. \quad (14.1)$$

We assume that $\mathbb{E}[|D|]$ is finite in order for $C(y)$ to be finite for every $y \in \mathbb{R}$, so that optimization Equation (14.1) is well-defined. *Assuming that* the DM knows the demand distribution, i.e., the probability distribution of the stochastic demand D , the DM can achieve the minimum in Equation (14.1) by setting the ordering amount as

$$y^* = \inf_{y \geq 0} \left\{ y : F(y) \geq \frac{b}{b+h} \right\}. \quad (14.2)$$

The function $F(y) = \Pr(D \leq y)$ is the Cumulative Distribution Function (CDF) of stochastic demand D . The quantity y^* is also known as the $\frac{b}{b+h}$ -quantile of the demand distribution.

To justify Equation (14.2), note that objective function $C(y)$ is convex on $\mathbb{R}$. At each $y \geq 0$, the function $C(y)$ has sub-gradient $\partial C(y) = [\partial^- C(y), \partial^+ C(y)]$, where

$$\partial^- C(y) = -b + (b+h)\Pr(D < y), \quad \partial^+ C(y) = -b + (b+h)\Pr(D \leq y)$$

are the left and right derivatives of convex function C . The optimality of y^* follows since $0 \in \partial C(y^*)$, by the definition of y^* in Equation (14.2).

The optimal ordering quantity y^* in Equation (14.2) crucially depends on cost parameters b, h , and the knowledge of the CDF F . However, in most real-life applications complete knowledge of the demand distribution is rarely available. Rather, the DM only has partial information about the latent demand distribution. In this survey chapter, we focus on the data-driven setting, where the partial information is a set of data relevant to D .

14.1.1 Sampling-Based Newsvendor via Sample Average Approximation

We first consider the newsvendor model in a sampling-based setting. In this setting, the DM does not know CDF F of stochastic demand D . Rather, the DM only has samples $d^{(1)}, \dots, d^{(N)}$, which are the realizations of N independently and identically distributed (iid) random variables with common CDF F .

In the sampling-based setting, a natural approach is the sample average approximation (SAA) method (Shapiro et al., 2009). The SAA method is a popular heuristic that is applicable to a wide class of sampling-based stochastic optimization problems, including the newsvendor problem. The SAA method involves solving the SAA problem, where the latent probability distribution in the problem instance is replaced by the empirical distribution defined by the samples.

In the context of the newsvendor problem, the SAA method is as displayed in Algorithm 14.1. The algorithm involves solving the optimization problem in Equation (14.1), with the latent demand distribution of D replaced with the empirical demand distribution based on the samples $d^{(1)}, \dots, d^{(N)}$. For $\hat{D}$, which follows the empirical demand distribution, we have $\Pr(D = d) = \frac{1}{N} \sum_{n=1}^N \mathbf{1}(d = d^{(n)})$. That is, the empirical demand distribution is the uniform distribution on the (multi-)set $\{d^{(1)}, \dots, d^{(N)}\}$. Mathematically, the SAA method in Algorithm 14.1 computes an optimal solution y^{SAA} to the SAA problem:

$$\min_{y \geq 0} C(y), \text{ where } C(y) = \frac{1}{N} \sum_{n=1}^N [h(y - d^{(n)})^+ + b(d^{(n)} - y)^+]. \quad (14.3)$$

In Algorithm 14.1, the equation on Line 2 follows from applying Equation (14.2) with the latent demand distribution replaced by the empirical demand distribution. Algorithm 14.1 can be efficiently implemented. For example, the identification of y^{SAA} can be done by sorting the N samples.

Algorithm 14.1 SAA method for sampling-based newsvendor

1. Inputs: samples $d^{(1)}, \dots, d^{(N)}$, per-unit overage cost $h > 0$, per-unit underage cost $b > 0$.

2. Identify
$$y^{\text{SAA}} = \min_{1 \leq j \leq N} \left\{ d^{(j)} : \frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq d^{(j)}) \geq \frac{b}{b+h} \right\}.$$

3. Output y^{SAA} as the ordering quantity.

By the Strong Law of Large Numbers, we know that for any $y \in \mathbb{R}$ the average $\frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq y)$ tends to $\Pr(D \leq y)$ almost surely, as N tends to infinity. Therefore, it is natural to expect that y^{SAA} is close to true optimal solution y^* when N is sufficiently large. Two crucial theoretical questions are: 1) How should we formalize the performance guarantee of y^{SAA} relative to optimal solution y^* , and 2) what is the number of samples needed for the SAA method to achieve a desired level of performance guarantee?

The research work by Levi et al. (2007) sheds light on these two questions with the following Theorem:

Theorem 14.1 (Levi et al., 2007) *For any error parameter $\varepsilon > 0$, the output y^{SAA} of the SAA method (Algorithm 14.1) satisfies*

$$\Pr \left[C(y^{\text{SAA}}) \leq (1 + \varepsilon) \min_{y \geq 0} C(y) \right] \geq 1 - 2 \exp \left(-\frac{2}{9} N \varepsilon^2 \left(\frac{\min\{h, b\}}{h + b} \right)^2 \right). \quad (14.4)$$

In Equation (14.4), the probability measure $\Pr$ is over the randomness of y^{SAA} , which depends on the iid samples $d^{(1)}, \dots, d^{(N)}$. The theorem illustrates a trade-off between the approximation ratio $(1 + \varepsilon)$ and the lower bound on the success probability in Equation (14.4). As error parameter ε decreases, the success probability lower bound on the right-hand side of Equation (14.4) decreases. The SAA method requires more iid samples to obtain a better approximation ratio with high probability. To quantify the number of samples needed for achieving a certain level of performance guarantee, it is useful to rephrase Theorem 14.1 as the following equivalent corollary.

Corollary 14.1 (Levi et al., 2007) *Consider any error parameter $\varepsilon > 0$ and confidence parameter $\delta \in (0, 1)$. Suppose the number of samples N satisfies*

$$N \geq \frac{9}{2\varepsilon^2} \cdot \left(\frac{h + b}{\min\{h, b\}} \right)^2 \cdot \log \frac{2}{\delta}. \quad (14.5)$$

Then the output y^{SAA} of the SAA method satisfies

$$\Pr \left[C(y^{\text{SAA}}) \leq (1 + \varepsilon) \min_{y \geq 0} C(y) \right] \geq 1 - \delta.$$

The sample bound in Equation (14.5) sheds light on how the performance guarantee of the SAA method depends on the parameters $h, b, \varepsilon, \delta$. First, the sample bound is linear in $1/\varepsilon^2$. When the approximation ratio $(1 + \varepsilon)$ becomes smaller, the SAA method requires more samples.

Second, the sample bound grows as the fraction $\frac{\min\{h, b\}}{h + b} \in (0, 1)$ decreases. To gain intuition, note that solving the newsvendor problem is equivalent to the identification of the $\frac{b}{b + h}$ -quantile y^* of the latent demand distribution. For identifying, it is desirable for the

DM to receive some samples larger than y^* , and some samples less than or equal to y^* . By the definition of y^* , these two events happen with probability $\frac{h}{h+b}$ and $\frac{b}{h+b}$, respectively. Consequently, it takes $\frac{h+b}{h} + \frac{h+b}{b} \in (\frac{h+b}{\min\{h,b\}}, \frac{2(h+b)}{\min\{h,b\}}]$ many samples in expectation for the agent to receive at least one sample $> y^*$ and at least one sample $\leq y^*$. Consequently, the sample bound increases with the fraction $\frac{h+b}{\min\{h,b\}}$.

Third, the sample bound is *distribution-free*, in the sense that the bound only depends on $h, b, \varepsilon, \delta$, but it is independent of the latent demand distribution on D . Theorem 14.1 and Corollary 14.1 hold as long as $\mathbb{E}[|D|]$ is finite, which is needed solely for the function $C(y)$ to be finite for any $y \in \mathbb{R}$. There is no need for D to have a bounded support, and Theorem 14.1 holds true even when the $\text{var}(D)$ is infinite. The rationale is that we only care about the ratio $C(y^{\text{SAA}})/C(y^*)$, and it turns out that the volatility of D in the numerator is nullified by that in the denominator. In fact, apart from the sampling-based newsvendor problem, Shmoys and Swamy (2006) and Charikar et al. (2005) demonstrate that a wide class of sampling-based covering problems also enjoy similar *distribution-free* sample bounds. These sampling-based covering problems include the stochastic set cover problem and the stochastic facility location problem in which the DM only has access to the samples but not the underlying probability distributions.

Interestingly, the number of samples needed for *optimization*, namely to achieve near optimality for the sampling-based newsvendor problem, could be less than that for *estimation*, such as estimating $\mathbb{E}[D]$. Indeed, for stochastic demand D with support $[0, \bar{D}]$, the number of samples needed for estimating $\mathbb{E}[D]$ within a multiplicative factor of $(1 + \varepsilon)$ scales with $\bar{D}^2 / \varepsilon^2$ in the worst case, which increases linearly with $\bar{D}^2$. In contrast, the number of samples needed for the sampling-based newsvendor problem remains the same, no matter how large $\bar{D}$ is. In fact, such a phenomenon where a data-driven optimization problem requires fewer samples than a related estimation problem not only occurs in the newsvendor problem, but also in pricing problems in auction settings (Huang et al., 2018).

We highlight the two major steps in the proof of Corollary 14.1:

1. With probability $1 - \delta$, the following two inequalities hold:

$$\partial^+ C(y^{\text{SAA}}) \geq -\frac{\varepsilon}{3} \min\{h, b\}, \quad \partial^- C(y^{\text{SAA}}) \leq \frac{\varepsilon}{3} \min\{h, b\}. \quad (14.6)$$

Consequently, there exists a sub-gradient $g \in \partial C(y^{\text{SAA}})$ such that $|g| \leq \frac{\varepsilon}{3} \min\{h, b\}$.

2. Conditioned on Equation (14.6), it can be deduced that $C(y^{\text{SAA}}) \leq (1 + \varepsilon) \min_{y \geq 0} C(y)$ with certainty. The argument crucially uses the fact that

$$C(y) \geq \bar{C}(y) = h(y - \mathbb{E}[D])^+ + b(\mathbb{E}[D] - y)^+,$$

which follows from Jensen's inequality. The non-negative function $\bar{C}(y)$ is a V-shape piece-wise linear function, which allows for the translation of the first-order approximation in Equation (14.6) to the zero-order approximation $C(y^{\text{SAA}}) \leq (1 + \varepsilon) \min_{y \geq 0} C(y)$.

Finally, the inequalities in Equation (14.6) can be deduced from the Massart inequality (Massart, 1990). By the definition of y^{SAA} , we have $\partial^-\hat{C}(y^{\text{SAA}}) \leq 0 \leq \partial^+\hat{C}(y^{\text{SAA}})$. Then, the inequalities in Equation (14.6) are shown by demonstrating that $|\partial^+\hat{C}(y^{\text{SAA}}) - \partial^+C(y^{\text{SAA}})| \leq \frac{\varepsilon}{3} \min\{h, b\}$ holds with a probability of at least $1 - \delta/2$, and similarly that $|\partial^-\hat{C}(y^{\text{SAA}}) - \partial^-C(y^{\text{SAA}})| \leq \frac{\varepsilon}{3} \min\{h, b\}$ holds with a probability of at least $1 - \delta/2$. Without loss of generality, let's focus on the bound involving $\partial^+\hat{C}(y^{\text{SAA}}), \partial^+C(y^{\text{SAA}})$. Now,

$$\begin{aligned}\partial^+C(y^{\text{SAA}}) &= -b + (h+b)\Pr(D \leq y^{\text{SAA}}), \\ \partial^+C(y^{\text{SAA}}) &= -b + (h+b) \cdot \frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq y^{\text{SAA}}).\end{aligned}$$

The closeness between $\partial^+C(y^{\text{SAA}}), \partial^+\hat{C}(y^{\text{SAA}})$ is shown by bounding the difference between $\Pr(D \leq y^{\text{SAA}})$ and $\frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq y^{\text{SAA}})$. Now, for any fixed y , the collection $\{\mathbf{1}(d^{(n)} \leq y)\}_{n=1}^N$ consists of N iid Bernoulli random variables with mean $\Pr(D \leq y)$. The difference can then be bounded by the Massart inequality (Massart, 1990):

Theorem 14.2 (Massart, 1990) *For any $\varepsilon > 0$, it holds that*

$$\Pr\left[\left|\Pr(D \leq y) - \frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq y)\right| > \varepsilon \text{ for all } y \in \mathbb{R}\right] \geq 1 - 2\exp\left(-\frac{\varepsilon^2}{2}\right). \quad (14.7)$$

We remark that the Massart inequality is a stronger version of the well-known Hoeffding inequality (Hoeffding, 1963), which only requires Equation (14.7) to hold for a fixed y .

A subsequent research work by Levi et al. (2015) revisits the SAA method for the sampling-based newsvendor problem, and provides a tighter sample bound:

Theorem 14.3 (Levi et al., 2015) *Fix accuracy parameter $\varepsilon > 0$ and confidence parameter $\delta \in (0, 1)$. Suppose the number of samples N satisfies*

$$N \geq \frac{18 + \varepsilon}{\varepsilon^2} \cdot \frac{h + b}{\min\{h, b\}} \cdot \log \frac{2}{\delta}. \quad (14.8)$$

Then the output y^{SAA} of the SAA method satisfies

$$\Pr\left[C(y^{\text{SAA}}) \leq (1 + \varepsilon) \min_{y \geq 0} C(y)\right] \geq 1 - \delta.$$

The sample bound in Equation (14.8) is an improvement upon the bound in Equation (14.5), modulo the absolute constants in these bounds. Indeed, the sample bound in Equation (14.5)

scales linearly with $\left(\frac{h+b}{\min\{h, b\}}\right)^2$, whereas the sample bound in Equation (14.8) grows linearly

only with $\frac{h+b}{\min\{h,b\}}$. In addition, both sample bounds in Equations (14.5) and (14.8) grow linearly with $\log(2/\delta)$, and in the regime when $\varepsilon \in (0,1]$ both grow linearly with $1/\varepsilon^2$. Levi et al. (2015) improve upon the analysis in Levi et al. (2007) essentially by utilizing the Bernstein inequality, which provides a sharper concentration than the Massart inequality when $\Pr(D \leq y^{\text{SAA}})$ is close to 0 or 1.

Perhaps surprisingly, the sample bound of the SAA method by Levi et al. (2015) has an optimal dependence on the model parameters, as demonstrated by the following result by Cheung and Simchi-Levi (2019) that quantifies the number of samples needed by *any* algorithm.

Theorem 14.4 (Cheung and Simchi-Levi, 2019) *Consider an algorithm A that inputs N iid samples $d^{(1)}, \dots, d^{(N)}$ from a latent demand distribution, the coefficients b, h , and outputs the ordering amount y^A . Fix $\varepsilon \in (0, 1/20)$ and $\delta \in (0, 1/4)$. Suppose that the algorithm satisfies*

$$\Pr \left[C(y^A) \leq (1 + \varepsilon) \min_{y \geq 0} C(y) \right] \geq 1 - \delta$$

for any latent demand distribution. Then, it holds that

$$N > \frac{1}{2000\varepsilon^2} \cdot \frac{(h+b)}{\min\{h,b\}} \cdot (1-4\delta). \quad (14.9)$$

The sample bound in Equation (14.9) shows that the dependence of the SAA sample bound in Equation (14.8) on h, b is tight, and the dependence is also tight on ε, δ when $\varepsilon \in (0, 1/20)$ and $\delta \in (0, 1/4)$. The main idea behind Theorem 14.4 is to construct two latent demand distributions F_1 and F_2 , which are discrete probability distributions with the common support $\mathcal{D} = \{0, 1, 2000 \cdot \max\{1, \frac{h}{b\varepsilon}\}\}$. Let's assume that $h \leq b$ for discussion's sake, while the case of $b < h$ can be similarly handled by interchanging the roles of h, b . The two distributions F_1 and F_2 are designed so that the following two properties hold:

1. F_1, F_2 have disjoint sets of $(1 + \varepsilon)$ -optimal solutions:

$$\{y : C(y) \leq (1 + \varepsilon) C(y^*) \text{ for } F_1\} \cap \{y : C(y) \leq (1 + \varepsilon) C(y^*) \text{ for } F_2\} = \emptyset.$$

2. F_1, F_2 have a small statistical distance, as their Kullback–Leibler (KL) divergence satisfies:

$$\text{KL}(F_1 \parallel F_2) = \sum_{d \in \mathcal{D}} \Pr_{D_1 \sim F_1} [D_1 = d] \log \frac{\Pr_{D_1 \sim F_1} [D_1 = d]}{\Pr_{D_2 \sim F_2} [D_2 = d]} \leq \frac{8h\varepsilon^2}{b+h}.$$

Property (1) implies the following. Suppose $\mathbb{A}$ is an algorithm that returns a $(1 + \varepsilon)$ -optimal solution with N samples under every of the demand distributions F_1, F_2 with a probability of $1 - \delta$. Then, $\mathbb{A}$ can be used to decide if the N samples are distributed as F_1 or F_2 . More

precisely, let the DM be given N iid samples, which are known to be distributed as F_i , where $i=1$ or $i=2$, but the identity of i is not known. Then the DM can use $\mathbb{A}$ to identify i correctly with a probability of at least $1 - 2\delta$.

Property (2) is the opposite of Property (1), in the sense that if two probability distributions have a small KL divergence, then it needs a large number of samples to differentiate between them. By applying an information-theoretic argument via Pinsker's inequality (Chapter 2 in Cover and Thomas, 2006), Cheung and Simchi-Levi (2019) show that Property (2) leads to

$$\Pr(\text{An algorithm identifies } i \text{ correctly with } N \text{ samples}) \leq \sqrt{\frac{(6400 \log 2) N h \epsilon^2}{h+b}}.$$

The conclusions from Properties (1) and (2) show that the number N of samples needed is at least $\frac{1}{2000\epsilon^2} \cdot \frac{(h+b)}{h} \cdot (1-4\delta)$, and finally the case of $b > h$ can be handled similarly.

Together, Levi et al. (2015) and Cheung and Simchi-Levi (2019) pin down the optimal dependence of samples bound on model parameters. It is important to note that the sample bounds in Theorems 14.3 and 14.4 are for the case of general demand distributions, which begs the question of if the SAA method could require fewer samples when the latent demand distribution satisfies the certain benign property. A subsequent work by Zhang et al. (2021) shows that the SAA method has a better dependence on the parameter ϵ if the latent demand distribution has an increasing failure rate (IFR).

Definition 14.1 A probability distribution has an IFR, if its CDF $F: \mathbb{R} \rightarrow [0,1]$ satisfies the property that the function $H(x) = \log(1 - F(x))$ is concave in the domain $\mathbb{R}$.

Many commonly used probability distributions, such as the (truncated) normal distribution, the exponential distribution, the uniform distribution, the logistics distribution, and the Gamma distribution with a shape parameter greater than 1, have IFR. The exposition by An (1996) provides an account of the properties of probability distributions with IFR.

Theorem 14.5 (Zhang et al., 2021) Assume that the latent demand distribution's CDF is absolute continuous and has an increasing failure rate. Fix error parameter $\epsilon > 0$ and confidence parameter $\delta \in (0,1)$. Suppose that the number of N samples satisfies

$$N \geq \frac{1}{2(1 + \sqrt{(\log 2 - 0.5)\epsilon})^2} \cdot \left(\frac{h+b}{\min\{h,b\}} \right)^2 \cdot \log \frac{2}{\delta}. \quad (14.10)$$

Then the output y^{SAA} of the SAA method satisfies

$$\Pr \left[C(y^{\text{SAA}}) \leq (1 + \epsilon) \min_{y \geq 0} C(y) \right] \geq 1 - \delta.$$

Theorem 14.5 demonstrates an improvement over Theorem 14.3 on the dependence on ϵ . In the regime $\epsilon \in (0,1]$, the sample bound in the former only scales as $1/\epsilon$, while the sample bound in the latter scales as $1/\epsilon^2$. However, the sample bounds in Theorems 14.3 and 14.5 are

incomparable. Indeed, the sample bound in Theorem 14.5 grows in proportion to $\left(\frac{h+b}{\min\{h,b\}}\right)^2$, which is worse than that for Theorem 14.3, which only grows in proportion to $\frac{h+b}{\min\{h,b\}}$.

Finally, it is interesting to note that the sample lower bound in Theorem 14.4 does not apply in the setting in Theorem 14.5. Indeed, Theorem 14.4 involves discrete demand distributions, which violate the continuity assumption on the CDF of the demand distribution in Theorem 14.5. Hence, there is no contradiction in the necessity of the scaling $1/\varepsilon^2$ in Theorem 14.4 and the scaling of $1/\varepsilon$ in Theorem 14.5.

14.1.2 Sampling-Based Newsvendor: Beyond SAA

While the SAA method is a natural heuristic for the sampling-based newsvendor problem, the research community also proposes other sampling-based algorithms. We highlight two research works (Liyanage & Shanthikumar (2005); Levi et al. (2015)) that propose refinements over the SAA method.

Liyanage and Shanthikumar (2005) consider the sampling-based newsvendor problem in a parametric setting. In addition to the iid samples $d^{(1)}, \dots, d^{(N)}$ over the latent demand distribution, the DM also knows that the latent demand distribution is an exponential distribution with an unknown mean θ . That is, $\Pr(D \leq y) = 1 - \exp(-y/\theta)$ for $y \geq 0$.

If θ were to be known, then the optimal solution y^* and the optimal value $C(y^*)$ can be explicitly expressed as

$$y^* = \theta \log\left(\frac{h+b}{h}\right), \quad C(y^*) = h\theta \left(\frac{h+b}{h} - 1 - \log\left(\frac{h+b}{h}\right)\right).$$

In the current setting when θ is not known, it is natural to follow an estimate-then-optimize method:

$$y^{ETO} = \hat{\theta} \log\left(\frac{h+b}{h}\right), \text{ where } \hat{\theta} = \frac{1}{N} \sum_{n=1}^N d^{(n)}, \quad (14.11)$$

which has an expected operational cost

$$\mathbb{E}[C(y^{ETO})] = h\theta \left(\frac{h+b}{h} - \frac{h+b}{h} \left(\frac{N}{N + \log(h+b/h)} \right)^N - \log\left(\frac{h+b}{h}\right) \right).$$

The expectation is taken over by y^{ETO} , which is random since y^{ETO} depends on the iid samples. Note that y^{ETO} is an unbiased estimator of y^* , i.e., $\mathbb{E}[y^{ETO}] = y^*$, but $C(y^{ETO})$ is not an unbiased estimator of $C(y^*)$, due to the non-linearity of C .

Liyanage and Shanthikumar (2005) propose the notion of operational statistics:

$$y^{\text{OS}} = N\hat{\theta} \left[\left(\frac{h+b}{h} \right)^{1/N+1} - 1 \right].$$

Note that y^{OS} is a biased (but asymptotically unbiased) estimator of y^* . The operational cost of the operational statistics can also be expressed in closed form:

$$\mathbb{E}[C(y^{\text{OS}})] = h\theta \left[\frac{h+b}{h} - 1 - (N+1) \left[\left\{ \frac{h+b}{h} \right\}^{\frac{1}{N+1}} - 1 \right] \right].$$

While it can be verified that $\lim_{N \rightarrow \infty} \mathbb{E}[C(y^{\text{OS}})] = \lim_{N \rightarrow \infty} \mathbb{E}[C(y^{\text{ETS}})]$, Liyanage and Shanthikumar (2005) show that the operational statistic is advantageous over the estimate-then-optimize approach and the SAA method:

Theorem 14.6 (Liyanage and Shanthikumar, 2005) *Consider the sampling-based newsvendor problem, where the latent demand distribution is known to be an exponential distribution. Then it holds that*

$$\mathbb{E}[C(y^{\text{OS}})] < \mathbb{E}[C(y^{\text{ETS}})], \quad \mathbb{E}[C(y^{\text{OS}})] < \mathbb{E}[C(y^{\text{SAA}})].$$

The inequality signifies that the separation of estimation and optimization is in fact sub-optimal for the sampling-based newsvendor problem, under the exponential distribution assumption. In addition, Liyanage and Shanthikumar (2005) also derive an explicit expression for $\mathbb{E}[C(y^{\text{SAA}})]$, and they also show that the operational statistic approach incorporates the distributional information and is able to outperform the SAA method. Finally, Liyanage and Shanthikumar (2005) also remarked that the methodology of operational statistics can be generalized to other classes of demand distributions, such as the Gamma distribution with unknown scale parameters and the normal distribution with unknown variance.

Levi et al. (2015) consider the sampling-based newsvendor problem in a non-parametric setting. They show that the dependence on ε in Theorem 14.3 can be improved, when the latent demand distribution belongs to a non-parametric family, which is characterized by a novel notion called *weighted mean spread*.

Definition 14.2 Consider a newsvendor problem instance with stochastic demand D , per-unit overage and underage costs h, b . The instance associated Weighted Mean Spread (WMS) is defined as $\Delta(y^*)f(y^*)$, where y^* is the optimal solution as stated in Equation (14.2), and $\Delta(y^*) = \mathbb{E}[D | D \geq y^*] - \mathbb{E}[D | D \leq y^*]$.

The WMS $\Delta(y^*)f(y^*)$ serves as a measure of the degree of variation of $C(y)$ in a neighborhood of the optimum y^* . The intuition is that a sampling-based newsvendor instance with a larger WBS is an easier instance, in the sense that it is easier to detect the decrease in $C(y)$ in a neighborhood of y^* .

Levi et al. (2015) consider the case when a lower bound v^* to the WBS of the latent demand distribution is known. They propose the following solution, which can be seen as a conservative variant of the SAA method:

$$y^{\text{WBS}} = \min_{1 \leq j \leq N} \left\{ d^{(j)} : \frac{1}{N} \sum_{n=1}^N \mathbf{1}(d^{(n)} \leq d^{(j)}) \geq \frac{b}{b+h} + \frac{\alpha}{2(h+b)} \right\}, \quad (14.12)$$

where $\alpha = \sqrt{2\epsilon b h v^*} + \kappa \epsilon$. In the definition of α , $\epsilon > 0$ is the desired error parameter, and κ is an absolute constant to be tuned. When α is set to be 0, we have $y^{\text{WBS}} = y^{\text{SAA}}$, but with the defined α being positive, in general we have $y^{\text{WBS}} \geq y^{\text{SAA}}$. Thus, the WBS-based solution y^{WBS} is more conservative than the SAA solution, in the sense that the former orders more inventory in anticipation of the stochastic demand. Levi et al. (2015) provide the following performance guarantee:

Theorem 14.7 (Levi et al., 2015) *Consider the sampling-based newsvendor problem. Suppose that the instance satisfies two assumptions: 1) the WBS of the latent demand distribution is at least $v^* > 0$, where v^* is known, and 2) the latent demand distribution has a continuous probability density function f , which is decreasing in the domain $[y^*, \infty)$. Then it holds that*

$$\Pr \left[C(y^{\text{WBS}}) \leq (1 + \epsilon) \min_{y \geq 0} C(y) \right] \geq 1 - 2K^*(\epsilon),$$

where y^{WBS} is defined in Equation (14.12), and $K^*(\epsilon)$ satisfies $\lim_{\epsilon \rightarrow 0} \frac{K^*(\epsilon)}{\exp(-0.25N\epsilon v^*)} = 1$.

The definition of $K^*(\epsilon)$ implies that, in order to achieve an $(1 + \epsilon)$ -approximation with probability at least $1 - \delta$, the WBS-based approach requires $\approx \frac{4}{\epsilon v^*} \log \frac{2}{\delta}$ many samples. In terms of ϵ , the sample bound is an improvement to Theorem 14.3 where the sample bound scales as $1/\epsilon^2$. Nevertheless, the two sample bounds are not directly comparable in full generality since the sample bound by Theorem 14.7 only applies when ϵ is close to 0, and it also involves the crucial parameter v^* .

Levi et al. (2015) demonstrate that a wide variety of probability distributions satisfies Equation 14.2 when $b > h$, which is often the case in real-life applications. Levi et al. (2015) also highlight that any log-concave demand distribution has WBS at least $v^* = \frac{h+b}{\min\{h, b\}}$.

Definition 14.3 A probability distribution is said to be log concave, if it has a continuous Probability Density Function (PDF) f , and the function $S(x) = \log f(x)$ is concave in x .

We remark that if a probability distribution is log-concave, then it also has an IFR. Many common probability distributions are log-concave, such as the normal distribution, the uniform distribution and the logistic distribution.

In comparison to the related result by Zhang et al. (2021) in Theorem 14.5, we first remark that Theorem 14.5 requires no assumption apart from the IFR assumption, whereas Theorem 14.7 requires f to be decreasing on $[y^*, \infty)$ on top of the log-concave assumption. In terms of sample bound, the bound in Theorem 14.7 only grows linearly with $\frac{h+b}{\min\{h, b\}}$ when ϵ is sufficiently close to zero, while the bound in Theorem 14.5 grows linearly with $\left(\frac{h+b}{\min\{h, b\}} \right)^2$, but the bound in Theorem 14.5 is valid for all $\epsilon > 0$.

14.1.3 Learning-Based Newsvendor Models

Motivated by the prevalence of data-rich environments and the popularity of machine learning, recent research works propose to study the learning-based newsvendor problem. Different from the original setting where the stochastic demand is modeled as a uni-variate random variable, in the learning-based setting stochastic demand $D(x)$ is a random function of the feature vector $x \in \mathbb{R}^p$. The feature vector $\mathbf{x}$ encodes exogenous explanatory variables, such as the seasonality, the weather and the consumer price index at the time when the ordering decision is made. The feature vector $\mathbf{x}$ is observed by the DM before the ordering decision is made. The overarching goal is to find a mapping from the feature vector $\mathbf{x}$ to the ordering decision that minimizes the expected operational cost. The learning-based newsvendor problem generalizes the original formulation in Equation (14.1).

To be precise, the dynamics of the learning-based newsvendor problem are as follows. As always, we assume the DM starts with no on-hand inventory:

1. The DM observes the feature vector $x \in \mathcal{X} \subseteq \mathbb{R}^p$, where $\mathcal{X}$ denotes the set of all possible feature vectors.
2. The DM decides to order y units of inventory, which could depend on $\mathbf{x}$.
3. The DM observes the realization of the stochastic demands $D(x)$.
4. The DM accounts for the operational costs, which also involve the per-unit overage cost h and the per-unit underage cost b .

The goal of the DM is to minimize the sum of the expected underage and overage costs. Let $\mathcal{Q} \subseteq \{q : \mathcal{X} \rightarrow \mathbb{R}^+\}$ be a set of mapping from a feature vector to an ordering decision. The DM's objective is to identify a mapping in $\mathcal{Q}$ that minimizes the total cost, which is to solve

$$\min_{q: q \in \mathcal{Q}} C_{\mathcal{X}}(q(\mathbf{x}) | \mathbf{x}), \text{ where } C_{\mathcal{X}}(q(\mathbf{x}) | \mathbf{x}) = \mathbb{E}[h(q(\mathbf{x}) - D(\mathbf{x}))^+ + b(D(\mathbf{x}) - q(\mathbf{x}))^+]. \quad (14.13)$$

It is clear that the classical newsvendor problem in Equation (14.1) is a specialization of the optimization Equation (14.13), by specializing $\mathcal{X}$ to be a singleton set, so that the collection $\mathcal{Q} = \{q : \mathcal{X} \rightarrow \mathbb{R}^+\}$ specializes to the set of possible ordering quantities, which is $\mathbb{R}^+$. The expectation in Equation (14.13) is taken over the randomness in demand $D(\mathbf{x})$, where the feature vector $\mathbf{x}$ is fixed and is deterministic.

At first sight, minimization over the set of mapping $\mathcal{Q}$ could appear unnecessary, since the optimization Equation (14.13) only concerns a particular feature vector $\mathbf{x}$, and it suffices to compute $q(\mathbf{x})$ instead of producing the whole function q . The rationale behind the formulation is that, even in a data-rich setting, it is highly unlikely that the DM has access to samples for each and every feature vector $\mathbf{x}$, since the feature set $\mathcal{X} \subseteq \mathbb{R}^p$ (if finite) typically has size exponential in p . Rather, it is desirable to follow a machine-learning approach, which is to generalize the ordering decisions contingent upon the feature vectors $\{\mathbf{x}_n\}_{n=1}^N$ in the data to a mapping q , which is from the whole of feature space $\mathcal{X}$ to ordering decisions $\mathbb{R}_{\geq 0}$.

Ban and Rudin (2019) consider the learning-based newsvendor problem in a data-driven setting, which shares similarity to the setting of a supervised learning problem. The DM is given a set of data $S = \{(\mathbf{x}^{(n)}, d^{(n)})\}_{n=1}^N$, where for each $n \in \{1, \dots, N\}$ we have $\mathbf{x}^{(n)} \in \mathcal{X}$, and $d^{(n)}$

is a realization of the random variable $D(\mathbf{x}^{(n)})$. Ban and Rudin (2019) propose to compute a mapping $q \in \mathcal{Q}$ by solving the associated Empirical Risk Minimization (ERM) problem:

$$\min_{q \in \mathcal{Q}} \hat{C}_{\mathcal{X}}(q | S), \text{ where } \hat{C}_{\mathcal{X}}(q | S) = \frac{1}{N} \sum_{n=1}^N [h(q(x^{(n)}) - d^{(n)})^+ + b(d^{(n)} - q(x^{(n)}))^+]. \quad (14.14)$$

The high-level idea behind the Equation (14.14) is to replace the latent distribution on $D(\cdot)$ by its empirical distribution, and it incorporates the feature vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$ in order to generalize the mapping learned on these sample feature vectors to whole set $\mathcal{X}$ of feature vectors. It can be verified that Equation (14.14) specializes in the SAA problem in Equation (14.3) when $\mathcal{X}$ is a singleton.

To gain tractability and insight, Ban and Rudin (2019) propose to focus on a case when $\mathcal{Q}$ consists of linear functions on $\mathbb{R}^p$. That is, for each $q \in \mathcal{Q}$, we have $q(\mathbf{x}) = \sum_{j=1}^p q_j x_j$ (for a vector $\mathbf{x} \in \mathbb{R}^p$, we denote x_j as the j -th coordinate of $\mathbf{x}$). Therefore, without loss of generality we identify $\mathcal{Q}$ as a subset of $\mathbb{R}^p$, and each mapping $q \in \mathcal{Q}$ is identified with a vector in $\mathbb{R}^p$. In addition, for each $\mathbf{x} \in \mathcal{X}$, we assume without loss of generality that $x_1 = 1$, that is, the first coordinate of each feature vector is equal to 1. Consequently, the ERM problem in Equation (14.14) can be formulated as the following linear program:

$$\begin{aligned} \min_{q \in \mathcal{Q}} \quad & \frac{1}{N} \sum_{n=1}^N h \cdot o^{(n)} + b \cdot u^{(n)} & (\text{LP-ERM}) \\ \text{s.t.} \quad & u^{(n)} \geq d^{(n)} - q_1 - \sum_{j=2}^p q_j x_j^{(n)} & \forall n \in \{1, \dots, N\} \\ & o^{(n)} \geq q_1 + \sum_{j=2}^p q_j x_j^{(n)} - d^{(n)} & \forall n \in \{1, \dots, N\} \\ & o^{(n)}, u^{(n)} \geq 0 & \forall n \in \{1, \dots, N\}. \end{aligned}$$

In the linear program (LP-ERM), the number of decision variables and the number of constraints are both linear in N , the number of samples. Ban and Rudin (2019) observe that, while (LP-ERM) yields algorithmically stable decisions, the resulting theoretical performance guarantee could be loose when p is large relative to sample size N . Consequently, Ban and Rudin (2019) propose to incorporate regularization into (LP-ERM) by formulating another optimization problem (Reg-ERM). The problem (Reg-ERM) has the same set of decision variables and feasible region as (LP-ERM), but the objective function of (Reg-ERM) is formulated as

$$\min_{q \in \mathcal{Q}} \lambda \cdot \|q\|_{\ell}^2 + \frac{1}{N} \sum_{n=1}^N h \cdot o^{(n)} + b \cdot u^{(n)}.$$

The parameter $\lambda \geq 0$ is a regularization parameter, and $\|\cdot\|_\ell$ is the ℓ -norm. When $\ell = 2$, the problem (LP-ERM) is a convex quadratic program. By specifying $\ell = 0$ or $\ell = 1$, it could lead to a sparse optimal solution, that is, a solution with few non-zero entries. While setting $\ell = 1$ still results in a convex program, setting $\ell = 0$ in general leads to a mixed-integer program.

Let's denote the optimal solutions to (LP-ERM) and (Reg-ERM) as q^{LP}, q^{Reg} , respectively. Ban and Rudin (2019) analyze the out-of-sample performances of these two solutions under the following set of assumptions:

1. The samples $(\mathbf{x}^{(1)}, d^{(1)}), \dots, (\mathbf{x}^{(N)}, d^{(N)})$ are iid, with common probability distribution $\mathcal{D}$. For $(\mathbf{X}, D(\mathbf{X}))$ distributed as $\mathcal{D}$, we have $\mathbf{X} \in \mathcal{X}$ with certainty, and the pair $(\mathbf{X}, D(\mathbf{X}))$ satisfies the following linear relation with certainty:

$$D(\mathbf{X}) = \beta^T \mathbf{X} + W. \quad (14.16)$$

The random variable W is independent of $\mathbf{X}$, and $W \in [\underline{D}, \bar{D}] \subset (0, \infty)$ almost surely.

2. For $(\mathbf{X}, D(\mathbf{X}))$ distributed as $\mathcal{D}$, the random vector $\mathbf{X}$ is normalized with certainty. That is, with certainty $\mathbf{X}_1 = 1$. In addition, for each $j \in \{2, \dots, p\}$, the uni-variate random variable $\mathbf{X}_j$ has mean 0 and standard deviation 1. Moreover, $\|\mathbf{X}\|_2 \leq \sqrt{p} X_{\max}$ for some absolute constant $X_{\max}$ almost surely.
3. $\mathcal{Q}$ is a compact convex set.

These assumptions are mild assumptions that are similar to standard assumptions in linear regression problems with random design. Ban and Rudin (2019) quantify the out-of-sample error of q^{LP} as stated in the following theorem:

Theorem 14.8 (Ban and Rudin, 2019) Consider a fresh sample $(\mathbf{x}^{(N+1)}, d^{(N+1)}) \sim \mathcal{D}$, where the DM only observes $\mathbf{x}^{(N+1)}$ but not the actual demand $d^{(N+1)}$ when the DM is about to make the ordering decision. Denote $y^{LP} = q^{LP}(\mathbf{x}^{(N+1)})$. The out-of-sample error $|\mathbb{E}[C_X(y^{LP} | \mathbf{x}^{(N+1)})] - \mathbb{E}[C_X(y^{LP} | S)]|$ is at most

$$\max\{h, b\} \bar{D} \left[\frac{2 \max\{h, b\}}{\min\{h, b\}} \frac{p}{N} + \left(\frac{4 \max\{h, b\}}{\min\{h, b\}} + 1 \right) \sqrt{\frac{\log(2/\delta)}{N}} \right] \quad (14.17)$$

$$+ \max\{h, b\} K \cdot \frac{\sqrt{\log N}}{N^{1/(2+p/2)}} \quad (14.18)$$

with probability $1 - \delta$. The parameter K is equal to $\sqrt{\frac{9(8+5p)}{4+p}} \frac{1}{(1-2^{-4/(4+p)})\lambda_2^*}$, and

$\lambda_2^* = \min_{d \in [\underline{D}, \bar{D}]} f_W(t)$, where f_W is the PDF of the noise term W defined in Equation (14.16).

The expectations are taken over by both S and $(\mathbf{x}^{(N+1)}, d^{(N+1)})$. The first term in Equation (14.17) accounts for the generalization error, which is the error due to the discrepancy between the training sample set $S = \{(\mathbf{x}^{(n)}, d^{(n)})\}_{n=1}^N$ and the fresh sample $(\mathbf{x}^{(N+1)}, d^{(N+1)})$. Even when $(\mathbf{x}_1, d_1), \dots, (\mathbf{x}_{N+1}, d_{N+1})$ are iid, there could still be a discrepancy between S and $(\mathbf{x}^{(N+1)}, d^{(N+1)})$.

due to their stochastic variations. The second term in Equation (14.18) accounts for the distance between ordering decision y^{LP} and the optimal ordering decision.

Ban and Rudin (2019) also quantify the out-of-sample error of q^{Reg} as stated in the following theorem:

Theorem 14.9 *Consider a fresh sample $(\mathbf{x}^{(N+1)}, d^{(N+1)}) \sim \mathcal{D}$, where the DM only observes $\mathbf{x}^{(N+1)}$ but not actual demand $d^{(N+1)}$ when the DM is about to make the ordering decision. Denote $y^{Reg} = q^{Reg}(\mathbf{x}^{(N+1)})$. The out-of-sample error $|\mathbb{E}[C_X(y^{Reg} | \mathbf{x}^{(N+1)})] - \mathbb{E}[\hat{C}_X(y^{Reg} | S)]|$ is at most*

$$\begin{aligned} & \max\{h, b\} \bar{D} \left[\frac{\max\{h, b\} X_{\max}^2}{\lambda} \frac{p}{N \underline{D}} + \left(\frac{2 \max\{h, b\} X_{\max}^2 p}{\lambda \bar{D}} + 1 \right) \sqrt{\frac{\log(2/\delta)}{N}} \right] \\ & + \max\{h, b\} \mathbb{E}[|y^{LP} - y^{Reg}|] + \max\{h, b\} K \frac{\sqrt{\log N}}{N^{1/(2+p/2)}} \end{aligned}$$

with probability $1 - \delta$. The parameter λ is the regularization parameter, and the parameter K is the same as that in Theorem 14.8.

The error terms have a similar interpretation to those in Theorem 14.8. We remark that apart from the linear regression-based approach highlighted previously, Ban and Rudin (2019) also propose a non-linear regression-based approach such as the kernel-optimization method, which is based on the Nadaraya–Watson kernel regression (Nadaraya (1964); Watson (1964)).

Finally, in addition to the research work by Ban and Rudin (2019), Oroojlooyjadid et al. (2020) propose a deep learning approach to the learning-based newsvendor problem, by training multi-layer perceptrons with the loss function being the newsvendor cost function. Altogether, the learning-based newsvendor problem embodies a fundamental problem that connects supervised learning methods with optimization problems formulation, which has been a flourishing line of research recently (for example, see Donti et al. (2017); Bertsimas and Kallus (2020); Elmachtoub and Grigas (2021); Gupta and Rusmevichientong (2021)).

14.2 MULTIPLE PERIOD SETTING: INVENTORY CONTROL MODEL

After surveying the existing works on the data-driven newsvendor model, which presents a single-period setting, we proceed to survey its various multi-period generalizations. We focus on the sampling-based setting, which is studied by the majority of the existing relevant research works.

We start with the sampling-based inventory control models with uncapacitated and capacitated settings, which were studied by Levi et al. (2007) and Cheung and Simchi-Levi (2019), respectively. Second, we review a sampling-based inventory control model with pricing decisions by Qin et al. (2019). Third, we review a sampling-based inventory control model on serial systems by Zhang et al. (2021). Finally, we survey a sampling-based model with censored demand data by Ban (2020).

14.2.1 Models for the Uncapacitated and Capacitated Settings

We first survey Levi et al. (2007) and Cheung and Simchi-Levi (2019), who, respectively, study an uncapacitated sampling-based inventory control problem and its capacitated generalization. We begin with the inventory model in the capacitated setting, and highlight the specialization to the uncapacitated case. The DM faces a finite time horizon with T discrete time periods, labeled as $1, \dots, T$. The agent starts with an inventory level $x_1=0$ at period 1. From period 1 to T , the DM performs the following actions:

1. The DM observes the starting inventory level x_t .
2. The DM orders $y_t - x_t$ units of inventory, where $0 \leq y_t - x_t \leq B_t$. The parameter B_t is the capacity of the inventory that can be ordered in the t^{th} period. In the uncapacitated setting Levi et al. (2007), we have $B_t = \infty$ for all t , while in the capacitated setting Cheung and Simchi-Levi (2019), each B_t is finite.
3. The DM observes the t^{th} period demand D_t .
4. If $y_t > D_t$, the DM incurs a linear holding cost of $h_t \times (y_t - D_t)$; else if $y_t \leq D_t$, the DM incurs a linear backlog cost $b_t \times (D_t - y_t)$. In the latter case, the unsatisfied demands are backlogged.
5. The DM proceeds to period $t+1$, with the starting inventory level being $x_{t+1} = y_t - D_t$.

In the model, unfulfilled demands are backlogged, and there is no lead time. The stochastic demands $D_1, \dots, D_T$ are independent, though not necessarily identically distributed. The DM aims to design a policy that minimizes her expected total operational cost

$$\mathbb{E} \left[\sum_{t=1}^T h_t (y_t - D_t)^+ + b_t (D_t - y_t)^+ \right], \quad (14.19)$$

subject to the capacity constraint in each period. For each time period $t = 1, \dots, T$, we assume that $h_t, b_t > 0$. When $T=1$ and $B_1 = \infty$, we recover the classical newsvendor problem.

14.2.1.1 (Modified) base-stock policies

Existing research works by Kapuściński and Tayur (1998) and Tayur (1993) show that the inventory control problem can be optimized by a *base-stock policy* with an uncapacitated setting, and by a *modified base-stock policy* (MBS) with a capacitated setting. We first define an MBS policy, and then illustrate how it specializes to a base-stock policy.

Definition 14.4 Under a MBS policy, which is parameterized by $(R_1, \dots, R_T)$, at period t the DM determines order-up-to level y_t in the following manner:

$$y_t = \begin{cases} x_t + B_t & \text{if } x_t \in (-\infty, R_t - B_t], \\ R_t & \text{if } x_t \in (R_t - B_t, R_t], \\ x_t & \text{if } x_t \in (R_t, \infty). \end{cases}$$

In other words, at each period t , the DM makes the inventory level y_t as close to R_t as possible, under capacity constraints. Under an MBS policy $(R_1, \dots, R_T)$, the decision made in period t is

only dependent on the amount of inventory x_t on hand and the modified base-stock R_t , but it does not depend on the other modified base-stocks and the observations made in the previous periods. In the uncapacitated case when $B_t = \infty$ for each t , the MBS policy specializes to a base-stock policy, which is also parameterized by $(R_1, \dots, R_T)$. At period t , the DM determines order-up-to level y_t in the following manner:

$$y_t = \begin{cases} R_t & \text{if } x_t \in (-\infty, R_t], \\ x_t & \text{if } x_t \in (R_t, \infty). \end{cases}$$

As in Kapuściński and Tayur (1998) and Tayur (1993), in the capacitated setting, there exists an optimal MBS policy $(R_1^*, \dots, R_T^*)$ under which the expected operational cost in Equation (14.19) is minimized. Similarly, with an uncapacitated setting, there exists an optimal base-stock policy under which the expected operational cost in Equation (14.19) is minimized. The derivation of the optimality of the (modified) base-stock policy is useful for the algorithm design and analysis in Levi et al. (2007) and Cheung and Simchi-Levi (2019). Thus, we review the derivation below. An optimal policy can be constructed by solving the following Bellman equations from $t=T$ to $t=1$:

$$V_t(x_t) = \min_{x_t \leq y_t \leq x_t + B_t} C_t(y_t) + \mathbb{E}[V_{t+1}(y_t - D_t)], \quad V_{T+1}(x_{T+1}) = 0.$$

The function $C_t(y_t) = \mathbb{E}[h_t(y_t - D_t)^+ + b_t(D_t - y_t)^+]$ is the t^{th} period expected operational cost. To facilitate the backward induction, we introduce the function U_t :

$$U_t(y_t) = C_t(y_t) + \mathbb{E}[V_{t+1}(y_t - D_t)]. \quad (14.19)$$

Thus, we have

$$V_t(x_t) = \min_{x_t \leq y_t \leq x_t + B_t} U_t(y_t). \quad (14.20)$$

The function $V_t(x_t)$ represents the expected operational cost over periods $t, \dots, T$ when the starting inventory level in period t is x_t , and the DM orders optimally in the periods $t, \dots, T$. The function $U_t(y_t)$ represents the expected operational cost over periods $t, \dots, T$ when the inventory level after ordering is y_t in period t , and the DM orders optimally in the periods $t+1, \dots, T$.

By a backward induction from $t=T$ to $t=1$, Aviv and Federgruen (1997) and Kapuściński and Tayur (1998) further show that

1. The functions U_t, V_t are convex for all t ,
2. The MBS policy $(R_1^*, \dots, R_T^*)$, where $R_t^* \in \operatorname{argmin}_{y_t \in \mathbb{R}} U_t(y_t)$, is optimal.

To this end, we remark that the convexity of U_t, V_t is not only important for the inventory control to be tractable, but it is also important for demonstrating the structural result that the optimum can be achieved by a modified base-stock policy.

14.2.1.2 Sampling-based setting

After describing the model and the optimal policy, we formally describe the sampling-based setting. The DM does not know the demand distributions. Rather, the DM only has access to a set of independent samples drawn from the latent demand distributions. More precisely, the DM has access to $N_t \geq 1$ iid samples $d_t^{(1)}, \dots, d_t^{(N_t)}$, which are identically distributed as the period t demand distribution. Similar to the sampling-based newsvendor setting, we assume that the expectation $\mathbb{E}[|D_t|]$ is finite for all $t \in \{1, \dots, T\}$, which is necessary for the problem to be well-defined. This is the only assumption made on $D_1, \dots, D_T$. In particular, we neither assume that the demand distributions are parametrized, nor assume that they have bounded supports.

14.2.2 Algorithms for the Uncapacitated Setting

Levi et al. (2007) construct an algorithm for the uncapacitated case ($B_t = \infty$ for all t), where the algorithm can be interpreted as a perturbed version of the SAA method. The algorithm returns a set of base-stocks $(\tilde{R}_1, \dots, \tilde{R}_T)$. The perturbation is applied in order to maintain the convexity of the cost-to-go functions in the sampling-based setting, where the convexity is crucial for demonstrating the near optimality of the output base-stocks.

The algorithm by Levi et al. (2007) involves solving the following *shadow dynamic program*, which loosely follows the derivation of the optimal base-stock policy in the previous section. In addition to the cost parameters $\{h_t, b_t\}_{t=1}^T$ and the samples, their algorithm also requires an error parameter ε as an input.

To this end, let's define

$$\hat{C}_t(y_t) = \frac{1}{N_t} \sum_{n=1}^{N_t} h_t(y_t - d_t^{(n)})^+ + b_t(d_t^{(n)} - y_t)^+, \quad (14.21)$$

which is the empirical operational cost in the t^{th} period. The base-stocks $\tilde{R}_T, \dots, \tilde{R}_1$ are constructed based on a backward induction, involving the construction of shadow cost-to-go functions $\tilde{U}_T, \dots, \tilde{U}_1$ and $\tilde{V}_T, \dots, \tilde{V}_1$. These functions mirror the cost-to-go functions $\{U_t, V_t\}_{t=1}^T$ in the derivation of the optimal base-stock policy. First, for T , the DM sets $\tilde{U}_T = \hat{C}_T$. Then, the DM sets

$$\tilde{R}_T = \min_{y \geq 0} \left\{ y : \partial^+ \tilde{U}_T(y) \geq \frac{\varepsilon}{4T} \right\}, \quad \tilde{V}_T(x_T) = \min_{y_T \geq x_T} \tilde{U}_T(y_T).$$

Inductively, assuming that the DM has constructed $\{\tilde{U}_\tau, \tilde{R}_\tau, \tilde{V}_\tau\}_{\tau=t}^T$, where $T \geq t \geq 2$, the DM constructs

$$\tilde{U}_{t-1}(y_{t-1}) = C_{t-1}(y_{t-1}) + \frac{1}{N_{t-1}} \sum_{n=1}^{N_{t-1}} \tilde{V}_t(y_{t-1} - d_{t-1}^{(n)}),$$

$$\tilde{R}_{t-1} = \min_{y \geq 0} \left\{ y : \partial^+ \tilde{U}_{t-1}(y) \geq \frac{\varepsilon}{4T} \right\}, \quad \tilde{V}_{t-1}(x_{t-1}) = \min_{y_{t-1} \geq x_{t-1}} \tilde{U}_{t-1}(y_{t-1}).$$

Finally, the DM collects the base-stocks $\tilde{R}_1, \dots, \tilde{R}_T$ obtained, and uses them as the base-stock policy. It is interesting to note that Levi et al. (2007) apply the perturbation in the computation of each $\tilde{R}_t$. Instead of trying to solve the empirical problem optimally by choosing $\tilde{R}_t = \min_{y \geq 0} \{y : \partial^+ \tilde{U}_t(y) \geq 0\}$, the choice of $\tilde{R}_t$ ensures that $\tilde{R}_t \geq \bar{R}_t$, which is the key to ensuring the convexity of their shadow cost-to-go functions.

After describing the algorithm by Levi et al. (2007), we provide the following theorem that demonstrates the near optimality of their algorithm. To this end, we say that a set of base-stocks $(R_1, \dots, R_T)$ is $(1 + \varepsilon)$ -optimal if the expected operational cost under the base-stock policy $(R_1, \dots, R_T)$ is at most $(1 + \varepsilon)$ times the optimal expected operational cost. For example, by definition, $(R_1^*, \dots, R_T^*)$ is 1-optimal.

Theorem 14.10 (Levi et al., 2007) *Consider the sampling-based inventory control problem, and fix error parameter $\varepsilon > 0$, confidence parameter $\delta \in (0, 1)$. Suppose for each t we have*

$$N_t \geq 72 \frac{T^2}{\varepsilon^2} \left(\frac{h_t + b_t}{\min\{h_t, b_t\}} \right)^2 \sum_{j=1}^t (T - j + 1)^2 \log \frac{2T}{\delta}.$$

Then the base-stock policy $(\tilde{R}_1, \dots, \tilde{R}_T)$ is $(1 + \varepsilon)$ -optimal with probability at least $1 - \delta$.

When T is set to be 1, we recover the sample bound in Theorem 14.1 for the sampling-based newsvendor problem, modulo the difference in the absolute constants.

14.2.3 Algorithms for the Capacitated Setting

Subsequently, Cheung and Simchi-Levi (2019) propose to study a capacitated generalization of Levi et al. (2007), where $B_1, \dots, B_T$ are now finite. They consider the traditional SAA approach, and directly analyze the SAA method. To this end, let's denote $SAA(T; N_1, \dots, N_T)$ as the *sample average approximation* counterpart of the original capacitated inventory control problem. The empirical problem $SAA(T; N_1, \dots, N_T)$ is the T -period problem where the t^{th} period demand distribution $\hat{D}_t$ is the empirical distribution for D_t constructed with the N_t samples:

$$\Pr(\hat{D}_t = d) = \frac{1}{N_t} \sum_{n=1}^{N_t} \mathbf{1}[d = d_t^{(n)}]. \quad (14.22)$$

The optimal cost of $SAA(T; N_1, \dots, N_T)$ is a random variable that depends on the random samples drawn.

The SAA method involves solving the SAA problem $SAA(T; N_1, \dots, N_T)$ for a set of modified base-stocks $(\hat{R}_t)_{t=1}^T$. For each t , the modified base-stock $\hat{R}_t$ is chosen to be the *smallest* minimizer of in the empirical cost-to-go function $\hat{U}_t$ for all t in the sample average problem $SAA(T; N_1, \dots, N_T)$, where $\hat{U}_t$ is defined in an analogous way to U_t :

$$\hat{U}_t(y_t) = \hat{C}_t(y_t) + \frac{1}{N_t} \sum_{n=1}^{N_t} \hat{V}_{t+1}(y_t - d_t^{(n)}), \quad (14.23)$$

$$\hat{V}_t(x_t) = \min_{x_t \leq y_t \leq x_t + B_t} \hat{U}_t(y_t).$$

In Equation (14.23), the function $\hat{C}_t(y_t)$ is equal to what is defined in (14.22).

Cheung and Simchi-Levi (2019) provide the following performance guarantee for the SAA method:

Theorem 14.11 (Cheung and Simchi-Levi, 2019) Fix error parameter $\varepsilon > 0$ and confidence parameter $\delta \in (0, 1)$. Consider the sample average approximation problem $SAA(T; N_1, \dots, N_T)$, where N_t satisfies

$$N_t \geq \max \left\{ (h_t + b_t)^2, \left(\sum_{s=t+1}^T h_s + b_s \right)^2 \right\} \left\{ \frac{144T^4}{\varepsilon^2 \min_{s \in \{1, \dots, T\}} \{ \min \{ h_s, b_s \} \}^2} \log \frac{4T}{\delta} \right\}. \quad (14.24)$$

With probability at least $1 - \delta$, the modified base-stock policy $(\hat{R}_1, \dots, \hat{R}_T)$ is $(1 + \varepsilon)$ -optimal to the original problem.

While Cheung and Simchi-Levi (2019) analyze a more general setting of a capacitated model than Levi et al. (2007), in Theorem 14.11 the number of samples needed for each period is proportional to T^6 , which is worse than that of T^5 in Theorem 14.10.

While Theorem 14.11 provides a sample bound to the SAA method, it is still not a full solution to the sampling-based capacitated inventory control problem. Indeed, it turns out that the underlying SAA problem $SAA(T; N_1, \dots, N_T)$ is computationally intractable:

Lemma 14.1 Consider the stochastic capacitated inventory control problem, where the demand distributions $D_1, \dots, D_T$ are explicitly given, and each of D_t has a discrete support $\{0, a_t\}$. If there is an algorithm that runs in time polynomial in T and returns an optimal modified base-stock policy, then $\mathbf{P} = \#\mathbf{P}$.

To overcome the intractability, Cheung and Simchi-Levi (2019) provide the following algorithm that sparsifies the cost-to-go functions in the SAA problem. Indeed, the intractability of the SAA problem is due to the fact that each cost-to-go function $\hat{U}_t$, while being piece-wise linear, could have 2^{T-t} linear pieces. Consequently, Cheung and Simchi-Levi (2019) propose to sparsify the cost-to-go function with Algorithm 14.2:

Algorithm 14.2 Algorithm SPARSIFY $(\eta, N_1, \dots, N_T)$

1. For each $t \in \{1, \dots, T\}$, draw N_t independent samples $d_t^1, \dots, d_t^{N_t}$ from D_t .

2. Construct the empirical distribution $\hat{D}_t; \Pr[D_t = d] = \frac{\sum_{i=1}^{N_t} \mathbf{1}[d = d_t^i]}{N_t}$.

3. Define $\tilde{V}_{T+1}^+(x) = 0$ for all x .

(Continued)

(Continued)

4. **for** $t = T, \dots, 1$ **do**5. Construct the right derivative function $\tilde{U}_t^+(y_t) = \hat{C}_t^+(y_t) + \mathbb{E}\tilde{V}_{t+1}^+(y_t - \hat{D}_t)$.6. By a binary search on the break points of $\tilde{U}_t^+$, compute the smallest $\tilde{R}_t \in \mathbb{R}$ such that $\tilde{U}_t^+(\tilde{R}_t) \geq 0$.7. Construct the following right derivative function $\hat{V}_t^+ : \mathbb{R} \rightarrow [-\sum_{s=t}^T b_s, \sum_{s=t}^T h_s]$

$$\hat{V}_t^+(x_t) = \begin{cases} \tilde{U}_t^+(x_t + B_t) & \text{if } x_t \in (-\infty, \tilde{R}_t - B_t) \\ 0 & \text{if } x_t \in [\tilde{R}_t - B_t, \tilde{R}_t) \\ \tilde{U}_t^+(x_t) & \text{if } x_t \in [\tilde{R}_t, \infty) \end{cases}.$$

8. (Sparsification) Now, for each x_t , define $\tilde{V}_t^+(x_t) = \eta \lfloor \frac{1}{\eta} \hat{V}_t^+(x_t) \rfloor$.9. **end for**10. Return the base-stocks $(\tilde{R}_1, \dots, \tilde{R}_T)$.

The parameter η controls the degree of sparsification. A larger η leads to a coarser sparsification of the cost-to-go functions. It turns out that the following choice of η leads to a computationally tractable algorithm while maintaining the near optimality:

$$\eta = \frac{\varepsilon \min_{t \in \{1, \dots, T\}} \{\min\{h_t, b_t\}\}}{6T^2}. \quad (14.25)$$

Theorem 14.12 (Cheung and Simchi-Levi, 2019) Fix $\varepsilon > 0, 0 < \delta < 1$. Suppose N_t satisfies the sample bound in Equation (14.25) for each t , and η is as defined in Equation (14.26). Then, the Algorithm SPARSIFY $(\eta, N_1, \dots, N_T)$ produces a set of $(1 + 2\varepsilon)$ -optimal modified base-stocks $(\tilde{R}_1, \dots, \tilde{R}_T)$ with probability $1 - \delta$. The algorithm has a running time polynomial in the quantities

$$\max_{t \in \{1, \dots, T\}} \left\{ \frac{h_t + b_t}{\min\{h_t, b_t\}} \right\}, T, \frac{1}{\varepsilon}, \log\left(\frac{1}{\delta}\right), \log(d_{\max} c^*),$$

where $d_{\max} = \max_{t, n_t} \{d_t^{(n_t)}\}$, and $c^* = \max_t \{\max\{h_t, b_t\}\}$.

Algorithm 14.2 still has a pseudo-polynomial running time in terms of the quantity $\max_{t \in \{1, \dots, T\}} \left\{ \frac{h_t + b_t}{\min\{h_t, b_t\}} \right\}$. To be a truly polynomial time algorithm, the proposed algorithm should have a running time logarithmic in $\max_{t \in \{1, \dots, T\}} \left\{ \frac{h_t + b_t}{\min\{h_t, b_t\}} \right\}$. Nevertheless, the pseudo-

polynomial dependence is not a limitation to Algorithm 14.2, since by the lower bound result in Theorem 14.4, even when $T = 1$ the number of samples needed for $(1 + \varepsilon)$ -optimality grows

at least linearly with $\frac{1}{\varepsilon^2} \cdot \frac{h_1 + b_1}{\min\{h_1, b_1\}}$.

14.2.4 Inventory Control with Pricing Decisions

Qin et al. (2019) consider a pricing and inventory control problem in a sampling-based setting. In addition to the cost parameters $\{h_t, b_t\}_{t=1}^T$, the DM is also endowed with a price range $[p_t^{\min}, p_t^{\max}]$ for each $t \in \{1, \dots, T\}$. Similar to before, the DM starts with no inventory, i.e., $x_1 = 0$. At period $t \in \{1, \dots, T\}$, the following occurs:

1. The DM observes the starting inventory level x_t .
2. The DM makes two decisions: 1) the ordering amount $y_t - x_t \in \mathbb{R}_{\geq 0}$, and 2) the price to charge $p_t \in [p_t^{\min}, p_t^{\max}]$.
3. The DM observes the t^{th} period demand $D_t = \Delta_t(p_t) + \eta_t$.
 - The function $\Delta_t : [p_t^{\min}, p_t^{\max}] \rightarrow \mathbb{R}_{\geq 0}$ is the mean demand function.
 - The quantity η_t is a zero mean continuous random variable. We assume $\eta_t \in [\omega_t^{\min}, \omega_t^{\max}]$ with certainty.
4. The DM earns a revenue of $p_t \times D_t(p_t)$.
5. If $y_t > D_t$, the DM incurs a linear holding cost of $h_t \times (y_t - D_t)$; else if $y_t \leq D_t$, the DM incurs a linear backlog cost $b_t \times (D_t - y_t)$. In the latter case, the unsatisfied demand is backlogged.
6. The DM proceeds to period $t+1$, with the starting inventory level being $x_{t+1} = y_t - D_t$.

The DM aims to maximize the expected profit, which is the difference between the expected revenue and the expected operational cost:

$$\mathbb{E} \left[\sum_{t=1}^T p_t \Delta_t(p_t) \right] - \mathbb{E} \left[\sum_{t=1}^T h_t (y_t - D_t)^+ + b_t (D_t - y_t)^+ \right].$$

In the sampling-based setting, the mean demand functions $\{\Delta_t\}_{t=1}^T$ and the random variables $\{\eta_t\}_{t=1}^T$ are latent. Rather, the DM only has the following partial information about the random demand at each t :

- For each t , the DM has access to N_t samples $\{(p_t^{(n)}, d_t^{(n)})\}_{n=1}^{N_t}$, where $d_t^{(n)} = \Delta_t(p_t^{(n)}) + \eta_t^{(n)}$. The noise terms $\{\eta_t^{(n)}\}_{n=1}^{N_t}$ are iid with probability distribution same as η_t .
- For each t , it holds that $\Delta_t(\cdot) = \sum_{k=1}^{K_t} \theta_{t,k}^* \Delta_t^{(k)}(\cdot)$. While the set $\Phi_t = \{\Delta_t^{(k)}\}_{k=1}^{K_t}$ is known to the DM, the coefficients $\{\theta_{t,k}^*\}_{k=1}^{K_t}$ are latent.

Qin et al. (2019) propose an algorithm $\pi^{\text{LR+SAA}}$ that combines linear regression on the latent coefficients $\{\theta_t^*\}_{t=1}^T$, with the SAA method. For each $t \in \{1, \dots, T\}$, they propose to compute the estimator $\hat{\theta}_t = (\hat{\theta}_{t,k})_{k=1}^{K_t}$ to θ_t^* by considering the minimization of the mean squared error:

$$\hat{\theta}_t \in \underset{\tilde{\theta}_t = (\tilde{\theta}_{t,k})_{k=1}^{K_t} \in \mathbb{R}^{K_t}}{\text{argmin}} \left\{ \sum_{n=1}^{N_t} \left(d_t^{(n)} - \sum_{k=1}^{K_t} \tilde{\theta}_{t,k} \Delta_t^{(k)}(p_t^{(n)}) \right)^2 \right\}.$$

After that, they replace the latent mean demand function Δ_t with $\hat{\Delta}_t = \sum_{k=1}^{K_t} \hat{\theta}_{t,k} \Delta_t^{(k)}$ to formulate the SAA problem. Finally, they apply the SAA method, and return the policy that optimizes for the SAA problem.

Under regularity conditions on the mean demand functions, Qin et al. (2019) demonstrate the following sample bound:

Theorem 14.13 (Qin et al., 2019) *Under regularity conditions, if $N_t \geq B \cdot \frac{T^6}{\varepsilon^2} \log \frac{T}{\delta}$ holds for all t , then the proposed algorithm π^{LR+SAA} satisfies*

$$\Pr\left(\left| \text{Expected Profit under } \pi^{LR+SAA} - \text{Expected Optimum} \right| \leq \varepsilon\right) \geq 1 - \delta,$$

where the probability measure $\Pr$ is over the training samples used for constructing π^{LR+SAA} .

Factor B depends on the latent demand distributions and the constants associated with the underlying regularity assumptions, which are detailed in Qin et al. (2019). Different from Levi et al. (2007), Cheung and Simchi-Levi (2019) and Qin et al. (2019) consider the absolute difference rather than the ratio between the expected profit of π^{LR+SAA} and the optimum. Indeed, the notion of approximation ratio is inapplicable when the objective value can be negative, which is the case in the pricing model (since the optimal profit could be negative), but not the models in Levi et al. (2007) and Cheung and Simchi-Levi (2019), where the objective value is the expected operational cost.

14.2.5 Inventory Control on Serial Systems

Zhang et al. (2021) consider inventory control in a serial inventory system with I stages, in a sampling-based setting. Zhang et al. (2021) focus on the infinite horizon setting, where the stochastic demands across different time periods are iid, different from Levi et al. (2007), Cheung and Simchi-Levi (2019), and Qin et al. (2019) who consider finite time horizons and non-identical demand distributions.

In each period t , stage $i \in \{1, \dots, I\}$ orders inventory from stage $i+1$, and stage I orders from an external supplier with ample inventory. There is a deterministic lead time for any stage to receive the shipment from its immediate upstream supplier. Stochastic demands only appear at stage 1, but not at any other stage.

At the beginning of period t , each stage receives its shipment due to the current period. After that, demand D_t is realized, and the demand is satisfied from stage 1's on-hand inventory as much as possible. Unfulfilled demands at stage 1 are fully backlogged. The DM incurs backlog cost at stage 1, in proportion to the amount of demands backlogged. The DM also incurs holding costs at each stage $i \in \{1, \dots, I\}$, in proportion to the amount of on-hand inventory at each stage. These operational costs are time-discounted by a fixed discounting factor $\alpha \in (0, 1)$. At the end of period t , each stage places an order to its immediate upstream stage, and then fills the order from its immediate downstream stage (if any) to the fullest extent with its on-hand inventory. The lead time at stage i is a deterministic number of periods $L_i \geq 1$. More precisely, stage i will receive the shipment at the start of period $t + L_i$, if at the end of period t , it places an order and stage $i+1$ has enough inventory.

Under the assumption that the latent demand distribution (faced in stage 1) is stationary across time, it is shown by Clark and Scarf (1960) and Federgruen and Zipkin (1984) that

the optimum can be achieved by a stationary echelon base-stock policy. A stationary echelon base-stock policy is defined by I indices $R_1, \dots, R_I$. For each i , the base-stock R_i concerns stage i 's echelon inventory position y_i , which is defined as the amount of inventory at or in-transit to stage i or its successor stages, minus backlog at stage 1. At each time step and at each stage i , the DM orders no inventory if $y_i \geq R_i$. Otherwise, the DM orders $y_i - R_i$ many units of inventory at stage i .

If the demand distribution were to be known, then optimal stationary echelon base-stock policy $(R_1^*, \dots, R_I^*)$ can be computed by inductively solving a sequence of optimization problems $R_i^* \in \arg\min_{y \geq 0} C_i(y)$ for $i = 1, \dots, I$, where $C_i(y) :=$

$$\begin{cases} \alpha^{L_1} \mathbb{E} \left[h_1(y - \zeta(L_1)) + (b + \sum_{i=1}^I h_i)(y - \zeta(L_1))^- \right], & \text{if } i = 1, \\ \alpha^{L_i} \mathbb{E}[h_i(y - \zeta(L_i) + C_{i-1}(\min\{R_{i-1}^*, [y - \zeta(L_i)]\}) - C_{i-1}(R_{i-1}^*)), & \text{if } i \in \{2, \dots, I\}. \end{cases}$$

where $\zeta(n)$ is a random variable that is identically distributed as $\sum_{i=1}^n D_i$, with $D_1, \dots, D_n$ iid as the latent demand distribution, and $(z)^- = \max\{-z, 0\}$.

In the sampling-based setting, the demand distribution (in stage 1) is not known to the DM. Rather, the DM only has access to N iid samples of the latent demand distribution. Zhang et al. (2021) propose to analyze the SAA method, which returns the set of base-stocks $(\hat{R}_1, \dots, \hat{R}_I)$ that optimizes the serial system when the latent demand is replaced by the empirical distribution constructed with the samples. for the sampling-based serial system problem:

Theorem 14.14 (Zhang et al., 2021) *For any $\varepsilon > 0$, $\delta \in (0, 1)$, the echelon base-stock policy under $(\hat{R}_1, \dots, \hat{R}_I)$ has an expected cost of at most $(1 + \varepsilon)$ times the optimum with probability $\geq 1 - \delta$, if the number of samples N satisfies $N \geq \frac{9}{2\varepsilon^2\eta^2} \log \frac{2}{\delta}$, where η is a constant that only depends on $h_1, \dots, h_I, b, \alpha, L_1, \dots, L_I$, but η is independent of T, ε, δ or the latent demand distribution.*

The exact expression of η can be found in Theorem 14.14 by Zhang et al. (2021). Since η is independent of the latent demand distribution, the sample bound in the theorem is a distribution-free bound, just like in the cases of Levi et al. (2007) and Cheung and Simchi-Levi (2019). To gain some insight into η , we remark that in the case of $I = L_1 = 1$, where the problem reduces to a single-stage inventory control problem with a stationary but latent demand distribution, the parameter η is equal to $\left(\frac{(h_1 + b_1)}{\min\{h_1, b_1\}} \right)^2$. In this special case, the sample bound in the theorem is the same as the newsvendor sample bound in Theorem 14.1. In addition to Theorem 14.14, Zhang et al. (2021) refine the dependence of $1/\varepsilon^2$ to $1/\varepsilon$ when the underlying demand has an increasing failure rate, hence generalizing Theorem 14.5.

14.2.6 Censored Demand Data

Ban (2020) takes a different look at the sampling-based inventory control model, by investigating the impact of censored demand data on a sampling-based inventory control problem

with fixed costs. While the model dynamics are the same as Levi et al. (2007) for the uncapacitated case (non-identical but independent demand distributions and backlog model), Ban (2020) considers a case with fixed order cost, which results in an additional operational cost of $K_t \times \mathbf{1}(y_t - x_t > 0) + c_t \times (y_t - x_t)$ at period t . The fixed order cost K_t is independent of the ordering quantity at period t , as long as the ordering quantity is positive.

It is shown by Scarf (1960) that the optimum can be achieved by a $\{(s_t, S_t)\}_{t=1}^T$ -type policy, where $s_t \leq S_t$ for each t . At period t , if the on-hand inventory x_t is less than s_t , then the DM raises the inventory level so that $y_t = S_t$. Otherwise, the DM does not order any inventory.

Ban (2020) considers the case when the demand distributions are latent, and the DM only has access to samples associated with the latent demand distributions. Ban (2020) constructs estimators for $\{s_t^*, S_t^*\}_{t=1}^T$, where $\{(s_t^*, S_t^*)\}_{t=1}^T$ is an optimal policy, which is latent since the parameters $\{s_t^*, S_t^*\}_{t=1}^T$ depend crucially on the latent demand distributions.

The DM has the access to N sets of sample paths, where a set of sample paths is either uncensored (indexed with $\mathcal{J}^c = \{1, \dots, N_0\}$) or censored (indexed with $\mathcal{J} \in \{N_0 + 1, \dots, N\}$), and N_0 can range from 0 to N . An uncensored sample path, indexed with $n \in \mathcal{J}^c$, consists of $d_1^{(n)}, \dots, d_T^{(n)}$, where $d_t^{(n)}$ is a sample drawn the latent demand distribution at time t , hence the sample is uncensored. A censored sample path, indexed with $n \in \mathcal{J}$, consists of $(y_1^{(n)}, d_1^{(n)}), \dots, (y_T^{(n)}, d_T^{(n)})$. For each t , the quantity $y_t^{(n)}$ is the amount of on-hand inventory at time t , and $d_t^{(n)}$ is the amount of inventory sold. When $d_t^{(n)} < y_t$, the sample $d_t^{(n)}$ reflects the true demand. When $d_t^{(n)} = y_t$, it could be the case that the true demand is higher than the amount of on-hand inventory y_t , in which case $d_t^{(n)}$ under-estimates the actual demand. Consequently, the case $d_t^{(n)} = y_t$ represents the censored case, where the actual demand can be higher than the sample value.

Ban (2020) studies three cases. In the first case, we have $N_0 = N$, meaning that all sample paths are uncensored. Ban (2020) considers the policy $\{(\hat{s}_t, \hat{S}_t)\}_{t=1}^T$ returned by the SAA method. Ban (2020) shows that $(\hat{s}_t, \hat{S}_t)$ is a consistent estimator to the optimal base-stocks (s_t^*, S_t^*) for each t , i.e., $\hat{S}_t$ converges in probability to S_t^* , and $\hat{s}_t$ converges in probability to s_t^* , as N tends to the infinity.

In the second case, when $1 \leq N_0 \leq N$, the DM has a mixture of uncensored and censored sample paths. While the DM could still construct an unbiased estimator for (s_t^*, S_t^*) for each t like in case 1 by discarding all the censored data, Ban (2020) shows that the censored data can provide additional information when pre-processed appropriately. Interestingly, Ban (2020) proposes a pre-processing method by considering the cost-to-go functions in the SAA problem, rather than directly trying to de-bias the censored demand data. Consequently, Ban (2020) is still able to construct a consistent estimator for the optimal base-stocks (s_t^*, S_t^*) for each t .

In the third case, Ban (2020) considers the fully censored case of $N_0 = 0$. The fully censored case presents unique challenges absent in the previous cases, which crucially harness the uncensored data. Alternatively, under certain regularity assumptions, Ban (2020) shows that it is still possible to construct a consistent estimator to the optimal (s_t^*, S_t^*) , with a gradient descent method that is different from the algorithms in the previous two cases.

Finally, Ban (2020) quantifies the confidence levels in the estimates, by investigating the variances of the limits $\lim_{N \rightarrow \infty} \sqrt{N}(\hat{s}_t - s_t^*)$, $\lim_{N \rightarrow \infty} \sqrt{N}(\hat{S}_t - S_t^*)$, which are shown to be normally distributed. It is shown that the variances are of the form $\sigma_u^2 + r \times \sigma_c^2$, where $r = N_0 / N$ is held constant in the limits. To conclude, we remark that data-driven inventory control with censored demand data is an interesting model that is also studied in various online settings (Huh & Rusmevichientong (2009); Besbes & Muharremoglu (2013); Gong & Simchi-Levi (2020)).

14.3 CONCLUSION

Statistical learning theory by Shalev-Shwartz and Ben-David (2014) and the theory of inventory control by Simchi-levi et al. (2014) are exciting research domains. They are of central importance in machine-learning and operations research. While this survey chapter has reviewed some of the contemporary developments in the intersection of these two domains, we believe that the research line of statistical learning in inventory management is still in its infancy. While we have reviewed the design and analysis of data-driven algorithms for classical inventory models, much more research has to be done to understand how data can be incorporated into decision-making in other inventory control models, particularly those motivated by contemporary applications (Qi et al., 2020). Among many of the interesting directions, it will be interesting to construct learning-based inventory control models in multiple period settings (a single-period specialization would be as shown by Ban and Rudin (2019) for example), and to understand the number of samples needed for achieving near optimality.

REFERENCES

- An, M. Y. (1996). *Log-concave probability distributions: Theory and statistical testing*. Game theory and information. University Library of Munich.
- Arrow, K. J., Karlin, S., & Scarf, H. (1960). Studies in the mathematical theory of inventory and production. *Mathematical Gazette*, 44(348).
- Aviv, Y., & Federgruen, A. (1997). Stochastic inventory models with limited production capacity and periodically varying parameters. *Probability in the Engineering and Informational Sciences*, 11(1), 107–135.
- Ban, G. (2020). Confidence intervals for data-driven inventory policies with demand censoring. *Operations Research*, 68(2), 309–326.
- Ban, G., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1), 90–108.
- Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3), 1025–1044.
- Besbes, O., & Muharremoglu, A. (2013). On implications of demand censoring in the newsvendor problem. *Management Science*, 59(6), 1407–1424.
- Charikar, M., Chekuri, C., & Pál, M. (2005). Sampling bounds for stochastic optimization. In C. Chekuri, K. Jansen, J. D. P. Rolim, & L. Trevisan (Eds.), *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005 and 9th International Workshop on Randomization and Computation, RANDOM 2005, Lecture Notes in Computer Science* (Vol. 3624, pp. 257–269). Springer.
- Cheung, W. C., & Simchi-Levi, D. (2019). Sampling-based approximation schemes for capacitated stochastic inventory control models. *Mathematics of Operations Research*, 44(2), 668–692.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory (Wiley series in telecommunications and signal processing)*. Wiley-Interscience.
- Donti, P. L., Amos, B., & Kolter, J. Z. (2017). Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*. (pp. 5490–5500). Curran Associates, Inc.
- Elmachtoub, A. N., & Grigas, P. (2021). Smart “predict, then optimize”. *Management Science*.
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836.

- Gong, X., & Simchi-Levi, D. (2020). Provably more efficient q-learning in the full-feedback/one-sided-feedback settings. *CoRR* abs/2007.00080. <https://arxiv.org/abs/2007.00080>
- Gupta, V., & Rusmevichientong, P. (2021). Small-data, large-scale linear optimization with uncertain objectives. *Management Science*, 67(1), 220–241.
- Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301), 13–30.
- Huang, Z., Mansour, Y., & Roughgarden, T. (2018). Making the most of your samples. *SIAM Journal on Computing*, 47(3), 651–674.
- Huh, W. T., & Rusmevichientong, P. (2009). A nonparametric asymptotic analysis of inventory planning with censored demand. *Mathematics of Operations Research*, 34(1), 103–123.
- Kapuściński, R., & Tayur, S. (1998). A capacitated production-inventory model with periodic demand. *Operations Research*, 46(6), 899–911.
- Levi, R., Perakis, G., & Uichanco, J. (2015). The data-driven newsvendor problem: New bounds and insights. *Operations Research*, 63(6), 1294–1306.
- Levi, R., Roundy, R. O., & Shmoys, D. B. (2007). Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 821–839.
- Liyana, L. H., & Shanthikumar, J. (2005). A practical inventory control policy using operational statistics. *Operations Research Letters*, 33(4), 341–348.
- Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Annals of Probability*, 18(3), 1269–1283.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9, 141–142.
- Oroojlooyjadid, A., Snyder, L. V., & Takáč, M. (2020). Applying deep learning to the newsvendor problem. *IIE Transactions*, 52(4), 444–463.
- Qi, M., Mak, H.-Y., & Shen, Z.-J. M. (2020). Data-driven research in retail operations—A review. *Naval Research Logistics (NRL)*, 67(8), 595–616.
- Qin, H., Simchi-Levi, D., & Wang, L. (2019). Data-driven approximation schemes for joint pricing and inventory control models. *Accepted by Management Science*. <https://ssrn.com/abstract=3354358>
- Scarf, H. (1960). The optimality of (s, s) policies in the dynamic inventory problem. *Mathematical Methods in Social Sciences*.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). *Lectures on stochastic programming*. Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9780898718751>
- Shmoys, D. B., & Swamy, C. (2006). An approximation scheme for stochastic linear programming and its application to stochastic integer programs. *Journal of ACM*, 53(6), 978–1012.
- Simchi-Levi, D., Chen, X., & Bramel, J. (2014). *The logic of logistics: Theory, algorithms, and applications for logistics and supply chain management* (3rd ed.). Springer.
- Tayur, S. R. (1993). Computing the optimal policy for capacitated inventory models. *Communications in Statistics. Stochastic Models*, 9(4), 585–598.
- Watson, G. S. (1964). Smooth regression analysis. *Indian Journal of Statistics, Series A (1961–2002)*, 26, 359–372.
- Zhang, K., Gao, X., Wang, Z., & Zhou, S. (2021). Sampling-based approximation for serial multi-echelon inventory system. *SSRN*. <https://ssrn.com/abstract=3859856>

15. Online learning in inventory and pricing optimization

Xiuli Chao, Boxiao Chen, and Huanan Zhang

15.1 INTRODUCTION

The gist of inventory management is balancing overage and underage costs, and the key to striking that balance lies in understanding customer demand. If there is no uncertainty in demand or the demand is deterministic, then the balancing can be achieved perfectly (given sufficient inventory supply capacity); otherwise, the crucial information necessary for optimal inventory management is the probability distribution of customer demand. When the selling price is also a decision, another piece of critical information for optimized decision is the dependency of customer demand on the selling price. When precise information of demand uncertainty and the demand–price relationship is known, the expected profit/cost for any given inventory/pricing decision problem can be derived, computational algorithms can be developed, and optimal inventory and pricing policy can be obtained.

The vast literature on inventory management assumes that the probability distribution of customer demand as well as the demand–price relationship are known *a priori*. As a result, the main tasks in the literature have been on the derivation of the objective function, characterization, and computation of the optimal inventory control and pricing decisions.

In practice, however, it is unlikely the case that the decision-maker has complete information about customer demand distribution and the exact relationship between demand and selling price. When such information is not available, it has to be estimated. If there is sufficient historical data, then one can use the available data to estimate the customer demand distribution and demand–price relationship. This is known as *offline learning*. For example, Levi et al. (2007) study a multi-period pure inventory control problem with independent and identically distributed (iid) demand, but the decision-maker has no prior knowledge about the demand distribution. Assuming that plenty of demand data are available and can be retrieved, the authors find the minimum data sample size required to achieve a given level of accuracy in minimizing total cost. In Huh et al. (2011), a pure inventory control problem with lost sales and censored demand is considered, and the Kaplan–Meier estimator is applied in an offline manner to aid inventory optimization.

Offline versus online learning. Typically, offline learning involves retrieving information from historical data and constructing a proxy objective function, which is then used to compute a (hopefully) near-optimal policy for subsequent implementation. Offline learning is based on two important assumptions. Besides the availability of *sufficient* demand data, it also assumes that the future demand pattern remains the same as that observed in the far past. In applications, both of these two assumptions may not be satisfied. For example, in the sales of new products, historical data may not be available at all. For mature products, even when there are plenty of historical data, the aged ones may not be very valuable. Indeed, it has been generally recognized that data in the far past may have very limited value in predicting the

future since many external factors, e.g., economic conditions and/or market environment, that affect or influence demand might have changed. As a result, it is only the relatively recent data that are most useful in predicting the demand in the near future.

It is important to note that some critical information for making inventory decisions may not be found in the offline data at all. To see this, consider the classic lost-sales inventory system with censored demand. In this problem, only sales data are available. The true demand, e.g., the demand data beyond the on-hand inventory level, is not observed. In such a scenario, if historical inventory decisions were less than the customer demand, then we would not have the required data to learn the true demand distribution in order to find the optimal inventory decision. It is clear that, the higher the starting inventory level in a period, the more demand information can be revealed. As a matter of fact, one important purpose of online learning is precisely to explore the action space in an effective manner so that necessary demand information can be obtained for making inventory decisions.

Online learning involves collecting and applying the most updated information in real time to learn the demand and make informed decisions. Most approaches in online demand learning have to carefully balance between exploration, which is used to learn the demand information, and exploitation, which utilizes the learned information to maximize revenue/profit or minimize cost. This chapter focuses on online learning. It is conceivable that an integrated approach, that combines some available offline data in online learning, can also be developed, and we refer the interested reader to Gong and Simchi-Levi (2020b), Bu et al. (2022), and the references therein.

Model classification. There are numerous ways to classify the models and problems for online learning in inventory control and pricing optimization. The first one is to classify the model according to whether it is purely an inventory optimization problem, or it is a joint pricing and inventory optimization problem. It is based on such a classification that we layout the discussion in this chapter. The second classification is based on whether demand data can be fully observed or only sales data are observed (censored demand). Clearly, censored demand introduces more complexity to learning as decisions can affect the quality of observed data, hence it has to be part of the algorithm design, while in the case that full demand data is observed, the decision can simply ignore exploration and only focus on optimizing the objective function (exploitation). As a result, simpler and more efficient algorithms can be developed for the latter case. The third classification is according to whether or not the demand model is parametric or nonparametric. For the first case, the demand distribution is specified up to some unknown parameters, while for the second, demand distribution is completely unknown. The fourth classification is based on the number of product types, leading to single- and multi-product models. In the latter case, the demands for different products are dependent on each other.

It is possible to classify the inventory problems according to other dimensions, such as whether the products are perishable or non-perishable (leading to perishable and non-perishable inventory models), or whether excess demand is lost or backlogged (leading to lost-sales inventory models and backlog inventory models), and others.

This chapter discusses some of the latest developments in online learning in inventory control and pricing optimization in supply chain management. The focus is on models and algorithms. For the numerical performance of the algorithms, we refer the interested reader to the relevant references. Throughout the chapter, the notation $f(x) = O(g(x))$ is defined as that there exists some constant $C > 0$ such that $f(x) \leq Cg(x)$ for all x large enough, $f(x) = \Omega(g(x))$ implies that there exists some constant $c > 0$ such that $f(x) \geq cg(x)$ for all x large enough, and $\tilde{O}(g(x))$ means that the logarithmic term in $g(x)$ has been hidden.

15.1.1 Organization of This Chapter

In the rest of this section, we briefly overview some research works in the literature on online inventory and pricing optimization, and discuss the commonly used metric, regret, for evaluating learning algorithms. We also elaborate on important approaches for developing learning algorithms. Using a simple example, Section 15.2 provides a pedagogical illustration of online optimization in inventory management. Then, Section 15.3 concentrates on demand learning in pure inventory control problems. Specifically, we first discuss the regret lower-bound results in Section 15.3.1, and then present learning algorithms for several classes of inventory systems, together with the upper bounds of their regrets. We discuss the classic periodic-review inventory systems in Section 15.3.2, perishable inventory systems in Section 15.3.3, lost-sales inventory systems with lead time in Section 15.3.4, multi-product systems with a warehouse-capacity constraint in Section 15.3.5, multi-product systems with substitution in Section 15.3.6, lot-sizing problems in Section 15.3.7, and finally dual-sourcing systems in Section 15.3.8. Section 15.4 focuses on learning in joint inventory and pricing optimization. Specifically, Section 15.4.1 discusses the backlog model while Section 15.4.2 studies the lost-sales model with censored demand, both focusing on a nonparametric setting, and Section 15.4.3 considers a parametric learning problem with a constraint on the number of price changes during the planning horizon. The chapter concludes with some discussions in Section 15.5.

15.1.2 Brief Historical Account

In this subsection, we briefly mention some relevant works on learning in inventory management, with no intention of a comprehensive review.

Learning via Bayesian updating. One classic approach taken in the inventory control literature on limited demand information is through Bayesian updating. Under such a framework, the demand distribution is assumed to be from a family known up to some parameters. The decision-maker has a prior on the parameters of the demand distribution, and in each period, the belief on these parameters is updated using Bayes' formula. For example, early papers such as Scarf (1959, 1960) and Iglehart (1964) consider cases where the demand distribution belongs to the exponential and range families. Other papers that incorporate the Bayesian approach into stochastic inventory models include Murray and Silver (1966), Azoury (1985), Lovejoy (1990), and Ding et al. (2002). One useful concept in Bayesian updating is conjugate prior, which is a class of probability distributions whose posterior belongs to the same class as the prior. As one can expect, when more demand data is observed one would have a better understanding of the parameters, enabling the decision-maker to make closer to optimal inventory decisions. The standard formulation of inventory control using Bayesian updating is dynamic programming, with the state of the system being the inventory level and the belief distribution of the unknown parameters, and the goal is to minimize expected total cost (or maximize expected total profit) for the rest of the planning horizon.

It is important to note that in the classic Bayesian framework, the objective function to optimize in each period is the expected total future cost computed using the belief of the demand distribution at that period. In other words, the evaluating criterion for optimization relies on the decision-maker's belief, so it changes from period to period. This is in sharp contrast with online demand learning popularized in the past decade, where the metric to evaluate a solution in any period is always computed using the true underlying demand distribution, which is to

be learned. Indeed, for online learning, the most commonly used criterion for optimization is regret, which is defined using the expected total cost in terms of the true underlying distribution. Refer to Section 15.1.3.

Online learning in pure inventory management. Borrowing ideas from statistical learning, machine learning, online convex optimization, and stochastic approximation, online learning in inventory control has gained much popularity in recent years. Applying results from multi-armed bandit problems, Chang et al. (2005) propose an adaptive sampling algorithm that approximates the optimal value of a finite-horizon Markov decision process (MDP) with finite state and action spaces, and applies it to an inventory control problem with unknown demand distribution and uncensored samples. The repeated newsvendor problem with unknown demand distribution and unobservable lost sales is studied by Godfrey and Powell (2001), which develops a learning algorithm that constructs a sequence of concave piecewise linear approximations using sample gradients. Kunnumkal and Topaloglu (2008) consider a number of inventory control problems for which the base-stock policies are known to be optimal and propose stochastic approximation methods that generate solutions approaching the clairvoyant optimal base-stock levels. Huh and Rusmevichientong (2009) study the classic periodic-review system and present a gradient-type learning algorithm and show that the regret is upper bounded by $O(\sqrt{T})$. Besbes and Muharremoglu (2013) formally study the lower bound of the regret for the repeated newsvendor problem, which is $\Omega(\sqrt{T})$ for a general problem and $\Omega(\log T)$ under some conditions. Huh et al. (2009) study online learning for the first inventory system beyond newsvendor type systems, the lost-sales inventory system with lead times, and it introduces a gradient-type learning algorithm with increasing cycle length, which achieves a regret upper bounded of $O(T^{2/3})$, compared with the *optimal base-stock* policy. Later Zhang et al. (2020) improve the rate to $\tilde{O}(\sqrt{T})$ by introducing a very different gradient-type learning algorithm. And then Agrawal and Jia (2019) introduce another search-based algorithm that not only achieves the $\tilde{O}(\sqrt{T})$ regret rate, but also improves the dependency on lead times from exponential to linear. Zhang et al. (2018) study the perishable inventory system and attain a $O(\sqrt{T})$ with a gradient-type algorithm. Shi et al. (2016) consider a multi-product inventory system with warehouse inventory constraint, and by utilizing the joint convexity and other system properties, the authors develop a learning algorithm with a tight convergence rate. Lugosi et al. (2022) study the structure of the hardness of the learning from censored demand in repeated newsvendor problem. Chen and Chao (2020) study the multi-product inventory system with substitutions and develop a phased exploration-and-exploitation algorithm. Yuan et al. (2021) study the lot-sizing problem with setup costs, and present a novel learning algorithm that integrates the gradient and the bandit methods to learn the optimal (s, S) policy, which achieves a $\tilde{O}(\sqrt{T})$ regret rate. Chen and Shi (2019) study the dual-sourcing problem, and developed a learning algorithm that converges to the *optimal tailored base-surge* policy with a $\tilde{O}(\sqrt{T})$ regret rate. Chen (2021) and Keskin et al. (2021b) explore the data-driven inventory control problem in non-stationary environments.

Reinforcement learning has been applied in a number of papers to study inventory control problems. For example, Oroojlooyjadid et al. (2017) develop a deep reinforcement learning-based algorithm for the beer game in supply chain management and show that the algorithm performs very well under both scenarios when its co-players act rationally or irrationally. The single-item, multi-period inventory management problem with fixed ordering cost is studied by Liu et al. (2021), where the authors develop a deep learning framework that directly outputs the optimal order timing and quantity with the given contextual information as the input.

Gijsbrechts et al. (2020) explore the lost sales, dual-sourcing, and multi-echelon inventory systems, they model each inventory problem as a Markov decision process, and apply the Asynchronous Advantage Actor Critic algorithm in a variety of parameter settings. Gong and Simchi-Levi (2020a) propose an Elimination-Based Half-Q-Learning algorithm as well as a Full-Q-Learning algorithm to solve inventory models with lead time and order capacity, and achieve a tight regret rate.

Online learning in joint inventory and pricing optimization. Burnetas and Smith (2000) is one of the earliest papers, if not the first one, to study joint pricing and inventory control with unknown demand distribution. The pricing mechanism is modeled as a multi-armed bandit problem, while the order quantity decision is based on a stochastic approximation procedure. Chen et al. (2019a) consider a nonparametric backlog system, and achieve a tight regret rate of $O(T^{1/2})$ by approximating the tangent line of the demand–price function using linear approximation and estimating the noise distribution using sample average approximation. The lost-sales system with censored demand is explored by Chen et al. (2021) and Chen et al. (2020c) with nonparametric learning. Assuming convex objective functions, Chen et al. (2021) apply spline approximation to learn the demand–price function and achieve a regret rate that almost matches the lower bound of $\Omega(T^{1/2})$. For the same setting, Chen et al. (2020c) improve the regret rate to $O(T^{1/2}(\log T)^2)$ by proposing a double bisection algorithm, and for nonconvex objective functions, Chen et al. (2020c) obtain a regret rate of $O(T^{3/5}(\log T)^2)$ by proposing a tournament-based algorithm and prove the regret lower bound of $\Omega(T^{3/5}/\log T)$ based on a generalization of the squared Hellinger distance. Chen and Chao (2019) and Chen et al. (2022) consider a limited number of price changes in the parametric backlog and lost-sales inventory systems, respectively. The regret of the algorithms from both papers achieves the tight rate of $O(T^{1/(m+1)})$ when the price is allowed to change $m \geq 1$ times during the planning horizon. The lower bound of regret is also established in Chen et al. (2020a). The classical backlog system with fixed ordering cost is considered in Chen et al. (2022), where UCB-based algorithms are developed through dynamic programming recursions to approach the optimal $(s, S, \mathbf{p})$ policy with a tight $O(T^{1/2} \log T)$ regret rate. Katehakis et al. (2020) consider the joint optimization problem with discrete backlogged demand in different settings with or without a leading price. Keskin et al. (2021a) study the joint pricing and inventory control problem in a changing environment under a parametric demand rate function and provide learning algorithms whose convergence rates match the theoretical lower bound.

15.1.3 Evaluation of Online Learning Algorithms

The mental model for online learning is that, even though the decision-maker has limited or even no demand information *a priori*, she still needs to make day-to-day decisions. Further, for any decision made by the decision-maker, we have to bear its consequence. For example, when a suboptimal decision is used, it leads to an expected profit that is lower than that of the true optimal solution, meaning that it leads to a loss compared with the true optimal solution. Indeed, there is a ground truth for customer demand that the clairvoyant knows. Hence, in online learning, the evaluation criterion is *regret*, which is defined as the difference between the value function under the learning algorithm and that of a clairvoyant who has prior information about demand probability distribution and the demand curve with respect to the price of the product. Specifically, suppose the goal is to minimize the total cost in a T -period problem,

and $C^\pi(T)$ denotes the expected cost using learning algorithm π , while $C^*(T)$ denotes the minimum expected total cost for the clairvoyant, then the regret is defined as

$$R^\pi(T) = C^\pi(T) - C^*(T). \quad (15.1)$$

It is typically the case that $R^\pi(T)$ is non-decreasing in T and sublinear in T . Hence, a learning algorithm is superior if $R^\pi(T)/T$ converges to zero faster.

Depending on the structure of the learning problem, there is usually a lower bound for $R^\pi(T)$, say $r(T)$, such that the regret of no learning algorithm can beat. That is, for any learning algorithm π , there exists a problem instance and time index t^* such that

$$R^\pi(T) \geq r(T), \quad T \geq t^*.$$

The lower bound for the regret of an online optimization problem, as well as its derivation, usually relies on the information-theoretic structure of the class of problem on hand.

Clearly, if an algorithm can be designed so that its regret has the same rate as $r(T)$, then it means that the algorithm achieves the best possible regret rate. Needless to say, in online learning for inventory control and pricing optimization problems, we aim to design algorithms whose regret rate matches that of the lower bound.

15.1.4 Common Approaches to Online Demand Learning

This subsection briefly discusses some of the most popular methods used in online learning for inventory and pricing optimization problems. The applications of several of these methods will be illustrated in the next subsection.

Statistical method. Statistics methods are typically classified as parametric and nonparametric methods. For parametric models, an extremely useful method for estimating parameters is the maximum likelihood estimation (MLE) method (see, e.g., Borovkov, 1998), but many other statistical methods have also been applied in learning, such as the least square method (and other loss function method) (see, e.g., Besbes and Zeevi (2015)). For nonparametric models, one simple approach is the sample average approximation (SAA), which uses demand samples to construct a proxy for the objective function, that is then used for optimization. In the case of censored data, the Kaplan–Meier estimator is a useful method for estimating the distribution and it has been used by Huh et al. (2011) in offline learning. Another method for the nonparametric problem is the spline method, which has been used in a number of papers, including Chen et al. (2019b) on network revenue management problems and Chen et al. (2021) on joint inventory control and pricing optimization problems. The spline method first explores a sufficient amount of time to collect data, and then applies that to construct a smooth estimate as a proxy of the objective function that is used to compute a solution for implementation in the exploitation phase.

Stochastic gradient descent method (SGD). Not knowing the demand distribution, the decision-maker is unable to compute the expected profit or cost, which is the objective function to be optimized. The stochastic gradient descent method, where the decision-maker chooses a feasible solution based on the gradient of realized sample objective in each iteration, has been proven to be a powerful method. The SGD method is derived from online convex optimization (see, e.g., Zinkevich (2003); Hazan et al. (2007); Agarwal et al. (2011)) and

stochastic approximation (see, e.g., Robbins and Monro (1951), Kiefer and Wolfowitz (1952), Lai and Robbins (1981)). This method is directly applicable when the problem in hand is, for a T -period problem, to optimize the sum of T -independent cost functions where the cost function in period t depends only on the decision in period t . However, in many problems in inventory control and supply chain management, a decision in one period may have a lasting impact, making the SGD method difficult to apply. There have been several interesting and innovation extensions that design the learning problem into learning cycles, so that the cost in one cycle depends only on the decision in the cycle and the problem can be solved by learning from cycle to cycle, e.g., Zhang et al. (2020).

Multi-armed bandit problems (MAB). Another extremely useful method in online learning is MAB. In MAB, multiple arms need to be played one at a time but with random rewards and an unknown mean value for each, and the goal is to find a policy to maximize the total expected reward up until any time T . For example, when each decision in the inventory and pricing optimization problem is treated as an arm, then it is seen that the joint inventory control and pricing optimization problem is reduced to a MAB. Out of the various algorithms developed for MAB, the upper confidence bound (UCB) method is particularly useful, and it has been widely applied in online learning of inventory control problems. In addition, MAB has been extended to continuum-armed bandit problems to handle uncountable many decisions (see, e.g., Kleinberg (2005); Auer et al. (2007); Cope (2009)).

Thompson sampling (TS) method. TS method is very similar to the Bayesian method in that it continuously updates the belief when making dynamic decisions. The key difference is that in TS, a parameter is randomly sampled based on the current belief, which is then used to compute the value function in search for a decision. Clearly, the TS method saves computational cost in computing the expected value function (with respect to the belief probability) in each period since it uses the cost function of a randomly sampled parameter. As in the Bayesian method, in TS the decision-maker typically chooses the prior that belongs to a class of conjugate distribution, so that the posterior, or updated, parameter distribution is easily computed. TS has been used in a number of papers in inventory and pricing optimization (see, e.g., Agrawal et al. (2017); Ferreira et al. (2018); Miao and Chao (2021)).

Other methods. A number of other methods have been adopted in online demand learning. For example, various optimization techniques have been successively applied in designing online algorithms, and among them the bisection method is one of the most useful. In such methods, data collected from exploration phases are used to evaluate the solution in each iteration, and that provides direction for improving the solution in the following step. Out of the machine learning methods used for inventory optimization, reinforcement learning has been used in several papers and the reader can refer to (e.g., Oroojlooyjadid et al. (2017); Liu et al. (2021); Gijbrecchts et al. (2020)). Since deep learning methods typically do not provide a theoretical performance guarantee (e.g., in terms of regret), we will not cover that stream of literature in this chapter.

15.2 AN ILLUSTRATIVE EXAMPLE

In this section, we use a simple example, the repeated newsvendor problem, to illustrate online learning in inventory management. The repeated newsvendor problem can also be considered as inventory control with perishable products in which the product has one period of lifetime, hence any product received in the period cannot be carried over to the next period. For ease of exposition, in this section we focus on pure inventory decisions.

Consider a planning horizon of T periods. Except in Section 15.2.6, the cost structure only assumes unit holding cost h , unit shortage cost b , and the purchasing cost is subsumed in holding and shortage cost and thus is assumed to be zero without loss of generality. The ordering lead time is zero. The demands over the T periods, denoted by $D_1, D_2, \dots$, and D_T , are iid with an unknown distribution. The objective is to minimize the expected total holding and shortage costs. Note that, for the repeated newsvendor problem, the inventory replenishment decision in period t , denoted by y_t , is always achievable since there is no carryover inventory, and the inventory level at the beginning of a period, before placing the order, is always 0.

The total cost for the repeated newsvendor problem can be written as, in terms of the decision y_t at the beginning period t ,

$$C(T) = \sum_{t=1}^T G(y_t), \quad (15.2)$$

where

$$G(y_t) = h\mathbb{E}[(y_t - D_t)^+] + b\mathbb{E}[(D_t - y_t)^+]. \quad (15.3)$$

Here, the expectation $\mathbb{E}[\cdot]$ is taken with respect to the true demand distribution, which is not known to the decision-maker *a priori*, and $x^+ = \max\{x, 0\}$ for any real number x .

In the subsequent subsections, we use various online learning methods to solve this problem, with the goal of minimizing regret, which is equivalent to minimizing the expected total cost. The lower bound of regret for this problem is $\Omega(\sqrt{T})$ (see, e.g., Huh and Rusmevichientong (2009)).

15.2.1 MLE for Online Learning

For parametric models, a useful and powerful method in online learning is the MLE method. Suppose, for example, the demand follows an exponential distribution. Its mean, say θ , is not known. For simplicity, we consider the case that the demand is observable. At any period t , the decision-maker has demand data from the previous $t-1$ periods, say $d_1, \dots, d_{t-1}$. The standard maximum likelihood method shows that the MLE estimate of θ is

$$\hat{\theta}_t = \frac{1}{t-1} \sum_{s=1}^{t-1} d_s.$$

Then, pretending that D_t has a mean value of $\hat{\theta}_t$, we maximize the standard newsvendor problem in Equation (15.3) to find the maximizer $\hat{y}_t$, and set the ordering decision as $\hat{y}_t$ for period t . From the newsvendor model, we have

$$\hat{y}_t = \hat{\theta}_t \log \left(1 + \frac{b}{h} \right),$$

and the true, but unknown, optimal inventory decision is

$$y^* = \theta \log \left(1 + \frac{b}{h} \right).$$

Assume that the optimal order-up-to level has a known upper bound, denoted by $\bar{y}$. The performance of the MLE method can be obtained from the following concentration inequality that (see, e.g., Proposition 2.8.3 of Vershynin (2018)), for some positive constant k , and ε small enough,

$$P\{|\hat{\theta}_t - \theta| > \varepsilon\} \leq 2e^{-2k(t-1)\varepsilon^2}. \quad (15.4)$$

Letting $\varepsilon = \sqrt{\log t / [k(t-1)]}$, we obtain

$$P\left\{|\hat{\theta}_t - \theta| > \sqrt{\log t / [k(t-1)]}\right\} \leq \frac{2}{t^2}.$$

For convenience, we define

$$\varepsilon_1 = \left\{|\theta_t - \theta| > \sqrt{\log t / [k(t-1)]}\right\}.$$

To derive the regret of the online algorithm that prescribed inventory decision $\hat{y}_t$ in period t , $t = 1, 2, \dots$, we first note the following fact:

$$G(\hat{y}_t) - G(y^*) \leq \max\{h, b\} |\hat{y}_t - y^*|.$$

We proceed as follows:

$$\begin{aligned} R(T) &= \mathbb{E} \left[\sum_{t=1}^T (G(y_t) - G(y^*)) \right] \\ &\leq \max\{h, b\} \mathbb{E} \left[\sum_{t=1}^T |y_t - y^*| \right] \\ &\leq \max(h, b) \sum_{t=1}^T \left(\mathbb{E}[|y_t - y^*| \varepsilon_1] P\{\varepsilon_1\} + \mathbb{E}[|y_t - y^*| \varepsilon_1^c] P\{\varepsilon_1^c\} \right) \\ &\leq \max(h, b) \sum_{t=1}^T \left(\mathbb{E}[|y_t - y^*| \varepsilon_1] P\{\varepsilon_1\} \right) + \log \left(1 + \frac{h}{b} \right) \mathbb{E}[|\theta_t - \theta| \varepsilon_1^c] P\{\varepsilon_1^c\} \\ &\leq \max(h, b) \sum_{t=1}^T \frac{2\bar{y}}{t^2} + \max(h, b) \log \left(1 + \frac{h}{b} \right) \sum_{t=2}^T \sqrt{\frac{\log t}{k(t-1)}} \\ &\leq O(1) + \frac{\max(h, b) \log(1 + h/b)}{\sqrt{k}} \sqrt{\log T} \sum_{t=1}^T 1/\sqrt{t} \\ &= O(\sqrt{T \log T}). \end{aligned}$$

This shows that online learning for repeated newsvendor problems with MLE has a regret rate $O(\sqrt{T \log T})$.

15.2.2 Thompson Sampling Method

For a parametric model, a routine approach is that the decision-maker uses some prior belief about the parameters, which takes the form of a probability distribution, and updates the belief from period to period. In any period, the decision-maker evaluates the objective using her current belief to choose a decision she believes to be optimal at the time. After the decision is implemented, the decision-maker uses the observed data to update her belief. In such a procedure, the decision-maker needs to compute the expected objective with respect to her belief on the parameter. This is the procedure adopted in the stream of literature on the Bayesian approach we reviewed in Section 15.1.2.

A closely related, but more efficient, algorithm is the TS method. First studied by Thompson in the 1930s (see Thompson (1933, 1935)), this method has gained momentum in recent years due to its effectiveness in online learning. The key difference between the Bayesian approach and the TS method is that, instead of computing the expected objective function using the belief probability distribution of the parameter, TS randomly samples a parameter (using the current belief probability distribution) and then uses the sampled parameter to compute the objective value function.

To illustrate, we assume that demand D_t follows an exponential distribution with an unknown parameter λ , which takes either a known value λ_1 or another known value λ_2 . For simplicity suppose demand is observable. The initial belief is that λ takes the two values with probabilities α_1 and $1 - \alpha_1$, respectively. In general, suppose in period t the belief probabilities for λ to be λ_1 and λ_2 are α_t and $1 - \alpha_t$, respectively. Then, TS randomly generates a λ according to this belief probability. Let the generated λ in period t be denoted by $\hat{\lambda}_t$. We compute the inventory decision in period t , denoted by $\hat{y}_t$ and it minimizes $G_t(y_t)$ of Equation (15.3) with D_t having an exponential distribution with parameter $\hat{\lambda}_t$. Thus

$$\hat{y}_t = \frac{1}{\hat{\lambda}_t} \log \left(1 + \frac{b}{h} \right).$$

After implementing $\hat{y}_t$, demand is realized and let it be d_t . Then, the belief probability for λ to be λ_1 and λ_2 are updated using the Bayesian formula

$$\alpha_{t+1} = P\{\lambda = \lambda_1 \mid D_t = d_t\} = \frac{\alpha_t \lambda_1 e^{-\lambda_1 d_t}}{\alpha_t \lambda_1 e^{-\lambda_1 d_t} + (1 - \alpha_t) \lambda_2 e^{-\lambda_2 d_t}}.$$

Then the process continues.

The TS method has been widely studied and applied in operations management problems. The standard evaluation criterion for TS is not the regret we defined earlier, but a weaker version called Bayesian regret, which is defined as the expected value of the regret in terms of the initial belief. We refer the readers to Russo and Van Roy (2014) and Russo et al. (2017).

15.2.3 SAA Method

Not knowing the demand distribution but having to make inventory decisions in each period, one approach is to use the up-to-date demand information to construct a proxy objective function using the sample average. Specifically, suppose the current period is t and an ordering decision y_t has to be made. If the probability distribution of demand D_t is known, then the decision maker wants to choose y_t so that Equation (15.3) is minimized. However, the expectation above cannot be computed as we do not know the probability distribution of D_t . Since we have demand data up to period $t-1$, denoted by $d_1, \dots, d_{t-1}$, we can estimate the expected cost by its sample mean, i.e., we replace $G(\cdot)$ by a proxy function given by

$$\hat{G}(y) = \frac{1}{t-1} \sum_{s=1}^{t-1} (h(y - d_s)^+ + b(d_s - y)^+),$$

and use a minimizer of $\hat{G}(\cdot)$, denoted by $\hat{y}_t$, as the inventory decision for period t . This leads to

$$\hat{y}_t = \inf \left\{ y : \frac{1}{t-1} \sum_{s=1}^{t-1} 1[d_s \leq y] \geq \frac{b}{h+b} \right\},$$

where $1[A]$ is the indicator function taking value 1 if the statement in A is true and 0 otherwise.

How good is this SAA method? We evaluate it using regret. Let y^* denote the clairvoyant optimal solution, then we compute

$$R(T) = \sum_{t=1}^T \mathbb{E}[G(\hat{y}_t) - G(y^*)].$$

For simplicity, suppose D_t is a continuous random variable with cumulative distribution function F , and it has a strictly positive density function on its compact domain lower bounded by $a > 0$. In the following, we show that the regret of the SAA method, under the condition of having a strictly positive density function, is $O(\log T)$. Since y^* is the newsvendor solution, one has

$$y^* = F^{-1} \left(\frac{b}{h+b} \right),$$

where F^{-1} is the inverse function of F . It is shown in Chen et al. (2019a) that, for $\theta > 0$,

$$P\left\{ |F(y_t) - F(y^*)| \geq \theta \right\} \leq 2e^{-2(t-1)\theta^2}.$$

This implies

$$P\left\{ |y_t - y^*| \geq \theta / a \right\} \leq 2e^{-2(t-1)\theta^2},$$

and

$$P\left\{ |y_t - y^*|^2 \geq \theta^2 / a^2 \right\} \leq 2e^{-2(t-1)\theta^2}.$$

For $t \geq 2$, the above inequality further implies

$$\mathbb{E}[(\hat{y}_t - y^*)^2] = \int_0^\infty P\{|\hat{y}_t - y^*|^2 \geq x\} dx \leq \frac{1}{a^2(t-1)}.$$

Let $r < \infty$ be the maximum value of $|G''|$ on the action space. Then, by Taylor expansion, the regret can be upper bounded as

$$\begin{aligned} R(T) &= \sum_{t=1}^T \mathbb{E}[G(\hat{y}_t) - G(y^*)] \leq \sum_{t=1}^T \frac{r}{2} \mathbb{E}[(\hat{y}_t - y^*)^2] \\ &= \frac{r}{2} \mathbb{E}[(\hat{y}_1 - y^*)^2] + \sum_{t=2}^T \frac{r}{2} \mathbb{E}[(\hat{y}_t - y^*)^2] \\ &\leq \frac{r}{2} \mathbb{E}[(\hat{y}_1 - y^*)^2] + \sum_{t=2}^T \frac{r}{2a^2(t-1)} \\ &= O(\log T). \end{aligned}$$

Note that because we assume the demand distribution has a strictly positive density function on its compact domain and it is lower bounded by a constant $a > 0$, the objective function $G(\cdot)$ is strongly convex. As a result, we are able to obtain a smaller regret $O(\log T)$ in this case. If $G(\cdot)$ is only convex but not strongly convex, then a regret of $O(\sqrt{T})$ should be expected.

15.2.4 Explore-First Algorithm

As discussed earlier, a key issue in online learning is to balance exploration and exploitation. A very straightforward and brute-force method is to first spend some time learning about the demand, and then use the collected information to make the decision for the rest of the planning horizon. There are several methods that employ this idea and here we discuss the simplest one, the explore-first algorithm.

Consider the case where the demand is discrete, and that there is no sufficient demand data that can be used for making inventory decisions at the beginning. To learn about the demand, the decision-maker can first use $T_0 < T$ periods of time for exploration, i.e., implement some heuristic inventory decision y_0 . Then, the collected demand data, denoted by $d_1, \dots, d_{T_0}$, is used to derive an inventory decision that is used until period T . Clearly, different methods can be utilized to make use of the collected data to obtain the inventory decision for the exploitation period, such as MLE and Least Square if it is a parametric model, or SAA if nonparametric. For illustration, let us assume that it is a parametric model where the demand is known to follow an exponential distribution but with an unknown mean value θ , and we use the least square method to estimate the parameter as follows:

$$\min \sum_{t=1}^{T_0} (\mathbb{E}[D_t] - d_t)^2.$$

Since $\mathbb{E}[D_t] = \theta$, the estimate of θ , denoted by $\hat{\theta}$, is

$$\hat{\theta} = \frac{1}{T_0} \sum_{t=1}^{T_0} d_t.$$

Then, for periods $T_0 + 1, T_0 + 2, \dots, T$, implement an inventory decision $\hat{y}$ which is obtained assuming that the demand is exponential with mean $\hat{\theta}$.

The regret for the explore-first algorithm depends on T_0 . In the following, we show that when $T_0 = T^{2/3}(\log T)^{1/3}$, the regret is $R(T) = O(T^{2/3}(\log T)^{1/3})$.

It follows from Bernstein's inequality (see, e.g., Vershynin (2018)) that Equation (15.4), with $\hat{\lambda}_t$ and λ replaced by $\hat{\theta}$ and θ respectively, continues to hold for large enough t . Letting $t - 1 = T_0$ and $\varepsilon = \sqrt{(\log T_0) / (kT_0)}$, we obtain

$$P\left\{|\theta - \theta| > \sqrt{\frac{\log T_0}{kT_0}}\right\} \leq 2 / T_0^2.$$

From newsvendor solution we have

$$\hat{y}_t = \hat{\theta} \log\left(1 + \frac{b}{h}\right), \quad y^* = \theta \log\left(1 + \frac{b}{h}\right).$$

Thus

$$G(y_t) - G(y^*) \leq \max\{h, b\} \log\left(1 + \frac{b}{h}\right) |\theta - \theta|.$$

For convenience, let

$$\varepsilon_2 = \left\{|\theta - \theta| > \sqrt{\frac{\log T_0}{kT_0}}\right\},$$

then

$$\begin{aligned} R(T) &= R(T_0) + \mathbb{E} \left[\sum_{t=T_0+1}^T (G(y_t) - G(y^*)) \right] \\ &= O(T_0) + \mathbb{E} \left[\sum_{t=T_0+1}^T (G(y_t) - G(y^*)) \mid \varepsilon_2 \right] P\{\varepsilon_2\} \\ &\quad + E \left[\sum_{t=T_0+1}^T (G(y_t) - G(y^*)) \mid \varepsilon_2^c \right] P\{\varepsilon_2^c\} \\ &\leq O(T_0) + O(1) \times (T - T_0) \frac{2}{T_0^2} + (T - T_0) \max(h, b) \left(1 + \frac{b}{h}\right) \sqrt{\frac{\log T_0}{kT_0}}. \end{aligned}$$

Letting $T_0 = T^{2/3}$, we obtain

$$R(T) = O(T^{2/3}(\log T)^{1/2}).$$

15.2.5 SGD Method

Now we assume that only sales data are observed, i.e., we have censored demand. From Equation (15.3), we want to find y_t that minimizes $G(y_t)$. However, we cannot compute $G(y_t) = \mathbb{E}[g(y_t, D_t)]$ due to the lack of information on the distribution of D_t , with

$$g(y_t, d_t) = h(y_t - d_t)^+ + b(d_t - y_t)^+,$$

where d_t is a realization of demand D_t , and the expectation is taken with respect to random variable D_t , whose distribution is unknown. However, the stochastic gradient of $g(y_t, D_t)$ with respect to decision y_t is easy to find, and it is given by h if $y_t \geq D_t$ and $-b$ if $D_t < y_t$. A key observation is that this gradient is known even with censored data: If there is a positive on-hand inventory level at the end of the period, then $D_t < y_t$ and gradient is h , and otherwise, the gradient is $-b$. This suggests that we can design an SGD algorithm as follows. Suppose $\bar{y}$ is an upper bound of the inventory decision, which is supposed to be known. Let y_1 be arbitrarily chosen (say $y_1 = \bar{y}/2$) and γ be a design parameter. The SGD method is, for $t \geq 1$,

$$\hat{y}_{t+1} = y_t - \frac{\gamma}{\sqrt{t}} \nabla g_t(y_t),$$

where $\nabla g_t(y_t) = h$ if $D_t < y_t$ (when positive on-hand inventory is left at the end of period t), and $\nabla g_t(y_t) = -b$ if $D_t \geq y_t$ (when no on-hand inventory is left at the end of period t). Then, set the inventory decision for period $t+1$ as

$$y_{t+1} = \min \{ \max \{ \hat{y}_{t+1}, 0 \}, \bar{y} \}.$$

It is shown by Huh and Rusmevichientong (2009) that the regret of the SGD algorithm against the clairvoyant optimal solution y^* is

$$R(T) = O(\sqrt{T}),$$

which is the same as the lower bound for this class of problems.

15.2.6 UCB Algorithm

For ease of discussion, in this subsection we transform the repeated newsvendor problem from minimizing the expected total cost to maximizing the expected total profit. We also introduce a unit purchasing cost c . At the beginning of every period t , the decision-maker decides an ordering quantity y_t , which costs cy_t . During the period, the demand is realized as d_t . The unit

sales price is denoted by p , thus the decision-maker collects a profit of $p \min\{y_t, d_t\}$ in period t . Then, the process proceeds to period $t+1$. For simplicity, assume that there is no holding cost or shortage cost at the end of each period, and that only sales, $\min(d_t, y_t)$, can be observed, giving us a problem with censored demand. So without knowing the demand distribution, the decision-maker aims to maximize the expected T -period profit.

Clearly, the single-period profit function is

$$G(y_t) = p\mathbb{E}[\min\{y_t, D_t\}] - cy_t, \quad (15.5)$$

where the expectation is with respect to random demand D_t , whose distribution is not known to the decision-maker. This is a nonparametric problem. Let the clairvoyant optimal solution be again denoted by y^* , which is the maximizer of Equation (15.5).

Consider the case that the demand follows a discrete distribution, and the decision-maker knows that the optimal inventory level cannot exceed some threshold $\bar{y}$. This problem can be considered as a multi-armed bandit problem, where there are $\bar{y}$ arms, each arm representing an integer-valued order-up-to level y .

The essence of the UCB algorithm is to explore each arm according to the upper confidence bound. This method ensures two things: First, it intends to spend more time playing the arm with a higher average historical reward, and second it allocates more exploration effort to those arms that have not been “sufficiently” explored.

We present a variation of the online UCB algorithm to solve the repeated newsvendor problem. The algorithm starts by implementing order-up-to level $\bar{y}$ in period 1, then the sales quantity $s = \min\{\bar{y}, d_1\}$ is observed and can be used to calculate the profit should another replenishment level $0 \leq i < \bar{y}$ was placed, which is equal to

$$r_i := p \min\{i, s\} - ic, \quad 0 \leq i \leq \bar{y}, \quad (15.6)$$

where “:=” stands for “defined as”. Let n_i denote the number of periods the inventory replenishment level i has been implemented. Then initially set $n_i := 1$ for $0 \leq i \leq \bar{y}$, and $t := 1$. For any arbitrary period $t \geq 1$, compute

$$UCB_i := \bar{r}_i + \max(h, p)\bar{y}\sqrt{2(\log t)/n_i},$$

where $\bar{r}_i$ denotes the average profit per ordering period received from replenishment decision i up to time $t-1$. Then in period t , replenish the inventory level to y_t , where

$$y_t := \arg \max_{0 \leq i \leq \bar{y}} UCB_i.$$

The observed sales data, i.e., $s = \min(d_t, y_t)$, is collected and used to compute the profit for all replenishment decisions $i \leq y_t$ using Equation (15.6). Then update $n_i := n_i + 1, i \leq y_t$, and update r_i for $i \leq y_t$. Finally, $t := t + 1$ and repeat the procedure.

The standard result from MAB shows that the regret for the UCB algorithm is $O(\sqrt{T \log T})$.

We remark that the UCB algorithm described above is more efficient than the standard UCB. In standard UCB, the reward for only one arm is collected in each iteration. In the

algorithm above, the reward for any arm $i \leq y_t$ is automatically obtained because of the structure of the problem at hand, hence it significantly reduces the exploration effort.

15.3 LEARNING OPTIMAL INVENTORY DECISIONS

In this section, we discuss the design of online learning algorithms for several important inventory control problems. We will see how some of the methods introduced in the previous section can be applied to each specific inventory model to achieve a near-optimal performance with a provable convergence rate.

15.3.1 Lower-Bound Results

The first paper that studies the lower bound of the regret for inventory learning problems is by Besbes and Muharremoglu (2013). This paper also proposes a learning-and-earning type algorithm, but in this subsection we focus on the lower-bound results presented in this paper. For both the discrete demand and continuous demand cases, it is assumed that there is a known upper bound of the optimal policy, $\bar{y}$.

Discrete demand cases. When the demand is discrete, it is shown by Besbes and Muharremoglu (2013) that if the demand is fully observable, the sample quantile policy, which always uses the empirical critical quantile as the ordering quantity in each period, has a regret of $O(1)$. However, if the demand is censored, a regret that grows over time is inevitable. It is shown that when the demand distribution $F(\cdot)$ satisfies the “minimal separation around optimal quantity” condition, i.e., $|F(x) - \beta| \geq \min\{\beta, 1 - \beta\} / 2$ for $x = x_F^*$ and $x_F^* - 1$, where β is the critical percentile $p / (h + p)$ and x_F^* is the optimal solution under $F(\cdot)$, the regret is lower bounded by $\Omega(\log T)$. When the minimal separation condition does not hold, the regret is lower bounded by $\Omega(\sqrt{T})$. For both cases, the regret is proved using the following example, by assigning different values to δ .

Example 15.1 For a T -period problem with a large enough T , consider the following two distributions:

$$F_a(i) = \begin{cases} \beta + \delta & \text{if } i = 0, \\ 1 & \text{if } i \geq 1; \end{cases} \quad F_b(i) = \begin{cases} \beta - \delta & \text{if } i = 0, \\ 1 & \text{if } i \geq 1. \end{cases}$$

We can see that the two distributions are very close to each other, but the optimal solutions are very different. The lower-bound results are based on this insightful example. We omit the details, and refer the interested reader to Besbes and Muharremoglu (2013).

Continuous demand cases. When demands are continuous, under the condition that $f(x) \geq \varepsilon > 0$ for all $x \geq 0$, the expected one-period cost is strongly convex, in this case, it is shown that the regret is lower bounded by $\Omega(\log T)$. This lower bound is proved by constructing a uniform distribution between $[\theta, 1]$, where θ varies between $[0, 1/2]$. When the strongly convex condition is not satisfied, we can use the distribution in Example 15.1, and add a small (continuous) noise term to make them continuous. In this case, the lower bound is still $\Omega(\sqrt{T})$, and the detailed proof is given by Zhang et al. (2020).

In summary, for both the discrete and continuous demand distributions, the lower bound under the general demand distribution is $\Omega(\sqrt{T})$, but it can be improved when the demand distribution satisfies some conditions.

15.3.2 Periodic-Review Inventory System with Censored Demand

In this subsection, we present the learning algorithm in Huh and Rusmevichientong (2009) for the periodic-review inventory system, where excess demand in each period is either lost or backlogged. Different from the repeated newsvendor problem, where the inventory expires at the end of each period, in many practical inventory control problems, the inventory would stay in the system at the end of each period. Although this is a natural feature of the inventory system, it creates a unique challenge to the design of learning algorithms for these systems, as each decision not only affects the current period, but also affects future periods. In this subsection, we start with the simple case where the replenishment lead time is zero.

Consider the periodic-review inventory system with an iid continuous demand in every period, denoted by D_t . If y_t is the inventory level at the beginning of period t after the ordering decision, then at the end of each period t , the inventory level is $(y_t - D_t)^+$, which stays in the system and becomes the starting inventory of the next period, denoted by x_{t+1} . When placing orders in each period t , the decision must satisfy $y_t \geq x_t$. The cost structure in each period is the same as the newsvendor problem defined in Equation (15.3).

With inventory carryover, we cannot freely select the inventory level in each period like in the repeated newsvendor model. Recall that the constraint in each period is $y_t \geq x_t$. With this constraint, we cannot freely decrease our base-stock level y_t , and sometimes it would be stuck in a higher level due to inventory carryover. But being forced to a higher inventory level is not going to hurt our demand observation, as a higher inventory level can reveal more demand information. Motivated by this observation, Huh and Rusmevichientong (2009) introduce the following gradient-based learning algorithm for this model, labelled as adaptive inventory management (AIM) algorithm.

Apart from the inventory dynamics in each period, the AIM algorithm maintains a *target inventory level* $\hat{y}_t$ in each period. The target inventory level follows the gradient-based update rule in the base model. With the inventory constraint, the actual inventory level y_t sometimes can be larger than $\hat{y}_t$. The whole dynamics can be summarized as follows, where $\mathbf{P}_{[a,b]}$ denotes the projection to set $[a,b]$,

- Learning Dynamics: $\hat{y}_{t+1} = \begin{cases} \mathbf{P}_{[0,\bar{y}]}(\hat{y}_t + \eta_t \cdot b), & \text{if } D_t < \hat{y}_t, \\ \mathbf{P}_{[0,\bar{y}]}(\hat{y}_t - \eta_t \cdot h), & \text{if } D_t \geq \hat{y}_t. \end{cases}$
- Inventory Dynamics: $x_{t+1} = (y_t - D_t)^+$, $y_{t+1} = \max\{y_{t+1}, x_{t+1}\}$.

Thus, the target inventory level $\hat{y}_t$ follows the base model, but it may not be achievable, so the actual inventory level could be higher in each period. From the Stochastic gradient descent results, we know that if we follow the $\hat{y}_t$, our regret will be $O(\sqrt{T})$. The only remaining piece is to show that $\mathbb{E}[\sum_{t=1}^T (y_t - \hat{y}_t)]$ is bounded by $O(\sqrt{T})$, as this will lead to a total worst-case regret of $O(\sqrt{T})$. Intuitively, as the step size η_t becomes smaller and smaller when t increases, the gap between $\hat{y}_t$ and y_t will also be smaller. In Huh and Rusmevichientong (2009), the

authors bound this gap by linking it with a GI/D/1 queue. We omit the details and refer the interested reader to their paper for the proof.

15.3.3 Perishable Inventory Systems

In the periodic-review inventory system discussed in the previous subsection, all the inventory units are assumed to have a lifetime of infinity, i.e., they will never expire. However, in practice, many products have finite lifetimes, like grocery and pharmaceutical products. Perishable inventory systems are important to our society, but managing perishable inventory products is known to be challenging even with full demand distribution information (see Nahmias (1975); Fries (1975)). In this subsection, we discuss the learning algorithm introduced by Zhang et al. (2018) for perishable inventory systems.

Model formulation. We present the system dynamics of the perishable inventory system as follows. We use $m \geq 2$ to denote the product lifetime for each product. Assume the demand in each period (D_t) is an iid continuous random variable. As the optimal policy is intractable, we focus on base-stock policies with a first-in-first-out (FIFO) issuing policy. The base-stock level is denoted as S , and under the FIFO issuing policy, we always use oldest inventory to meet the demand first. The lead time is assumed to be 0. To keep track of on-hand products with different remaining lifetimes, we need to use a m -dimensional vector. We use

$$\mathbf{x}_t = [x_{t,1}, \dots, x_{t,i}, \dots, x_{t,m-1}, x_{t,m}]$$

as the inventory vector at the beginning of period t , where $x_{t,i}$ represents the total inventory in period t with remaining lifetime $\leq i$. We will also use $\mathbf{0}$ to denote the vector with all elements being 0. Under the base-stock policy, we would have $x_{t,m-1} = S$ when there is no overshoot from the previous period.

For the cost parameters, we first assume that the unmet demand is lost and censored. Every period, apart from the holding and shortage cost, all the products with remaining lifetimes one that fail to meet the demand will expire at the end of the period, with a unit outdated cost of θ . The unit purchasing cost is sunk to 0 without loss of generality. The total cost in period t , C_t , can be computed as

$$C_t = h(x_{t,m} - D_t)^+ + b(D_t - x_{t,m})^+ + \theta(x_{t,1} - D_t)^+.$$

And the system dynamics from period t to $t+1$ is

$$x_{t+1,j} = \left(x_{t,j+1} - D_t - (x_{t,1} - D_t)^+ \right)^+, \quad \text{for } 1 \leq j \leq m-1,$$

$$x_{t+1,m} = \min(S, x_{t+1,m-1}).$$

Because of the complexity of perishable inventory system, we focus on finding best base-stock policy, which is known to perform well and is widely adopted in practice. Refer to Bu et al. (2023). The optimal base-stock policy S^* can be computed from

$$S^* = \operatorname{arginf}_S \left\{ \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T C_t^{\pi(S)} \right] \right\},$$

where $\pi(S)$ denotes the base-stock policy with base-stock level S . Our goal is to develop a learning algorithm that only uses the censored demand information to converge to S^* , with a regret of $O(\sqrt{T})$, which matches the lower bound of the regret.

Learning algorithm design. For the perishable inventory system, the cost in each period is determined by the decision—the base-stock level S , together with the starting inventory vector $\mathbf{x}_t$. And the decision in each period affects the starting inventory vectors in the following period. Hence we cannot directly apply the SGD method to this problem and update the decision in each period accordingly. To overcome this challenge, Zhang et al. (2018) introduce a novel idea of updating the decision in cycles, where the beginning of each cycle is defined as the period after the lost sales of the system.

First, it was shown that this would ensure convexity: For the perishable inventory system operated under a base-stock policy $\pi(S)$, if the system begins with empty inventory, then for any realization of demand $\omega = (d_1, d_2, \dots)$, the T -period total cost is convex in S for any $T \geq 1$. With this convexity result, we see that when there is a stockout in period t , all the inventory would be cleared, and the total cost from period $t+1$ to any future period is convex in the base-stock level S . We introduce the Cycle-Update Policy as follows. The algorithm assumes that there is a known compact set $[0, \bar{S}]$ that contains the optimal policy S^* .

- Initialization: Set $\tau_1 = 1$. Arbitrarily initialize the base-stock level S_1 within $(0, \bar{S})$. Let $x_{1,m-1} = S_1$, and the cycle counter to $k=1$.
- For each period t ,
 - If there is no stockout in period $t-1$, keep the same base-stock level as period $t-1$ in this period.
 - If there is stockout in period $t-1$, i.e., $\mathbf{x}_t = \mathbf{0}$, set $\tau_{k+1} = t$ as the beginning of a new cycle $k+1$. Update the base-stock level S_{k+1} by

$$S_{k+1} = \mathbf{P}_{[0, \bar{S}]}(S_k - \eta_k \nabla_1 G(S_k, (\tau_k, \tau_{k+1}); \omega)),$$

where η_k is the step size defined by $\eta_k = \gamma/\sqrt{k}$ for some positive constant γ , and $\nabla_1 G(S_k, (\tau_k, \tau_{k+1}); \omega)$ is the gradient of the k -th cycle cost with respect to S_k , and with fixed τ_k and τ_{k+1} . Apply base-stock level S_{k+1} and proceed to the next period.

With censored demand information, we can compute the gradient of the cycle cost with respect to the base-stock level. We refer the interested reader to Zhang et al. (2018) for details. We can see that the major difference between the CUP algorithm and the conventional SGD algorithm is the introduction of cycles with random cycle lengths.

Regret analysis. We define the T -period regret of CUP, $\mathcal{R}_T^{\text{CUP}}(\omega)$, as

$$\mathcal{R}_T^{\text{CUP}}(\omega) = \mathbb{E} \left[\sum_{t=1}^T \left(C_t^{\pi(S_t)}(\omega) - C_t^{\pi(S^*)}(\omega) \right) \right],$$

where S_t is the base-stock level under the CUP algorithm, and S^* is the clairvoyant optimal base-stock policy (OPT for short). It is shown by Zhang et al. (2018) that the regret of the CUP

algorithm is $O(\sqrt{T})$, where the constant term in the big O notation is in the form of $k_1/\gamma + k_2\gamma$, for some constant terms k_1 and k_2 that depend on other problem parameters.

The main challenge of the regret analysis in comparing the CUP system and the OPT system is that at the beginning of each cycle, the CUP system has no inventory, but the OPT system could still have inventory in the system, which complicates the comparison. To overcome this challenge, Zhang et al. (2018) introduce a bridging policy, called the replacement of old inventories (ROI for short). The ROI system also keeps the base-stock level S^* . But at the beginning of each cycle, the ROI system replaces all the inventory with brand-new inventory, so that the ROI system will have the same holding and lost-sales cost as the OPT system, but pays a lower outdated cost. Indeed, it is shown that for each problem instance of the perishable inventory system, given any sample path $\omega = \{d_1, d_2, \dots\}$ and any $T \geq 1$, the total cost incurred by the bridging policy ROI is less than or equal to the total cost incurred by the optimal base-stock policy $\pi(S^*)$.

As the ROI system gives a lower cost in each sample path than the OPT system, when bounding the regret, we can replace the OPT cost with the ROI cost. The remaining task is to analyze the cost difference between the CUP system and the ROI system. Recall that within each cycle, the costs under both systems are convex functions of the respective base-stock levels. This cost gap can be bounded following standard SGD proofs, and we refer the interested reader to Zhang et al. (2018) for the details.

15.3.4 Lost-Sales Model with Positive Lead Time

Lead times, the time between an order is placed and arrival, are common in real-world supply chain systems. Lost-sales inventory systems with lead times are ubiquitous, but notoriously difficult to solve, as it is a high-dimensional problem and the optimal policy is known to be complicated. We refer the interested reader to Zipkin (2008) for the state-of-art results for lost-sale inventory systems with positive lead times and complete information of demand distribution.

In this subsection, we discuss three online learning algorithms designed for this model when demand distribution is not known *a priori*. As the full-information optimization problem is known to be hard to solve, all of these three algorithms focus on finding the optimal *base-stock* policy, instead of the true optimal policy. The *base-stock* policy is not optimal, but it is widely adopted due to its simplicity and near-optimal performance. Refer to Huh et al. (2009).

We first rigorously define the system dynamics for the lost-sales inventory system under the base-stock policy. The demands are still assumed to be iid across periods and are continuous. The lead times are assumed to be L periods. For every period t , the starting inventory state is denoted by

$$x_t = [q_{t-1}, \dots, q_{t-L+1}, I_t],$$

where I_t is the on-hand inventory at the beginning of period t , and q_k is the order placed in period k . Following the base-stock policy with base-stock level S , we have $q_t = (S - I_t - \sum_{k=t-L+1}^{t-1} q_k)^+$. Let $y_t = [q_t, \dots, q_{t-L+1}, I_t]$ denote the inventory vector after the ordering. The demand d_t is realized after the order q_t is placed. The one-period cost is $h \cdot (I_t - d_t)^+ + b \cdot (d_t - I_t)^+$, and the starting inventory state in the next period becomes $y_t = [q_t, \dots, q_{t-L+2}, I_{t+1} = q_{t-L+1} + (I_t - d_t)^+]$. We assume the starting inventory vector is $\mathbf{0}$, and the lost sales are not observable.

Challenges in learning. For this problem, due to the high-dimensional inventory state and complicated inventory dynamics, every order affects the future inventory and costs. In order to learn the impact of each base-stock level, one cannot simply test it for one period, or even L periods to check the cost. An effective algorithm has to carefully design the testing cycles, where within each cycle, the system maintains the same base-stock level. We will see that in all three algorithms, the cycle idea is involved.

A preliminary convexity result. Before introducing the algorithms, there is a preliminary convexity result on the base-stock policy for the lost-sales inventory system with lead times introduced in Janakiraman and Roundy (2004). For ease of presentation, we present a slight variant of their result as follows: For any fixed starting inventory state and a given demand sample path, the total N -period cost is convex in the base-stock level S .

Gradient-based method with increasing cycle lengths. The first learning algorithm introduced for this model is by Huh et al. (2009). The algorithm is a gradient-based algorithm. We discuss the main idea of the algorithm as follows.

The algorithm runs in cycles. The cycle length of cycle k is set to be $\lceil \sqrt{k} \rceil$. Within each cycle k , the algorithm maintains the same base-stock level S_k . At the end of the cycle, the algorithm uses the gradient information in the *last* period of the cycle to update the base-stock level. Note that the gradient information can be computed using censored demand data.

In their paper, the authors show that as the cycle length increases, the bias of using the last period gradient will converge to zero. This ensures the convergence of the algorithm. However, the drawback of using increasing cycle lengths for a gradient-based method is the slow convergence rate. As the cycle length increases, the worst-case regret rate of the algorithm becomes $O(T^{2/3})$, which is not tight.

Gradient-based method with stationary cycle lengths. The second paper that studies this problem is by Zhang et al. (2020). They propose another gradient-based algorithm that can achieve a $\tilde{O}(\sqrt{T})$ regret rate. As the algorithm itself contains many technical details, we have to omit some of them and focus on the main ideas. Different from the algorithm in Huh et al. (2009) that adopts cycles with fixed and increasing lengths, the algorithm in Zhang et al. (2020) uses cycles with random cycle lengths, but they are stationary and non-increasing. With stationary cycle lengths, one has to evaluate the cycle gradient in an unbiased manner, without relying on the natural convergence from increasing cycle lengths. However, this is challenging when we keep changing the base-stock policies between cycles, and the cost within each cycle depends on the policies used in the previous cycles. One has to break these dependencies to be able to evaluate the true cycle gradient in this case.

Similar to Huh et al. (2009), it is assumed that the decision-maker knows an upper and lower bound of the optimal base-stock level, denoted as $\bar{S}$ and $\underline{S}$. An important observation is that in the same sample path, if between period $t-L$ to period $t-1$, the system that follows base-stock level $\underline{S}$ ($\underline{S}$ system in short) has no lost sales for these L consecutive periods, then in period t all the systems with a base-stock level higher than $\underline{S}$ will have the same pipeline inventory, and the only difference will be on the on-hand inventory, i.e., the inventory vector after ordering will be $[d_{t-1}, \dots, d_{t-L}, S - \sum_{k=t-L}^{t-1} d_k]$. We say that a triggering event occurs whenever there is no lost-sale for L consecutive periods in the $\underline{S}$ system, and we call the period right after the triggering event a triggering period. Denote the period that k -th time the triggering event happens as τ_k . We can see that if we increase the base-stock level from S_1 to S_2 right after τ_k , then in the next triggering period τ_{k+1} , the inventory vector of the system will be $[d_{\tau_{k+1}-1}, \dots, d_{\tau_{k+1}-L}, S_2 - \sum_{k=\tau_{k+1}-L}^{\tau_{k+1}-1} d_k]$, which is no longer affected by S_1 . Following this idea, we are able to break the dependencies between cycles.

However, to detect the triggering event, we need to maintain a correct simulation of the $\underline{S}$ system throughout the planning horizon using the sales data of the learning system. A sufficient condition for achieving the correct simulation of the $\underline{S}$ -system is to ensure that the system under the learning algorithm always has no lower on-hand inventory than the simulated $\underline{S}$ -system. Denote the system under the learning algorithm as the π system. Suppose the states of the learning system and the simulated $\underline{S}$ -system at the beginning of period t are $(q_{t-1}^a, q_{t-1}^a, \dots, q_{t-L+1}^a, I_t^a)$, $a = \pi, \underline{S}$, respectively. Then the on-hand inventory level at the beginning of period $t+1$ would be

$$I_{t+1}^a = q_{t-L+1}^a + (I_t^a - d_t)^+, \quad a = \pi, \underline{S}.$$

In general, we may not be able to simulate the $\underline{S}$ system using only the sales quantity $\min(I_t^\pi, d_t)$. However, if $I_t^\pi \geq I_t^{\underline{S}}$, then the $\underline{S}$ -system can be correctly simulated because

$$I_{t+1}^{\underline{S}} = q_{t-L+1}^{\underline{S}} + (I_t^{\underline{S}} - d_t)^+ = q_{t-L+1}^{\underline{S}} + [I_t^{\underline{S}} - \min(d_t, I_t^\pi)]^+.$$

This shows that, under the condition $I_t^\pi \geq I_t^{\underline{S}}$ for all t , the $\underline{S}$ -system can be correctly simulated by pretending that the demand in period t is equal to the sales quantity in the π -system.

The algorithm proposed by Zhang et al. (2018) maintains enough inventory in the π -system through a very careful design on transitions between cycles when the inventory level is updated. We omit these parts, and interested readers can check the algorithm in the paper for details. In summary, the algorithm makes use of the triggering periods to break the dependency between cycles, and maintains enough inventory to simulate the $\underline{S}$ -system to detect the triggering periods. As the cycle time is stable, the expected T -period regret of the algorithm is $O(\sqrt{T})$. However, by the definition of the triggering periods, the expected cycle length has an exponential dependence on the lead time L . The worst-case regret also exponentially increases in L . The next paper improves this dependency by following a very different approach.

An advanced line search method with UCB/LCB. The third paper that studies this problem is by Agrawal and Jia (2019). Unlike the previous two methods that utilize the gradient methods for this problem, this paper proposed an advanced line search method. The classic line (one-dimensional) search methods, such as the golden section search and the Fibonacci search, are for deterministic problems. In these methods, the goal is to find the maximum (minimum) point of a one-dimensional unimodal function over a compact interval $[L, U]$. We can efficiently shrink this interval by evaluating two points and comparing their value in each iteration. In this inventory problem, the convexity/concavity of the expected cost/profit ensures the unimodality. With censored demand observation, it is easy to work on the expected profit. However, the challenge of the random demand remains. If one can evaluate the expected profit under a certain base-stock level, then we can directly apply the existing methods. This is not achievable under finite time, because of 1) the demand randomness, and 2) the initial state effect.

Agrawal and Jia (2019) propose a method to overcome these challenges. First, they show that the effect of the initial state on the profit is upper bounded by a constant which is linear in the lead time L . This means one does not need to “calibrate” the initial state at the beginning of each cycle, like in Zhang et al. (2020), as the error can be bounded. Second, they propose a novel idea to avoid evaluating the expected profit $Q(S_1)$ and $Q(S_2)$ under two base-stock levels

S_1 and S_2 . Instead, they construct the upper and lower confidence bounds to help to shrink the interval. The main idea is that if the LCB of the profit at a base-stock level S_1 , denoted by $LB(S_1)$, is higher than the UCB of the profit at another base-stock level S_2 , denoted by $UB(S_2)$, then we can conclude that with a very high chance, $Q(S_1)$ is higher than $Q(S_2)$. We refer interested readers to their paper for detailed algorithm construction. The worst-case regret of the algorithm is $\tilde{O}(\sqrt{T})$ and the constant term is linear in the lead time L .

15.3.5 Multiple Products with Warehouse-Capacity Constraint

With the recent growth of fast delivery services offered by online platforms, companies aim at utilizing warehouses that are in close proximity to the end customers. Often that will lead to warehouses with limited storage capacity. Shi et al. (2016) consider the multi-product inventory model with warehouse-capacity constraint, and propose a gradient-based nonparametric learning algorithm to achieve a $O(\sqrt{T})$ worst-case regret.

Inventory model formulation and clairvoyant optimal policy. The inventory model considered by Shi et al. (2016) is a multi-product inventory system, where each product follows a periodic-review inventory control with censored demand observation. The total number of products is n , and the warehouse-capacity constraint is denoted as M , where the feasible region is defined as $\Gamma \triangleq \{\mathbf{y}_t \in \mathbb{R}_+^n : \sum_{i=1}^n y_t^i \leq M\}$, i.e., the total inventory cannot exceed M . The goal of the optimization problem is to minimize the expected cost $\mathcal{C}(\pi)$, where

$$\mathcal{C}(\pi) = \sum_{t=1}^T \mathbb{E}[\mathbf{c} \cdot \mathbf{y}_t + (\mathbf{h} - \mathbf{c}) \cdot (\mathbf{y}_t - \mathbf{D}_t)^+ + \mathbf{b} \cdot (\mathbf{D}_t - \mathbf{y}_t)^+],$$

and the starting inventory is assumed to be empty. We have $\mathbf{c}$, $\mathbf{h}$, and $\mathbf{b}$ representing the per-unit purchasing, holding, and lost-sales penalty cost vector, while $\mathbf{y}_t$ and $\mathbf{D}_t$ represent the inventory vector after purchase, and the demand vector, in period t . The demands are assumed to be continuous, and independent across time and products, and with a positive mean. It is also assumed that the CDF of the demand of each product is differentiable, with a non-zero density function on $[0, M]$ to ensure the (joint) convexity of the expected cost $\mathcal{C}(\pi)$, and there is a unique minimizer of $\mathcal{C}(\pi)$ within $\mathbf{y} \in \Gamma$. Denote this minimizer as $\mathbf{y}^*$. Despite the interactions between products imposed by the warehouse-capacity constraint, it can be shown that a myopic minimizer $\mathbf{y}^*$ is the optimal base-stock level for the finite-period optimization problem.

Challenges of the learning algorithm design. The goal of the learning algorithm is to use the censored demand observation to converge to the optimal policy. If there is no inventory carryover, then as the cost is jointly convex in each period, we can update the base-stock levels $\mathbf{y}_t$ following the exact SGD method. However, with inventory carryover and warehouse-capacity constraint, (1) we may not be able to decrease the base-stock level of a certain product i due to excess inventory of i ; and (2) we may not be able to increase the base-stock level of i due to excess inventory of other products, because of the warehouse-capacity constraint. As a result, we may not observe the gradient information due to the lack of inventory and censored demand information.

Learning algorithm design and analysis. To overcome the challenges, Shi et al. (2016) propose the Data-Driven Multi-product Algorithm (DDM). DDM maintains a vector triplet of sequences $(\mathbf{z}_t, \hat{\mathbf{y}}_t, \mathbf{y}_t)_{t \geq 0}$, where $(\mathbf{z}_t)_{t \geq 0}$ represents the *constraint-free* target inventory level, $(\hat{\mathbf{y}}_t)_{t \geq 0}$ represents the *target* inventory level, and $(\mathbf{y}_t)_{t \geq 0}$ represents the *actual implemented* inventory level.

The algorithm proceeds as follows:

- Initialization: Set $t=0$, and arbitrarily set $\mathbf{y}_0 = \hat{\mathbf{y}}_0 = \mathbf{z}_0$ be any values within Γ .
- Determining $\mathbf{z}_{t+1}$ and $\hat{\mathbf{y}}_{t+1}$:
 1. At the beginning of each period t , if $\mathbf{y}_t \geq \hat{\mathbf{y}}_t$ (i.e., $y_t^i \geq \hat{y}_t^i$ for all $i = 1, \dots, n$), the algorithm updates the constraint-free target inventory level $\mathbf{z}_{t+1}$ by

$$\mathbf{z}_{t+1} = \hat{\mathbf{y}}_t - \frac{\eta}{\sqrt{t}} \mathbf{G}_t(\hat{\mathbf{y}}_t),$$

where η is a positive constant, and the i^{th} component of $\mathbf{G}_t$ is defined as

$$G_t^i(\hat{\mathbf{y}}_t) = \begin{cases} h^i & \text{if } \hat{y}_t^i > d_t^i, \\ -(b^i - c^i) & \text{if } \hat{y}_t^i \leq d_t^i. \end{cases}$$

Then solve the target inventory level $\hat{\mathbf{y}}_t$ as $\hat{\mathbf{y}}_t = \arg \min_{\mathbf{w} \in \Gamma} \|\mathbf{w} - \mathbf{z}_{t+1}\|_2$.

2. Otherwise, keep the constraint-free and constrained target inventory levels unchanged, i.e., $\mathbf{z}_{t+1} = \mathbf{z}_t$ and $\hat{\mathbf{y}}_{t+1} = \hat{\mathbf{y}}_t$.
- Determining $\mathbf{y}_{t+1}$: Define the set J and its complement as

$$J \triangleq \{i : x_{t+1}^i > \hat{y}_{t+1}^i\}, \quad \bar{J} \triangleq \{i : x_{t+1}^i \leq \hat{y}_{t+1}^i\}.$$

1. For each product $i \in J$, set $y_{t+1}^i = x_{t+1}^i$.
2. For the remaining products, determine the $\mathbf{y}_{t+1}$ by solving the following optimization problem:

$$\min \sum_{i \in \bar{J}} (y_{t+1}^i - \hat{y}_{t+1}^i)^2 \quad \text{s.t.} \quad \sum_{i \in \bar{J}} y_{t+1}^i \leq M - \sum_{j \in J} x_{t+1}^j, \quad y_{t+1}^i \geq x_{t+1}^i, \quad \forall i \in \bar{J}. \quad (15.7)$$

We see that the learning component of the DDM algorithm follows the SGD method. To overcome the challenges imposed by inventory carryover and warehouse-capacity constraints, Shi et al. (2016) introduce the constraint-free target inventory levels $(\mathbf{z}_t)_{t \geq 0}$ and the target inventory levels $(\hat{\mathbf{y}}_t)_{t \geq 0}$.

After the initialization, the first step is the gradient step. In this step, to ensure unbiasedness, we can only use the demand observation from base-stock level $\mathbf{y}_t$ if $\mathbf{y}_t \geq \hat{\mathbf{y}}_t$. The gradient information is used to update the constraint-free target inventory levels $\mathbf{z}_{t+1}$, then it is projected back to Γ to form the target inventory levels $\hat{\mathbf{y}}_t$. If the actual inventory level is less than the target inventory level, we keep $\mathbf{z}$ and $\hat{\mathbf{y}}$ unchanged.

In the second step of constructing $\mathbf{y}_{t+1}$ from $\hat{\mathbf{y}}_{t+1}$, the challenge is the warehouse-capacity constraint. We may not be able to reach $\hat{\mathbf{y}}_{t+1}$, for example, when there is a product with a lot of excess inventory carryover from the last period. DDM algorithm solves the optimization problem in Equation (15.7) to set the $\mathbf{y}_{t+1}$ to be as close to $\hat{\mathbf{y}}_{t+1}$ as possible while retaining the feasibility.

The regret analysis of DDM is separated into two parts, $\mathbb{E}\left[\sum_{t=1}^T \Pi(\hat{\mathbf{y}}_t) - \sum_{t=1}^T \Pi(\mathbf{y}^*)\right]$, and $\mathbb{E}\left[\sum_{t=1}^T \Pi(\mathbf{y}_t) - \sum_{t=1}^T \Pi(\hat{\mathbf{y}}_t)\right]$. The first part is the usual SGD regret. Note that as $\hat{\mathbf{y}}_t$ is not updated every period, the convergence is slowed down, but only by a constant, and hence the first part is still $O(\sqrt{T})$. The constant term in the big O notation is in the form of $k_1/\eta + k_2\eta$ for some constant terms k_1 and k_2 that are dependent on the problem parameters. For the second part, the regret is bounded by using a GI/G/1 queue as a bridge, extending the analysis in Huh and Rusmevichientong (2009) to this multi-product system. The analysis for this part is more involved and we refer the interested reader to Shi et al. (2016) for the detailed proof.

15.3.6 Multiple Products with Substitution

When a product runs out of stock and a demand arises, it is possible that the demand will buy some substitutable products. If both demand distribution and substitution probabilities are unknown to the decision-maker *a priori*, how to make inventory replenishment decisions? This subsection is based on Chen and Chao (2020) who study the inventory control problem for multiple products with stockout substitutions. We will review the model, learning algorithm, and regret rate.

Demand process. A firm selling $K \geq 2$ substitutable products at fixed prices $p^1, \dots, p^K$ over T periods. Let $\mathbf{x}_t = (x_t^1, \dots, x_t^K)$ and $\mathbf{y}_t = (y_t^1, \dots, y_t^K)$ denote the inventory levels before and after replenishment, respectively, at the beginning of period t . For each product k , the primary demands D_t^{0k} , $t = 1, \dots, T$, are iid discrete random variables with unknown probability mass function $\mathbf{w}^k$. Let $\mathbf{w} = (\mathbf{w}^1, \dots, \mathbf{w}^K)$.

In each period t , the primary demand D_t^{0k} is satisfied as much as possible by on-hand inventory y_t^k , and when a stockout occurs, an unsatisfied customer switches to product j as a substitute with probability a^{kj} , where $a^{kk} = 0$ and $\sum_j a^{kj} \leq 1$. But if this customer experiences stockout again at product j , then she leaves the system. Let $a^{k0} = 1 - \sum_j a^{kj}$ be the probability that an unsatisfied primary customer leaves the system without attempting a substitutable product. Suppose that primary demands are satisfied first. Let $D_t^{kj}(a^{kj}, y_t^k)$ denote the number of customers who face a stockout at k then attempt to substitute j , and note that it depends on the substitution probability as well as the inventory level of product k . Let $\mathbf{a}^k = (a^{k1}, \dots, a^{kK})$ and $\mathbf{a} = (\mathbf{a}^1, \dots, \mathbf{a}^K)$.

We denote the aggregate demand for product k by $D_t^k(\mathbf{w}, \mathbf{a}, \mathbf{y}_t) = D_t^{0k} + \sum_j D_t^{jk}(a^{jk}, y_t^j)$, which depends on the inventory levels of all the products due to the stockout substitutability feature among multiple products. The firm knows neither $\mathbf{w}$ nor $\mathbf{a}$ *a priori*. We suppress the dependence of $D_t^k(\mathbf{w}, \mathbf{a}, \mathbf{y}_t)$ on $\mathbf{w}, \mathbf{a}, \mathbf{y}_t$ when no confusion may arise.

Information structure and system dynamics. The firm observes the sales quantity of each product k , which is the truncated total demand, i.e., $s_t^k = \min\{D_t^k, y_t^k\}$. From the sales data, the firm *cannot* differentiate if one customer belongs to the primary demand k or substitution demand from another product j .

At the beginning of period t , the firm observes the starting inventory levels $\mathbf{x}_t$. The firm determines an ordering quantity for each product to bring the inventory levels to $\mathbf{y}_t \geq \mathbf{x}_t$. The unit ordering cost for each product is, without loss of generality, normalized to 0. The primary demand D_t^{0k} for each product arrives and is satisfied by on-hand inventory as much as possible. If stock-out occurs, substitution may take place. Unsatisfied demands are *lost* and *unobservable*. At the end of period t , the firm collects a profit of $\sum_{k=1}^K (p^k \min\{D_t^k, y_t^k\} - h^k (y_t^k - D_t^k)^+ - b^k (D_t^k - y_t^k)^+)$, where h^k is the per-unit holding cost for inventory k and b^k is the shortage cost every time a demand for product k is rejected. State transition is $x_{t+1}^k = (y_t^k - D_t^k)^+$ for $1 \leq k \leq K$.

The firm aims to construct an admissible policy $\{y_t, t=1, \dots, T\}$ with $y_t^k \geq x_t^k$ and $x_t^k = (y_{t-1}^k - D_{t-1}^k(\mathbf{w}, \mathbf{a}, \mathbf{y}_{t-1}))^+$, such that the expected T -period total profit is maximized.

Clairvoyant optimal policy. If the firm had complete knowledge of the primary demand distributions $\mathbf{w}$ and the substitution probabilities $\mathbf{a}$ *a priori*, by Sobel (1981), a myopic policy is optimal for this problem, and to maximize the T -period total profit, it suffices to focus on the single-period problem $\max_{\mathbf{y}} Q(\mathbf{w}, \mathbf{a}, \mathbf{y})$, where

$$Q(\mathbf{w}, \mathbf{a}, \mathbf{y}) = \sum_{k=1}^K \left(p^k \mathbb{E} \left[\min\{D^k(\mathbf{w}, \mathbf{a}, \mathbf{y}), y^k\} \right] - h^k \mathbb{E} \left[(y^k - D^k(\mathbf{w}, \mathbf{a}, \mathbf{y}))^+ \right] - b^k \mathbb{E} \left[(D^k(\mathbf{w}, \mathbf{a}, \mathbf{y}) - y^k)^+ \right] \right).$$

Let $\mathbf{y}^*$ be the optimal solution. The clairvoyant optimal policy for the T -period problem is to implement order-up-to levels $\mathbf{y}^*$ in every period. Regret of a learning algorithm π is defined as

$$R^\pi(T) = V^*(T) - V^\pi(T),$$

where $V^*(T)$ is the optimal profit of the clairvoyant solution over T periods and $V^\pi(T)$ is the profit of algorithm π .

When $(\mathbf{w}, \mathbf{a})$ are not known, $Q(\mathbf{w}, \mathbf{a}, \mathbf{y})$ cannot be evaluated, and $\mathbf{y}^*$ cannot be computed. Therefore, the firm needs to learn the demand structure using data.

Learning algorithm. Next, we briefly review the outline of the learning algorithm by Chen and Chao (2020).

- The algorithm divides the planning horizon into learning cycles, and further splits each cycle into an exploration phase followed by an exploitation phase. The exploration phase contains $K+1$ intervals, with the length of each interval exponentially increasing in the cycle index. The length of the exploitation phase double exponentially increases in the cycle index. Out of the $K+1$ exploration intervals, the first K ones are the cyclic exploration intervals, and the last one is the benchmark interval.
- During the k -th cyclic exploration interval, the algorithm does not order new inventory for the k -th product, while keeping the inventory level of other products high. As a consequence, most stockouts happening during the k -th cyclic interval is from the k -th product to other products.
- During the benchmark interval, the algorithm keeps the inventory level of all products high, so that the number of stockouts is significantly reduced and the majority of customers are satisfied with their primary product.

- At the end of the exploration phase, the algorithm uses sales data collected during the benchmark interval to approximate the demand distribution for each product and obtain the estimator $\hat{\mathbf{w}}$. The algorithm then compares the sales data during the k -th, $k = 1, \dots, K$, cyclic interval with that during the benchmark interval to calculate the difference, which is considered as the surrogate stockouts and used to estimate the substitution probability from the k -th product to other products. Let the estimator for the substitution probability be $\hat{\mathbf{a}}$.
- With the estimators $\hat{\mathbf{w}}$ and $\hat{\mathbf{a}}$, the algorithm maximizes the objective function $Q(\hat{\mathbf{w}}, \hat{\mathbf{a}}, \mathbf{y})$ and solves for the optimizer $\hat{\mathbf{y}}$.
- The algorithm implements $\hat{\mathbf{y}}$ in every period of the exploitation phase.

Regret. The upper bound for the regret of the learning algorithm is

$$O((\log T)(\log \log T)^2).$$

This matches the theoretical lower bound of $\Omega(\log T)$ (proved in the paper) up to the lower order $\log \log T$ term.

Finally, we add a remark that the learning algorithm proposed by Chen and Chao (2020) can be applied to general, Markov chain-based, stockout substitution models in which a customer attempts to substitute (when facing stockout) multiple times according to a Markovian fashion before leaving the system. They allow the number of substitution attempts to be either finite or infinite, and the product allocation rule can be very general such as (1) customers are satisfied on a first-come-first-served basis, or (2) primary customers are satisfied before substitution customers, and substitution demands (from various sources) are satisfied randomly. It is proved that the regret rate remains the same for this more general scenario.

15.3.7 Learning the (s, S) Policy for the Lot-Sizing Problem

It is well-known that the optimal policy for an inventory problem with a setup cost is a (s, S) policy. A (s, S) policy places no order when the starting inventory x_t is no less than s , and will place an order to raise the inventory level up to S otherwise. Yuan et al. (2021) develop a learning algorithm for inventory problems with setup cost when demand distribution is not known *a priori*.

Transforming the objective. Consider the lost-sales model with censored demand, and zero ordering lead time. Yuan et al. (2021) first transform the problem into a “pseudo cost” minimization problem, which is essentially the profit maximization problem. The key step in this transformation is to transform the shortage cost $b(d_t - x_t - q_t)^+$ to $bd_t - b \min(x_t + q_t, d_t)$, and then drop the bd_t term as it is not affected by the decision. After the transformation, the pseudo cost is defined as

$$\tilde{C}_t(x_t, q_t, d_t) = K \cdot 1[q_t > 0] + cq_t + h(x_t + q_t - d_t)^+ - b \min(x_t + q_t, d_t).$$

The $-\tilde{C}_t(x_t, q_t, d_t)$ represents the one-period profit.

To facilitate the discussion, we use the (δ, S) representation of the (s, S) policy, where $\delta := S - s$. For a fixed policy, use τ_i to denote the period that the i -th order is placed. We call

the time between τ_i and τ_{i+1} a cycle. Denote the expected cycle length of policy (δ, S) as $\mathbb{E}[L(\delta, S)]$. Define the cycle pseudo cost of cycle i as

$$G_i(\delta, S) := c(x_{\tau_i} - x_{\tau_{i+1}}) + \sum_{t=\tau_i}^{\tau_{i+1}-1} \tilde{C}_t.$$

Following this definition, apart from the second term which accounts for the usual cycle cost, we also have the first term to account for the inventory level difference at τ_i and τ_{i+1} on the purchasing cost. Let $\mathbb{E}[G(\delta, S)]$ be the expected cycle pseudo cost for a time-generic cycle. It can be shown that the optimal policy π^* can be defined as follows using the Renewal Reward Theorem (see, e.g., Ross, 1996):

$$\pi^* = (\delta^*, S^*) = \arg \min_{(\delta, S)} \frac{\mathbb{E}[G(\delta, S)]}{\mathbb{E}[L(\delta, S)]}.$$

Let $V(\delta, S) = \mathbb{E}[G(\delta, S)] / \mathbb{E}[L(\delta, S)]$, and denote the optimal value under the optimal policy as V^* . The goal of the learning algorithm is to converge quickly to V^* .

Properties of the transformed objective. Before introducing the learning algorithm, it is necessary to look at the properties of the transformed objective $V(\delta, S)$. The paper considers the iid demand, with a bounded pdf, i.e., $f(d) < \rho$ for all $d \geq 0$ and a constant ρ . Unlike the previous models including those with positive lead times or warehouse-capacity constraints, the (joint) convexity no longer holds. Although $V(\delta, S)$ is partially convex along the S dimension, it is not convex along the δ dimension. We highlight the important system properties as follows while omitting some of the technical details.

- For a fixed δ , $V(\delta, S)$ is Lipschitz continuous and convex in S , and the Lipschitz constant can be independent of δ .
- For a fixed δ , given a cycle with length L and ending inventory level x_{L+1} , the stochastic (S -partial) gradient of $V(\delta, S)$, $\tilde{\nabla}$, can be calculated as

$$\tilde{\nabla} \begin{cases} hL & \text{if } x_{L+1} > 0, \\ h(L-1) - b + c & \text{if } x_{L+1} = 0. \end{cases}$$

- Let $V^*(\delta) := \min_S V(\delta, S)$. We have that $V^*(\delta)$ is Lipschitz continuous in δ .

Learning algorithm sketch and regret results. We see that the lack of joint convexity gives rise to a new level of challenge. To overcome these, Yuan et al. (2021) combine the gradient methods and bandit methods in a novel way. The feasible region of the policies is assumed to be $0 \leq \delta \leq S \leq \beta$, where β is a constant. The algorithm is parameterized by parameters N, J, η_n , and Δ_n . We discuss the tuning of these parameters after the algorithm. We present the main steps of the algorithm as follows.

- Initialization
 - Discretize the feasible region of δ , $[0, \beta]$, into J equally spacing set with J gaps, $\{\delta_1, \dots, \delta_J\}$, where $\delta_j = j\beta / J$.

- For each δ_j , arbitrarily set the associated S_j^1 within $[\delta_j, \beta]$.
- Maintain an active set $\mathcal{A}^n$ throughout the learning algorithm, and initialize $\mathcal{A}^1 = \{1, \dots, J\}$. For each element j in the active set $\mathcal{A}^n$, also maintain a cumulative cycle pseudo cost $\hat{G}_j^{n-1}$, and a cumulative cycle length $\hat{L}_j^{n-1}$. Initialize $\hat{G}_j^0 = \hat{L}_j^0 = 0$ for all $j = 1, \dots, J$.
- Execution step
 - At the beginning of each epoch $n, n = 1, \dots, N$, among all the active policies (δ_j, S_j^n) in the active set $\mathcal{A}^n$, find the policy with the largest S . Denote the index of the policy with the largest S as j^n . The policy to implement in this epoch is $(\delta^n, S^n) = (\delta_{j^n}, \max(x^n, S_{j^n}))$, where x^n is the starting inventory level at the beginning of epoch n .
 - Find the policy with the largest δ among all the active policies, and denote it as $\bar{\delta}^n$. Run policy (δ^n, S^n) for some complete cycles until the cumulative demand is larger than $\bar{\delta}^n$.
 - Simulate all the active policies once using the censored demand data obtained from executing policy (δ^n, S^n) . Collect the simulated cycle pseudo cost G_j^n , cycle length L_j^n , and the gradient $\tilde{V}_j^n$.
- Prune the active set
 - For each $j \in \mathcal{A}^n$, update $S_j^{n+1} = \mathbf{P}_{[\delta_j, \beta]}(S_j^n - \eta_n \tilde{V}_j^n)$, and update the empirical information $\hat{G}_j^n = \hat{G}_j^{n-1} + G_j^n$ and $\hat{L}_j^n = \hat{L}_j^{n-1} + L_j^n$.
 - Update the active set by

$$\mathcal{A}^{n+1} = \left\{ j \in \mathcal{A}^n : \frac{\hat{G}_j^n}{\hat{L}_j^n} - \min_{j' \in \mathcal{A}^n} \frac{\hat{G}_{j'}^n}{\hat{L}_{j'}^n} \leq \Delta^n \right\},$$

and proceed to the next epoch with $n := n+1$. Terminate the algorithm when $n=N$.

This algorithm is named as the (δ, S) learning algorithm by Yuan et al. (2021). We can see that the (δ, S) algorithm has two layers to learn the optimal policy. For δ , it adopts the idea of the Action Elimination algorithm, which is an algorithm designed for bandit problems. Along with the action elimination to find a good candidate set of δ 's, the second layer adopts the gradient methods to improve the S_j for each δ_j . Intuitively, when the associated S_j for each δ_j is not very close to the optimal one, the cost could be higher. But as each active policy is tested or simulated at least once in each epoch, the S_j will also converge to the optimal S_j^* simultaneously.

In Yuan et al. (2021), it is proved that when the parameters are tuned as

$$J = \lfloor \sqrt{N} \rfloor, \quad \eta_n = k_1 \frac{1}{\sqrt{N}}, \quad \Delta^n = k_2 \frac{\log(8N^2)}{\sqrt{N}},$$

where k_1 and k_2 are constant terms determined by parameters associated with assumptions on demand, then the regret over N epochs is bounded by $O(\log N \sqrt{N})$. One caveat of this result is that it depends on N rather than the planning horizon T , which may not be known *a priori*. To overcome this, one can adopt the so-called *doubling trick* to remove this requirement and create an “anytime” algorithm, which will give a $O(\log T \sqrt{T})$ worst-case expected total regret. We refer the interested reader to their paper for detailed discussions and proofs.

15.3.8 Tailored Base-Surge Policy for Dual-Sourcing System

To mitigate supply risk, companies normally order from multiple suppliers to replenish inventories. Dual sourcing is a particularly important problem with two suppliers, typically one with a short lead time but high cost while the other with a long lead time but low cost. For this class of problems, it has been shown that tailored base-surge (TBS) policy performs very well (Janakiraman et al. (2015); Xin and Goldberg (2018)). Chen and Shi (2019) study this problem when demand is not known *a priori*.

Demand and supply process, and system dynamics. Consider a firm selling one product over T periods. We denote the demand in period t by D_t , and assume that D_t across time periods are iid continuous random variables with expectation μ . We will use D to denote the time-generic demand. There are two suppliers, an expedited supplier and a regular supplier. Denote the order lead times for the expedited supplier and the regular supplier by l^E and l^R , respectively, with $l^E < l^R$. Denote the per-unit ordering costs for the expedited supplier and the regular supplier by c^E and c^R , respectively. Let $c = c^E - c^R > 0$ and $l = l^R - l^E > 0$. Without loss of generality, assume that $l^E = 0$ and $c^R = 0$. We use h to denote the per-unit holding cost and b to denote the per-unit shortage penalty cost.

At the beginning of period t , the firm observes the on-hand inventory x_t and all the inventories in the pipeline ordered from the regular supplier. The inventory order placed l periods ago from the regular supplier for q_{t-l}^R units is delivered, so the on-hand inventory level reaches $x_t + q_{t-l}^R$. The firm determines an ordering quantity q_t^E from the expedited supplier that will arrive immediately and pays cq_t^E for it. The on-hand inventory level reaches $x_t + q_{t-l}^R + q_t^E$. The firm then determines an ordering quantity q_t^R from the regular supplier that will arrive at the beginning of period $t+l$. Demand D_t realizes to be d_t and is satisfied as much as possible by the on-hand inventory. Unsatisfied demand is lost and unobservable. As such, the firm only observes the sales quantity $\min(x_t + q_{t-l}^R + q_t^E, d_t)$, rather than the true realized demand d_t . Overage and underage costs are then assessed. The total cost for period t is given by

$$C_t(x_t, q_t^E, q_t^R, d_t) = cq_t^E + h(x_t + q_{t-l}^R + q_t^E - d_t)^+ + b(d_t - x_t - q_{t-l}^R - q_t^E)^+.$$

Leftover inventories are carried over to the beginning of the next period, i.e., $x_{t+1} = (x_t + q_{t-l}^R + q_t^E - d_t)^+$. The firm would like to make the optimal decisions about q_t^E and q_t^R , so that the T -period total cost is minimized.

TBS policy. For the inventory control problem in the dual-sourcing system, the structure of the optimal policy remains poorly understood. However, it is known that TBS policy performs very well when lead time l is large. As such, Chen and Shi (2019) focus on learning the class of TBS policies, rather than the exact optimal policies. A TBS policy is specified by two parameters, Q and S . In each period t , the firm will order $q_t^R \equiv Q$ from the regular supplier, and order $q_t^E = (S - (x_t + Q))^+$ from the expedited supplier, i.e., ordering up to the target level S . We define the so-called *overshoot* O_t to be the quantity by which the inventory level exceeds S , i.e.,

$$O_t = (x_t + Q - S)^+,$$

and the recursion for the stochastic process of the overshoot random variable $\{O_t\}$ follows

$$O_{t+1} = (O_t - D_t + Q)^+.$$

For a given Q , let $O_\infty(Q)$ denote the steady-state overshoot random variable, which is defined as the steady-state version of $\{O_t\}$. The existence of the stationary distribution is guaranteed by Loynes' lemma (see Loynes (1962)). Given (Q, S) , the long-run average cost of the TBS policy is

$$\begin{aligned}\mathbb{E}[V(Q, S)] &= c(\mathbb{E}[\min\{S + O_\infty(Q), D\}] - Q) + h\mathbb{E}[(S + O_\infty(Q) - D)^+] \\ &\quad + b\mathbb{E}[(D - S - O_\infty(Q))^+] \\ &= c(\mu - Q) + h\mathbb{E}[(S + O_\infty(Q) - D)^+] + (b - c)\mathbb{E}[(D - S - O_\infty(Q))^+].\end{aligned}$$

Given Q , it is clear that $\mathbb{E}[V(Q, S)]$ is convex in S . Let

$$S^*(Q) = \max_{S \in [Q, S^h]} \mathbb{E}[V(Q, S)].$$

Note that $\mathbb{E}[V(Q, S)]$ is not necessarily jointly convex in (Q, S) .

Transforming the objective. Due to demand censoring, the lost-sales quantity $(D - S - O_\infty(Q))^+$ in $\mathbb{E}[V(Q, S)]$ cannot be observed. Therefore, we carry out a simple transformation of the cost function as follows. Let the pseudo cost

$$G(Q, S) = -cQ + h(S + O_\infty(Q) - D)^+ - (b - c) \min\{S + O_\infty(Q), D\},$$

and the long-run average cost of the (Q, S) policy equals

$$\mathbb{E}[V(Q, S)] = \mathbb{E}[G(Q, S)] + b\mu,$$

where the first term on the RHS becomes observable, and the second term is not observable but independent of any control policies. It then makes sense to focus on minimizing $\mathbb{E}[G(Q, S)]$ instead of $\mathbb{E}[V(Q, S)]$. It can be seen that $\mathbb{E}[G(Q, S)]$ is convex in S with minimizer $S^*(Q)$. Moreover, $\mathbb{E}[G(Q, S^*(Q))]$ is convex in Q . Let the optimal TBS policy be parameterized by (Q^*, S^*) .

When the demand distribution is not known *a priori*, the firm cannot directly compute the optimal TBS policy parameterized by (Q^*, S^*) . We aim to develop a learning algorithm π , for which the regret is minimized.

Learning algorithm. Chen and Shi (2019) propose an algorithm that combines the power of bisection with SGD. The algorithm updates and prescribes two parameters (Q, S) for every period, i.e., the ordering quantity Q from the regular source, and the order-up-to level S from the expedited source. The outline is presented below.

- The learning algorithm proceeds in epochs $n = 1, 2, \dots$, where each epoch contains a random number of rounds and each round contains an exponentially increasing number of

intervals. The algorithm can be essentially viewed as having an inner layer and an outer layer.

- The inner layer, the operations inside each interval belonging to some epoch n , updates the order-up-to level S using the projected SGD based on the current estimate Q_n .
- The outer layer, the operations inside each epoch and each round, updates the high probability range of Q_{n+1} for the next epoch $n+1$ using the bisection search method, by adaptively tracking the performance of three candidate points Q_n^l, Q_n^c, Q_n^r .
- Due to incomplete demand information as well as not reaching the steady state of overshoot, one cannot observe the true costs $G(Q_n^j, S^*(Q_n^j))$, $j = l, c, r$, and thus we develop SGD-based cost approximations G_{nm}^j , $j = l, c, r$, to carry on with the bisection procedure.

Regret. Regret of the learning algorithm is upper bounded by

$$O(\sqrt{T}(\log T)^3(\log \log T)^2).$$

Note that Besbes and Muharremoglu (2013) and Zhang et al. (2020) have established lower bounds $\Omega(\sqrt{T})$ for the repeated newsvendor problem (with a single source and no inventory carryover). Therefore, the regret upper bound is tight up to the logarithmic factor.

15.4 LEARNING OPTIMAL JOINT INVENTORY AND PRICING DECISIONS

Inventory replenishment and pricing are important levers and they can be integrated to better match demand with supply. For example, a firm may want to offer a markdown when the inventory level is high and raise the price when the inventory level is low. Joint pricing and replenishment decisions are challenging, and failure to tackle this issue can directly affect the bottom line of a company. In the academic literature, the model of joint pricing and inventory control with known demand distribution is studied by Federgruen and Heching (1999), and it has been extended to different settings. For a comprehensive review, see survey papers by Petruzzi and Dada (1999), Elmaghraby and Keskinocak (2003), Yano and Gilbert (2005), and Chen and Simchi-Levi (2012).

The traditional literature assumes that the demand–price relation and demand distribution are known to the firm, which in general may not hold in practice. In this section, we discuss learning models for the joint inventory and pricing optimization problem where the demand distribution and the demand–price relationship are not known *a priori*. We consider a periodic-review inventory system in which the firm (e.g., a retailer) sells a non-perishable product over a planning horizon of T periods. At the beginning of each period t , the firm observes on-hand inventory x_t and determines an inventory order-up-to level y_t as well as a price p_t , where $y_t \geq x_t$, $y_t \in \mathcal{Y} = [y^l, y^h]$ and $p_t \in \mathcal{P} = [p^l, p^h]$. For simplicity, we assume that the system is initially empty, i.e., $x_1 = 0$. Demand for period t , denoted by $D_t(p_t)$, is stochastic and price dependent. Demand is satisfied as much as possible by on-hand inventory. There might be a mismatch between supply and demand. If $y_t > D_t(p_t)$, any leftover inventories will be carried over to the next period, and for each unit the firm pays a holding cost h . If $y_t < D_t(p_t)$,

excessive demands are not fulfilled, and the firm pays a penalty cost b for each unit of stock-out. Per-unit ordering cost is normalized to 0 without loss of generality. The firm's objective is to maximize the T -period total profit.

If the distribution of $D_t(p_t)$ is known *a priori* to the firm (complete information scenario), then the optimization problem the firm wishes to solve is

$$\max_{\substack{(p_t, y_t) \in \mathcal{P} \times \mathcal{Y} \\ y_t \geq x_t}} \sum_{t=1}^T Q(p_t, y_t),$$

where $G(p_t, y_t)$ is the expected one-period reward in period t . Let $V^*(T)$ represent the maximum T -period expected profit generated from the optimal policy should the firm have complete information.

When the demand distribution is unknown *a priori*, the firm needs to learn the demand information on the fly to maximize its profit. Specifically, the firm needs to prescribe pricing and ordering decisions for each period based on the information up to that period. An admissible policy is represented by a sequence of prices and order-up-to levels, $\{(p_t, y_t), t \geq 1\}$, where (p_t, y_t) depends only on realized data and decisions made prior to period t , and $y_t \geq x_t$, i.e., (p_t, y_t) is adapted to the filtration generated by $\{(p_s, y_s), o_s; s = 1, \dots, t-1\}$. Here o_s represents the observable data of demand. Ideally, $o_s = D_s(p_s)$, meaning that demand is fully observable, but in some cases demand data is censored, which yields $o_s < d_s$. The firm's objective is to find an admissible policy to (i) learn the unknown demand distribution, and (ii) generate as much profit as possible. The regret for algorithm π is given by

$$R^\pi(T) = V^*(T) - \mathbb{E}[\sum_{t=1}^T Q(p_t, y_t)].$$

In this chapter, we will discuss a number of models under the framework of joint inventory and pricing. These models differ in the following three dimensions.

- Backlog versus lost sale: In a backlog system, if $y_t < D_t(p_t)$, any unsatisfied demands will be backlogged and served in future periods, and $x_{t+1} = y_t - D_t(p_t)$. In a lost-sale system, unmet demands will leave the market without any purchases, and $x_{t+1} = (y_t - D_t(p_t))^+$.
- Unlimited price changes versus limited price changes: We will discuss one model where the firm is not allowed to make price changes more than a certain number of times. The rest of the models allow an unlimited number of price changes.
- With versus without setup cost: If a setup cost is present, a fixed amount of fee will be charged whenever a positive amount of inventory is ordered.

In Section 15.4.1 and Section 15.4.2, we discuss the classic joint inventory and pricing problem with backlogged demand and lost sales, respectively. In Section 15.4.3, we consider scenarios with a limited number of price changes. In Section 15.4.4, we discuss the joint pricing and inventory control problem with setup cost.

15.4.1 Nonparametric Learning with Backlogged Demand

Chen et al. (2019a) study the classical joint inventory and pricing problem with backlogged demand. The demand in period t is either $D_t(p_t) = \lambda(p_t) + \varepsilon_t$ (additive) or $D_t(p_t) = \lambda(p_t) \varepsilon_t$ (multiplicative), where $\lambda(\cdot)$ is a strictly decreasing deterministic function and $\varepsilon_t, t = 1, 2, \dots, T$, are iid random variables with unknown probability density function denoted by $f(\cdot)$ and cumulative distribution function denoted by $F(\cdot)$. Unsatisfied demands are backlogged thus one has $x_{t+1} = y_t - D_t(p_t)$ for all $t = 1, \dots, T$. The reward in period t is

$$Q(p_t, y_t) = p_t \mathbb{E}[D_t(p_t)] - h \mathbb{E}[y_t - D_t(p_t)]^+ - b \mathbb{E}[D_t(p_t) - y_t]^+.$$

By Sobel (1981), myopic policy is optimal for this problem. Therefore, to optimize the T -period problem, it suffices to solve the single-period problem

$$\max_{(p,y) \in \mathcal{P} \times \mathcal{Y}} Q(p, y), \quad (15.8)$$

where

$$Q(p, y) = p \mathbb{E}[D(p)] - h \mathbb{E}[y - D(p)]^+ - b \mathbb{E}[D(p) - y]^+.$$

In the case of complete information on demand distribution, the optimal policy is to institute the single-period solution in each and every period t of the T -period problem.

However, the firm knows neither the function $\lambda(\cdot)$ nor the distribution of random variable ε_t . In the backlog system, true demand realizations can be observed. Therefore, $o_t = D_t(p_t)$, and an admissible policy (p_t, y_t) is adapted to the filtration generated by $\{(p_s, y_s), D_s(p_s); s = 1, \dots, t-1\}$.

Learning algorithm. The following are the main steps of the learning algorithm developed by Chen et al. (2019a).

- The algorithm proceeds in exponentially increasing cycles, and the length of the i -th cycle, $i = 1, 2, \dots$, is $I_i = I_0 v^i$ for input $I_0 > 0$ and $v > 1$.
- The i -th cycle is split into half and half. For the first half, a fixed pair of price-inventory $(\hat{p}_i, \hat{y}_{i,1})$ is implemented in every period; while for the second half, a fixed pair $(p_i + \delta_i, y_{i,2})$ is implemented in every period, where $\delta_i = \rho(2I_i)^{-1/4}$ for input $\rho > 0$.
- Collect realized demand data during the i -th cycle, based on which (1) conduct a linear approximation to estimate $\lambda(\cdot)$; (2) conduct a sample average approximation to estimate the distribution of error ε .
- Construct an empirical objective function based on the linear approximation and sample average approximation obtained in the previous step, and solve for the optimizer $(\hat{p}_{i+1}, \hat{y}_{i+1,1})$. Then get the second price as $\hat{p}_{i+1} + \delta_{i+1}$, and the second inventory order-up-to level is the optimal y corresponding to $p = \hat{p}_{i+1} + \delta_{i+1}$.
- The algorithm proceeds from cycle i to $i+1$.

Regret. Regret of the learning algorithm outlined above is upper bounded by $O(T^{1/2})$. The lower bound for regret is $\Omega(T^{1/2})$, which is implied by Keskin and Zeevi (2014). This shows that the regret rate for the algorithm is tight.

Intuition of proof. The algorithm learns the demand curve and the distribution of demand error in exponentially increasing cycles. Note that during cycle i , two distinct prices are implemented, based on which demand data is generated. The two prices are different by δ_i , which decreases to 0 as i increases. Therefore, the two prices are getting closer, and the linear function yielded by linear approximation approaches the tangent line of $\lambda(\cdot)$, providing gradient information for future decisions. Refer to Chen et al. (2019a) for detailed proof.

15.4.2 Nonparametric Learning with Lost Sales and Censored Demand

This subsection considers the classical joint inventory control and pricing problem with lost sales and censored demand. The main challenge for this problem arises from the fact that neither the objective value (zeroth-order feedback) nor the derivative (first-order feedback) is observable to the firm. The objective value is inaccessible because the lost sales are not observable, and the derivative with respect to the price decision cannot be observed because of the unknown demand–price relationship. The discussion in this subsection is based on Chen et al. (2021) and Chen et al. (2020c).

We focus on the additive demand model $D_t(p_t) = \lambda(p_t) + \varepsilon_t$ with $\lambda(\cdot)$ being a non-increasing deterministic function and ε_t , $t = 1, 2, \dots, T$, being iid random variables with $\mathbb{E}[\varepsilon_t] = 0$. For notational convenience, we use ε_t and ε interchangeably because of the iid assumption. Demands are satisfied as much as possible by on-hand inventory, and unsatisfied demands are lost and unobservable. For system dynamics one has $x_{t+1} = (y_t - D_t(p_t))^+$. The instantaneous reward for period t is

$$\begin{aligned} Q(p_t, y_t) &= p_t \mathbb{E}[\min\{y_t, D_t(p_t)\}] - b \mathbb{E}[D_t(p_t) - y_t]^+ - h \mathbb{E}[y_t - D_t(p_t)]^+ \\ &= p_t \mathbb{E}[D_t(p_t)] - (b + p_t) \mathbb{E}[D_t(p_t) - y_t]^+ - h \mathbb{E}[y_t - D_t(p_t)]^+. \end{aligned}$$

The firm knows neither the function $\lambda(p_t)$ nor the distribution of the random term ε_t *a priori*, which must be learned from censored demands collected over time while maximizing the cumulative profit. In this system, demand is censored, therefore, $o_t = \min\{D_t(p_t), y_t\}$. For an admissible policy, (p_t, y_t) is adapted to the filtration generated by $\{(p_s, y_s), \min\{D_s(p_s), y_s\} : s = 1, \dots, t-1\}$ under censored demand.

If the underlying demand–price function $\lambda(p)$ and the distribution of the error term ε_t were known *a priori*, the clairvoyant optimal policy for this problem is a myopic policy (refer to Sobel (1981)). Define the single-period problem by

$$Q(p, y) = p \mathbb{E}[D_1(p)] - (b + p) \mathbb{E}[D_1(p) - y]^+ - h \mathbb{E}[y - D_1(p)]^+.$$

To find the optimal pricing and inventory decisions, it suffices to maximize the single-period revenue $Q(p, y)$, which can be expressed as

$$\max_p \left\{ p \lambda(p) - \min_y \left\{ (b + p) \mathbb{E}[\lambda(p) + \varepsilon - y]^+ + h \mathbb{E}[y - \lambda(p) - \varepsilon]^+ \right\} \right\}.$$

Hence, we rewrite the clairvoyant problem as

$$\max_{p, y} Q(p, y) = \max_p G(p),$$

where

$$G(p) = p\lambda(p) - \min_y \left\{ (b+p)\mathbb{E}[\lambda(p) + \varepsilon - y]^+ + h\mathbb{E}[y - \lambda(p) - \varepsilon]^+ \right\}.$$

Under the assumption that $G(\cdot)$ is concave, both Chen et al. (2021) and Chen et al. (2020c) study this problem and obtain near-optimal regret rates (Section 15.4.2.1). Moreover, Chen et al. (2020c) also consider the case when $G(\cdot)$ is nonconcave, and prove that the regret of their learning algorithm matches the theoretical lower bound (Section 15.4.2.2).

15.4.2.1 Algorithms and results for concave $G(\cdot)$

In this section, we will discuss two algorithms that are proposed in the literature for the case with concave $G(\cdot)$. The first algorithm is a spline approximation-based algorithm, and the second one is a bisection-based algorithm. These two algorithms are totally different in both design and analyses.

Spline approximation-based algorithm. Chen et al. (2021) assume $G(\cdot)$ is concave, and provide a learning algorithm based on spline approximation, which we briefly describe as follows.

- The learning algorithm follows an exploration–exploitation framework, with the length of exploration being roughly on the order of $\sqrt{T}$, followed by an exploitation phase.
- During the exploration phase, the algorithm keeps the inventory level high, so that it can observe more demand realizations. For pricing decisions, the algorithm explores uniformly in the pricing space $[p^l, p^h]$.
- After collecting sales data from the exploration phase, the algorithm constructs a spline approximation, $\hat{\lambda}(\cdot)$, of the demand function $\lambda(\cdot)$, and builds a sample average approximation to estimate the distribution of error ε .
- The algorithm formulates an empirical objective function by replacing $\lambda(p)$ with the spline approximation and replacing the true error distribution with the sample average approximation, and solves for the optimal $\hat{p}, \hat{y}$.
- $(\hat{p}, \hat{y})$ is implemented for every period during the exploitation phase.

Regret. The regret rate of the spline approximation-based algorithm is upper bounded as $O(T^{1/2+\varepsilon}(\log T)^3 \log \log T)$, where $\varepsilon = 1/\sqrt[3]{\log T} + 0.25/\sqrt{\log T}$. Here note that for any constant $c > 0$, one has $\log \log T / \log T < \varepsilon < c$ (or equivalently, $\log T < T^\varepsilon < T^c$), for large enough T . Since the regret lower bound for this problem is $\Omega(\sqrt{T})$, theoretically the spline approximation-based algorithm almost matches the lower bound up to T^ε .

Bisection-based algorithm. A different algorithm is proposed by Chen et al. (2020c) for concave $G(\cdot)$, which approaches the optimal y using bisection and optimal p using trisection. The key idea of the algorithm is to construct a “difference estimator” that estimates the difference between the rewards at two price points instead of trying to estimate the reward value at a single price point. Below is the outline of their learning algorithm.

- The algorithm proceeds in trisection iterations, and the length of iteration increases exponentially in its index.
- During each iteration i , the algorithm explores two price points, $\alpha_i < \beta_i$, which are the trisection points of the current price interval.

- Using bisection to obtain $\hat{y}_i$ and $\hat{y}'_i$, the estimated optimal inventory level for α_i and β_i , respectively.
- Based on the knowledge of $\hat{y}_i$ and $\hat{y}'_i$, estimate the reward differences at α_i and β_i (the difference estimator).
- Eliminate suboptimal prices using trisection. If the reward at α_i is smaller than that at β_i , eliminate all prices smaller than α_i ; otherwise, eliminate all prices larger than β_i . The algorithm proceeds to iteration $i+1$.

Regret. The regret rate of the bisection-based algorithm for concave $G(\cdot)$ is upper bounded by $O(\sqrt{T}(\ln T)^2)$. This regret upper bound improves the one for the spline approximation-based algorithm and almost matches the theoretical lower bound of $\Omega(\sqrt{T})$.

15.4.2.2 Algorithms and results for nonconcave $G(\cdot)$

In this section, we will discuss a learning algorithm proposed for the setting with nonconcave $G(\cdot)$ as well as its regret convergence results.

Learning algorithm. For nonconcave $G(\cdot)$, Chen et al. (2020c) still apply bisection to search for the optimal y , but for p the previous trisection framework cannot be applied anymore due to loss of concavity. They design an active tournament algorithm based on the difference estimator to search for the optimal p .

- The algorithm proceeds in iterations, the length of which is roughly exponentially increasing. The interval of eligible prices is discretized to be a grid.
- For every iteration i , the algorithm keeps an active set, which is a subset of the grid. For each price within the active set, the algorithm estimates its corresponding order-up-to level using bisection. Then by comparing the reward of each price in pairs using the difference estimator, the algorithm obtains the best-performing price of the current active set.
- The algorithm eliminates from the current active set all the prices whose reward is worse than the best-performing price by more than Δ_i , where Δ_i decreases exponentially in i . A new active set for the next iteration is thus formed.

Regret. The regret rate for nonconcave $G(\cdot)$ is upper bounded by $O(T^{3/5}(\ln T)^2)$. Chen et al. (2020c) then prove the lower bound for nonconcave $G(\cdot)$ and shows that the upper bound matches the lower bound. They prove that there exists a problem instance such that for any learning-while-doing policy π and the sequential decisions $\{p_t, y_t\}_{t=1}^T$ the policy π produces, it holds for sufficiently large T that $\sup_{\lambda} \mathbb{E} \left[V^*(T) - \sum_{t=1}^T Q(p_t, y_t) \right] \geq C \times T^{3/5} / \ln T$ for some constant $C > 0$. The lower bound is established by a novel information-theoretical argument based on generalized squared Hellinger distance, which is significantly different from conventional arguments that are based on Kullback–Leibler divergence.

15.4.3 Parametric Learning with Limited Price Changes

In practice, a firm may be constrained from making frequent price changes. Cheung et al. (2017) discuss several practical reasons for not allowing frequent price changes, including customers' negative responses (e.g., that may cause confusion and affect the seller's brand reputation) and the cost associated with such changes (e.g., due to changing price labels in brick-and-mortar stores, etc.). Clearly, such a constraint limits the firm's ability to learn demand.

Chen and Chao (2019) and Chen et al. (2020a) study the joint pricing and inventory control problem with limited price changes, under the backlog and lost-sale system, respectively.

Demand in period t , $t \in \{1, 2, \dots, T\}$, is random and depends on the selling price p_t and its distribution function belongs to some family parameterized by $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^k$, $k \geq 1$, where $\mathcal{Z}$ is a compact and convex set. Let $D_t(p_t, \mathbf{z})$ be the demand in period t with probability mass function $f(\cdot; p_t, \mathbf{z})$ and support $\{d^l, d^l + 1, \dots, d^h\}$. The firm knows $f(\cdot; p_t, \mathbf{z})$ up to the parameter vector $\mathbf{z}$, which has to be learned from sales data.

This subsection will be mainly devoted to discussing algorithms and results in Chen et al. (2020a), where unsatisfied customers are lost and the firm can only observe sales data but not the actual demand when stockout occurs. Therefore $o_t = \min\{D_t(p_t, \mathbf{z}), y_t\}$, and (p_t, y_t) is adapted to the filtration generated by $\{(p_s, y_s), o_s : s = 1, \dots, t-1\}$ under censored demand. Let $p_t \in \mathcal{P} = [p^l, p^h]$ and $y_t \in \mathcal{Y} = \{y^l, y^l + 1, \dots, y^h\}$, where the bounds of support $0 \leq p^l \leq p^h < +\infty$ and $0 \leq y^l \leq y^h < +\infty$ are known. The state transition is $x_{t+1} = (y_t - D_t(p_t, \mathbf{z}))^+$.

The expected total profit over the planning horizon, given an admissible policy $\phi = ((p_1, y_1), (p_2, y_2), \dots, (p_T, y_T))$, is

$$V^\phi(T, \mathbf{z}) = \sum_{t=1}^T \{p_t \mathbb{E}[\min\{D_t(p_t, \mathbf{z}), y_t\}] - \{h \mathbb{E}[y_t - D_t(p_t, \mathbf{z})]^+ + b \mathbb{E}[D_t(p_t, \mathbf{z}) - y_t]^+\} \} \quad (15.9)$$

and, given an integer $m > 0$, the prices need to satisfy the *limited price change constraint*:

$$\sum_{t=1}^{T-1} \mathbf{1}(p_t \neq p_{t+1}) \leq m, \quad (15.10)$$

where $\mathbf{1}(A)$ is the indicator function taking value 1 if statement A is true and 0 otherwise.

The single-period objective function is

$$Q(p, y, \mathbf{z}) = p \mathbb{E}[D(p, \mathbf{z})] - h \mathbb{E}[y - D(p, \mathbf{z})]^+ - (b + p) \mathbb{E}[D(p, \mathbf{z}) - y]^+, \quad (15.11)$$

where $D(p, \mathbf{z})$ is a generic random demand when the true parameter is $\mathbf{z}$ and the price is $p \in \mathcal{P}$. For the underlying system parameter vector $\mathbf{z}$, let (p^*, y^*) be a maximizer of $Q(p, y, \mathbf{z})$. If $\mathbf{z}$ is known, then the firm could set (p^*, y^*) every period without changing the price, and this is the clairvoyant solution for the T -period problem.

Demand models are categorized into two groups, (1) the well-separated case, and (2) the general case. Two probability mass functions are said to be identifiable if they are not identically the same.

15.4.3.1 Algorithms and results for well-separated demand

The family of distributions $\{f(\cdot; p, \mathbf{z}) : \mathbf{z} \in \mathcal{Z} \subset \mathbb{R}\}$ is called well-separated if for any $p \in \mathcal{P}$ the class of probability mass functions $\{f(\cdot; p, \mathbf{z}) : \mathbf{z} \in \mathcal{Z}\}$ is identifiable, i.e., $f(\cdot; p, \mathbf{z}_1) \neq f(\cdot; p, \mathbf{z}_2)$ for $\mathbf{z}_1 \neq \mathbf{z}_2 \in \mathcal{Z}$.

If a family of distributions is well-separated, then no matter what selling price p is charged, the sales data will allow the firm to learn about the parameter $\mathbf{z}$. This shows that, in the

well-separated case, pricing exploration can be a side benefit from exploitation, thus no active pricing exploration is necessary.

Two scenarios of limited price constraint are considered for well-separated demand. The first scenario is that the number of price changes is restricted to be no more than a given integer $m \geq 1$ that is independent of the length of planning horizon T , while for the second scenario, the number of allowed price changes is at most $\beta \log T$ for the T -period problem for some constant $\beta > 0$.

Algorithm under the constraint of m price changes. The main idea of the algorithm is to estimate the known parameter z by MLE based on censored demand. The outline of the algorithm is as follows.

- The algorithm divides the planning horizon T into $m+1$ stages, with $I_i = \lceil T^{i/(m+1)} \rceil$ being the length of the i -th stage, $i = 1, \dots, m$, and $I_{m+1} = T - \sum_{i=1}^m I_i$.
- The same price is applied within each stage. At the end of each stage i , an estimator $\hat{z}_i$ of z is constructed based on MLE using censored demand data from the current stage.
- Then using $\hat{z}_i$ to replace the unknown z , a data-driven optimization problem is solved to obtain p_{i+1} , which is the price implemented during stage $i+1$.

Regret under the constraint of m price changes. Regret of the algorithm above is upper bounded by $O(T^{1/(m+1)})$, and regret lower bound is $\Omega(T^{1/(m+1)})$ for well-separated demand with no more than m price changes.

One fundamental challenge to proving this lower bound is that the times of price changes are dynamically determined, i.e., they are increasing random stopping times. Chen et al. (2020a) construct an adversarial parameter class, among which a policy needs to identify the true parameter. The parameter class is constructed in a hierarchical manner such that when going further down the hierarchy the parameters are harder to distinguish. A delicate information-theoretical argument is employed to prove the lower bound.

In practice, it may happen that T is not clearly specified at the beginning. The firm requires that the price change cannot be too often, but it usually allows more price changes for a longer planning horizon. Chen et al. (2020a) propose a learning algorithm where the number of price changes is restricted to $\beta \log T$ for some constant $\beta > 0$.

Algorithm under the constraint of $\beta \log T$ price changes. The algorithm runs very similarly to the one for m price changes, except that now the number of periods in i is given by $I_i = \lceil I_0 v^i \rceil$, $i = 1, 2, \dots, N$, and there is a total of $O(\log T)$ iterations.

Regret under the constraint of $\beta \log T$ price changes. The regret upper bound for the learning algorithm with no more than $\beta \log T$ price changes is $O(\log T)$. The regret lower bound is $\Omega(\log T)$ for $T \geq 1$.

15.4.3.2 Algorithms and results for general demand

Now we consider the more general case that the parameters in probability mass function $f(\cdot; p, \mathbf{z})$ is a k -dimensional vector, i.e., $\mathbf{z} = (z_1, \dots, z_k) \in \mathcal{Z} \subset \mathbb{R}^k$ for some integer $k \geq 1$. For a set of given prices $\mathbf{p} = (p_1, \dots, p_k) \in \mathcal{P}^k$, and correspondingly realized demands $\mathbf{d} = (d_1, \dots, d_k) \in \{d^l, d^l + 1, \dots, d^h\}^k$, define

$$\mathcal{Q}^{p, \mathbf{z}}(\mathbf{d}) = \prod_{j=1}^k f(d_j; p_j, \mathbf{z}).$$

The family of distributions $\{Q^{p,z}(\cdot) : \mathbf{z} \in \mathcal{Z}\}$ is said to belong to the general case if there exist k price points $\bar{\mathbf{p}} = (\bar{p}_1, \dots, \bar{p}_k) \in \mathcal{P}^k$ such that the family of distributions $\{Q^{\bar{\mathbf{p}}, \mathbf{z}}(\cdot) : \mathbf{z} \in \mathcal{Z}\}$ is identifiable, i.e., $Q^{\bar{\mathbf{p}}, \mathbf{z}_1}(\cdot) \neq Q^{\bar{\mathbf{p}}, \mathbf{z}_2}(\cdot)$ for any $\mathbf{z}_1 \neq \mathbf{z}_2$ in $\mathcal{Z}$.

Suppose we are allowed to make up to m price changes during the planning horizon. We consider the case of $m \geq k$ in this section, as in the case of $m < k$ no algorithm will be able to identify the k unknown parameters and therefore the regret would be linear in T .

Algorithm for general demand. The algorithm follows an exploration–exploitation framework, and the unknown parameter vector $\mathbf{z}$ is estimated by MLE.

- During the exploration phase, the algorithm tests each of the prices $p_1, \dots, p_k$ for $\lceil T^{1/2} / k \rceil$ periods and collects the censored demand data.
- At the end of the exploration phase, the algorithm constructs an MLE estimator $\hat{\mathbf{z}}$.
- The algorithm then solves an empirical objective function with the unknown parameter $\mathbf{z}$ replaced by $\hat{\mathbf{z}}$, obtaining the optimal decision $(\hat{p}, \hat{y})$, which is applied to every period of the exploitation phase.

Regret for general demand. The regret upper bound for the general demand case is provided as follows: if the demand is unbounded $d^h = +\infty$, then the regret for general demands is upper bounded by $O(T^{1/2} \log T)$; if the demand is bounded $d^h < +\infty$, then the regret for general demands is upper bounded by $O(T^{1/2})$. The theoretical lower bound for this problem is $\Omega(T^{1/2})$, which is established by Broder and Rusmevichientong (2012) for a dynamic pricing problem with infinite initial inventory.

15.4.4 Backlog System with Fixed Ordering Cost

In this section, we consider joint inventory and pricing optimization in a periodic-review inventory system with fixed ordering cost. With complete information on the demand distribution, this problem has been studied by Chen and Simchi-Levi (2004a) and Chen and Simchi-Levi (2004b). When demand distribution is not known *a priori*, this problem is studied in Chen et al. (2020b), where demand is modeled as $D = D_0(p) + \beta$, and $D_0 : [0, 1] \rightarrow [\underline{d}_0, \bar{d}_0]$ is the (expected) demand function of price p and β is the random noise with 0 mean. The authors consider both linear models and generalized linear models for $D_0(p)$ with *unknown* parameters θ_0 . The distribution for β is unknown in the nonparametric sense. During every period, unsatisfied demand is backlogged. Let $K > 0$ be the fixed ordering cost, $c > 0$ be the variable ordering cost of ordering one unit of inventory, and $h : \mathbb{R} \rightarrow \mathbb{R}^+$ be the holding cost (when the remaining inventory level is positive) or the backlogging cost (when the remaining inventory level is negative). The instantaneous reward for period t is

$$\tau_t = -K \times \mathbf{1}\{y_t > x_t\} - c(y_t - x_t) + p_t(D_0(p_t) + \beta_t) - h(y_t - D_0(p_t) - \beta_t),$$

and the firm aims to maximize the T -period total reward.

With known demand curve D_0 and noise distribution μ_0 , the work of Chen and Simchi-Levi (2004a) proves that, under mild conditions, for both the average and discounted profit criterion there exists an $(s, S, \mathbf{p})$ policy that is optimal in the long run. Under an $(s, S, \mathbf{p})$ -policy, the retailer will only order new inventories when $x_t < s$, and after the ordering of new inventories

maintain $y_t = S$. The function $\mathbf{p}$ prescribes the pricing decision that depends on the initial inventory level of the same period.

The performance of a particular $(s, S, \mathbf{p})$ policy can be evaluated as follows. Define $H_0(x, p; \mu)$ as the *expected* immediate reward of pricing decision p at inventory level x and noise distribution μ , without ordering new inventories. It is easy to verify that

$$H_0(x, p; \mu) = -\mathbb{E}_\mu[h(x - D_0(p) - \beta)] + pD_0(p) - cD_0(p).$$

For a certain $(s, S, \mathbf{p})$ policy, define quantities $I(s, x, \mathbf{p}; \mu)$ and $M(s, x, \mathbf{p}; \mu)$ as follows:

$$I(s, x, \mathbf{p}; \mu) = \begin{cases} H_0(x, \mathbf{p}(x); \mu) + \mathbb{E}_\mu[I(s, x - D_0(\mathbf{p}(x)) - \beta, \mathbf{p}; \mu)], & x \geq s, \\ 0, & x < s; \end{cases}$$

$$M(s, x, \mathbf{p}; \mu) = \begin{cases} 1 + \mathbb{E}_\mu[M(s, x - D_0(\mathbf{p}(x)) - \beta, \mathbf{p}; \mu)], & x \geq s, \\ 0, & x < s; \end{cases}$$

Define $r(s, S, \mathbf{p}; \mu)$ as

$$r(s, S, \mathbf{p}; \mu) = \frac{-K + I(s, S, \mathbf{p}; \mu)}{M(s, S, \mathbf{p}; \mu)}.$$

When $I(s, S, \mathbf{p}; \mu_0)$ and $M(s, S, \mathbf{p}; \mu_0)$ are bounded, Lemma 2 from Chen and Simchi-Levi (2004a) show that the average reward of the $(s, S, \mathbf{p})$ policy over T periods approaches $r(s, S, \mathbf{p}; \mu_0)$ when T grows.

Learning algorithm. The learning algorithm proposed by Chen et al. (2020b) is based on an $(s, S, \mathbf{p})$ -policy with evolving inventory levels (s, S) and pricing strategies $\mathbf{p}$. A regularized least-squares estimation is used to estimate θ_0 , and a sample average approximation approach is used to construct an empirical distribution for β . Below we present the outline of their algorithm.

- The T time periods are partitioned into epochs, conveniently labeled as $\mathcal{B}_1, \mathcal{B}_2, \dots$, such that the re-stocking only occurs at the first time period of each epoch $\mathcal{B}_b$, $b \in \{1, 2, \dots\}$.
- Each epoch $\mathcal{B}_b$ is also associated with inventory levels (s_b, S_b) and pricing strategy $\mathbf{p}_b$, such that for the first time period $t_b \in \mathcal{B}_b$, the re-stocked inventory level is $y_{t_b} = S_b$.
- Epoch $\mathcal{B}_b$ terminates whenever $x_t < s_b$, and for all $t \in \mathcal{B}_b \setminus \{t_b\}$, $y_t = x_t$ and $p_t = \mathbf{p}_b(x_t)$.
- At the beginning of epoch b , the algorithm constructs the regularized least square estimator $\hat{\theta}_b$ and the empirical noise distribution $\hat{\mu}_b$, based on which it constructs the upper confidence bounds for $D_0(\cdot)$ and $H_0(\cdot; \mu)$, denoted as $\bar{D}_b(\cdot)$ and $\bar{H}_b(\cdot; \hat{\mu}_b)$, respectively.
- Policy for epoch b , $(s_b, S_b, \mathbf{p}_b)$, can be computed as follows. For any $s \in [s, S]$, $S \in [S, \bar{S}]$, $r \in \mathbb{R}$, demand function $\bar{D}_b : [0, 1] \rightarrow [\underline{d}, \infty)$, noise distribution $\hat{\mu}_b$ and the associated $\bar{H}_b : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}$, define

$$\phi^{(s, S)}(x; \bar{D}_b, r, \mu_b) = \begin{cases} \sup_{p \in [0, 1]} H(x, p; \mu_b) - r + \mathbb{E}_{\mu_b}[\phi^{(s, S)}(x - \bar{D}_b(p) - \beta; \bar{D}_b, r, \mu_b)], & x \geq s; \\ 0, & x < s. \end{cases}$$

For every (s, S) , define

$$\bar{r}_b(s, S) = \inf\{r \in \mathbb{R} : \phi^{(s, S)}(S; \bar{D}_b, r, \mu_b) = K\}$$

and let the pricing strategy $\mathbf{p}$ (associated with inventory levels s, S) be the optimal solution to the $\phi^{(s, S)}(\cdot; \bar{D}_b, \bar{r}_b(s, S), \mu_b)$ dynamic programming; that is, $\mathbf{p}(x)$ is defined such that

$$\begin{aligned} \phi^{(s, S)}(x; \bar{D}_b, \bar{r}_b(s, S), \mu_b) &= \bar{H}_b(x, \mathbf{p}(x); \mu_b) - \bar{r}_b(s, S) \\ &+ \mathbb{E}_{\mu_b}[\phi^{(s, S)}(x - \bar{D}_b(\mathbf{p}(x)) - \beta; \bar{D}_b, \bar{r}_b(s, S), \mu_b)] \end{aligned}$$

for all x . Select $(s_b, S_b) = \arg \max_{s, S} \bar{r}_b(s, S)$ and let $\mathbf{p}_b$ be the optimal pricing decisions associated with dynamic programming $\phi^{(s_b, S_b)}(\cdot; \bar{D}_b, \bar{r}_b(s_b, S_b), \mu_b)$.

Regret. The regret of the algorithm described above is upper bounded as $\tilde{O}(\sqrt{T})$. In the $\tilde{O}(\cdot)$ notation, we omit polynomial dependency on $\log T$ and other problem parameters. With $K = c = 0$ and $h(\cdot) \equiv 0$, the problem becomes a pure pricing problem with unknown linear demand functions. As long as $\tau > 1$, the work of Broder and Rusmevichientong (2012) proves an $\Omega(\sqrt{T})$ lower bound for any admissible pricing policies. Therefore, the $\tilde{O}(\sqrt{T})$ regret established here is optimal.

In the previous algorithm, a dynamic programming needs to be carried out after each epoch b to obtain a new policy $(s_b, S_b, \mathbf{p}_b)$. Because each epoch lasts at most $\bar{S} / \underline{d} = O(1)$ selling periods, the algorithm requires $\Omega(T)$ DP calculations which can be computationally expensive. Chen et al. (2020b) then propose an improved algorithm that only needs $O(\tau \log T)$ DP calculations to achieve virtually the same regret, which is much more computationally efficient.

Learning algorithm with infrequent DP updates. The outline of this algorithm is presented below.

- Define the determinant of the sample covariance

$$\Lambda_b := I_{\tau \times \tau} + \sum_{t \in \mathcal{H}_{b-1}} \eta(p_t) \eta(p_t)^\top, \text{ where } \eta(\cdot) \text{ (is a known feature map of the demand model.)}$$

- A new $(s, S, \mathbf{p})$ policy is computed only if 2^i , $i \in \{1, 2, \dots\}$ epochs are met, or Λ_b doubles.
- This greatly reduces the number of DP calculations from $O(T)$ to $O(\tau \log T)$.

Regret for infrequent DP updates. For the algorithm with infrequent DP updates, the regret is upper bounded as $\tilde{O}(\sqrt{T})$.

15.5 CONCLUDING REMARKS

This chapter reviews some of the latest developments in the field. With the increasing availability of data and ease of collection, data-driven optimization will play a central role in making operations decisions. Since most systems would have some data available, even though data

has decreasing value in predicting the future as it becomes older, we believe the integration of recent offline data and online learning to be particularly useful for inventory decisions.

In the review of this chapter, we focus on the case that the holding cost and shortage cost (either backlog or lost-sales) are linear, or the so-called V-shaped inventory cost. Yang and Shi (2023) consider the case with more general convex inventory cost with discrete demand, and they show that the lower bound for the regret of that more general case is $\Theta(T^{2/3})$. They also present learning algorithm that achieve $O(T^{2/3})$ regret rate. Such results hold true for the case with or without ordering setup cost.

Online learning inventory control and pricing optimization is a fast-developing area. Therefore, in this chapter, it was not possible to be either comprehensive or complete. In particular, it was noted that data-driven optimization has been taking shape in many companies, especially online retailers such as Amazon, JD.com, and Walmart. We expect there will be explosive growth in both online and offline (and integration of the two) optimization in inventory control and pricing optimization of supply chain management in the near future.

ACKNOWLEDGMENT

The authors thank Prof. Wang Chi Cheung and Mr. Zhongzhu Chen for their valuable comments on an earlier version of this chapter.

REFERENCES

- Agarwal, A., Foster, D. P., Hsu, D. J., Kakade, S. M., & Rakhlin, A. (2011). Stochastic convex optimization with bandit feedback. *Advances in Neural Information Processing Systems*, 24, 1035–1043.
- Agrawal, S., Avadhanula, V., Goyal, V., & Zeevi, A. (2017). Thompson sampling for the MNL-bandit. In Satyen Kale and Ohad Shamir (Eds.), *Conference on Learning Theory* (pp. 76–78). PMLR.
- Agrawal, S., & Jia, R. (2019). Learning in structured MDPs with convex cost functions: Improved regret bounds for inventory management. In *Proceedings of the 2019 ACM conference on economics and computation* (pp. 743–744).
- Agrawal, S., & Jia, R. (2022). Learning in structured MDPs with convex cost functions: Improved regret bounds for inventory management. *Operations Research*, 70(3), 1646–1664.
- Auer, P., Ortner, R., & Szepesvári, C. (2007). Improved rates for the stochastic continuum-armed bandit problem. In Nader H. Bshouty and Claudio Gentile (Eds.), *Proceedings of the 20th international conference on learning theory (COLT)* (pp. 454–468).
- Azoury, K. (1985). Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science*, 31(9), 1150–1160.
- Besbes, O., & Muharremoglu, A. (2013). On implications of demand censoring in the newsvendor problem. *Management Science*, 59(6), 1407–1424.
- Besbes, O., & Zeevi, A. (2015). On the surprising sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, 61(4), 723–739.
- Borovkov, A. (1998). *Mathematics statistics*. Gordon and Breach Science Publishers.
- Broder, J., & Rusmevichientong, P. (2012). Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4), 965–980.
- Bu, J., Gong, X., & Chao, X. (2023). Asymptotic optimality of base-stock policies for perishable inventory systems. *Management Science*, 69(2), 846–864.
- Bu, J., Simchi-Levi, D., & Xu, Y. (2022). Online pricing with offline data: Phase transition and inverse square law. *Management Science*, 68(12), 8568–8588.

- Burnetas, A. N., & Smith, C. E. (2000). Adaptive ordering and pricing for perishable products. *Operations Research*, 48(3), 436–443.
- Chang, H. S., Fu, M. C., Hu, J., & Marcus, S. I. (2005). An adaptive sampling algorithm for solving Markov decision processes. *Operations Research*, 53(1), 126–139.
- Chen, B. (2021). Data-driven inventory control with shifting demand. *Production and Operations Management*, 30(5), 1365–1385.
- Chen, B., & Chao, X. (2019). Parametric demand learning with limited price explorations in a backlog stochastic inventory system. *IISE Transactions*, 51(6), 605–613.
- Chen, B., & Chao, X. (2020). Dynamic inventory control with stockout substitution and demand learning. *Management Science*, 66(11), 5108–5127.
- Chen, B., Chao, X., & Ahn, H.-S. (2019a). Coordinating pricing and inventory replenishment with nonparametric demand learning. *Operations Research*, 67(4), 1035–1052.
- Chen, B., Chao, X., & Shi, C. (2021). Nonparametric learning algorithms for joint pricing and inventory control with lost-sales and censored demand. *Mathematics of Operations Research*, 46(2), 726–756.
- Chen, B., Simchi-Levi, D., Wang, Y., & Zhou, Y. (2022). Dynamic pricing and inventory control with fixed ordering cost and incomplete demand information. *Management Science*, 68(8), 5684–5703.
- Chen, B., Chao, X., & Wang, Y. (2020a). Data-based dynamic pricing and inventory control with censored demand and limited price changes. *Operations Research*, 68(5), 1445–1456.
- Chen, B., & Shi, C. (2019). Tailored base-surge policies in dual-sourcing inventory systems with demand learning. Available at: SSRN 3456834.
- Chen, B., Simchi-Levi, D., Wang, Y., & Zhou, Y. (2020b). Dynamic pricing and inventory control with fixed ordering cost and incomplete demand information. *Management Science* (in press).
- Chen, B., Wang, Y., & Zhou, Y. (2020b). Optimal policies for dynamic pricing and inventory control with nonparametric censored demands. Available at: SSRN 3750413.
- Chen, Q., Jasin, S., & Duenyas, I. (2019b). Nonparametric self-adjusting control for joint learning and optimization of multiproduct pricing with finite resource capacity. *Mathematics of Operations Research*, 44(2), 601–631.
- Chen, X., & Simchi-Levi, D. (2004a). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, 52(6), 887–896.
- Chen, X., & Simchi-Levi, D. (2004b). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, 29(3), 698–723.
- Chen, X., & Simchi-Levi, D. (2012). Pricing and inventory management. *The Oxford handbook of pricing management*, 1, 784–824.
- Cheung, W. C., Simchi-Levi, D., & Wang, H. (2017). Dynamic pricing and demand learning with limited price experimentation. *Operations Research*, 65(6), 1722–1731.
- Cope, E. W. (2009). Regret and convergence bounds for a class of continuum-armed bandit problems. *IEEE Transactions on Automatic Control*, 54(6), 1243–1253.
- Ding, X., Puterman, M., & Bisi, A. (2002). The censored newsvendor and the optimal acquisition of information. *Operations Research*, 50(3), 517–527.
- Elmaghraby, W., & Keskinocak, P. (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10), 1287–1309.
- Federgruen, A., & Heching, A. (1999). Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3), 454–475.
- Ferreira, K. J., Simchi-Levi, D., & Wang, H. (2018). Online network revenue management using Thompson sampling. *Operations Research*, 66(6), 1586–1602.
- Fries, B. (1975). Optimal ordering policy for a perishable commodity with fixed lifetime. *Operational Research*, 23(1), 46–61.
- Gijsbrechts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. (2020, October 6). Can deep reinforcement learning improve inventory management? Performance on dual sourcing, lost sales and multi-echelon problems. *Manufacturing & Service Operations Management*, 24(3), 1349–1368.
- Godfrey, G. A., & Powell, W. B. (2001). An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with applications to inventory and distribution. *Management Science*, 47(8), 1101–1112.

- Gong, X.-Y., & Simchi-Levi, D. (2020a). Provably efficient reinforcement learning for episodic stochastic inventory control models. Available at: SSRN.
- Gong, X.-Y., & Simchi-Levi, D. (2020b). Provably more efficient q-learning in the one-sided-feedback/full-feedback settings. <https://arxiv.org/abs/2007.00080>
- Hazan, E., Agarwal, A., & Kale, S. (2007). Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2–3), 169–192.
- Huh, W. H., & Rusmevichientong, P. (2009). A non-parametric asymptotic analysis of inventory planning with censored demand. *Mathematics of Operations Research*, 34(1), 103–123.
- Huh, W. H., Rusmevichientong, P., Levi, R., & Orlin, J. (2011). Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research*, 59(4), 929–941.
- Huh, W. T., Janakiraman, G., Muckstadt, J. A., & Rusmevichientong, P. (2009). An adaptive algorithm for finding the optimal base-stock policy in lost sales inventory systems with censored demand. *Mathematics of Operations Research*, 34(2), 397–416.
- Iglehart, D. (1964). The dynamic inventory problem with unknown demand distribution. *Management Science*, 10(3), 429–440.
- Janakiraman, G., & Roundy, R. O. (2004). Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Operations Research*, 52(5), 795–803.
- Janakiraman, G., Seshadri, S., & Sheopuri, A. (2015). Analysis of tailored base-surge policies in dual sourcing inventory systems. *Management Science*, 61(7), 1547–1561.
- Katehakis, M. N., Yang, J., & Zhou, T. (2020). Dynamic inventory and price controls involving unknown demand on discrete nonperishable items. *Operations Research*, 68(5), 1335–1355.
- Keskin, N. B., Li, Y., & Song, J.-S. J. (2022). Data-driven dynamic pricing and ordering with perishable inventory in a changing environment. *Management Science*, 68(3), 1938–1958.
- Keskin, N. B., Min, X., & Song, J.-S. J. (2021). The nonstationary newsvendor: Data-driven nonparametric learning. Available at: SSRN 3866171.
- Keskin, N. B., & Zeevi, A. (2014). Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62(5), 1142–1167.
- Kiefer, J., & Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics*, 23(3), 462–466.
- Kleinberg, R. (2005). Nearly tight bounds for the continuum-armed bandit problem. In L. K. Saul, Y. Weiss, and L. Bottou (Eds.), *Advances in Neural Information Processing Systems* (Vol. 17, pp. 697–704). MIT Press.
- Kunnumkal, S., & Topaloglu, H. (2008). Using stochastic approximation methods to compute optimal base-stock levels in inventory control problems. *Operations Research*, 56(3), 646–664.
- Lai, T., & Robbins, H. (1981). Consistency and asymptotic efficiency of slope estimates in stochastic approximation schemes. *Probability Theory and Related Fields*, 56(3), 329–360.
- Levi, R., Roundy, R. O., & Shmoys, D. B. (2007). Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 821–839.
- Liu, M., Qi, M., & Shen, Z.-J. M. (2021). End-to-end deep learning for inventory management with fixed ordering cost and its theoretical analysis. <http://doi.org/10.2139/ssrn.3888897>
- Lovejoy, W. (1990). Myopic policies for some inventory models with uncertain demand distributions. *Management Science*, 36(6), 724–738.
- Loynes, R. M. (1962). The stability of a queue with non-independent inter-arrival and service times. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 58, No. 3, pp. 497–520). Cambridge University Press.
- Lugosi, G., Markakis, M., & Neu, G. (2022). On the hardness of learning from censored demand. Available at SSRN 3509255.
- Miao, S., & Chao, X. (2021). Dynamic joint assortment and pricing optimization with demand learning. *Manufacturing and Service Operations Management*, 23(2), 525–545.
- Murray, G., & Silver, E. (1966). A Bayesian analysis of the style goods inventory problem. *Management Science*, 12(11), 785–797.
- Nahmias, S. (1975). Optimal ordering policies for perishable inventory-II. *Operational Research*, 23(4), 735–749.

- Oroojlooyjadid, A., Nazari, M., Snyder, L., & Takáč, M. (2017). A deep q-network for the beer game: A deep reinforcement learning algorithm to solve inventory optimization problems. arXiv preprint arXiv:1708.05924.
- Petruzzi, N. C., & Dada, M. (1999). Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2), 183–194.
- Robbins, H., & Monro, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics*, 22(3), 400–407.
- Ross, S. M. (1996). *Stochastic processes* (Vol. 2). Wiley.
- Russo, D., & Van Roy, B. (2014). Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4), 1221–1243.
- Russo, D., Van Roy, B., Kazerouni, A., Osband, I., & Wen, Z. (2017). A tutorial on Thompson sampling. arXiv preprint arXiv:1707.02038.
- Scarf, H. (1959). Bayes solution to the statistical inventory problem. *Annals of Mathematical Statistics*, 39(2), 490–508.
- Scarf, H. (1960). Some remarks on Bayes solutions to the inventory problem. *Naval Research Logistics Quarterly*, 7(4), 591–596.
- Shi, C., Chen, W., & Duenyas, I. (2016). Technical note—Nonparametric data-driven algorithms for multiproduct inventory systems with censored demand. *Operations Research*, 64(2), 362–370.
- Sobel, M. J. (1981). Myopic solutions of Markov decision processes and stochastic games. *Operations Research*, 29(5), 995–1009.
- Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4), 285–294.
- Thompson, W. R. (1935). On the theory of apportionment. *American Journal of Mathematics*, 57(2), 450–456.
- Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*. Cambridge University Press.
- Xin, L., & Goldberg, D. A. (2018). Asymptotic optimality of tailored base-surge policies in dual-sourcing inventory systems. *Management Science*, 64(1), 437–452.
- Yang, J., & Shi, J. (2023). Discrete-item inventory control involving unknown censored demand and convex inventory costs. *Production and Operations Management*, 32(1), 45–64.
- Yano, C. A., & Gilbert, S. M. (2005). Coordinated pricing and production/procurement decisions: A review. In Amiya K. Chakravarty and Jehoshua Eliashberg (Eds.), *Managing business interfaces* (pp. 65–103).
- Yuan, H., Luo, Q., & Shi, C. (2021). Marrying stochastic gradient descent with bandits: Learning algorithms for inventory systems with fixed costs. *Management Science*, 67(10), 6089–6115.
- Zhang, H., Chao, X., & Shi, C. (2018). Technical note—Perishable inventory systems: Convexity results for base-stock policies and learning algorithms under censored demand. *Operations Research*, 66(5), 1276–1286.
- Zhang, H., Chao, X., & Shi, C. (2020). Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Science*, 66(5), 1962–1980.
- Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent. In Tom Fawcett and Nina Mishra (Eds.), *Proceedings of the 20th international conference on machine learning (ICML-03)* (pp. 928–936).
- Zipkin, P. (2008). On the structure of lost-sales inventory models. *Operations Research*, 56(4), 937–944.

16. Inventory models with financial flows

Kevin H. Shang and Jing-Sheng Jeannette Song

16.1 INTRODUCTION

The goal of supply-chain management is to match supply with demand effectively by coordinating the activities of multiple firms involved in the production, distribution, and sales of a physical good. The performance of a supply chain depends on how well the material flows, information flows, and financial flows within the supply chain are coordinated. The intertwined relationships between these flows have made this coordination process challenging in practice. While a supply chain includes financial flows, they are seldom modeled explicitly in the inventory literature until recently. One reason for this is that many classical models originated in a centralized system (e.g., a vertically integrated firm), in which finding an efficient inventory policy to govern the material flow to meet demand in a satisfactory fashion is of primary concern, and the capital involved in managing the inventory of a specific item is not particularly constrained. Another justification is perhaps due to the MM theorem (Modigliani & Miller, 1958), which states that when the financial markets are perfect and efficient, a firm's value is independent of its capital structure, i.e., the specific mix of debt and equity used to finance a company's assets and operations. This is often referred to as *the capital structure irrelevance principle*. The majority of the inventory literature has been focused on maximizing a firm's profit, or equivalently, the present value of its future profits, generated through inventory decisions. Nevertheless, the financial markets are hardly perfect and efficient. For example, during the 2008 financial crisis, many supply chains were disrupted because upstream firms failed to maintain their normal operations due to financial illiquidity. Also, for entrepreneurs and small and medium-sized enterprises (SMEs), capital constraints are non-negligible. Thus, it is crucial to understand how financial flows impact inventory decisions, and how to obtain the optimal joint inventory and financial decisions for a supply chain.

There is a long debate about a firm's valuation in the literature. Different perspectives regarding a firm's value have been proposed. For example, Baye and Prince (2014) define the value of a firm as the present value of the firm's current and future profits. This perspective originated from Modigliani and Miller. As we shall see later, maximizing a firm's value under this definition is the same as maximizing a firm's equity or net worth in a finite time horizon inventory model. We shall follow this definition for the inventory models introduced in this chapter. That said, in some corporate finance textbooks, the value of a firm is defined as the present value of total dividend payments to the shareholders. Note that the dividend payment cycle for most firms is on a quarterly or bi-annual basis, whereas the inventory order decision is mostly on a daily or weekly basis. Thus, from the operations' lens, the objective of maximizing the net worth (equity) of a firm does not contradict the latter perspective as the firm can maximize its net worth and thereafter design an appropriate dividend policy (i.e., either paying dividends with cash, keeping cash as retained earnings or both) at the end of the planning horizon.

The literature that studies the interface of operations and finance mainly considers issues originated from the relaxation of the perfect market assumptions in Modigliani and Miller: absence of taxes, no bankruptcy costs, no transaction costs on trading securities, information symmetry, and the borrowing rate the same as the return rate. Many game-theoretic models have been published to address financing issues under information asymmetry between borrowers and lenders and under the existence of bankruptcy costs. We refer the reader to Babich and Kouvelis (2018) for a review of papers published in a special issue of *Manufacturing & Service Operations Management*. Most of these game theoretical models focus on one-period models that intend to reveal insights into the interface problems. The goal of this chapter is to complement this literature by focusing on multi-period inventory models with financial considerations.

There are two issues to address for the finite-horizon models. The first is to obtain the optimal joint inventory and payment decisions when there is a financial constraint for inventory replenishment. One way to model the financial constraint is to assume a hard budgetary constraint that restricts the inventory order quantity. The other way is to assume that the firm can create short-term debt by borrowing funds from an external source. In Section 16.2, we review papers using this modeling approach. In such a setting, it is practical to assume that the interest rate for the debt (borrowing rate) is higher than that of returns on the investment (return rate), which relaxes the perfect market assumption in the MM theorem. Clearly, when the borrowing rate is extremely large, the optimal joint inventory and payment decision would be close to that of the budgetary constraint model.

The second issue is about payment timing for supply chain firms. This issue matters even when there is no financial constraint for firms. As we shall demonstrate in Section 16.2, when the financial market is perfect, i.e., the borrowing rate is equal to the return rate, the financial cost rate of holding inventory is equivalent to the unit purchase cost times the return rate. Consequently, in the standard inventory model, which assumes that the inventory order and payment occur at the same time, one can show that the holding cost rate is the sum of the physical holding cost rate and the financial holding cost rate. In practice, however, some firms with stronger bargaining power may delay the payment time, which causes supply-chain inefficiency because of the increased financial holding cost for some supply-chain partners. Broadly, this is related to the allocation of the cost of capital in a supply chain.

In Section 16.3, we present a continuous-time framework that evaluates the financing inventory cost in a series inventory system under wholesale price contracts and any given payment timing scheme that specifies when to pay the upstream stage and when to collect the payment from the downstream stage for any unit flowing through the system. We then review how this framework can be applied to design payment timing contracts to improve inventory decisions in decentralized supply chains.

In Section 16.4, we briefly summarize other related research. Finally, in Section 16.5, we point out several innovative research directions related to blockchain and crowdfunding, and conclude the chapter.

16.2 MODELS WITH FINANCIAL CONSTRAINTS

We start with a single-stage inventory model that sets the stage for the problem. In Section 16.2.1, we consider a finite-horizon inventory model with cash flows triggered by inventory decision and demand orders. We provide a condition under which inventory decisions and cash flows

can be decoupled. This result leads to the standard inventory model in the literature. This condition satisfies the no-friction assumption of the MM theorem. On the other hand, when this condition does not hold, the cash flow does influence the inventory decision. We show that the system working capital is an important state that determines the optimal joint inventory decision. In Section 16.2.2, we consider a centralized supply chain owned by a firm with a financial constraint. We shall characterize the optimal joint inventory and cash policy that maximizes the working capital for the vertically integrated firm.

16.2.1 Single-Stage System

Consider a firm facing nonstationary random demand in a finite horizon. At the beginning of each period, an order is placed to meet the uncertain demand. The objective is to maximize the expected net worth (equity) at the end of the horizon. For simplicity, let us assume the firm has no investment activities other than investing in inventory and in the money markets with the cash it accumulates. Also, we do not consider long-term assets and liabilities. Thus, the objective is equivalent to maximizing the expected working capital at the end of the horizon.

We assume that the order lead time is zero. Define the following parameters for the system.

- h_p = physical holding cost rate (\$/unit/period);
- b_p = physical backorder cost rate (\$/unit/period);
- r = interest return rate for investing in the money market;
- d = borrowing rate from the capital market, where $d \geq r$;
- p = unit selling price;
- c = unit purchase cost.

Here, the physical holding cost rate h_p refers to the costs of carrying physical inventory, which does *not* include the financial opportunity cost due to holding inventory. The physical back-order cost rate b_p should be viewed the same way – it is the tangible, monetary penalty costs related to backlogging, e.g., expediting production and delivery costs.

We count the time forward, i.e., $t = 1, 2, \dots, t, t+1, \dots, T$, where T is the end of the horizon. Let D_t be the demand that occurred in period t . To examine the system dynamics, we introduce the following state variables:

- x_t = net inventory level at the beginning of period t ;
- w'_t = net cash level at the beginning of period t ;
- w_t = working capital at the beginning of period $t = w'_t + cx_t$.

Let y_t be the inventory position after ordering, and the order quantity is $(y_t - x_t)$. Define the inventory-related cost and the cash-related gain in period t :

$$G_t(y_t) = E[h_p(y_t - D_t)^+ + b_p(y_t - D_t)^-],$$

$$R_t(y_t) = r(w'_t - c(y_t - x_t))^+ - d(w'_t - c(y_t - x_t))^-,$$

where $x^+ = \max\{x, 0\}$, $x^- = -\min\{x, 0\}$, and $E[\cdot]$ is the expected value over the random demand. The first term $(y_t - D_t)^+$ in the G_t function is the on-hand inventory at the end of the period, whereas the second term $(y_t - D_t)^-$ is the backorder level. The G_t function represents the inventory holding and backorder cost, or the inventory-related cost in short in period t . Similarly, the term $(w'_t - c(y_t - x_t))$ in the R_t function is the net cash level after inventory payment. It yields an interest gain r if positive and an interest loss d if negative (representing the borrowing rate from financial institutions in order to pay the ordered inventory). The R_t function is the cash-related gain in period t .

The transitions of the inventory and cash states between periods are as follows:

$$x_{t+1} = y_t - D_t, \quad (16.1)$$

$$\begin{aligned} w'_{t+1} &= w'_t - c(y_t - x_t) + pD_t + R_t(y_t) - G_t(y_t) \\ &= (1+r)(w'_t - c(y_t - x_t))^+ \\ &\quad - (1+d)(w'_t - c(y_t - x_t))^- + pD_t - G_t(y_t). \end{aligned} \quad (16.2)$$

Equation (16.2) states that the next period's net cash is the result of total cash inflows and outflows of the current period. Here, we assume that the customer will pay on order so the revenue is pD_t . If there is debt, i.e., $(w'_t - c(y_t - x_t))$ is negative, it will incur an interest loss and this debt will carry over to the next period.

The net working capital is

$$\begin{aligned} w_{t+1} &= (1+r)(w_t - cy_t)^+ - (1+d)(w_t - cy_t)^- + pD_t - G_t(y_t) + c(y_t - D_t) \\ &= (1+r)w_t + (p-c)D_t - rcy_t - G_t(y_t) - (d-r)(w_t - cy_t)^- \end{aligned} \quad (16.3)$$

$$= w_t + (p-c)D_t - G_t(y_t) + r(w_t - cy_t)^+ - d(w_t - cy_t)^-. \quad (16.4)$$

So far, we have considered a very practical and general environment for a firm that can invest and borrow at different rates.

16.2.1.1 Perfect financial markets

Let's consider a special case in which the financial market is *perfect*. In our model, this is equivalent to assuming that $r=d$, i.e., the interest return rate is equal to the borrowing rate. In this case, the firm can freely borrow cash (as there is no penalty for borrowing excess cash) so cash is no longer a concern. This is exactly what the classic inventory model assumes. With this assumption, Equation (16.4) becomes

$$\begin{aligned} &(1+r)w_t + (p-c)D_t - rcy_t - G_t(y_t) \\ &= (p-c(1+r))D_t + (1+r)w_t - (h_p + rc)(y_t - D_t)^+ - (b_p - rc)(y_t - D_t)^-. \end{aligned} \quad (16.5)$$

Notice that D_t is an exogenous random variable, and w_t is the initial system state at the beginning of period t . Thus, to maximize the expected working capital in period $t+1$, one only needs to minimize the expected cost in Equation (16.5), i.e.,

$$E[(h_p + rc)(y_t - D_t)^+ + (b_p - rc)(y_t - D_t)^-]. \quad (16.6)$$

Equation (16.6) is the single-period inventory-related cost in the classic inventory model. We have a sound economic meaning for the cost parameters. The holding cost rate h is $(h_p + rc)$, which is the sum of the physical holding cost rate h_p and the opportunity cost of capital rc due to inventory investment. We refer to rc as the *financial* holding cost rate. The backorder cost rate b is $(b_p - rc)$, which is the physical backorder cost minus the opportunity cost of capital rc .¹ It can be shown that a base-stock policy is optimal. More specifically, one can view the holding cost rate h and backorder cost rate b as follows:

$$h = h_p + rc;$$

$$b = b_p - rc.$$

If the demand is iid, the optimal base-stock level s^* can be obtained from the well-known optimality equation as shown in the inventory teaching note: Finding s^* such that

$$P(D \leq s^*) = \left(\frac{b_p - rc}{b_p + h_p} \right) = \left(\frac{b}{b + h} \right).$$

For simplicity, let's write

$$s^* = F^{-1} \left(\frac{b}{b + h} \right),$$

where F is the cdf of the demand distribution and F^{-1} is the inverse cdf function.²

We pose a discussion on estimating the cost parameters in practice here. According to the above analysis, when the financial market is perfect, the inventory holding cost rate is composed of the physical holding cost rate h_p and the financial holding cost rate rc . Recall that r is the interest rate due to the cash investment in the capital market. Broadly speaking, if a firm conducts investments by financing through debt and equity, the required expected return would be WACC, the weighted average cost of capital. In most OM textbooks, WACC is often recommended to estimate the financial holding cost rate. To estimate the physical holding cost rate, however, is a tough task. The physical holding cost comes from, for example, managing and maintenance expenses, storage, insurance, shrinkage, obsolescence, etc. The list is very long and often business specific. Technically speaking, one should sum up all these costs in a time period and allocate the cost to each inventory unit sold during this period. There are a number of alternative cost accounting systems that can be relevant for some purposes while being inadequate for others. Thus, it is neither always possible nor economical to keep track of all costs, or to split them and allocate them properly. Fortunately, the physical holding cost rate is often small.

As for the backorder rate, this is the penalty cost incurred for an arriving demand that cannot be satisfied immediately due to stock out. Some physical (tangible) penalty costs may occur, e.g., expedited shipping costs, overtime production, etc. Let b_p denote the physical backorder cost. Then, the backorder cost rate is $b = b_p - rc$. This is because a unit of inventory shortage

implies that additional cash c was invested in the capital market, gaining the return rate r . Thus, the actual tangible backorder cost rate is b , which is less than b_p .

16.2.1.2 Imperfect financial markets

Let's turn to a more realistic scenario where the financial market is not perfect, i.e., $d > r$. In this case, Equation (16.4) suggests that to maximize the expected working capital in period $t+1$, one has to minimize $E[G_t(y_t)] - R_t(y_t)$, or equivalently,

$$E[h_p(y_t - D_t)^+ + b_p(y_t - D_t)^-] - r(w'_t - c(y_t - x_t))^+ + d(w'_t - c(y_t - x_t))^- \quad (16.7)$$

One can see that the inventory decision y_t is affected by cash level w'_t through the last two terms of Equation (16.7). Thus, the inventory decision cannot be decoupled from the financial flow.

Luo and Shang (2019) show that the optimal inventory policy depends on the working capital level. More specifically, recall $w_t = w'_t + cx_t$. Define

$$\underline{s} = F^{-1}\left(\frac{b_p - dc}{b_p + h_p}\right) = F^{-1}\left(\frac{b_p - dc}{b + h}\right),$$

$$s = F^{-1}\left(\frac{b_p - rc}{b_p + h_p}\right) = F^{-1}\left(\frac{b}{b + h}\right).$$

Clearly, $\underline{s} \leq s$. The optimal policy is executed as follows. When working capital level w_t is greater (less, respectively) than cs ($c\underline{s}$, respectively), one should order up to base-stock level s ($\underline{s}$, respectively). If the working capital is between cs and $c\underline{s}$, one should order up to the total working capital level w_t . We refer to this optimal policy as the $(\underline{s}, s)$ policy.

16.2.1.3 Trade credit

The single-stage model can be generalized by incorporating two-level trade credit, i.e., the firm offers trade credit to its customers while receiving credit from its supplier. The trade credit is a one-part (net-term) contract, that is, the payment is due in a certain time period after the invoice is issued. The firm receives sales revenue, or accounts receivable (A/R), after a delayed *collection period* defined as n , following the demand, and pays for the ordered inventory, or accounts payable (A/P), after a *payment period*, define as m , following the delivery of goods. The other settings are the same as those of the single-stage model define before.

The sequence of events is as follows: At the beginning of period t , (1) an inventory order decision is made and a new A/P is generated; (2) shipment arrives; (3) payment due in this period (corresponding to the inventory ordered in period $t-m$) is made to the supplier; (4) a deficit penalty cost is incurred in case of insufficient payment (a negative cash level) or an interest return is gained in case of a positive cash level; (5) demand is realized during the period and a new A/R is generated. Customer payment due in this period (corresponding to the sales in period $t-n$) is collected; at the end of the period, all inventory-related costs are calculated.

The problem can be formulated as a multi-state dynamic program that keeps track of inventory level and cash balance, as well as different ages of A/P and A/R within the payment and collection periods, respectively. Define the state variables at the beginning of period t :

$\mathbf{P}_t = (P_{t-m}, \dots, P_{t-1})$: m -dimensional vector of accounts payable;

$\mathbf{R}_t = (R_{t-n}, \dots, R_{t-1})$: n -dimensional vector of accounts receivable.

Here, P_{t-i} and R_{t-j} denote the A/P and A/R created in period $t-i$ and $t-j$, respectively, for $i = 0, 1, \dots, m$ and $j = 0, 1, \dots, n$. So P_{t-m} and R_{t-n} are the most aged A/P and A/R, while P_t and R_t are A/P and A/R created in the current period.

It turns out that the exact optimal policy cannot be characterized because of the curse of dimensionality in the Bellman equation: the inventory decision results in the inventory-related and cash-related costs in *each* of the collection and payment periods, which, in turn, affects the cash level as well as the working capital in the objective function. Consequently, Luo and Shang (2019) propose a simplified model by streamlining the cash dynamics. Under this simplification, they show that the two-parameter, working capital-dependent policy is near-optimal for the exact system. The caveat is to define *effective* working capital to control the inventory decision.

The effective working capital $\underline{w}_t$ is defined differently based on the order of m and n . Specifically, when $m \leq n$,

$$\underline{w}_t = w_t - \sum_{k=1}^{n-m} R_{t-k},$$

which is equal to the working capital in period t excluding the known accounts receivable in periods $t-n+m, \dots, t-1$. On the other hand, when $m > n$, the effective working capital is

$$\bar{w}_t = w_t + p \sum_{k=1}^{(m-n)} E[D_k],$$

where the second term is the *expected* A/R within $(m-n)$ periods. Using the effective capital, a two-parameter control policy similar to the one introduced in Section 16.2.1.2 can be derived. The heuristic policy resembles practical working capital management under which a firm makes inventory decisions according to the working capital level. A numerical study suggests that the heuristic is effective.

16.2.2 Series System

Luo and Shang (2015) consider the joint inventory order and cash payment problem in a supply chain. They focus on a periodic-review, two-stage serial supply chain in which stage 1 orders from stage 2, which orders from an outside ample vendor. The supply chain is owned by a single corporation, with stage 1 being the headquarter and stage 2 the subsidiary. The logistics of this supply chain are fairly standard: Stage 1 faces stochastic customer demand D_t in period t . The demands are independent between periods, but the demand distributions may differ from period to period. We assume that the material lead time is one period for both stages (without loss of generality). In each period, each stage reviews its local inventory position (= inventory on order + inventory on hand – backorders) and places an order from its upstream stage. Unsatisfied demands are fully backlogged.

The headquarter creates a corporate master account that manages the cash of the entire supply chain. In each period, after receiving the customer's payment, the headquarter decides the amount of cash used for external investments, such as money and bond markets, facility expansion, or R&D, etc. The remaining cash will be used for operations, that is, paying for inventory ordered for the outside vendor. Here and in the sequel, we use prime to indicate local (stage-specific) variables and parameters. The external investment portfolio has a return rate of η' . Since holding cash for operations has a zero return rate, η' can be viewed as a *cash holding cost rate*, which represents the opportunity cost of holding cash. Moreover, the headquarters can liquidate its portfolio assets to assist with inventory payment, if necessary. Nevertheless, how much cash can flow into the cash account depends on an exogenous market condition or the type of invested assets described by a limit $K'(\geq 0)$ in each period. (For example, some R&D investments may not be liquidated. In this case, K' would be zero.) Let β'_i and β'_o denote the unit transaction cost charged on the cash transferred to and from the cash account, respectively. In practice, these transaction costs can be regarded as brokerage fees. Here, β'_i , β'_o and K' represent the level of easiness of liquidating the portfolio assets into cash. Note that when $\beta'_o = 0$, $K' = 0$, β'_i corresponding to the borrowing rate d , and η' corresponding to the investment return rate r , the model will reduce to the single-stage model introduced in Section 16.2.1 if there is only one stage.

We now introduce the other cost parameters. Following the inventory literature, we charge a linear local holding cost h'_i for each unit of inventory held at stage i in each period, and backorder cost b for each unit of backorder incurred at stage 1 in each period. Here, we assume that $h'_1 > h'_2 > \eta'c$, i.e., holding a unit of inventory at downstream is more costly than that at upstream, and holding a unit of inventory is more costly than holding the same value amount of cash. The latter is generally true since inventory holding cost consists of both the financial opportunity cost and the physical shelf cost.

The inventory replenishment and cash retention decision is made centrally by the headquarters. The sequence of events in a period is as follows: At the beginning of the period, (1) shipments are received at both stages; (2) payment is made to the outside vendor; (3) cash retention decision is made; (4) orders are placed at both stages. During this period, demand is realized and sales revenue is collected. At the end of the period, all inventory and cash-related costs are calculated. The planning horizon is T periods, and the objective is to maximize the expected working capital for the supply chain at the end of the horizon.

We now define state and decision variables. For stage $i = 1, 2$ and period t , let

$x'_{1,t}$ = net inventory level at stage 1 after Event (1);

$x'_{2,t}$ = on hand inventory level at stage 2 after Event (1);

w'_t = cash balance in the pooled account after Event (2);

v_t = amount of cash transferred into the pooled account in Event (3);

$z_{i,t}$ = order quantity for stage i made in Event (4);

Note that v_t^+ is the cash amount that flows into the pooled account and v_t^- is the cash amount that flows out for investment. Clearly, v_t cannot exceed K' . Let p_1 be the unit selling price

to the end customer and c be the unit procurement cost from the outside vendor. We assume $c < p_1$ to ensure profitability. The system dynamics are shown below:

$$x'_{1,t+1} = x'_{1,t} + z_{1,t} - D_t, \quad (16.8)$$

$$x'_{2,t+1} = x'_{2,t} + z_{2,t} - z_{1,t}, \quad (16.9)$$

$$w'_{t+1} = w'_t + v_t - cz_{2,t} + p_1 D_t. \quad (16.10)$$

For the cash dynamics in Equation (16.10), we assume that the actual payment transaction to the outside vendor occurs upon the receipt of shipments. That is, the vendor will not receive the payment determined in period t until period $t+1$, when stage 2 receives the shipment (placed in period t). This payment practice is similar to a letter of credit (LC). In other words, we can view that there is a one-period lead time for the cash payment. (Our analysis holds for an alternative assumption of payment on an order by slightly changing the dynamics.) As for the payment received at stage 1, we assume that the customer will pay at the order epoch. This assumption is reasonable as all demands will be filled under the backorder model. We do not include inventory holding and backorder costs in Equation (16.10) because of tractability; see Section 16.2.1.3.

Define $\mathbf{x}' = (x'_1, x'_2)$, and $\mathbf{z} = (z_1, z_2)$. The constraint set in each period is

$$\hat{S}(x'_2, w') = \{\mathbf{z}, v \mid 0 \leq z_1 \leq x'_2, 0 \leq z_2 \leq (w' + v) / c, v \leq K'\}. \quad (16.11)$$

The first constraint states that stage 1's order quantity cannot exceed stage 2's on-hand inventory; the second constraint states that stage 2's order quantity is constrained by the cash balance in the pooled account, which also implies that the investment amount in each period cannot exceed its on-hand cash level, i.e., $v \geq -w'$. Finally, the last constraint imposes a limit K' on the amount of cash that can be injected into the pooled cash account.

It can be shown that maximizing the expected system net worth (or working capital in this model) is equivalent to minimizing the expected total cost which includes inventory-related costs and cash-related costs. Thus, we focus on the latter expression. Specifically, the single-period expected cost function is

$$\begin{aligned} \hat{G}_t(\mathbf{x}', w', z_2, v) = & E_{D_t} \left[h'_1(x'_1 - D_t)^+ + b(x'_1 - D_t)^- \right] + h'_2 x'_2 + cz_2 \\ & + \eta' E_{D_t} (w' + v + p_1 D_t) + \beta'_1 v^+ + \beta'_0 v^-. \end{aligned} \quad (16.12)$$

The first line in the cost function is the inventory-related cost, which includes inventory holding, backlogging and procurement costs. By convention, we charge h'_2 to the pipeline inventory so $h'_2 x'_2$ is the cost for the inventories held at stage 2 plus those in the pipeline. The second line is the cash-related cost, which includes cash holding and transaction costs. As shown, we charge η' for $w' + v + p_1 D_t$ because the inventory payment to the outside vendor is held until the receipt of goods.

Let α be the single-period discount rate. Denote $\hat{J}_t(\mathbf{x}', w', \mathbf{z}, v)$ as the expected cost over period t to $T+1$, given states and decisions $(\mathbf{x}', w', \mathbf{z}, v)$. Denote $\hat{V}_t(\mathbf{x}', w')$ as the minimum expected cost over period t to $T+1$ over all feasible decisions. The dynamic program is

$$\hat{J}_t(\mathbf{x}', w', \mathbf{z}, v) = \hat{G}_t(\mathbf{x}', w', z_2, v) + \alpha E_{D_t} \left[\hat{V}_{t+1}(x'_1 + z_1 - D_t, x'_2 + z_2 - z_1, w' + v - cz_2 + p_1 D_t) \right], \quad (16.13)$$

$$\hat{V}_t(\mathbf{x}', w') = \min_{\mathbf{z}, v \in \hat{S}(x'_2, w')} \hat{J}_t(\mathbf{x}', w', \mathbf{z}, v), \quad (16.14)$$

with $\hat{V}_{T+1}(\mathbf{x}', w') = 0$.

The local formulation in Equation (16.13) and Equation (16.14) is difficult to solve. Specifically, one can show the joint convexity of $\hat{J}_t(\cdot)$ and derive a state-dependent global minimum solution. However, computing the solution is quite hard due to the curse of dimensionality. Luo and Shang (2015) show that the problem can be solved efficiently by converting the system into an echelon perspective, and the optimal policy has a simple structure.

16.2.2.1 Echelon formulation

The original two-stage system can be transformed into a three-stage serial model by introducing new system variables. First, define the following *echelon* variables:

$$x_1 = x'_1, x_2 = x'_1 + x'_2, w = x'_1 + x'_2 + w' / c.$$

Let $\mathbf{x} = (x_1, x_2)$. We refer to $\mathbf{x}$ as the *echelon net inventory level*, and w as the *net working capital level* measured in inventory units, which is obtained by converting cash to inventory at the value of c . This state transformation explicitly treats cash as inventory. More specifically, the financial flow in the system can be seen as an extension of the material flow after “flipping” the corporate master account to upstream. We define the corresponding echelon decision variables:

$$y_1 = x'_1 + z_1, y_2 = x'_1 + x'_2 + z_2, r = x'_1 + x'_2 + (w' + v) / c.$$

Let $\mathbf{y} = (y_1, y_2)$. With this transformation, the cash account becomes stage 3 in the new system, directly supplying stage 2. We hereby refer to echelon 3 (with state variable w) as the system working capital.

Similar to the multi-echelon inventory model, we derive the echelon holding cost rate as follows: $\eta = \eta'c$, $h_2 = h'_2 - \eta'c$, and $h_1 = h'_1 - h'_2$. Since $h'_1 > h'_2 > \eta'c$ by assumption, we have $h_1 > 0$ and $h_2 > 0$. Furthermore, let $\beta_i = \beta'_i c$, $\beta_o = \beta'_o c$, $\theta = p_1 / c - 1 > 0$, and $K = K' / c$. With these echelon terms, the state dynamics in Equations (16.8)–(16.10) become

$$x_{1,t+1} = y_{1,t} - D_t, \quad x_{2,t+1} = y_{2,t} - D_t, \quad w_{t+1} = r_t + \theta D_t,$$

and the constraint set becomes

$$S(\mathbf{x}, w) = \{\mathbf{y}, r | x_1 \leq y_1 \leq x_2 \leq y_2 \leq r \leq w + K\}. \quad (16.15)$$

We further specify the holding and backorder costs associated with each echelon:

$$H_{1,t}(x_1) = E_{D_t}[(h_1 + h_2 + \eta + b)(D_t - x_1)^+ + h_1(x_1 - D_t)],$$

$$H_{2,t}(x_2) = E_{D_t}h_2(x_2 - D_t),$$

$$H_{3,t}(r) = E_{D_t}\eta(r + \theta D_t).$$

Then, we can rewrite the dynamic program in Equations (16.13) and (16.14) as follows:

$$J_t(\mathbf{x}, w, \mathbf{y}, r) = G_t(\mathbf{x}, w, y_2, r) + \alpha E_{D_t} V_{t+1}(y_1 - D_t, y_2 - D_t, r + \theta D_t), \quad (16.16)$$

$$V_t(\mathbf{x}, w) = \min_{\mathbf{y}, r \in \mathcal{S}(\mathbf{x}, w)} J_t(\mathbf{x}, w, \mathbf{y}, r), \quad (16.17)$$

where the single-period cost function can be shown as

$$\begin{aligned} G_t(\mathbf{x}, w, y_2, r) &= H_{1,t}(x_1) + H_{2,t}(x_2) + H_{3,t}(r) \\ &\quad + c(y_2 - x_2) + \beta_i(r - w)^+ + \beta_o(r - w)^-. \end{aligned}$$

16.2.2.2 The optimal policy

Here, we only state the optimal joint policy for the cash pooling model in Equations (16.16) and (16.17), which includes two types of decisions made through four control parameters (y_1^*, y_2^*, l^*, u^*) in each period. For the inventory ordering decisions, each stage implements an echelon base-stock policy. That is, stage i reviews its x_i at the beginning of each period. If $x_i < y_i^*$, it orders up to y_i^* or as close as possible if its upstream does not have sufficient stock; otherwise, it does not order. For the cash retention decision, stage 1 reviews w : if $w > u^*$, it disposes of cash down to the maximum of u^* and x_2 ; if $w < l^*$, it retrieves cash up to l^* or as close as possible (due to the upper bound K); otherwise, it does not transfer cash.

The derivation of the optimal policy extends the decomposition framework of Clark and Scarf (1960), where the notion of *induced penalty cost function* is introduced. Simply speaking, the induced penalty cost function is the cost charged to an upstream stage when it fails to satisfy the order of its immediate downstream stage. In the current cash pooling model, a *new* notion of the penalty cost function is introduced. This new penalty function serves an opposite role to the induced penalty function, in that it is charged to a *downstream* stage if it stocks too much, making the system working capital exceed an ideal interval. We refer to Luo and Shang (2015) for the detailed derivation.

It is interesting to compare the cash pooling model with the Clark-Scarf model when the financial holding cost rate (or the external investment rate) η increases. The Clark-Scarf model can be viewed as a special cash pooling system in which $\beta_i = \beta_o = 0$, and $K = \infty$, that is, no financial constraint in the model. In the Clark-Scarf model, when η increases, the local holding cost rates $h_2' (= h_2 + \eta)$ and $h_1' (= h_1 + h_2 + \eta)$ increase. Under such a condition, it is known that the optimal base-stock level for stage 1, y_1^o , will increase whereas the optimal system stock level, y_2^o , will decrease (Shang & Song, 2003). This is intuitive – since it is more costly to hold inventory for the entire system, it makes sense to reduce the total system stock; however, since the “marginal” holding cost at stage 1 (i.e., $h_1 = h_1' - h_2'$) remains the same, the system tends to push more stock to stage 1 in order to maintain the same service level.

The conclusion is different from that of the cash pooling model, where Luo and Shang (2015) show that the system stock, i.e., y_2^* , increases in η . A higher η implies a better external investment return. Intuitively, the firm should dispose of more cash to external investment, which leads to a low cash level. Thus, it is important for the divisions to stock more to prevent inventory shortage as it might be costly to transfer cash for purchasing inventory when the transaction costs are sufficiently large. This effect is more prominent when the demand is increasing because more cash is needed. Luo and Shang (2015) provide a refinement to traditional wisdom: A firm should stock less when the inventory holding cost increases. Our finding suggests that this statement is true only when there are no costs or restrictions for cash transfers.

16.3 MODELS WITH PAYMENT TIMES

The timing with which a firm pays its supplier and when it receives payment from customers affects how much funding each firm needs. In turn, this affects how the financing costs are allocated across the supply chain. Changing the timing of payments may also incentivize individual firms to follow different physical inventory management policies. For example, by paying a supplier earlier, a retailer may not only reduce the supplier's financial inventory costs, but also incentivize them to carry more inventory to provide the retailer with a higher service level. Thus, it is desirable to have an analytical modeling tool that can simultaneously evaluate supply-chain members' financing costs under different payment timing arrangements and also derive their relationship with physical inventory policies at each stage. Tong et al. (2020) make an initial attempt to develop such a tool for a general n -stage supply chain. They also apply this tool to demonstrate how payment timing can be used to improve coordination in a two-stage supply chain. We now briefly introduce this framework.

The framework builds on and extends the literature on dynamic multi-echelon inventory systems, in particular on the serial system with stationary demand, linear ordering costs, and full backlogging. The development in Sections 16.3.1–16.3.3 is related to the stream of literature focusing on performance evaluation and optimization within a given type of inventory policy for a continuous-review centralized system; see Axsäter (2003) for a review. This stream takes a continuous-time discrete-demand formulation and studies the steady-state system behavior. We adopt this setting because it offers a convenient and natural way to trace and track the ordering (information flow) and shipping/delivery (material flow) of each unit that goes through the system, and characterize the payment process (cash flow) of this unit due to payment triggers associated with those events. This unit analysis (pegging each unit with its customer demand) is inspired by the approach taken by Axsäter (1990) for the Poisson demand process and Muharremoglu and Tsitsiklis (2008) for even more general serial inventory systems. Chapter 6 of this handbook provides a comprehensive review of this approach.

16.3.1 Physical and Financial Flows

Consider a serial supply chain of size n . The most upstream stage n procures from an outside supplier with ample supply and sells to stage $n-1$ who, in turn, sells to stage $n-2$. This process continues until stage 1 sells to the final customer at stage 0. We denote stages with the subscript j , $j = 0, 1, \dots, n$. There is a constant lead time L_j between stages $j+1$ and

j . The final demand process at stage 1, $\mathbf{D}_1 = \{D_1(t) : t \geq 0\}$ is an exogenous counting process with stationary, independent increments, where $D_1(t)$ is the cumulative demand in $[0, t]$, $t \geq 0$. Let λ be the mean demand rate, i.e., $E[D_1(t)] = \lambda t$, $t \geq 0$. Unmet orders are fully backordered and orders are fulfilled in the order that they are received. Each stage uses continuous review.

The following are standard measures to describe the physical flow of the system:

$I_j(t)$ = local physical inventory at stage j at time t

$B_j(t)$ = local physical backorders at stage j at time t

$IO_j(t)$ = inventory on order at stage j at time t

$IT_j(t) = IO_j(t) - B_{j+1}(t)$ = inventory in transit from stage $j+1$ to stage j at time t

$IN_j(t) = I_j(t) - B_j(t)$ = net inventory at stage j at time t

$ITP_j(t) = IN_j(t) + IT_j(t)$ = inventory-transit position at stage j at time t

$IOP_j(t) = IN_j(t) + IO_j(t)$ = inventory-order position at stage j at time t .

By studying these physical inventory measures, researchers have developed a fairly thorough understanding of the material flow in serial supply chains and how to manage it. For example, it is well-known that under linear order, holding, and backorder costs, an echelon base-stock policy is optimal in terms of minimizing the long-run average system cost – that is, there is a target level at each stage, and it is optimal for each stage to order enough units to bring inventory in transit to, located at, or downstream from that stage up to the specified target. If those echelon targets are weakly increasing in stage number j , then there exists an equivalent local base-stock policy consisting of base-stock levels $\mathbf{y} = (y_1, \dots, y_n)$ (Axsäter & Rosling, 1993). Under a local base-stock policy $\mathbf{y}$, stage j monitors its stage- j inventory order position $IOP_j(t)$ continuously. Whenever $IOP_j(t)$ falls below the target level y_j , an order is placed from stage $j+1$ to bring it back to this target. There exist efficient policy evaluation and optimization procedures for echelon base-stock policies (and thus their equivalent local base-stock policies); see, e.g., Zipkin (2000).

However, these physical flow measures do not track the times that payments are made between firms, and hence do not capture the *financial* flow of the system. Payments in practice are usually determined according to the times orders are placed, shipments are initiated, or inventory is received. To describe the financial flow, we introduce several fundamental supply-chain processes associated with these events:

$\{R_j(t), t \geq 0\}$ = cumulative units received at stage j (from stage $j+1$),

$\{D_j(t), t \geq 0\}$ = cumulative units demanded at stage j (from stage $j-1$),

$\{S_j(t), t \geq 0\}$ = cumulative units shipped at stage j (to stage $j-1$), $j = 1, 2, \dots, n$,

where $n+1$ refers to the outside supplier and 0 refers to the final customer. Assume that each of these processes are counting processes that start at zero at time 0–.

Assuming $I_j(0-) = B_j(0-) = IO_j(0-) = IT_j(0-) = D_j(0-) = 0$, we have

$$\begin{aligned} B_j(t) &= D_j(t) - S_j(t), \quad I_j(t) = R_j(t) - S_j(t), \\ IO_j(t) &= D_{j+1}(t) - R_j(t), \quad IT_j(t) = S_{j+1}(t) - R_j(t), \quad \text{for all } t \geq 0. \end{aligned}$$

Consequently,

$$\begin{aligned} R_j(t) &= IN_j(t) + \sum_{i=1}^{j-1} IOP_i(t) + D_1(t) \\ D_j(t) &= \sum_{i=1}^{j-1} IOP_i(t) + D_1(t) \\ S_j(t) &= -B_j(t) + \sum_{i=1}^{j-1} IOP_i(t) + D_1(t). \end{aligned} \tag{16.18}$$

The final demand process is exogenously given and the behavior of physical inventory measures can be characterized by leveraging the existing literature. We next show that payment processes follow the fundamental processes, hence, by the above relationships, we can study financial flows through physical inventory measures.

16.3.2 Payment Triggers and Processes

We focus our attention on payment terms of a certain type: a wholesale price payment at one stage is always made in full to its immediate upstream stage when the associated physical unit reaches a certain reference event corresponding to an increment in one of the fundamental processes in the supply chain, not necessarily at its own stage.

For $j = 0, 1, 2, \dots, n$, let w_j be the per unit wholesale price paid by stage j to stage $j+1$. (Here, w_0 is the price to the final customer.) Define τ_j as the *payment trigger between stage j and stage $j+1$* . Then $\tau_j \in \mathcal{I} = \{r_i, d_i, s_i, i = 0, 1, \dots, n\}$, where r_i = *receipt at stage i* , d_i = *demand at stage i* , and s_i = *shipment at stage i* . Note that it is possible that stage j pays its upstream according to a reference event of another stage i . For example, under $\tau_1 = s_2$, stage 1 pays stage 2 when stage 2 initiates a shipment to stage 1. (It is straightforward to extend our payment triggers to include constant payment delays, such as a payment that is due ten days after a reference event. These types of extensions do not have an effect on inventory decision-making based on the average-cost model.)

Each possible choice of payment trigger τ_j , in combination with the physical inventory flow, generates a specific payment process for stage j . Specifically, denote the payment process under any $\tau_j \in \mathcal{I}$ as

$$N_j(t | \tau_j) = \text{cumulative number of payments made by stage } j \text{ by time } t.$$

Then, the outgoing payment process at stage $j \in \{1, \dots, n\}$ (or, equivalently, the incoming payment process at stage $j+1$) can be expressed as a fundamental process at some stage $i \in \{1, \dots, n\}$:

$$N_j(t | \tau_j) = \begin{cases} R_i(t) & \text{if } \tau_j = r_i \\ D_i(t) & \text{if } \tau_j = d_i \\ S_i(t) & \text{if } \tau_j = s_i. \end{cases} \quad (16.19)$$

Observe that the payment process for stage j need not equal a fundamental process at its own stage; it may equal a fundamental process at another stage $i \neq j$. For example, if $\tau_1 = d_2$, then stage 1's outgoing payment process is simply $N_1(t | \tau_1 = d_2) = D_2(t)$. Thus, payment processes can be expressed in terms of fundamental processes.

A *payment scheme* for an n -stage supply chain is defined as the collection of the relevant $n+1$ payment triggers between each pair of consecutive stages, $\tau = (\tau_n, \dots, \tau_0) \in \mathcal{I}^{n+1}$.

16.3.3 Financial Inventory-Related Cost Assessment

For stage j , the relevant payment triggers are (τ_j, τ_{j-1}) . For every unit passing through, the stage pays w_j according to τ_j and receives w_{j-1} according to τ_{j-1} . While these triggers remain fixed for all units, the times at which they are activated will naturally be different for different units flowing through the system, and it will be convenient to track those specific times. To that end, label the unit that is used to fulfill the k th demand as the k th unit. (Because of the FCFS rule, the same physical unit will be the k th unit at all stages.) Then we denote d_1^k the time epoch of the k th demand at stage 1 (exogenous), τ_j^k the time epoch at which stage j pays for the k th unit, and τ_{j-1}^k the time epoch at which stage j collects for the k th unit.

A unit of inventory is financed by stage j at time t if at time t it has been paid for (to the upstream) but its payment from the downstream has not yet been collected, which will be true if $\tau_j^k \leq t < \tau_{j-1}^k$. On the other hand, a unit is negatively financed or “floated” by stage j at time t if its payment from the downstream has been collected before it is paid to the upstream, which holds if $\tau_{j-1}^k \leq t < \tau_j^k$. At any given time t , the total (net) amount of *financed inventory* for stage j is defined as

$$\begin{aligned} I_j^f(t | \tau_j, \tau_{j-1}) &= \sum_k \left[1\{\tau_j^k \leq t < \tau_{j-1}^k\} - 1\{\tau_{j-1}^k \leq t < \tau_j^k\} \right] \\ &= N_j(t | \tau_j) - N_{j-1}(t | \tau_{j-1}). \end{aligned} \quad (16.20)$$

Similarly, we say a unit of *final* customer demand is financed by stage j at time t if at time t the final customer demand has arrived for that unit, but stage j has not yet collected payment from the downstream for that unit, i.e., $d_1^k \leq t < \tau_{j-1}^k$. On the other hand, demand is negatively financed or “floated” by stage j at time t if payment has been collected for the unit that will satisfy that final customer demand at some point in the future, i.e., $\tau_{j-1}^k \leq t < d_1^k$. At any given time t , the total (net) number of *financed demands* for stage j is defined as

$$\begin{aligned} B_j^f(t | \tau_{j-1}) &= \sum_k \left[1\{d_1^k \leq t < \tau_{j-1}^k\} - 1\{\tau_{j-1}^k \leq t < d_1^k\} \right] \\ &= D_1(t) - N_{j-1}(t | \tau_{j-1}). \end{aligned} \quad (16.21)$$

The definition of financed demands here is always in relationship with the final customer demand at stage 1. In particular, under many common payment triggers it may be rare for a

supplier to have positive financed demands according to this definition: suppliers would often collect payments from the retailer for a unit before it is demanded by the final customer. In such cases, the financed demands are negative for the supplier.

Let α_j reflect stage j 's interest rate; more generally, it approximates the time value of money for stage j . Note that stage j incurs costs at a rate $\alpha_j w_j$ for each unit of financed inventory, and at a rate $\alpha_j (w_{j-1} - w_j)$ for each unit of financed demands. Therefore we define the total controllable (by payment timing) financial inventory cost rate function for stage j is

$$C_j^f(t | \tau_j, \tau_{j-1}) = \alpha_j w_j I_j^f(t | \tau_j, \tau_{j-1}) + \alpha_j (w_{j-1} - w_j) B_j^f(t | \tau_{j-1}).$$

For the supply chain, the total controllable financial inventory cost rate is simply the sum of the individual stages' costs

$$C^f(t | \tau) = \sum_{j=1}^n C_j^f(t | \tau_j, \tau_{j-1}).$$

For a given vector of payment triggers τ , let the vector $\mathbf{y}$ denote a base-stock policy with base-stock levels $\mathbf{y}$. Let $I_j^f(\mathbf{y})$, $B_j^f(\mathbf{y})$, and $C_j^f(\mathbf{y})$ denote the expected values of the limiting distributions of financed inventory, financed demands, and financial inventory cost rate, respectively, at stage j under policy $\mathbf{y}$. We can then find the optimal base-stock policy by minimizing the average cost

$$C_j^f(\mathbf{y}) = \alpha_j w_j I_j^f(\mathbf{y}) + \alpha_j (w_{j-1} - w_j) B_j^f(\mathbf{y}).$$

From Equations (16.18), (16.20), and (16.21), we can express the financial inventory cost rates for each stage and the entire system in terms of physical inventory measures. This enables combining inventory-related financial inventory costs with physical holding or backorder costs that are based on physical inventory measures – these latter costs would simply be captured by additional terms in the total cost rate function. This way of formulating costs has the added advantage of separating financial and physical costs, which are typically lumped together in the existing literature, thus allowing one to study the impacts of these costs separately.

Finally, this cost accounting method has the following intuitive cost-conservation property: If every stage has the same interest rate, then for any fixed inventory policy the total supply-chain financial inventory cost rate can be obtained by summing all the single-stage costs resulting from local payment triggers $(\tau_n, \tau_{n-1}), (\tau_{n-1}, \tau_{n-2}), \dots, (\tau_1, \tau_0)$ as above, or by treating the entire supply chain as a single unit and using the supply chain's payment triggers (τ_n, τ_0) . In other words, the timing of internal payments between stages does not directly affect the total supply chain's financial inventory costs (for any given ordering policies).

The framework introduced in this section assumes that stages are not financially constrained and does not explicitly model cash management decisions. However, differences in cash flow generated by inventory policies and varied payment timing arrangements still matter because of the time value of money.

16.3.4 Coordinating Supply Chains through Payment Timing Contracts

The payment timing framework offers potential tools to study incentives to coordinate a decentralized supply chain via payment timing contracts. This falls into the general topic of supply-chain finance.

Indeed, Tong et al. (2020) apply these tools to analyze the behavior of a two-stage decentralized inventory system under wholesale price payment timing contracts. Similar to Cachon and Zipkin (1999), they study competitive base-stock policies for each stage in a setting with no expediting so the stages cannot be decoupled. They enrich this setting by explicitly modeling how various wholesale price and payment timing arrangements between stages affect the financial holding and backorder costs assigned to each player due to the time value of money, and how those arrangements might be leveraged to achieve or approach system coordination.

Specifically, they consider a two-stage supply chain consisting of two independent decision makers – the supplier and the retailer – and inventory is managed using simultaneously-chosen local base-stock policies y_1 and y_2 as in Cachon and Zipkin (1999). For convenience, denote the wholesale prices $w_2 = c$, $w_1 = w$, $w_0 = p$.

16.3.4.1 Standard payment timing contracts

First, assume all payments are made to the upstream upon shipment, $\tau = (s_3, s_2, s_1)$, and refers to it as a “standard” wholesale price payment timing contract because it reflects payment timing that has historically been common in practice. This contract also has the natural property that each stage pays for its own physical inventory.

The authors show that, in a decentralized supply chain under a standard wholesale price and time contract with $w \in (c, p)$, $\tau = (s_3, s_2, s_1)$, the retailer chooses a base-stock level that is smaller than the centralized optimal level, while the supplier chooses a base-stock level that is larger than the centralized optimal level. These deviations from the centralized optimal levels are increasing in w .

16.3.4.2 Consignment payment timing contracts

Because the payment scheme $\tau = (s_3, s_2, s_1)$ leads to too little inventory at the retailer and too much inventory at the supplier, one may hypothesize that delaying the retailer’s payment to the supplier may help to alleviate the coordination problem because it should increase the supplier’s financial inventory costs and decrease the retailer’s financial inventory costs. Therefore, it is reasonable to consider the payment scheme $\tau = (s_3, s_1, s_1)$, under which the retailer need not pay the supplier the wholesale price until she also receives payment from the final customer. This is a form of consignment contract commonly seen in practice.

Tong et al.’s paper shows that, in a decentralized supply chain under a wholesale price with full consignment timing, $w \in (c, p)$, $\tau = (s_3, s_1, s_1)$, and an upper bound Y on base-stock levels, the retailer holds less inventory than in the centralized policy. This deviation from the centralized optimal inventory level is decreasing in w .

Thus, even though under consignment the retailer bears no financial inventory holding costs, the decentralized supply chain still does not maintain sufficiently high levels of inventory at the retailer.

16.3.4.3 Coordination via partial consignment payment timing contracts

Finally, Tong et al. (2016) show that a wholesale price with a partial consignment timing contract can achieve coordination. Under partial consignment, the retailer pays γw , for some $\gamma \in (0, 1)$ according to s_2 and $(1 - \gamma)w$ according to s_1 .

Specifically, it is shown that, in a decentralized supply chain under a wholesale price with a partial consignment timing contract, $\tau = \gamma(s_3, s_2, s_1) + (1 - \gamma)(s_3, s_1, s_1)$, and $\gamma = \frac{c(p - w)}{w(p - c)}$, the optimal centralized inventory level is a Nash equilibrium.

Thus, for any wholesale price, a single partial consignment parameter is able to simultaneously incentivize the retailer to increase its base-stock level and incentivize the supplier to decrease its base-stock level to *exactly* the right levels. Note that the proportion of up-front payment that achieves coordination is a decreasing function of the wholesale price. This is consistent with earlier results that the efficiency loss of standard timing is small for low wholesale prices, but the efficiency loss of consignment timing is small for high wholesale prices.

This coordination result is attractive because it implies that one can manipulate the profit-margin split by choosing a desired wholesale price while at the same time still achieving supply-chain coordination by choosing a coordinating level of partial consignment. The γ component of the contract can be used to achieve the centralized optimal inventory levels (increasing the size of the pie) while the w can be used to adjust the split in the profit margin. As long as one can overcome a firm's potential increase in the financial inventory costs in the optimal timing arrangement by giving them a larger share of the profit margin, Pareto improvement over standard or full consignment wholesale price contracts is possible.

16.3.5 The Role of 3PLs in Supply-Chain Finance

Some elements of the payment timing framework introduced in this section are applied by Chen et al. (2019) to study an innovative practice of some third-party logistics providers (3PLs) in the domain of supply-chain finance. That is, in addition to their traditional transportation services, these 3PLs provide procurement and financial assistance to buyers in the supply network, especially SMEs in developing countries that have limited access to bank loans. In particular, the 3PLs can often obtain payment delay arrangements from the financially stronger manufacturers via a special trade credit term (ranging from 30–60 days) and/or a letter of credit (typically 30 days). When the innovative 3PL delivers the products to a buyer, it collects both the purchase payment and logistics fee from the buyer. In this way, the buyers do not need to communicate directly with the manufacturer; the 3PL serves as the intermediary for both ordering and payments. In doing so, the 3PL can also partially extend favorable credit terms to the buyers.

Motivated by this practice, Chen et al. (2019) develop a game-theoretic model of a three-player supply chain, consisting of one manufacturer, one 3PL, and one buyer. They later extend the model to include multiple buyers. The buyer faces a single selling season with stochastic demand and a fixed market price. The supply chain is governed by wholesale price-only contracts for both procurement and shipping. They compare two scenarios. In the first, traditional scenario, the 3PL only ships the products from the manufacturer to the buyer. The buyer pays the manufacturer when ordering and the 3PL upon delivery, respectively. In the second, procurement service scenario, the 3PL takes orders from the buyer and procures the items from the manufacturer, allowing the buyer to pay both the purchasing cost and the logistics service fee when the order is delivered. Meanwhile, the 3PL transfers the order payment to the manufacturer after a grace period, according to their agreement. Within each model, with a given transportation time and payment grace period, the players decide the wholesale price, the shipping fee, and the order quantity, respectively, each to maximize its own expected profit. By comparing the equilibria of the two models, the authors characterize the conditions under which this innovation benefits all parties in the supply chain so that the new business model is sustainable.

In most decentralized supply-chain models, the role of a 3PL is absent; see Cachon (2003) and Özer (2011) for reviews of related literature. This is largely due to the fact that, traditionally, the only contribution of the 3PL is shipping. In order to study the new 3PL procurement

service practice, the models in Chen et al. (2019) explicitly capture the cash-flow dynamics by tracing the payment timing in relationship with the milestone events such as ordering, shipping, and shipment receipt.

In addition to identifying the Pareto optimality conditions, the authors also show that the supply-chain profit can be higher under leadership by the 3PL than by the manufacturer; the intermediary role of the 3PL is crucial; and the benefit is more likely to occur with more buyers.

16.4 RELATED LITERATURE

We now summarize other recent papers related to dynamic inventory and cash-flow control for interested readers. For the single-stage models, Buzacott and Zhang (2004) incorporate asset-based financing into production decisions. They demonstrate the importance of joint consideration of production and financing decisions for capital-constrained firms. Gupta and Wang (2009) consider a stochastic inventory system where the trade credit term is modeled as a non-decreasing holding cost rate according to an item's shelf age. Under the assumption that the full payment is made when the item is sold, they prove that a base-stock policy is optimal. Chao et al. (2008) consider a self-financed retailer who replenishes inventory in a finite horizon with iid demand. They study a lost-sales model and the available cash forms a hard constraint on the inventory order quantity. They show that a capital-dependent base-stock policy is optimal. Li et al. (2013) study a dynamic model in which inventory and financial decisions are made simultaneously in order to maximize the firm's value – the expected present value of dividends minus total capital subscriptions. Katehakis et al. (2016) analyze a dynamic model where a firm can finance its inventory with bank loans. On-hand cash earns deposit interest and short-term debt bears loan interest. Their objective is also to maximize the expected value of the firm's capital at the end of the planning horizon. Bendavid et al. (2017) consider a firm with two-level trade credit whose replenishment decisions are constrained by the working capital requirement. They conduct a simulation study under the base-stock policy subject to a working capital constraint.

For the multi-echelon models, Hu and Sobel (2007) analyze a serial inventory model with the objective of optimizing the expected present value of dividends. They show that an echelon base-stock policy is no longer optimal with financial constraints. Shang et al. (2009) provide a framework of supply-chain finance that demonstrates how inventory decisions can be coordinated between supply-chain partners through payment transfers in a serial system with fixed order costs. Protopappa-Sieke and Seifert (2010) conduct a simulation study on a two-stage supply chain to reveal qualitative insights on the allocation of working capital between the supply chain partners.

16.5 FUTURE RESEARCH

We would like to point out a few directions which have the potential for future research in the interface of operations and finance. The first one is the adoption of blockchain for enterprises. Blockchain promises to provide transparency and security for commercial activities many innovative smart contracts are created based on blockchain technology. Chod et al. (2020)

consider a firm that can signal its operational capabilities through either inventory transactions or loan requests to its lenders. The authors refer to the former as inventory signaling and the latter as cash signaling. There is a potential information distortion under cash signaling as the firm may create a loan more than needed for inventory investment. They show that signaling through inventory transactions is more efficient than cash signaling. Blockchain technology plays a key role in inventory signaling as it can provide inventory transaction information to the lender with minimal monitoring costs. One interesting research question is an inventory financing model based on blockchain technology. Many SMEs have difficulty financing from third-party financial institutions and they usually require a partnership with a large firm to secure bank financing under supply-chain finance schemes (e.g., purchase order financing, factoring, etc.). Now, with blockchain technology, financial institutions may be able to design a smart contract over time that continuously monitors SMEs' inventory or operational performance. What would be the smart contract form? How does a firm react to the smart contract for the inventory decision in a finite horizon? These are open questions that will significantly benefit SMEs.

The second potential research direction is crowdfunding, which can be an important source of financing for entrepreneurs who want to start a new business. However, because of information asymmetry, the backers (those who supply funds in the crowdfunding platform) are often subject to two risks: funds misappropriation and performance opacity. The former means the entrepreneur may run away with backers' money, whereas the latter means the product specification may be misrepresented. Belavina et al. (2020) consider this issue and show that each of these issues can impact crowdfunding efficiency. To mitigate these risks, they propose two mechanisms based on deferred payments. They show that early stopping (i.e., stopping the fundraising campaign once the goal is achieved) dominates escrow (i.e., escrowing any excess funds higher than the goal as insurance). In addition to the information asymmetry issue, inventory management under crowdfunding may have the potential for research. For example, how much capacity to reserve before the campaign starts? After all, the backers also expect to have faster delivery for their backed products. How to work with the suppliers? How to ration the products between the backers and regular consumers if the campaign is successful? What are the relationships between the pledged price, selling price, initial production quantity and the inventory policy if the product is successful? These are open questions that should be addressed in a finite-horizon model.

NOTES

1. It can be shown that the discount rate α is equal to $1/(1+r)$.
2. If lead time is positive, say L , the demand should be replaced with the total demand during lead time and the review period.

REFERENCES

- Axsäter, S. (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research*, 38(1), 64–69.
- Axsäter, S. (2003). Supply chain operations: Serial and distribution inventory systems. In S.C. Graves, and A.G. de Kok (Eds.) *Handbooks in operations research and management science* (Vol. 11, pp. 525–559).

- Axsäter, S., & Rosling, K. (1993). [Notes]: Installation vs. echelon stock policies for multilevel inventory control. *Management Science*, 39(10), 1274–1280.
- Babich, V., & Kouvelis, P. (2018). Introduction to the special issue on research at the interface of finance, operations, and risk management (iFORM): Recent contributions and future directions.
- Baye, M. R., & Prince, J. (2014). *Study guide for managerial economics and business strategy*. McGraw Hill.
- Belavina, E., Marinesi, S., & Tsoukalas, G. (2020). Rethinking crowdfunding platform design: Mechanisms to deter misconduct and improve efficiency. *Management Science*, 66(11), 4980–4997.
- Bendavid, I., Herer, Y. T., & Yücesan, E. (2017). Inventory management under working capital constraints. *Journal of Simulation*, 11(1), 62–74.
- Buzacott, J. A., & Zhang, R. Q. (2004). Inventory management with asset-based financing. *Management Science*, 50(9), 1274–1292.
- Chao, X., Chen, J., & Wang, S. (2008). Dynamic inventory management with cash flow constraints. *Naval Research Logistics*, 55(8), 758–768.
- Cachon, G. P. (2003). Supply chain coordination with contracts. *Handbooks in Operations Research and Management Science*, 11, 227–339.
- Cachon, G. P., & Zipkin, P. H. (1999). Competitive and cooperative inventory policies in a two-stage supply chain. *Management Science*, 45(7), 936–953.
- Chen, X., Cai, G., & Song, J.-S. (2019). The cash flow advantages of 3PLs as supply chain orchestrators. *Manufacturing and Service Operations Management*, 21(2), 435–451.
- Chod, J., Trichakis, N., Tsoukalas, G., Aspegren, H., & Weber, M. (2020). On the financing benefits of supply chain transparency and blockchain adoption. *Management Science*, 66(10), 4378–4396.
- Clark, A., & Scarf, K. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- Gupta, D., & Wang, L. (2009). A stochastic inventory model with trade credit. *Manufacturing and Service Operations Management*, 11(1), 4–18.
- Hu, Q. J., & Sobel, M. J. (2007). Echelon base-stock policies are financially sub-optimal. *Operations Research Letters*, 35(5), 561–566.
- Katehakis, M. N., Melamed, B., & Shi, J. (2016). Cash-flow based dynamic inventory management. *Production and Operations Management*, 25(9), 1558–1575.
- Li, L., Shubik, M., & Sobel, M. J. (2013). Control of dividends, capital subscriptions, and physical inventories. *Management Science*, 59(5), 1107–1124.
- Luo, W., & Shang, K. (2015). Joint inventory and cash management for multi-divisional supply chains. *Operations Research*, 63(5), 1098–1116.
- Luo, W., & Shang, K. H. (2019). Managing inventory for firms with trade credit and deficit penalty. *Operations Research*, 67(2), 468–478.
- Modigliani, F., & Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, 48(3), 261–297.
- Muharremoglu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, 56(5), 1089–1103.
- Özer, Ö. (2011). Inventory management: Information, coordination, and rationality. In Karl G. Kempf, Pinar Keskinocak, and Reha Uzsoy (Eds.) *Planning production and inventories in the extended enterprise* (pp. 321–365). Springer.
- Protopappa-Sieke, M., & Seifert, R. W. (2010). Interrelating operational and financial performance measurements in inventory control. *European Journal of Operational Research*, 204(3), 439–448.
- Shang, K. H., & Song, J.-S. (2003). Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science*, 49(5), 618–638.
- Shang, K. H., Song, J.-S., & Zipkin, P. H. (2009). Coordination mechanisms in decentralized serial inventory systems with batch ordering. *Management Science*, 55(4), 685–695.
- Tong, J., DeCroix, G., & Song, J.-S. (2020). Modeling payment timing in multiechelon inventory systems with applications to supply chain coordination. *Manufacturing and Service Operations Management*, 22(2), 346–363.
- Zipkin, P. (2000). *Foundations of inventory management*. McGraw Hill.

17. Behavioral inventory management

Andrew M. Davis and Jordan D. Tong

17.1 INTRODUCTION

While the majority of inventory-management research consists of prescriptive models that derive optimal policies, in the past few decades researchers have developed a stream of research that leverages ideas and methodologies commonly used in behavioral economics and psychology to make different kinds of contributions to inventory management. This stream, which we will refer to as “behavioral inventory management,” places emphasis on describing human decision behavior – explicitly acknowledging how it may deviate from optimal policies.

Given the rise of digitization and artificial intelligence, a natural first question is: is inventory management still influenced by human decision-making? Our answer is a resounding “yes” – but with the qualification that there is significant heterogeneity across organizations, and that the role which humans play is evolving (see Section 17.4.3). For example, Zhao et al. (2021) conduct interviews with supply-chain managers and administer a survey asking how inventory decisions are made. Out of 54 respondents, none indicated that inventory decisions are made without any human involvement, but 35 suggested that they rely on software systems that are adjusted by human judgment.

Behavioral inventory-management research is not inherently at odds with more traditional prescriptive research methodologies. Rather, we believe that the two approaches are complementary (e.g., see Section 17.3.2), and that the development of a stream of behavioral research is a natural component of the evolution of the field. The role that behavioral economics has played in the field of economics provides an insightful perspective (see Thaler & Ganser, 2015). Initially, it received significant resistance, but over time it improved the field’s impact and validity. Similar to behavioral economics, behavioral operations initially focused on documenting violations of “standard theory” in laboratory settings and congregated around a select few problems – and has received criticisms for doing so. However, as behavioral economics matured, researchers began to develop more useful behavioral models and identify more compelling field evidence – which ultimately led to a significant impact on practice. We believe there is potential for behavioral inventory-management research to contribute in a similar fashion, and that the sub-field is well on its way in this regard.

This chapter continues in Section 17.2 with a review of behavioral research by inventory-model setting. We start by describing two early streams of laboratory experiments that helped build momentum in behavioral inventory-management research – the newsvendor problem (Section 17.2.1) and inventory control in a serial supply chain (Section 17.2.2). Because they serve as a type of “foundation” for the sub-field, we describe a few of these influential papers in more detail. In Section 17.2.3, we briefly survey papers in alternative inventory settings to provide a feel for the breadth of research in the sub-field of behavioral inventory management.

Whereas Section 17.2 reviews past work – organized by inventory setting – Section 17.3 differs by discussing how behavioral science can advance inventory research beyond documenting deviations from rationality – organized by pathways to improve inventory performance. Section 17.3.1 provides examples of ways to improve via the changing of human decision behavior. Section 17.3.2 highlights ways to improve by responding optimally to others' suboptimal behaviors. Section 17.3.3 considers improvements by designing processes and systems.

Finally, we conclude in Section 17.4 with a high-level discussion of future research opportunities by (a) going beyond newsvendor and beer game settings (as discussed in Section 17.2), (b) going beyond documenting deviations from the rational benchmark as discussed in Section 17.3, (c) going beyond directly studying inventory decisions by studying other decisions humans make that impact inventory management, and (d) going beyond the laboratory.

17.2 REVIEW OF BEHAVIORAL INVENTORY RESEARCH BY MODEL SETTING

While behavioral inventory research spans a wide range of methodologies – laboratory experiments, field experiments, empirical studies, and behavioral models – its early development primarily involved laboratory experiments. Therefore, we focus on these papers in this section. While the section mentions many papers, our primary objective is not to be “comprehensive” but rather to provide some depth in two key sub-streams, along with some breadth across different settings.

17.2.1 The Newsvendor Problem

There is a vast literature within behavioral operations management that investigates inventory orders in a “one-shot” environment when facing random demand, referred to as the newsvendor problem. To provide a theoretical overview of the newsvendor problem more formally, we follow Schweitzer and Cachon (2000). The task involves setting an inventory order, z , prior to observing realized demand. Let D represent this single-period realized stochastic demand, F the cumulative distribution function of demand, and f the density function of demand (D is continuous, differentiable, and strictly increasing). The decision-maker knows demand distribution F , sells each unit at a price p , and purchases each unit at cost c , where $p > c$. If $z > D$ then each leftover unit is worth a per unit salvage value $s > c$. Let the decision-maker's realized profit be given by:

$$\pi(z, D) = (p - s) \min(z, D) - (c - s)z,$$

and the expected profit be given by:

$$E[\pi(z, D)] = (1 - F(z)) \pi(z, z) + \int_0^z f(w) \pi(z, w) dw.$$

Let $z^* = \arg \max E[\pi(z, D)]$. The solution to the newsvendor problem, often referred to as the “critical fractile” is given by:

$$F(z^*) = \frac{p - c}{p - s}.$$

In this chapter, we will refer to the product as being “high profit” when the critical fractile is greater than 1/2 (i.e., the optimal inventory order is above the mean for a symmetric demand distribution), specifically:

$$\frac{1}{2} < \frac{p - c}{p - s},$$

and “low profit” when it is less than 1/2 (i.e., the optimal inventory order is below the mean for a symmetric demand distribution). Last, as is common in many behavioral newsvendor experiments, without loss of generality, we assume that $s = 0$.

Given that human decision-makers, who may be responsible for making newsvendor inventory decisions in practice, are susceptible to behavioral biases (e.g., Davis, 2019), many behavioral inventory papers involve controlled human-subject experiments and investigate how newsvendor order decisions are *actually* made relative to how they *should* be made. One important result from this literature is that human decision-makers set inventory orders that exhibit a “pull-to-center” (PTC) effect: average inventory orders are set between the normative theoretical prediction, outlined above, and the mean of the demand distribution. In this subsection, we provide a brief history of the experimental newsvendor literature by focusing on three papers which study the PTC effect: one which identifies the effect, another that tests the robustness of the effect, and a third which investigates the external validity of the effect. For those interested in comprehensive survey articles of the behavioral literature on the newsvendor problem, please see Becker-Peth and Thonemann (2019), Zhang and Siems (2019), and Perera et al. (2020).

17.2.1.1 Identifying the PTC effect

The seminal paper on behavioral inventory management on the newsvendor problem is by Schweitzer and Cachon (2000).¹ They recruited 34 MBA students to partake in a newsvendor experiment which consisted of 30 rounds of consecutive inventory order decisions. Consistent with the standard newsvendor setting, participants were always shown the selling price per unit, the cost per unit, and the demand distribution (the salvage value per unit was 0). While the selling price was always 12 “francs” per unit, the cost was 3 francs per unit in the “high-profit” condition and 9 francs per unit in the “low-profit” condition. Participants were tasked with making 15 inventory-order decisions in each of these two manipulations. The demand distribution in each round followed a discrete uniform distribution between 1 and 300, resulting in an expected-profit-maximizing inventory order of 225 in the high-profit condition and 75 in the low-profit condition.

Before making a decision in each round, participants were provided with the ability to “test” different inventory orders. For instance, for a test inventory order, participants could see the distribution of profits, the probability that demand will be higher/lower than the order, the break-even demand level, and more. By providing such decision support, Schweitzer and Cachon minimize the likelihood that bounded rationality (i.e., that participants simply could not understand the problem) is the primary driver of any observed outcomes. After making an order decision in a round, participants would then see the actual demand and their realized

profit. Last, to ensure that participants took the task seriously, Schweitzer and Cachon paid one randomly-selected participant cash which was proportional to their profit in the game.

After averaging the inventory orders for each participant in each condition, and conducting a series of hypothesis tests, Schweitzer and Cachon found evidence of the PTC effect. In the high-profit condition, the average observed inventory order was 176.68, significantly lower than the normative prediction of 225, but higher than the mean demand of 150. Conversely, in the low-profit condition, the average observed inventory order was 134.06, significantly higher than the normative prediction of 75, but lower than the mean demand of 150. Figure 17.1 depicts the observed demand draws, the optimal inventory order, and the average observed inventory order, by round. As one can see, the PTC effect was not only present in the first round but also persists over time.²

In an effort to understand what drives the PTC effect, Schweitzer and Cachon investigate alternative behavioral models. For brevity, we note that Schweitzer and Cachon's experiment allowed for the possibility of losses to occur (e.g., a high inventory order and low demand realization). Thus, "prospect theory" (Kahneman & Tversky, 1979), which predicts that human decision-makers are risk averse in the domain of gains and risk seeking in the domain of losses, could account for the PTC effect. To formally test this conjecture, they conducted a second experiment which included a uniform demand distribution from 901 to 1200, such that losses could not occur. Yet, the results from this experiment demonstrated that the PTC effect continued to persist. Ultimately, after evaluating a number of utility-maximizing models and decision heuristics, Schweitzer and Cachon concluded that two plausible explanations for the PTC effect were (1) a mean-anchoring (and insufficient adjustment) bias and (2) a preference for minimizing *ex-post* inventory errors.

To provide details of how one can incorporate such behavioral biases into the standard newsvendor theory, we begin with the mean-anchoring bias. A mean-anchoring bias posits that a decision-maker, rather than making an inventory order that corresponds to the normative theoretical prediction, instead chooses a behavioral inventory order, z_m , given by the following equation:

$$z_m = (1 - \eta)F^{-1}\left(\frac{p-c}{p}\right) + \eta\lambda, \quad (17.1)$$

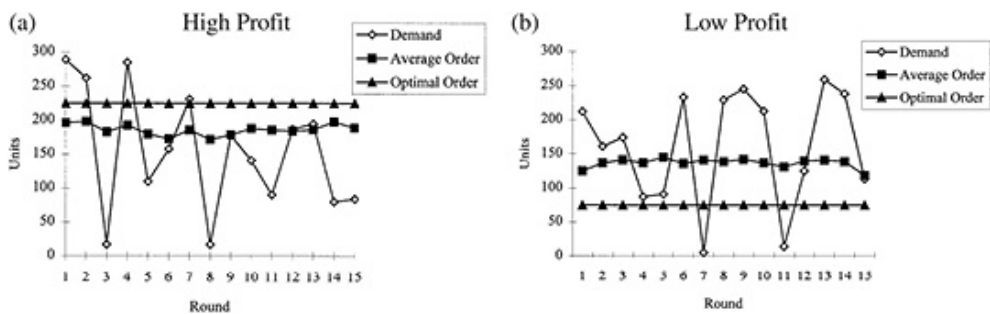


Figure 17.1 Inventory order results in Schweitzer and Cachon (2000)

where λ represents the mean of the demand distribution and $\eta \in [0,1]$ represents the degree of the mean-anchoring bias. As $\eta \rightarrow 0$ the decision-maker sets an inventory order closer to optimal z^* and when $\eta \rightarrow 1$ they set an inventory order closer to the mean of the demand distribution λ . Hence, it can capture the PTC effect.

Now consider the preference of minimizing *ex-post* inventory errors. Intuitively, this means that a decision-maker cares about reducing the absolute deviation between their inventory order and realized demand. Thus, they attempt to maximize an expected utility function, $E[u(z, D)]$, comprised of their standard expected newsvendor profit along with a psychological disutility term that is increasing in the degree of *ex-post* inventory errors (we assume the decision-maker is risk neutral):

$$E[u(z, D)] = E[\pi(z, D)] - \int_0^\infty f(w) \delta(|z - w|) dw, \quad (17.2)$$

where $\delta(\cdot)$ captures the disutility from *ex-post* inventory errors ($\delta' > 0$ and $\delta(0) = 0$). Intuitively, if there is a minimal degree of *ex-post* inventory error bias, a decision-maker sets an inventory order closer to the optimal value z^* . But, if there is a significant degree of *ex-post* inventory error bias, then the decision-maker sets an inventory order closer to λ . More formally, let z_e^* represent the inventory order that maximizes Equation (17.2). Schweitzer and Cachon (2000) prove that, under mild conditions, for a high-profit product, $\lambda \leq z_e^* \leq z^*$. Conversely, for a low-profit product, $\lambda > z_e^* > z^*$.

For those interested in more details about behavioral biases that can capture the PTC effect, and how they are incorporated into the newsvendor model, please see Becker-Peth and Thonemann (2019).

17.2.1.2 Evaluating the robustness of the “pull-to-center” effect

After Schweitzer and Cachon (2000), additional laboratory experiments aimed to determine the robustness of the PTC effect in the newsvendor problem. For example, Benzion et al. (2008) found that the PTC effect persists under alternative demand distributions and Lurie and Swaminathan (2009) observed that more frequent feedback can actually lead to worse inventory decisions. Another noteworthy paper that evaluates the robustness of the PTC effect is by Bolton and Katok (2008). They conducted a series of experiments which investigated newsvendor decisions with many rounds of decisions, restricted decision spaces, detailed feedback, and augmented decision support. Here we provide a summary of their work.

One of the criticisms of Schweitzer and Cachon’s experiments is that participants were not given an opportunity to learn. Specifically, participants made 15 decisions under the high-profit condition and 15 decisions under the low-profit condition. Bolton and Katok (2008) remedy this by including 100 inventory decisions, with a uniform demand distribution. Further, they employed a between-subjects design: one group of participants made 100 decisions in the high-profit treatment (referred to as “high-safety stock”) and a different group of participants made 100 decisions in the low-profit treatment (referred to as “low-safety stock”). A benefit of this between-subjects approach is that there is no need to determine whether “order effects” are influencing outcomes (i.e., allowing a participant to play one manipulation before another).

For their first experiment, Bolton and Katok noted that the expected-profit function of the newsvendor problem is relatively flat around the optimal order, which may also impede learning. To provide a direct test of this hypothesis, they contrasted their baseline experimental

treatment with two other variants which restricted the decision space. In one, they restricted the decision space to nine potential inventory orders (versus 100 in the baseline treatment). In another, they restricted the decision space to three inventory orders. Therefore, their initial experiment compared three treatments to one another, all of which included 100 decision periods, but differed in the number of possible inventory decision orders: “100-option,” “9-option,” and “3-option.”

Bolton and Katok (2008) find supporting evidence of the PTC effect in all three experimental treatments. Regarding the effect of learning, they find that participants do make better inventory orders over time, but that the effect is small: average orders increase by only 0.126 units per round in the high-safety stock condition and decreased by only 0.038 units per round in the low-safety stock condition. Even in the final ten rounds of decisions (91–100), inventory orders are weakly significantly different from the optimal order in the high-safety stock condition (p -value = 0.089) and significantly different from the optimal order in the low-safety stock condition (p -value = 0.002).

Bolton and Katok also calculate the proportion of maximum expected profit that is achieved by the participants’ decisions. This information is depicted in Figure 17.2. What is noteworthy in this figure is that all three treatments are similar, within the high-safety stock and low-safety stock conditions. In fact, within each safety-stock condition, there are no significant differences across treatments (and no significant differences in the last ten rounds of decisions).

Bolton and Katok (2008) proceed by investigating whether the PTC effect exists when additional information about foregone options is provided. Specifically, they designed an experimental treatment which showed participants their profits for a chosen inventory order and foregone profits for inventory orders not chosen, referred to as “FORE.” They also designed a variant where they showed this same information based on ten-round moving averages, referred to as “MAVG.” To minimize complexity in these new experimental treatments, they used the same inventory-order decision space as the previous three-option treatment (i.e., three-option is a baseline for FORE and MAVG). Turning to results, Bolton and Katok once again find similar order decisions and a lack of differences across the three treatments.

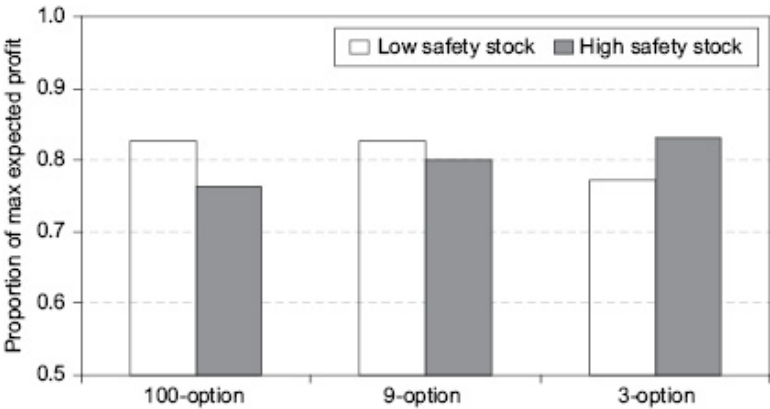


Figure 17.2 Proportion of maximum expected profit in Bolton and Katok’s (2008) 100-Option, 9-Option, and 3-Option Treatments

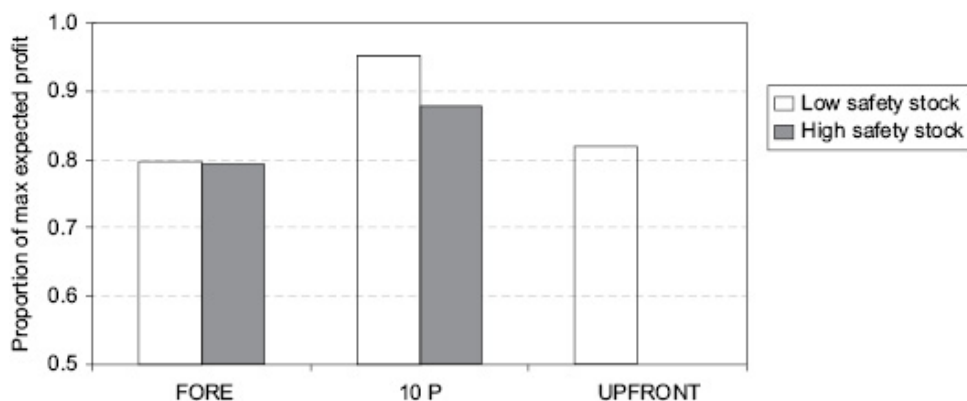
After including 100 rounds of decisions, restricting decision spaces, and providing information about foregone options, Bolton and Katok went even further in their last set of experiments. Specifically, they ran two final treatments which built on the previous FORE treatment (which, recall, is based on the three-option treatment). In one, they forced participants' inventory-order decisions to remain fixed for ten consecutive decision rounds. In this treatment, referred to as "10P," participants made 100 rounds of inventory-order decisions, but they effectively participated in 1000 demand draws. In the other treatment, they relaxed this ten-period aspect, but provided extensive decision support by detailing both actual profits (for different values of demand) and expected profits, for each of the three potential inventory orders, referred to as "UPFRONT."

The proportion of maximum expected profit observed in the FORE (which is a baseline for these final treatments), 10P, and UPFRONT treatments is illustrated in Figure 17.3. One can discern that the 10P manipulation leads to a higher expected profit over the FORE treatment in both the high-safety and low-safety stock conditions (p -value=0.021 and p -value=0.001) and over the UPFRONT treatment in the low-safety stock condition (p -value=0.015; they did not run high-safety stock in UPFRONT).

Overall, Bolton and Katok's (2008) work demonstrates the robustness of the PTC effect across a variety of scenarios. The only treatment which resulted in a marked improvement in newsvendor profits was their 10P treatment, which included (1) 100 rounds of decisions, (2) a three-option restricted inventory-order decision space, (3) profit information about foregone options, and (4) fixed orders such that a decision-maker participated in 10 demand draws for each order.

17.2.1.3 Investigating the external validity of the "pull-to-center" effect

A common objection to laboratory experiments regards the use of student participants. Despite a number of studies in experimental economics demonstrating that student participants make decisions that are similar to managers (Plott, 1987; Cooper et al., 1999), it is possible that the



Note: UPFRONT was not run in the high-safety stock condition.

Figure 17.3 Proportion of maximum expected profit in Bolton and Katok's (2008) FORE, 10P, and UPFRONT treatments

PTC effect may not be present among individuals that have extensive training in inventory decisions and/or are experienced inventory managers.

Bolton et al. (2012) directly examine this hypothesis through a 2x3 between-subjects experiment. The first factor that they manipulated was the amount of training provided before making decisions. In one variant they provided basic game instructions, similar to existing newsvendor experiments (referred to as “Basic”). In the other variant, participants watched a 60-minute video that explained the optimal order quantity calculation in the newsvendor (referred to as “Trained”). The video also explained that decision-makers have a tendency to set orders erroneously toward mean demand.

In the second factor of their 2x3 experimental design, Bolton et al. considered three types of participants: freshmen business students, master’s degree students who have taken at least one course in operations management and completed their undergraduate studies in business administration, and managers who have at least one year of experience making newsvendor-type decisions in practice.

All six treatments focused on a high-profit product and included 100 rounds of inventory-order decisions. However, these 100 rounds were divided into three phases. In the first phase, rounds 1–40, participants were shown 50 draws of customer demand. In the second phase, rounds 41–80, participants were told that demand follows a discrete uniform distribution between 1 and 100. In the third phase, rounds 81–100, participants were shown a graph that depicted the expected profit with respect to the inventory order, including the optimal inventory order (75 in all treatments).

The average orders for all six treatments are shown in Figure 17.4. Beginning with the Basic condition (solid lines), one can discern a clear PTC effect for freshmen, master’s students, and managers, across all three phases of the experiment. Further, while the PTC effect is present in all phases, it diminishes during the third phase (i.e., when participants are shown a graph with the optimal order). Comparing the three groups of participants to one another in the Basic condition, there are no significant differences in the first and second phases, but in the third phase, managers exhibit the strongest PTC effect – to the extent that master’s students make significantly better inventory-order decisions than managers (p -value=0.016). Turning

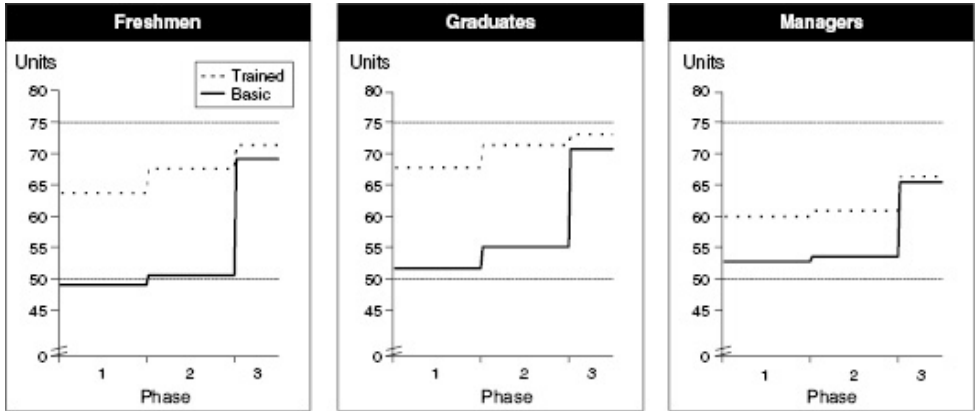


Figure 17.4 Average orders by treatment and phase in Bolton et al. (2012)

to the “Trained” condition (dashed lines), all three groups make better decisions (i.e., less PTC effect) compared to the “Basic” condition in the first and second phases. Also, in the first two phases of the “Trained” condition, master’s students make significantly better inventory decisions than managers. In summary, Bolton et al. (2012) show that students and managers both exhibit a PTC effect and that on-the-spot training improves the decisions of students and managers to a similar degree.

17.2.2 Serial Supply Chain

The history of behavioral inventory-management research in the serial supply-chain setting has been dominated by experimental papers that examine the bullwhip effect: “the phenomenon where orders to the supplier tend to have a larger variance than sales to the buyer (i.e., demand distortion), and the distortion propagates upstream in an amplified form (i.e., variance amplification)” (Lee et al., 1997, p. 546). Sterman (1989), who documented the bullwhip effect using the popular classroom exercise called the “beer game,” is widely considered one of the seminal papers in the behavioral operations management sub-field. Therefore, we begin by describing its experimental setting and summarizing its main results.

17.2.2.1 Experimental examination of the bullwhip effect

Sterman (1989) considers a four-stage serial supply chain, where each human participant manages the inventory at their assigned stage $j = 1, 2, 3, 4$ using periodic review. In every period t , each player decides an order quantity z_t^j to place to the upstream. Exogenous demand D_t occurs at stage 1, and stage 4 procures from an outside supplier with ample supply. There is a four-period lead time at stages 1–3 ($L_j = 4$, $j = 1, 2, 3$) – it takes two periods for orders to reach the upstream, and two periods to deliver shipments from the upstream. There is a three-period lead time at stage 4, ($L_4 = 3$). At each of the four stages, the local holding cost is $h = 0.5$ per unit/period and the local backorder cost is $b = 1$ per unit/period. The total supply-chain cost through period T is

$$C(T) = \sum_{j=1}^4 \sum_{t=1}^T [h \max\{I_t^j, 0\} - b \min\{I_t^j, 0\}]$$

where I_t^j denotes the net inventory (inventory on-hand minus backorders) after shipping and receiving.

The events in each period are as follows: (1) shipments arrive from upstream, (2) new orders arrive from downstream, (3) orders are filled or backlogged if necessary, and (4) new orders are placed. To describe the inventory dynamics, we first define S_t^j as the quantity shipped from j to its downstream in period t as:

$$S_t^j = \begin{cases} \min\{D_t, \max\{I_{t-1}^j + S_{t-2}^{j+1}, 0\}\} & \text{for } j = 1 \\ \min\{z_{t-2}^{j-1}, \max\{I_{t-1}^j + S_{t-2}^{j+1}, 0\}\} & \text{for } j = 2, 3 \\ \min\{z_{t-2}^{j-1}, \max\{I_{t-1}^j + z_{t-3}^j, 0\}\} & \text{for } j = 4. \end{cases}$$

So, the inventory updates as follows:

$$I_t^j = \begin{cases} I_{t-1}^j + S_{t-2}^{j+1} - D_t & \text{for } j = 1 \\ I_{t-1}^j + S_{t-2}^{j+1} - z_{t-2}^{j-1} & \text{for } j = 2, 3 \\ I_{t-1}^j + z_{t-3}^{j+1} - z_{t-2}^{j-1} & \text{for } j = 4. \end{cases}$$

Demand is non-stationary and participants were not informed how demand would be generated. Demand for all teams was four units in periods 1–4, then eight units in periods 5 or later. The game was initialized in period 0 with 12 units of inventory on-hand at each stage, $4 \times L_j$ units of inventory on-order, and no backorders. This initialization is “in equilibrium” in the sense that nothing would change if demand were constant at 4 each period and each participant ordered four in each period. Participants knew these initial states, but throughout the game, a participant at each stage could not observe the inventory nor order decision information at other stages.

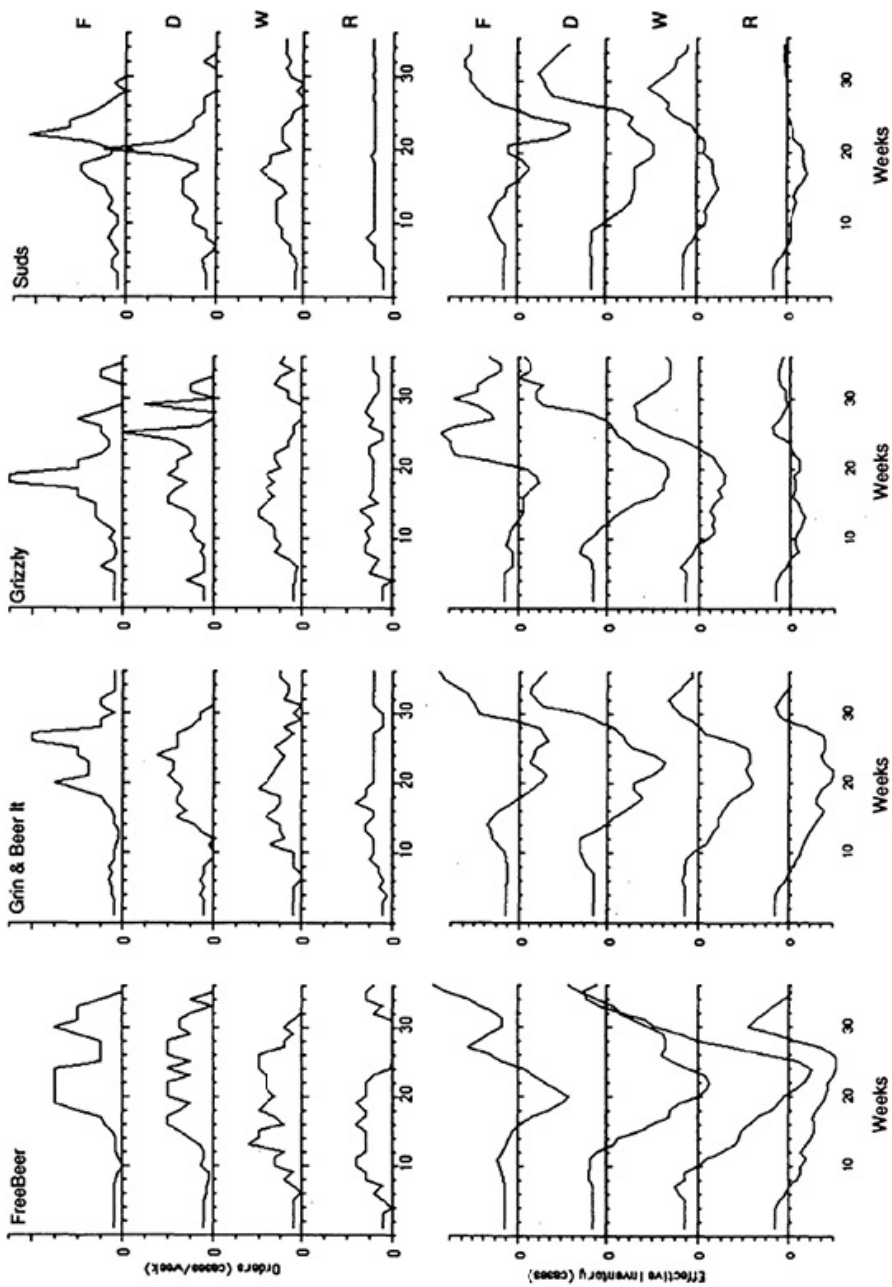
Although order decisions are decentralized, the incentive structure of the experiment is based on the entire supply chain’s total inventory holding and backorder costs over the horizon $T=36$. (To avoid end-of-game effects, subjects are told the game will last 50 periods.) Sterman (1989) conducted the experiment with several supply-chain teams playing simultaneously, and implemented a winner-take-all monetary reward for the team with the lowest total costs. Sterman’s original sample included 11 teams of “undergraduate, MBA, and Ph.D. students at MIT’s Sloan School of Management, executives from a variety of firms participating in short courses on computer simulation, and senior executives of a major computer firm” (p. 328).

Figure 17.5, from Sterman (1989), illustrates typical results by showing the order decisions and inventory levels (on-hand inventory minus backorders) for four supply chains. Among the several metrics that Sterman reports, perhaps the most straightforward measure of the bullwhip effect is the order variance at each stage. While the variance of the demand rate is only $1.6 \text{ (cases/week)}^2$, the average variance of the order rates are 13, 23, 45, and 72 for $j = 1, 2, 3, 4$, respectively. These averages reflect a 77% variance amplification between stages 1 and 2, 96% amplification between 2 and 3, and 60% amplification between stages 3 and 4 (where recall that $L_4=3$, while $L_1, L_2, L_3 = 4$).

Sterman’s primary explanation for the bullwhip effect is a cognitive one. He argues that participants under-account for the on-order inventory, a phenomenon referred to as *supply-line under-weighting*. For example, if a participant under-weights a full supply line, they will underestimate their inventory position and may repeatedly place too large of orders (until these units arrive and cause excessive on-hand inventory). As evidence of this supply-line under-weighting, he estimates the following model for a participant j ’s order quantity decision in period t :

$$z_t^j = \max\{0, f_t^j + \alpha(\bar{s} - I_t^j - \beta IO_t^j) + \varepsilon\}$$

where f_t^j is the participant’s demand forecast (assumed to update according to exponential smoothing), $\bar{s}$ serves a similar function as a base-stock level (discussed further below), IO_t^j is the current outstanding on-order inventory, ε is assumed to be Gaussian noise, and α, β



Note: Top rows are orders; bottom rows are inventory (from bottom to top, retailer, wholesaler, distributor, factory). Tick marks denote ten units.

Figure 17.5 Experimental results for four typical chains in Sterman (1989)

are constants. Thus, there are four parameters to be estimated: $\bar{s}, \alpha, \beta$, and an exponential forecasting smoothing parameter embedded in f_t^j . The parameter associated with supply-line under-weighting is β . Sterman estimates the average value of β across participants to be 0.34. In other words, on average, participants reduce their order quantity for on-order inventory by only about 1/3 as much as they do for on-hand inventory.

17.2.2.2 Isolating the “behavioral bullwhip effect”

Lee et al.’s (1997) paper on the bullwhip effect, though not an experimental paper, influenced experimental papers that followed. In it, Lee et al. examine several *rational* causes of the bullwhip effect – including demand signal processing, inventory rationing, order batching, and price variations – using simple mathematical models that assume strategic and cost-minimizing agents. Though it is reasonable to conclude that Sterman’s results are primarily driven by behavioral drivers and not rational ones cited by Lee et al. (1997), it is not possible to cleanly tease the two apart in Sterman’s experiment.

Croson and Donohue (2006) modified Sterman’s experimental design to help address this issue by eliminating any rational, or what they call “operational,” causes of the bullwhip effect. The key difference was to make the demand process a known stationary distribution (in this case, a uniformly-distributed random integer between 0 and 8, inclusive). Doing so eliminates demand signal processing as a plausible rational driver of the bullwhip effect, and also allows for a clear optimal policy to be established. Namely, a local base-stock policy is optimal under a known stationary demand (Chen, 1999) and there is no bullwhip effect with rational agents. However, if demand is unknown or non-stationary, even fully rational agents will exhibit a bullwhip effect (Lee et al., 1997; Graves, 1999). Croson and Donohue (2006) also implemented their experimental design with a computer-based system, thereby eliminating accounting errors, and implemented a continuous compensation scheme, which incentivizes performance evenly for losing teams.

Figure 17.6 shows the variance amplification across the 11 groups tested in Study 1 of Croson and Donohue (2006). Similar to Sterman (1989), the average of these variances reflects a 73% variance amplification between stages 1 and 2, 111% amplification between 2 and 3, and 48% amplification between stages 3 and 4. Thus, it demonstrates a “behavioral bullwhip effect” isolated from any “operational” causes of the bullwhip effect.

To test for supply-line under-weighting, Croson and Donohue (2006) estimate the following regression:

$$z_t^j = \max\{0, \alpha_0 + [\alpha_I I_{t-1}^j + \alpha_S S_t^j + \alpha_R R_t^j] + \alpha_{IO} IO_t^j + \alpha_t + \varepsilon\}$$

where I_{t-1}^j is the on-hand inventory of the previous period, S_t^j is the shipment received from the upstream, R_t^j is the order received from the downstream, and IO_t^j is the on-order inventory. Thus, this model captures a base-stock policy with base-stock level α_0 if $\alpha_I = \alpha_S = \alpha_{IO} = -1$, $\alpha_R = 1$, and $\alpha_t = \varepsilon = 0$. With this approach, Croson and Donohue test for supply-line under-weighting by comparing the coefficient for on-hand inventory (−0.23) with the coefficient for on-order inventory (−0.03). This difference implies that participants did not reduce their orders nearly as much to account for supply-line inventory relative to on-hand inventory. In other words, supply-line under-weighting exists even with known stationary demand.

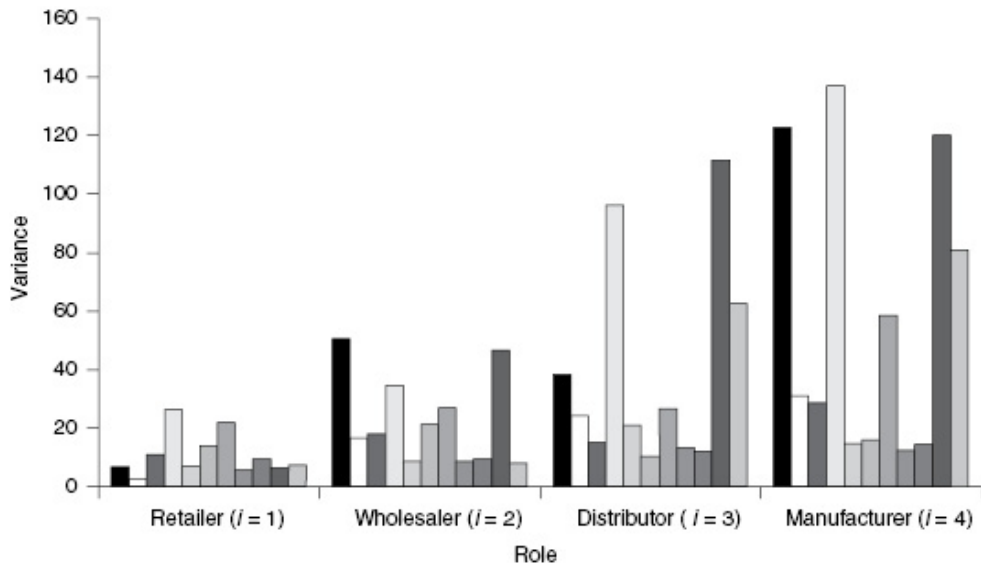


Figure 17.6 Variance of orders in Croson and Donohue (2006)

17.2.2.3 Behavioral bullwhip moderating factors and mitigation strategies

Beyond the identification and isolation of the “behavioral bullwhip” effect, researchers have investigated several environmental factors and interventions that may mitigate or exacerbate it (see Narayanan & Moritz, 2015; Chen & Wu, 2019 for reviews). For example, researchers have found evidence that sharing real-time on-hand inventory information (Croson & Donohue, 2006) or sharing real-time point-of-sale information (Croson & Donohue, 2003) partially mitigates the bullwhip effect when demand follows a known stationary random process (i.e., when there are no rational causes of the bullwhip effect). However, Steckel et al. (2004) find that point-of-sale sharing may actually increase supply-chain costs when demand follows certain demand patterns (under which rational causes of the bullwhip effect are not eliminated). Wu and Katok (2006) show that hands-on experience plus an opportunity to discuss collaboratively with supply-chain partners can lead to significant improvements. Narayanan and Moritz (2015) find that people who tend to score higher on a “cognitive reflection test” display less supply-line-under-weighting bias.

While supply-line under-weighting is the most well-known driver of the behavioral bullwhip effect, human behavior is more complex than this single bias. Croson et al. (2014) argue that beyond supply-line under-weighting, another driver of the behavioral bullwhip effect is coordination risk: people strategically hedge against the uncertain behavior of their supply-chain partners. Thus, they suggest injecting some “coordination stock” into the system to buffer against this internal coordination risk. Most recently, Oroojlooyjadid et al. (2021) use deep reinforcement learning to optimize the inventory decisions of one player, given the complex human decision behavior of supply-chain partners. In general, they suggest that deep reinforcement learning may be a fruitful way to study how a player should behave optimally given the complex and irrational human decision-making behavior of others in the system.

17.2.3 Other Inventory Settings and Extensions

While much of the behavioral operations literature is devoted to investigating inventory-order decisions in a newsvendor or serial supply-chain setting, more recent experiments have examined other settings. Here we briefly summarize a selection of these papers.

17.2.3.1 Economic order quantity and review periods

Despite the use of the economic order quantity (EOQ) model in practice, few papers have examined it from a behavioral standpoint. One paper that does investigate an issue related to the EOQ model setting is Stangl and Thonemann (2017). They present participants with one of two inventory metrics, days-of-supply versus inventory-turnover rate. Participants must then determine the total ordering costs of three products, each with different holding costs, in an EOQ environment. Despite the two inventory metrics containing the same information, Stangl and Thonemann observe that the inventory-turnover metric leads to 60–80% higher estimated ordering costs (i.e., lower quantities) compared to the days-of-supply metric. Another paper that studies an EOQ-related topic is Katok et al. (2008), who examine the role of review periods on base-stock level decisions. They consider a service-level agreement setting with random demand, where a bonus is awarded for satisfying a target fill rate (total units sold/total demand) aggregated over a certain number of periods T . They find that a larger T leads to higher base-stock levels.

17.2.3.2 Censored-demand/unobservable lost sales

When the demand is censored by the inventory level, the prescriptive analytical research has focused on demand estimation (e.g., see Nahmias, 1994) and the joint demand-learning and inventory ordering problem (e.g., see Ding et al., 2002). Whereas this research establishes ways to produce unbiased estimates of demand, and suggests that one should inflate the inventory levels early on to “stalk information” (Lariviere & Porteus, 1999), Feiler et al. (2013) found in experiments that managers tend to underestimate mean demand when it is censored by the inventory level, which leads to biased-low inventory decisions. In fact, Rudi and Drake (2014) find that, even if subjects are explicitly provided with the demand distribution, censored-demand feedback can still bias order decisions low.

17.2.3.3 Supply risk and dual sourcing

There is an emerging literature within behavioral inventory management that investigates supply risk (both yield and/or disruption risk). These studies often examine environments where a buyer has the ability to sole-or-dual source from suppliers. For instance, Gurnani et al. (2014) and Kalkanci (2017) study a buyer who is tasked with ordering inventory from two suppliers, which differ in their cost and risk profiles. Both papers find evidence of buyers over-diversifying their inventory-order decisions across suppliers, relative to the normative predictions. In contrast to this, Goldschmidt et al. (2021) find that buyers have a tendency to under-diversify in their inventory decisions. Their paper differs from Gurnani et al. (2014) and Kalkanci (2017) in that they assume suppliers are homogeneous, the cost of a disruption is marginally increasing in missed demand, and disruptions are low-probability high-impact events.

17.2.3.4 Price-setting and revenue management

In some settings, a manager may be able to make a joint decision about both the price and the inventory quantity (e.g., see Petruzzi & Dada, 1999). Ramachandran et al. (2018) experimentally examine such a newsvendor setting. In their experiments, they find that people tend to under-price and over-order (relative to the optimal joint solution). In revenue management settings, a manager may not be able to change the beginning inventory quantity available, but can control the price at which they are willing to sell over the course of the season. Bearden et al. (2008) study decision behavior in such a revenue management experiment. They find that people tend to demand too high of a price when they have many units left to sell and too low of a price when they have only a few left to sell. Caro and de Tejada Cuenca (2023) conduct a field study at Zara to examine manager overriding behavior of decision-support-system markdown pricing recommendations.

17.2.3.5 Multi-location retailer inventory sharing

Inventory quantities are often aggregated to satisfy demand for multiple retailers. Ho et al. (2010) conduct a newsvendor experiment and find evidence of the PTC effect when a single inventory order is used to fulfill demand across two retailers. There are also behavioral papers which consider how inventory is shared among retailers, after demand is realized, through transshipments. In such a setting, Villa and Castañeda (2018) analyze initial inventory orders and observe that face-to-face interactions between retailers can lead to higher profits. Zhao et al. (2021) extend this line of research by allowing retailers to request or fulfill transshipped units, and find that retailers underestimate the benefits of transshipment. Other behavioral papers on transshipments permit retailers to set the transfer price per unit, while also analyzing inventory-order decisions. For instance, Li and Chen (2020) conduct an experiment that manipulates whether transfer prices are set before or after demand occurs and whether retailers can choose to transship units or not. Last, Davis et al. (2022) study centralized (a single joint quantity) versus decentralized (transshipments) retailer inventory-sharing strategies in a two-stage supply chain and find that inventory orders are quite close to the normative predictions.

17.3 BEYOND DOCUMENTING DEVIATIONS FROM RATIONALITY IN BEHAVIORAL INVENTORY-MANAGEMENT RESEARCH

Is behavioral inventory research's sole purpose to highlight the flaws in rational models? Has the field run out of significant behavioral inventory contributions now that the newsvendor setting and bullwhip effect have been explored? While these are natural questions, we believe that the answer is "no" to both. In this section, we aim to outline multiple roles that behavioral science can play in advancing inventory management beyond documenting deviations from rationality, and discuss specific examples in detail.

Because we organized the previous sections by inventory-model setting, a natural conjecture is that the primary contribution of behavioral research is to put human decision-makers in an inventory setting and document their deviations from a rational benchmark. Yet, we believe this view is too narrow. Figure 17.7 provides a framework for how understanding human behavior can lead to improving inventory performance (broadly defined). Perhaps the most

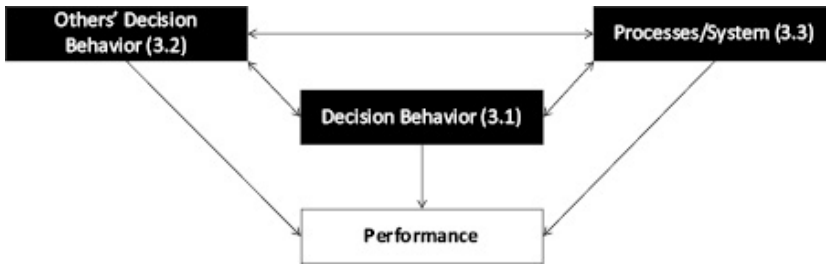


Figure 17.7 Potential pathways to impact inventory performance

obvious manner to improve performance is to identify suboptimal decision behavior and then directly address it. In other words, attempt to directly eliminate any human decision biases. Indeed, such a direct strategy is a fruitful area of research, which we discuss in Section 17.3.1. Of course, identification of when, how, and why humans display the most severe suboptimal behavior is key to informing this (and most all) strategies to improve.

Beyond direct debiasing strategies, there are also more indirect ways of improving performance. These become accessible once one better understands (and can model) human behavior. Being able to anticipate how competitors or partners will behave should change the way one optimally responds. In other words, researchers can study how to optimally compete or cooperate with other players in the system *conditional* on their behavioral regularities. We discuss this behaviorally informed optimal response angle in Section 17.3.2.

Furthermore, instead of trying to directly change a decision-maker, researchers can also take the perspective of a centralized system planner to design the environment or system *conditional* on patterns of human behavior. Such design choices can be at the micro (e.g., within the organization) or macro (e.g., supply-chain or market design) level. We discuss this behavioral system-design angle in Section 17.3.3.

17.3.1 Improving Decision Behavior by Addressing “Behavioral Problems”

Correcting costly behavioral decision biases can represent a substantial opportunity to improve profits. Therefore, naturally, a stream of research within the field of behavioral operations has focused on investigating ways to mitigate (mis)behavior in various inventory settings.

Of course, an immediate strategy for correcting behavioral biases is to rely 100% on automated systems. Yet, as noted in the Introduction (Section 17.1) of this chapter, it is often infeasible or undesirable to do so. With this in mind, behavioral solutions for addressing observed inventory biases can be practical and efficient. These include, among others, improved training, debiasing techniques, and decision framing. Because we detailed an example of improved training in Section 17.2.1 (e.g., with the experiments in Bolton et al., 2012), in this subsection we summarize two papers which investigate debiasing techniques and one which investigates an alternative decision-making frame.

Ren and Croson (2013) argue and provide experimental evidence that overconfidence, notably overprecision, can account for the standard PTC effect in newsvendor order decisions. Specifically, overprecision is a well-known phenomenon in the psychology and economics literature which, in a newsvendor context, translates into a decision-maker underestimating

the standard deviation of the demand distribution. This, in turn, leads to inventory orders that exhibit a PTC effect.

In addition to showing this result theoretically, Ren and Croson (2013) conduct a human-subject newsvendor experiment with a normal demand distribution, with the aim of demonstrating that a decision-maker's degree of overprecision correlates with their degree of the PTC effect. To this end, participants were initially asked to complete a questionnaire which measured, among others, levels of overprecision. Participants then made 50 rounds of newsvendor decisions. Ren and Croson (2013) included two experimental treatments, one with a high-margin product (75% critical fractile) and one with a low-margin product (25% critical fractile). In terms of results, they find support for their hypotheses in both treatments and write "more overprecise individuals exhibit greater order bias in the experiment" (p. 2507).

Ren and Croson (2013) then proceed to investigate how to de-bias the overprecision effect and improve inventory-order decisions. To do so, they adopt a tool developed by Haran et al. (2010) called "SPIES," which can reduce overprecision by facilitating decision-makers to consider the likelihood of events occurring in the tails of a probability distribution. If effective in a newsvendor task, this should reduce a decision-maker's tendency to underestimate the standard deviation of demand and mitigate the PTC effect. Therefore, Ren and Croson (2013) conduct an additional experiment where, in one of the manipulations, they required decision-makers to complete a SPIES task every five rounds. The SPIES task involved estimating the probability that demand in the following round will be in an interval that is one, two, or three standard deviations from the mean. Turning to the results, Ren and Croson (2013) observe that inventory decisions between the baseline setting and SPIES setting increase from 102.76 to 108.05 in the high-margin setting (optimal 120), and decrease from 95.19 to 90.35 (optimal 80). Overall, Ren and Croson (2013) not only identify a bias that is a plausible driver of the PTC effect, but they also identify a way to mitigate it, resulting in improved decisions and higher profitability.

As mentioned in Section 17.2.3.2, past research has shown that human demand beliefs are systematically lower when demand is censored by the inventory level than when demand is uncensored, which in turn leads to biased-low inventory orders. With this in mind, Tong et al. (2018) consider a censored-demand newsvendor setting and investigate ways to mitigate this "censorship bias." In particular, they posit that a simple debiasing mechanism (referred to as "REDO") which encourages decision-makers to adjust each observation in the demand sample and then assess its mean, as opposed to assessing a mean of a biased sample and making an adjustment, can lead to improved decisions.

Tong et al. (2018) conduct a series of human-subject experiments to test their conjecture. They first administered an experiment with three treatments, all with an unknown normal demand distribution, a critical fractile of 50% (to remove any PTC effect), and 30 rounds of decisions. In one treatment, participants played a newsvendor game where demand is censored if a stockout occurred. In a second treatment, the REDO condition, they required participants to answer "What is your best guess of what the exact demand was today?" (after a stockout) and "What was the exact demand?" (after not stocking out). In the third treatment, participants played a newsvendor game without censored demand. In all three treatments, after completing the task, participants were asked to estimate mean demand.

The results of these experiments show that the REDO manipulation is successful in debiasing beliefs. Inventory-order decisions in the REDO treatment are (1) significantly better than in the censored-demand treatment and (2) not significantly different than in the

uncensored-demand treatment. Tong et al. (2018) then explore a second set of experiments which considered a high-margin product (67%) and added a fourth manipulation: a censored-demand newsvendor task where participants had to estimate mean demand after each decision (referred to as “EMD”). By including this EMD treatment, Tong et al. (2018) are able to determine if the success of their REDO manipulation is robust to a high-margin product setting. Also, they can determine if the improvements from the REDO mechanism are due to helping participants think *better* about demand, as opposed to merely thinking about demand more often.

Consistent with their first experiment, the results of the second experiment demonstrated that the REDO treatment leads to higher (i.e., better) inventory orders than in the baseline censored-demand setting or the EMD treatment. Figure 17.8 illustrates the resulting mismatch costs across all four treatments in their second experiment. As with the work of Ren and Croson (2013), Tong et al. (2018) are able to show that a simple debiasing technique is effective in improving inventory decisions.

To briefly provide an example of how an alternative decision-making frame can lead to improved inventory decisions, Kremer et al. (2010) conduct an experiment where a decision-maker can choose among seven different options (with seven possible payoffs for each option). In one set of treatments, the decision-maker faced a standard newsvendor context (referred to as “OPERATIONS”), where they set an inventory level in the face of a demand distribution with seven discrete values, {500,550,...,800}. In another set of treatments, there was no operations context (referred to as “NEUTRAL”), and the decision-maker merely chose a letter, A–G, where each letter had seven possible random payoffs, {500,550,...,800}. Importantly, the payoff matrices between the OPERATION condition and the NEUTRAL condition were identical. Kremer et al. administered these two treatments with both high-margin and low-margin settings.

The average orders in the OPERATIONS treatments, and their equivalent decisions in the NEUTRAL treatments, are shown in Figure 17.9. The PTC effect is diminished in the NEUTRAL context. Aside from helping identify what may be driving the PTC effect (i.e., teasing out those theories that only apply in an operations context), the work of Kremer et al. (2010) indicates that re-framing an inventory problem can lead to improved decisions.

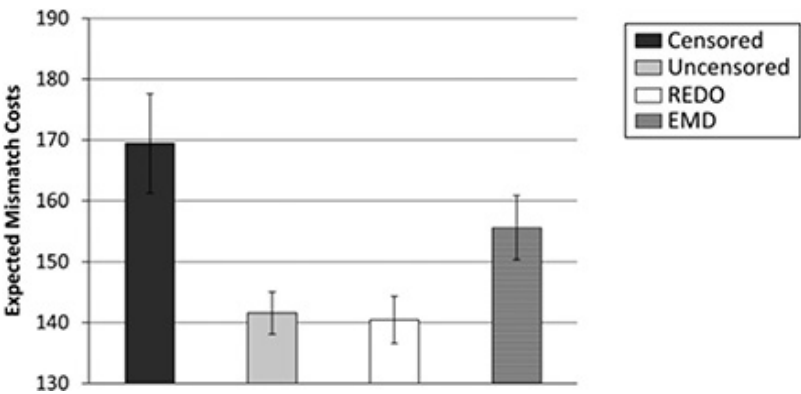


Figure 17.8 Expected mismatch costs in Tong, Feiler, and Larrick’s (2018) Study 2

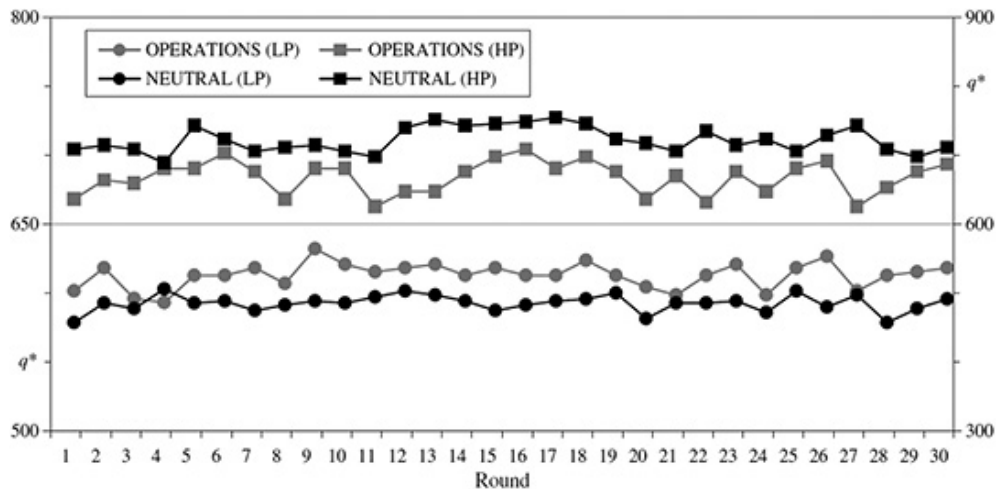


Figure 17.9 Order quantities and decisions in Kremer, Minner, and Van Wassenhove (2010)

17.3.2 Responding Optimally to Others' Predictable Irrationality

The previous subsection provided examples of how to leverage knowledge of human behavior and improve performance by directly changing decision behavior. Another pathway for improving performance is more indirect: developing behavioral models, for partners or competitors, that are superior to rational ones in terms of predictive accuracy, then leveraging them to respond optimally.

To provide a concrete example, we describe the approach in Becker-Peth et al. (2013), one of the early papers in behavioral operations to employ this approach. The authors construct and estimate a behavioral model of newsvendor order decisions under a buyback contract. They then take the perspective of a large seller who has the power to design the contracts for multiple newsvendor buyers. They use their behavioral model to choose contract parameters that maximize channel profit – showing that they can calibrate such a model to design contracts that perform better than assuming the buyers will behave rationally.

Their first step was to build upon the literature on judgment and decision-making and the existing evidence on newsvendor decision-making behavior to develop specific hypotheses of how humans may respond differently to buyback contracts than a rational decision-maker. Given an order quantity z and demand realization d , the profit is:

$$\pi(z, d) = (p - w) \min(z, d) - (w - b)(z - \min(z, d))$$

where w is the wholesale price, b is the buyback price, and p is the sales price. In contrast, Becker-Peth et al. (2013) assume the individual perceives the following value function:

$$v(z, d) = (p - w) \min(z, d) - \beta(w - \gamma b)(z - \min(z, d)).$$

Here, $\beta > 1$ captures loss-aversion (Kahneman & Tversky, 1979): the authors conjecture that the cost associated with leftover inventory is perceived as a loss and that this loss is more

painful than an equivalent gain from sales. Further, γ captures a type of mental accounting: the authors conjecture that decision-makers are not indifferent to the source of the income – they may “feel” the income from returns more than do the income from sales ($\gamma > 1$) or vice versa ($\gamma < 1$). They hypothesize $\gamma > 1$, in part because of the result from the following simple scenario survey that they executed, which is similar to those used by Thaler (1985):

Experiment 2. Mr. A bought 200 newspapers at 1 euro each, sold 100 of them for 4 euros each, and returned the 100 unsold newspapers to the publisher, receiving no additional compensation and netting 200 euros in profit. Mr. B bought 200 newspapers at 1 euro each, sold 100 of them for 3 euros each, and returned the remaining 100 unsold newspapers to the publisher, receiving 1 euro for each and netting 200 euros in profit. Who is happier? Mr. A (6), Mr. B (33), no difference (7).

Even though Mr. A and Mr. B end up with the same profit, people generally prefer Mr. B’s experience – which is consistent with weighting the income stream from returns more than sales (see also Chen et al., 2013 for similar findings).

Considering the value function, the authors propose the following behavioral order quantity:

$$\hat{z} = (1 - \eta) F^{-1} \left(\frac{p - w}{p - w + \beta(w - \gamma b)} \right) + \eta \lambda$$

The term $F^{-1} \left(\frac{p - w}{p - w + \beta(w - \gamma b)} \right)$ maximizes the value function. Motivated by an anchoring and adjustment heuristic (Tversky & Kahneman, 1974) and informed by past newsvendor experimental results (Schweitzer & Cachon, 2000; Bolton & Katok, 2008), the authors predict that decision-makers will anchor on the mean demand λ and only partially adjust toward the value-maximizing quantity (see also Equation (17.1)).

In their main experiment, Becker-Peth et al. (2013) examined participants’ order decisions under 28 different sets of contract parameters, in random order. Figure 17.10a shows the mean

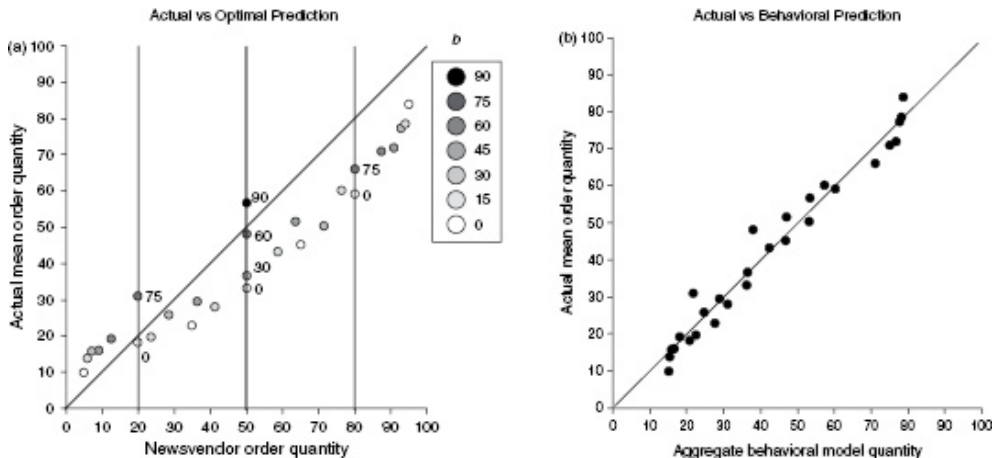


Figure 17.10 Mean orders and predictions from the main experiment in Becker-Peth et al. (2013)

order quantities under each contract relative to the optimal expected-profit-maximizing newsvendor order quantity. It illustrates that (1) actual order quantities are biased relative to the optimal order quantity, and (2) even when the optimal order quantity is the same, the combination of (w , b) matters such that a higher b leads to a higher order. Figure 17.10b shows that the behavioral model is able to predict these average order quantities better than the rational model.

Of course, the behavioral model has more parameters and embeds the rational model when $\gamma=\beta=1$ and $\eta=0$, so it will always fit data better *ex-post*. However, are the behaviors it captures robust enough to make more accurate predictions, that can then be used to design contracts more effectively? To test this idea, the authors conduct a validation experiment in which they (1) calibrate the behavioral model based on data from 19 contracts that have been designed to maximize channel profit using the standard rational model (Phase 1), (2) fit the behavioral model by estimating each of the model parameters γ , β , and η , then (3) solve for the w , b that induces the participant to choose an order quantity that maximizes channel profit, for eight more settings with varying critical ratios (Phase 2).

The table in Figure 17.11 summarizes the results of this validation experiment. The key observation is that the channel profits increase significantly, by roughly 10%, moving from Phase 1 (choosing contract parameters to optimize channel profit assuming a rational model) to Phase 2 (choosing contract parameters to optimize channel profit leveraging the fitted behavioral model). In other words, the behavioral model is able to predict the buyer's order decisions sufficiently well to make better contract parameter decisions.

While we discussed Becker-Peth et al. (2013) in detail, it is not unique in its approach within the behavioral operations management literature. For example, Ovchinnikov et al. (2015) take a similar approach in a newsvendor competition setting. Namely, they show that by calibrating a behaviorally informed model of a newsvendor competitor (instead of following a standard Nash equilibrium strategy), one can solve for the optimal response and significantly improves one's performance. Advances in machine learning also hold some promise with this approach. For example, as mentioned above in Section 17.2.2, Oroojlooyjadid et al. (2021) take the perspective of a single player within a serial supply chain where the other members of the supply chain are subject to behavioral biases. They then take a reinforcement learning approach to derive how to best respond to their human supply-chain partners to minimize the total cost of the supply-chain.

17.3.3 Designing Processes and Systems with Human Behavior Considerations

Instead of optimally responding to a competitor or supply-chain partner's behavior, the last major improvement pathway is to design processes or systems conditional on these human

Validation experiment	Phase 1	Phase 2		Phase 3
		Individual	Aggregate	Trained
Individual behavioral contracts	1,365	1,533 (+12.3%)		
Aggregate vs. individual behavioral contracts	1,389	1,546 (+11.3%)	1,519 (+9.4%)	
Training vs. individual behavioral contracts	1,426	1,579 (+10.7%)		1,503 (+5.4%)

Figure 17.11 Expected profits from the validation experiment of Becker-Peth et al. (2013)

behavioral regularities. In other words, take the perspective of the central planner who influences performance not by directly manipulating inventory decision behavior, but by designing the system so that it performs well when human decision-makers are in it.

Feiler and Tong (2022) represent a recent example of this approach. The authors consider an organization's process for forecasting, product selection/design, and inventory production decisions in the context of a new product. In a behavioral model, they first incorporate two "behavioral features": random noise in the forecast and naive (non-regressive) statistical thinking. They then use this model to explain why an individual's forecasts (and subsequent production decisions) for their chosen product tend to be too high.

Figure 17.12 provides the intuition from their model. For each product, the human can obtain predictive information about the true demand mean for each product. However, due to cognitive limitations, humans add random noise to their interpretation of this information. If they were statistically sophisticated and self-aware of their cognitive limitations, they would make a mean-reverting correction to account for such noise (figure on the right). However, they do not make such corrections. Now, even if such noise is unbiased across all products, the selection process of trying to pick the "best" product systematically chooses the products for which the individual errors are on the high side. In other words, erring high for a product causes them to be more likely to choose it.

The authors next provide experimental evidence that the modeled behavioral mechanism indeed leads to biased-high forecasts and over-production for chosen products. They show that (1) forecasts and production decisions are biased high for chosen products vs. randomly-selected products, and (2) this bias is larger in more complex information settings under which the human adds more noise (even though the underlying value of information is the same).

Finally, the authors leverage their behavioral model to generate and test insights into the implications of the organizational process design. Namely, a process design implication of their behavioral model is as follows: you can (at least partially) avoid the mechanism driving the biased-high forecasting described above by getting a second independent forecast for the chosen product from a different individual (see proposition 2). The intuition is that a second person does not suffer from the selection effect in which a person chooses the product because they randomly erred high with their beliefs.

Figure 17.13 shows a screenshot of the experiment the authors use to test this process design implication. In the experiment, there are two conditions. Under "Choice," the same person

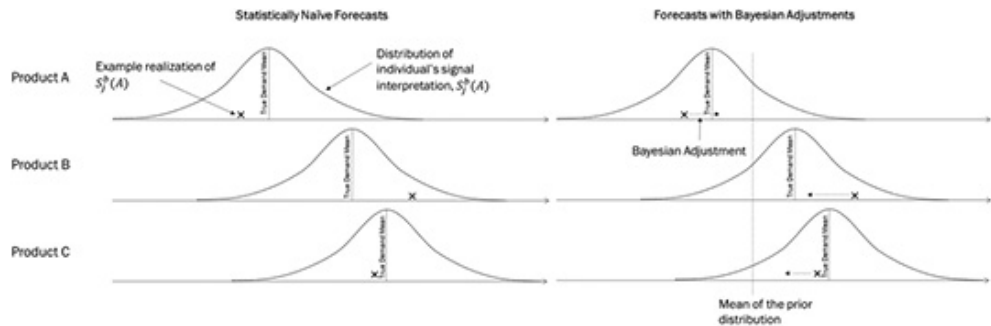


Figure 17.12 Intuition of the behavioral model in Feiler and Tong (2022)

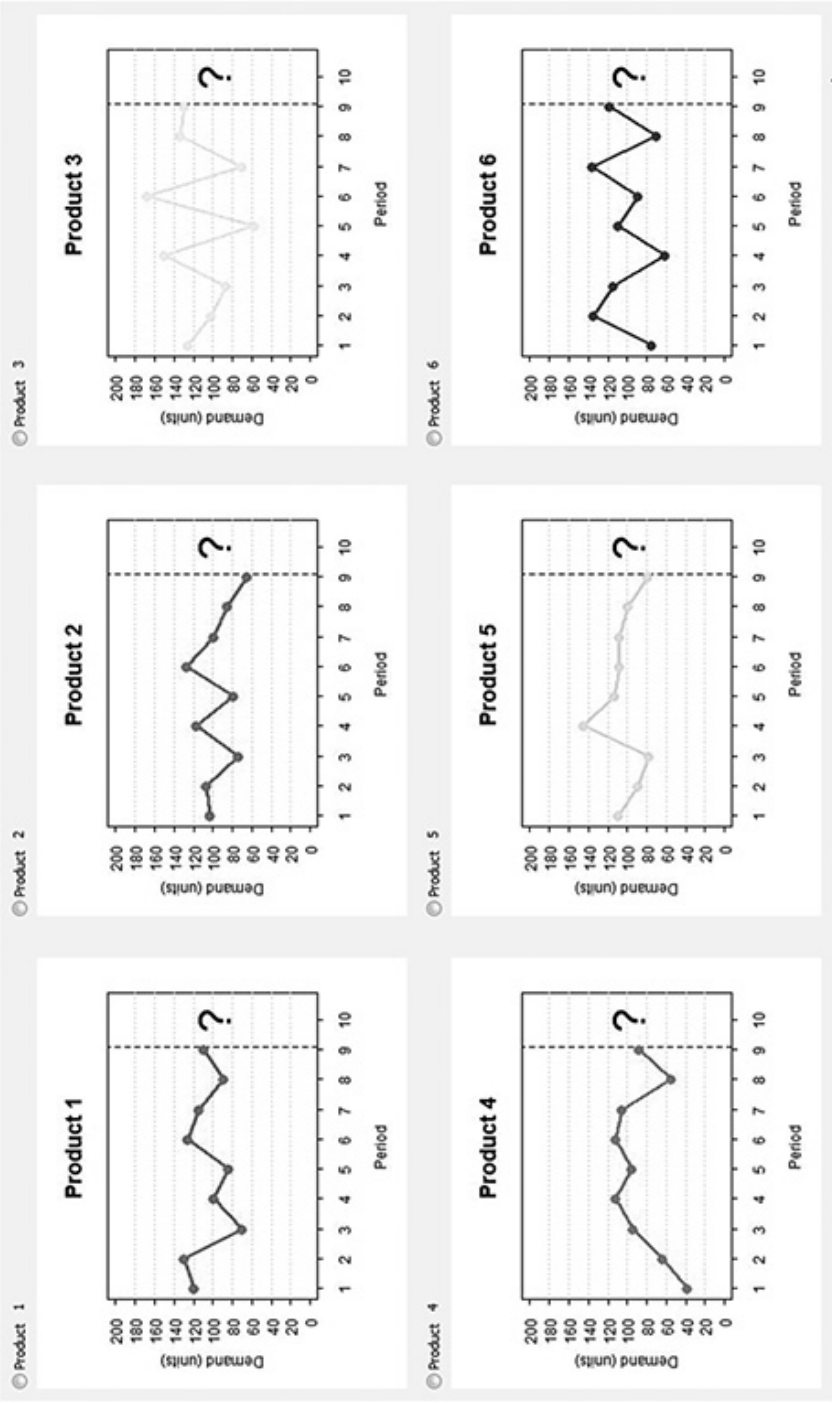


Figure 17.13 Sample experimental screenshot in Feiler and Tong (2022)

chooses which one of the six products they want to sell in period 10, then decides how much inventory to stock (where overage cost = underage cost = \$1.). Under “Independent,” an independent subject takes the product choice from a person under Choice and then decides only how much inventory to stock. Figure 17.14 shows that, indeed, inventory orders are biased about 12 units when the same person chooses the product and places the inventory-order decision. In contrast, the inventory orders are about seven units lower when an independent person makes the inventory decision for the product chosen by a participant in the Choice condition.

Davis and Hyndman (2019) provide another recent example of how altering a process or system can improve outcomes, in the context of supply-chain contract design. Most behavioral inventory studies consider a setting where an upstream supplier makes a one-shot wholesale price offer to a downstream retailer, who then sets an inventory order. Much of this research also assumes that the retailer incurs the cost of any unsold inventory. However, in practice, companies often negotiate wholesale prices and inventory orders. Further, suppliers often perform demand-fulfillment duties for retailers (e.g., e-commerce), and thus incur the cost of any unsold inventory. With this motivation, Davis and Hyndman (2019) conduct a human-subjects experiment and investigate how supply-chain outcomes are affected by alternative bargaining processes between buyers and suppliers, and, which party incurs the cost of unsold inventory (i.e., inventory risk).

The experimental design of Davis and Hyndman (2019) includes the standard one-shot ultimatum setting as a baseline condition (referred to as “Ult”). Specifically, a proposing party made a one-shot take-it-or-leave-it wholesale price offer to a responding party, who then unilaterally made an inventory-order decision (or rejects the offer). While such an interaction has been explored in past studies on inventory risk (e.g., Davis et al., 2014), Davis and Hyndman (2019) consider two additional manipulations that varied how the two parties arrive at a contract. In one, the two parties engaged in a dynamic back-and-forth bargaining process over the wholesale price. After this, the party incurring the inventory risk unilaterally made the inventory-order decision (referred to as “Neg-W”). In the other setting, the two parties again engaged in a back-and-forth bargaining process, but both the wholesale price *and* inventory order were jointly negotiated (referred to as “Neg-WQ”). This last manipulation is useful in

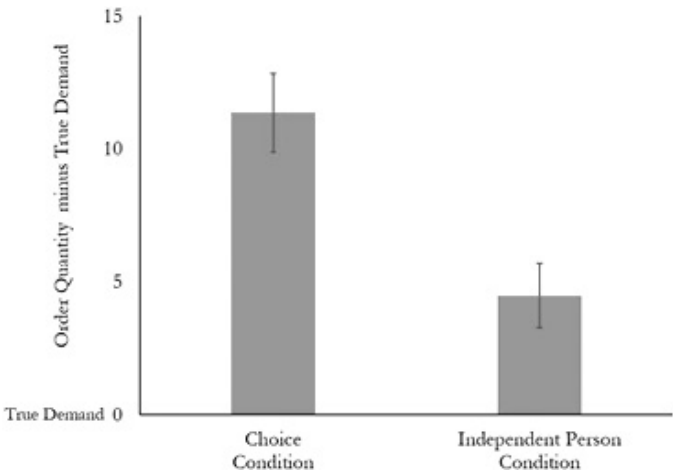


Figure 17.14 Inventory order bias in Feiler and Tong (2022)

that the normative theory predicts that the two parties should agree on the first-best inventory order, resulting in 100% supply-chain efficiency (and an equal split of expected profits). In short, changing the process in which inventory orders are set can, even theoretically, lead to different profit outcomes.

In the other dimension of their experimental design, Davis and Hyndman (2019) varied where the inventory risk resided in the supply chain. In one set of treatments, the retailer incurred the risk. In a second variant, the supplier incurred the risk. And in a third variant, the location was endogenous such that the retailer and supplier bargained over who would incur the inventory risk.

The expected supply-chain efficiency, retailer profit, and supplier profit, from Davis and Hyndman (2019) are depicted in Figure 17.15. Comparing the Neg-W treatments with the Neg-WQ treatments, for both inventory risk locations, one can immediately see that efficiency increases, from 74.91 to 90.64% in retailer risk, and from 82.49 to 92.39% in supplier risk. This improvement is driven by the fact that inventory orders are set closer to first-best levels. Further, these more favorable inventory orders are driven by more equitable wholesale prices. As a consequence, changing the negotiation process from Neg-W to Neg-WQ (i.e., including the inventory order in the negotiation) leads to a Pareto improvement: both the retailer and supplier earn higher profits.

From a methodological perspective, many behavioral inventory papers implement laboratory experiments and behavioral models. But, one need not pursue such a combination in the same paper. For example, in Su’s (2008) influential paper on incorporating decision noise into inventory decisions, he leverages the prior newsvendor experimental data from Bolton and Katok (2008) to validate his model. Then, in his supply-chain design applications he derives results such as showing that the value of inventory pooling is higher under human decision

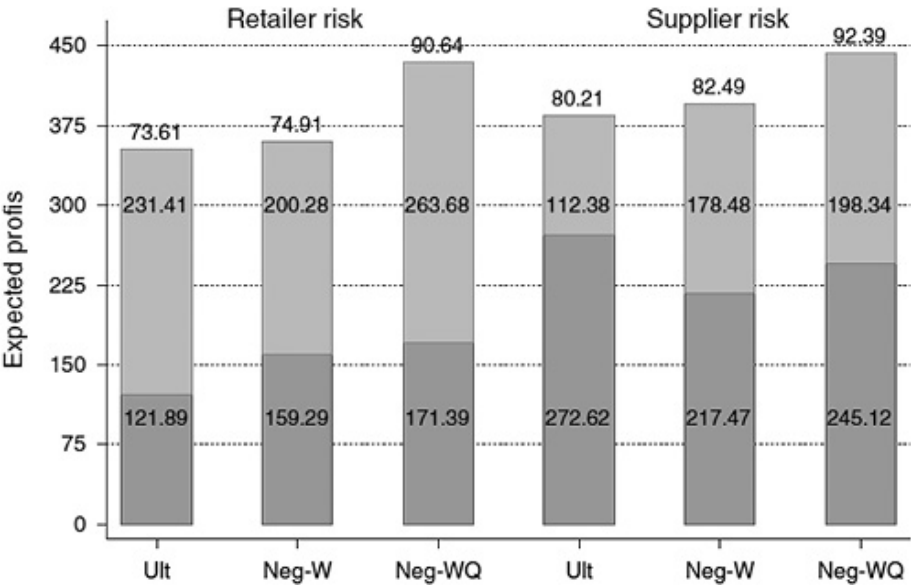


Figure 17.15 *Expected supply chain efficiency and profits in Davis and Hyndman (2019)*

noise than with rational decision-makers. Another example is Li et al. (2017). In their paper, the authors leverage the experimental findings on overconfidence in newsvendor decisions from Ren and Croson (2013) to inform a model of overprecision – a form of underestimating the demand variance. They then analytically derive equilibrium results in the competitive newsvendor setting. Among other results, they show that such overconfidence can lead to a Pareto improvement for two newsvendors under competition. A final example is Chen et al. (2021), who utilize data from Bolton and Katok (2008), Ren and Croson (2013), and Ockenfels and Selten (2014), in validating a model of forecast anchoring that can capture both the mean *and* variance of observed inventory-order decisions in a newsvendor environment. In short, new behavioral insights can lead to a variety of actionable micro- and macro-level design prescriptions by incorporating the predictable patterns of human behavior.

17.4 CONCLUSION AND PERSPECTIVES ON FUTURE RESEARCH OPPORTUNITIES

We conclude with a high-level discussion on future research opportunities.

17.4.1 Beyond Newsvendor and Beer Games

Incorporating human behavior into inventory management has led to important contributions, as evidenced by the influential streams of research in newsvendor decision behavior and the beer game. Documenting deviations from rationality in these settings, while still important, is a relatively saturated approach (see Section 17.2.1 & Section 17.2.2). Of course, behavioral inventory decisions are important in other inventory decision-making settings beyond the newsvendor and beer game settings, too, as reviewed in Section 17.2.3. While seeking to document deviations from rationality in other inventory decision-making settings remains an opportunity for future research, we believe this approach to be somewhat limiting given the current maturity of the area.

17.4.2 Beyond Documenting Deviations from Rationality

In Section 17.3, we argued against the misconception that contributions in behavioral inventory management are limited only to documenting deviations from rationality. Understanding human behaviors can generally serve as a building block toward impacting inventory performance through intervention, training, debiasing, or nudging techniques (Section 17.3.1), responding optimally to predictable behaviors in partners or competitors (Section 17.3.2), and designing system/supply chains with behavioral considerations (Section 17.3.3). We believe these approaches have examples in the literature yet overall still represent substantial future opportunities. They can also potentially be methodologically diverse: lab or field empirical methods, analytical or algorithmic modeling methods.

17.4.3 Beyond Direct Inventory-Order Decisions

The majority of the papers discussed in this chapter have directly considered inventory-order decision behavior. Such an approach is both natural and important. At the same time, current

technological developments and the evolution of the field bring new opportunities to study how human behavior may impact inventory management in other, more indirect, ways.

In some settings, the impact that humans have on inventory orders is primarily through the inputs to a decision policy (e.g., humans may influence the demand forecast, but perhaps not the conversion to inventory orders), or through adjustments to the outputs of an algorithm (e.g., humans can override algorithms). For example, Kesavan and Kushwaha (2020) study decision-makers' overriding behavior of an algorithm in an automobile replacement parts retailer and find that managers' overriding behavior reduces profitability by over 5% overall (but increases profitability for some types of product categories).

Even in settings that appear "fully automated," we note that human behavior's importance typically has not been eliminated, but rather moved to a different stage. For example, human decision-makers may be tasked with choosing among fully automated policies and evaluating their performance. Human behavior at this policy choice and evaluation stage then becomes a critical issue. For example, Stangl and Thonemann (2017) consider a scenario in which a manager is supposed to choose among two inventory improvement initiatives based on their projected impact on inventory metrics. In this setting, the type of inventory metric used (days-of-supply versus inventory-turn rate) can significantly impact the choice for behavioral reasons. Similarly, in Kesavan and Kushwaha (2020) mentioned above, managers tend to mostly rely on the system to automatically make replenishment decisions (as there are so many of them). Instead, they pay more attention to the decision of whether to stop replenishing items in general at the SKU-store level.

Thus, as algorithmic and technological advancements change the nature of humans' impact on inventory management, there is opportunity for future behavioral research to study these different types of interactions.

17.4.4 Beyond the Laboratory

Finally, we view providing field evidence as a substantial future research opportunity. With a few exceptions (e.g., Caro & de Tejada Cuenca, 2023; Kesavan & Kushwaha, 2020), the empirical evidence reviewed in this chapter is from laboratory experiments. On one hand, it is the natural evolution for behavioral biases to be first documented and explored in controlled laboratory settings before being tested or validated by field evidence. On the other hand, there is now a sizable body of laboratory experiments in the area, yet field evidence appears to still be hard to find. For example, despite being arguably the most well-known and replicated laboratory finding, field testing of the pull-to-center effect appears to be limited to testing managers in the lab (Bolton et al., 2012); there is no direct field evidence to our knowledge.

Access to data and other practical challenges present some barriers to field data. For example, estimating the bullwhip effect itself with industry data is already a significant empirical challenge (e.g., see Cachon et al., 2007; Bray & Mendelson, 2012), so attempting to further isolate behavioral factors or manipulating the environment through a field experiment may require substantial creativity or access to novel data. Similarly, newsvendor-type decisions in the field are often in settings where the true overage and underage costs are difficult to estimate (e.g., see Olivares et al., 2008), and may be challenging to tease apart from the demand forecasting, revenue management, and other complicating factors. Nevertheless, we believe that the lack of field testing is primarily due to the lack of researcher attention, not because these types of challenges are prohibitively difficult to overcome. Field research also need not always test effects already documented in the lab: it can be a source of identifying new

decision tasks and behaviors, and illuminating the types of human decisions that are most relevant for inventory researchers to understand.

NOTES

1. Although there is earlier related research in the field of accounting (e.g., Hoskin, 1983).
2. Subsequent studies will evaluate whether 15 rounds, in a specific newsvendor setting, is sufficient for participants to “learn” the optimal order.

REFERENCES

- Bearden, J. N., Murphy, R. O., & Rapoport, A. (2008). Decision biases in revenue management: Some behavioral evidence. *Manufacturing and Service Operations Management*, 10(4), 625–636.
- Becker-Peth, M., Katok, E., & Thonemann, U. W. (2013). Designing buyback contracts for irrational but predictable newsvendors. *Management Science*, 59(8), 1800–1816.
- Becker-Peth, M., & Thonemann, U. W. (2019). Behavioral inventory decisions. In K. Donohue, E. Katok, & S. Leider (Eds.), *Behavioral operations management*, chap. 11 (pp. 393–432). Wiley.
- Benzion, U., Cohen, Y., Peled, R., & Shavit, T. (2008). Decision-making and the newsvendor problem: An experimental study. *Journal of the Operational Research Society*, 59(9), 1281–1287.
- Bolton, G. E., & Katok, E. (2008). Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing and Service Operations Management*, 10(3), 519–538.
- Bolton, G. E., Ockenfels, A., & Thonemann, U. W. (2012). Managers and students as newsvendors. *Management Science*, 58(12), 2225–2233.
- Bray, R. L., & Mendelson, H. (2012). Information transmission and the bullwhip effect: An empirical investigation. *Management Science*, 58(5), 860–875.
- Cachon, G. P., Randall, T., & Schmidt, G. M. (2007). In search of the bullwhip effect. *Manufacturing and Service Operations Management*, 9(4), 457–479.
- Caro, F., & de Tejada Cuenca, A. S. (2021). Believing in analytics: Managers adherence to price recommendations from a DSS [Working paper]. University of California Los Angeles.
- Caro, F., & de Tejada Cuenca, A. S. (2023). Believing in analytics: Managers’ adherence to price recommendations from a DSS. *Manufacturing and Service Operations Management*, 25(2), 524–542.
- Chen, F. (1999). Decentralized supply chains subject to information delays. *Management Science*, 45(8), 1076–1090.
- Chen, K.-Y., & Wu, D. (2019). Buyer-supplier interactions. In K. Donohue, E. Katok, & S. Leider (Eds.), *The handbook of behavioral operations*, chap. 13 (pp. 459–487). Wiley.
- Chen, L., Davis, A. M., & Kim, D. (2021). Predicting mean and variance in inventory order decisions [Working paper]. Cornell University. <https://doi.org/10.2139/ssrn.2310619>
- Chen, L., Gürhan Kök, A., & Tong, J. D. (2013). The effect of payment schemes on inventory decisions: The role of mental accounting. *Management Science*, 59(2), 436–451.
- Cooper, D. J., Kagel, J. H., Lo, W., & Gu, Q. L. (1999). Gaming against managers in incentive systems: Experimental results with Chinese students and Chinese managers. *American Economic Review*, 4(4), 781–804.
- Croson, R., & Donohue, K. (2003). Impact of pos data sharing on supply chain management: An experimental study. *Production and Operations Management*, 12(1), 1–11.
- Croson, R., & Donohue, K. (2006). Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, 52(3), 323–336.
- Croson, R., Donohue, K., Katok, E., & Stermann, J. (2014). Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, 23(2), 176–196.
- Davis, A. M., & Hyndman, K. (2019). Multidimensional bargaining and inventory risk in supply chains: An experimental study. *Management Science*, 65(3), 1286–1304.

- Davis, A. M., Huang, R., & Thomas, D. J. (2022). Retailer inventory sharing in two-tier supply chains: An experimental investigation. *Management Science*, 68(12), 8773–8790.
- Davis, A. M., Katok, E., & Santamaría, N. (2014). Push, pull, or both? A behavioral study of how the allocation of inventory risk affects channel efficiency. *Management Science*, 60(11), 2666–2683.
- Ding, X., Puterman, M. L., & Bisi, A. (2002). The censored newsvendor and the optimal acquisition of information. *Operations Research*, 50(3), 517–527.
- Feiler, D. C., Tong, J. D., & Larrick, R. P. (2013). Biased judgment in censored environments. *Management Science*, 59(3), 573–591.
- Feiler, D., & Tong, J. (2022). From noise to bias: Overconfidence in new product forecasting. *Management Science*, 68(6), 4685–4702.
- Goldschmidt, K., Kremer, M., Thomas, D. J., & Craighead, C. W. (2021). Strategic sourcing under severe disruption risk: Learning failures and under-diversification bias. *Manufacturing and Service Operations Management*, 23(4), 761–780.
- Graves, S. C. (1999). A single-item inventory model for a nonstationary demand process. *Manufacturing and Service Operations Management*, 1(1), 50–61.
- Gurnani, H., Ramachandran, K., Ray, S., & Xia, Y. (2014). Ordering behavior under supply risk: An experimental investigation. *Manufacturing and Service Operations Management*, 16(1), 61–75.
- Haran, U., Moore, D. A., & Morewedge, C. K. (2010). A simple remedy for overprecision in judgment. *Judgment and Decision Making*, 5(7), 467–476.
- Ho, T.-H., Lim, N., & Cui, T. H. (2010). Reference dependence in multilocation newsvendor models: A structural analysis. *Management Science*, 56(11), 1891–1910.
- Hoskin, R. E. (1983). Opportunity cost and behavior. *Journal of Accounting Research*, 21(1), 78–95.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
- Kalkanci, B. (2017). Supply risk mitigation via supplier diversification and improvement: An experimental evaluation [Working paper]. Georgia Institute of Technology.
- Katok, E., Thomas, D., & Davis, A. (2008). Service-level agreements as coordination mechanisms: The effect of review periods. *Manufacturing and Service Operations Management*, 10(4), 609–624.
- Kesavan, S., & Kushwaha, T. (2020). Field experiment on the profit implications of merchants' discretionary power to override data-driven decision-making tools. *Management Science*, 66(11), 5182–5190.
- Kremer, M., Minner, S., & Van Wassenhove, L. N. (2010). Do random errors explain newsvendor behavior? *Manufacturing and Service Operations Management*, 12(4), 673–681.
- Larivière, M. A., & Porteus, E. L. (1999). Stalking information: Bayesian inventory management with unobserved lost sales. *Management Science*, 45(3), 346–363.
- Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4), 546–558.
- Li, M., Petrucci, N. C., & Zhang, J. (2017). Overconfident competing newsvendors. *Management Science*, 63(8), 2637–2646.
- Li, S., & Chen, K.-Y. (2020). The commitment conundrum of inventory sharing. *Production and Operations Management*, 29(2), 353–370.
- Lurie, N. H., & Swaminathan, J. M. (2009). Is timely information always better? The effect of feedback frequency on decision making. *Organizational and Human Decision Processes*, 108(2), 315–329.
- Nahmias, S. (1994). Demand estimation in lost sales inventory systems. *Naval Research Logistics (NRL)*, 41(6), 739–757.
- Narayanan, A., Arunachalam, & Moritz, B. B. (2015). Decision making and cognition in multi-echelon supply chains: An experimental study. *Production and Operations Management*, 24(8), 1216–1234.
- Ockenfels, A., & Selten, R. (2014). Impulse balance in the newsvendor game. *Games and Economic Behavior*, 86, 237–247.
- Olivares, M., Terwiesch, C., & Cassorla, L. (2008). Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Science*, 54(1), 41–55.
- Ovchinnikov, A., Moritz, B., & Quiroga, B. F. (2015). How to compete against a behavioral newsvendor. *Production and Operations Management*, 24(11), 1783–1793.
- Perera, H. N., Fahimnia, B., & Tokar, T. (2020). Inventory and ordering decisions: A systematic review on research driven through behavioral experiments. *International Journal of Operations and Production Management*, 40(7/8), 997–1039.

- Petruzzi, N. C., & Dada, M. (1999). Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2), 183–194.
- Plott, C. R. (1987). Dimensions of parallelism: Some policy applications of experimental methods. In A. E. Roth (Ed.), *Laboratory experimentation in economics: Six points of view*. Cambridge University Press.
- Ramachandran, K., Tereyagoglu, N., & Xia, Y. (2018). Multidimensional decision making in operations: An experimental investigation of joint pricing and quantity decisions. *Management Science*, 64(12), 5461–5959.
- Ren, Y., & Croson, R. (2013). Overconfidence in newsvendor orders: An experimental study. *Management Science*, 59(11), 2502–2517.
- Rudi, N., & Drake, D. (2014). Observation bias: The impact of demand censoring on newsvendor level and adjustment behavior. *Management Science*, 60(5), 1334–1345.
- Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3), 404–420.
- Stangl, T., & Thonemann, U. W. (2017). Equivalent inventory metrics: A behavioral perspective. *Manufacturing and Service Operations Management*, 19(3), 472–488.
- Steckel, J. H., Gupta, S., & Banerji, A. (2004). Supply chain decision making: Will shorter cycle times and shared point-of-sale information necessarily help? *Management Science*, 50(4), 458–464.
- Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3), 321–339.
- Su, X. (2008). Bounded rationality in newsvendor models. *Manufacturing and Service Operations Management*, 10(4), 566–589.
- Takác, M., Oroojlooyjadid, A., Nazari, M. R., & Snyder, L. V. (2021). A deep q-network for the beer game: Deep reinforcement learning for inventory optimization. *Manufacturing and Service Operations Management*, 24(1), 285–304.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199–214.
- Thaler, R. H., & Ganser, L. J. (2015). *Misbehaving: The making of behavioral economics*. WW Norton.
- Tong, J., Feiler, D., & Larrick, R. (2018). A behavioral remedy for the censorship bias. *Production and Operations Management*, 27(4), 624–643.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Villa, S., & Castañeda, J. A. (2018). Transshipments in supply chains: A behavioral investigation. *European Journal of Operational Research*, 269(2), 715–729.
- Wu, D. Y., & Katok, E. (2006). Learning, communication, and the bullwhip effect. *Journal of Operations Management*, 24(6), 839–850.
- Zhang, Y., & Siemsen, E. (2019). A meta-analysis of newsvendor experiments: Revisiting the pull-to-center asymmetry. *Production and Operations Management*, 28(1), 140–156.
- Zhao, H., Xu, L., & Siemsen, E. (2021). Inventory sharing and demand-side underweighting. *Manufacturing and Service Operations Management*, 23(5), 1217–1236.

PART III

CONTEXT SPECIFIC MODELS AND METHODS

18. Healthcare inventory management

Turgay Ayer, Chelsea C. White III, and Can Zhang

18.1 INTRODUCTION

The healthcare industry is one of the largest and fastest-growing industries in the world. For example, in 2019, the United States (US) spent \$3.8 trillion (which consists of 17.7% of the US GDP), or approximately \$11,582 per person, on healthcare, and the nation's healthcare spending is predicted to reach \$6.2 trillion by 2028 (Centers for Medicare and Medicaid Services, 2021). With such high and rapidly growing spending, a critical question faced by various policy-/decision-makers in the healthcare system is how to effectively manage the resources to ensure the delivery of high-quality care while minimizing the costs.

Providing quality care to patients requires that the necessary medical products and supplies and the essential medicines are available on hand. As such, the effective management of healthcare product inventory is a vital element in the process of healthcare delivery. Our objective in this chapter is to provide a brief discussion on how inventory management theory and concepts can be applied to the healthcare sector by considering the unique opportunities and challenges of the sector.

In particular, two important features of healthcare products make their inventory management especially challenging: First, most healthcare products are perishable. For example, many blood products have a short shelf-life of a few days to a few weeks. Virtually all pharmaceuticals have an expiration date. Similarly, most vaccines have a short shelf-life of up to a few months (in fact, the Pfizer COVID-19 vaccine is stable for only about six hours at room temperature). Such perishability imposes substantial challenges to inventory management, especially for critical life-saving healthcare products. Because of the high stock-out costs for critical healthcare products, a high inventory level is often necessary to ensure their availability. However, due to perishability, a high inventory level can result in substantial wastages of critical and scarce resources (e.g., blood products). The combination of both high shortage and wastage costs makes effective management of critical healthcare products particularly important.

Second, another important feature of healthcare products that makes their inventory management more challenging is the multi-stakeholder nature of the healthcare system. For most healthcare products, the payer is often neither the consumer (patients) nor the healthcare provider, but a third party (e.g., the government, the employer, or private insurance). Such a unique structure of the healthcare product supply chain creates new opportunities and challenges for supply chain and inventory management of the sector. In addition to a third-party payer, other roles in a healthcare product supply chain that are more or less unique to the sector include the group purchasing organizations (GPOs) who negotiate purchase contracts with manufacturers on behalf of their healthcare provider members, and pharmacy benefits managers (PBMs) who help employers and other health plan holders to design their insurance coverage and negotiate prices and rebates with drug manufacturers. In such a multi-stakeholder environment, each stakeholder

has its own objectives and incentives, and the inventory management of many healthcare products is directly affected by such incentives and the contracts among the stakeholders.

To understand the effect of the aforementioned two features of healthcare products on inventory management, in the rest of this chapter, we will focus on discussing the inventory management for blood products and pharmaceuticals (in §18.2 and §18.3, respectively), since blood products are highly perishable and a pharmaceutical chain involves a multitude of stakeholders. We remark that our goal in this chapter is not to provide an exhaustive review of the literature on healthcare inventory management. Instead, in each of the next two sections, we highlight the key challenges and trade-offs by presenting some base models from a few representative papers, followed by a brief discussion of the key findings from several recent papers. Finally, we conclude this chapter with a few directions for future research (in §18.4).

18.2 INVENTORY MANAGEMENT FOR BLOOD PRODUCTS

Every day, more than 36,000 units of blood products are needed in the United States. Because of the high prevalence of chronic diseases and the aging population, the demand for blood transfusions has been continuously high. For example, the American Cancer Society estimates that more than 1.8 million people will be diagnosed with cancer in a year, and many of them will need (daily) blood transfusion during their chemotherapy treatment (American Red Cross, 2021).

Unlike many commercial products, blood products cannot be manufactured. Instead, the supply of blood products relies on limited volunteer donations (Ayer et al. 2019). In the United States, each year, only about 3% of the population donates blood (American Red Cross, 2021). Due to both the high demand and the limited supply, blood shortages have been frequently reported, and especially so during the COVID-19 pandemic (Memorial Sloan Kettering Cancer Center, 2021). Such shortages of blood can cause serious problems and increase morbidity and mortality of patients.

What further complicates the problem is the perishability of the blood products. Many blood products, such as red blood cells (RBCs) and platelets, have short shelf lives. As a result, although blood shortages have been a major concern, a significant portion of the blood products has been wasted. For example, in the United States, approximately 10% of the platelet inventory is outdated every year (Jones et al. 2020).

Given the medical importance of blood products, their limited supply, and their perishable nature, effective management of blood inventory is extremely important for ensuring patient safety and for keeping the wastage at a low level. Accordingly, there is a relatively large body of literature on inventory management for blood products (see overviews of the literature in Beliën and Forcé (2012), Deniz et al. (2010), Karaesmen et al. (2011), Nahmias (2011), and Pierskalla (2005)). In this section, we discuss three important inventory management problems for blood products: (a) optimal inventory decisions for single-location systems, (b) optimal inventory decisions for multi-location systems, and (c) optimal issuing decisions.

18.2.1 Optimal Inventory Decisions for Single-Location Systems

In this section, we first present a baseline inventory management model for blood and other perishable products in a single-location system based on two seminal papers on perishable

inventory management, Fries (1975) and Nahmias (1975), and discuss the key structural properties of the optimal inventory policy. We then briefly discuss the key findings from a few recent papers that study blood/perishable inventory management in single-location systems.

Consider a single location (e.g., a hospital) that manages the inventory of blood products with an L -period lifetime. At the beginning of each day, the hospital purchases fresh (i.e., age 0) blood products from a supplier with a unit ordering cost c . The ordered products arrive at the hospital with zero lead time. Products are used to meet demand in a first-in-first-out (FIFO) manner. If the demand exceeds the available inventory at the hospital, unmet demand is satisfied from outside sources (i.e., there is no backlog of demand) and the hospital incurs a unit shortage cost p for each unit of stock-out. On the other hand, if the demand is lower than the inventory level, for each unit of excess inventory, the hospital incurs a unit holding cost h . At the end of each period, all on-hand inventory ages by one, and products reaching age L are disposed of, for which the hospital incurs a unit outdating cost w .

Let T denote the planning horizon. For $i = 1, \dots, L-1$ and $t = 1, \dots, T+1$, let $x_{i,t}$ denote the inventory level of age i at the beginning of period t , and let $\mathbf{x}_t = (x_{1,t}, \dots, x_{L-1,t})$ denote the inventory vector. Let z_t denote the ordering quantity at period t , and let $y_t = \sum_{i=1}^{L-1} x_{i,t} + z_t$ denote the total inventory level after ordering. Let D_t denote the (random) demand at period t . Then, the single-period cost is defined as follows:

$$G(\mathbf{x}_t, z_t) = cz_t + p(D_t - y_t)^+ + h(y_t - D_t)^+ + w(x_{L-1,t} - D_t)^+.$$

The objective is to find an optimal ordering policy that minimizes the expected (discounted) total cost during the planning horizon. Let $C_t(\mathbf{x}_t)$ denote the optimal expected cost-to-go function at period t given the system state $\mathbf{x}_t$. Let α denote the discount factor. Let $C_{T+1}(\mathbf{x}_{T+1}) = -c \sum_{i=1}^{L-1} x_{i,T+1}$. Then, the optimality equation for this problem is as follows:

$$C_t(\mathbf{x}_t) = \min_{z_t \geq 0} \{ \mathbb{E}[G(\mathbf{x}_t, z_t) + \alpha C_{t+1}(\mathbf{x}_{t+1})] \},$$

where $x_{i,t+1} = (x_{i-1,t} - (D_t - \sum_{k=i}^{L-1} x_{k,t})^+)^+$, $i = 1, \dots, L-1$, and $x_{0,t} = z_t$.

Theoretically, the optimal ordering policy for this problem can be obtained by using standard backward dynamic programming. However, given that the state space is multi-dimensional, determining an optimal policy is computationally intractable for realistic-size problems. Therefore, the existing literature on blood and perishable inventory management primarily focuses on deriving structural properties of the optimal ordering policy and developing heuristic/approximation policies. We start by presenting a key structural property of the optimal ordering policy below.

Theorem 18.1 *The optimal ordering quantity z_t^* is decreasing in the inventory levels. Moreover,*

$$-1 \leq \frac{\partial z_t^*}{\partial x_{1,t}} \leq \frac{\partial z_t^*}{\partial x_{2,t}} \leq \dots \leq \frac{\partial z_t^*}{\partial x_{L-1,t}} \leq 0.$$

Theorem 18.1 characterizes how the optimal ordering quantity in a blood/perishable inventory system changes in the inventory levels of different ages. In particular, Theorem 18.1

states that the optimal ordering quantity is decreasing in the inventory levels of all ages, and is more sensitive to the inventory levels of fresher products. The key reason why the optimal ordering quantity is more sensitive to fresher inventory is that products of similar ages have a stronger substitution relationship. As a result, an increase in younger inventory has a greater effect on reducing the optimal ordering quantity (of age-0 product). In contrast, an increase in older inventory has a smaller effect because these products will be outdated soon (if not used to meet demand).

Theorem 18.1 also states that the optimal ordering quantity has a bounded sensitivity in the inventory levels (i.e., $\frac{\partial z_t^*}{x_{i,t}} \geq -1, i = 1, \dots, L-1$), which implies that the optimal inventory level after ordering is non-decreasing in the initial inventory level.

We remark that while Fries (1975) and Nahmias (1975) primarily focus on stationary (or stochastically non-decreasing) demand, the result in Theorem 18.1 continues to hold under general nonstationary demand. This is later proven using L^1 -convexity or multimodularity (Chen et al., 2014; Li and Yu, 2014).

When demand is independent and identically distributed (iid), Fries (1975) and Nahmias (1975) also establish the condition on when the ordering quantity is strictly positive. Let $\Phi(\cdot)$ denote cumulative distribution function (cdf) of D_t , and let $B = \Phi^{-1}\left(\frac{p-c}{p+h-\alpha c}\right)$. Then:

Theorem 18.2 *Suppose demand over time is iid. Then, if $\sum_{i=1}^{L-1} x_{i,t} \geq B$, we have $z_t^* = 0$. On the other hand, if $\sum_{i=1}^{L-1} x_{i,t} < B$, we have $z_t^* > 0$ and $y_t^* \leq B$.*

Theorem 18.2 defines an ordering region (i.e., $\sum_{i=1}^{L-1} x_{i,t} < B$) for the perishable inventory management problem when demand is stationary. Specifically, whenever the total inventory level $\sum_{i=1}^{L-1} x_{i,t}$ exceeds a certain threshold B (where B is defined as the optimal inventory level of a nonperishable system with the same cost parameters as the current system), the optimal policy is to order nothing regardless of the age distribution. On the other hand, if the total inventory level falls within the ordering region, the optimal ordering quantity is strictly positive, and after ordering, the total inventory level still falls within the ordering region (i.e., $y_t^* \leq B$).

Collectively, Theorems 18.1 and 18.2 characterize important structural properties of the optimal inventory policy for a blood/perishable inventory system, which deepens the understanding of the structure of the problem. In the meantime, these results also imply that the optimal inventory policy can be a complex function of the inventory levels of different ages, which can be difficult to compute and implement. Given the complex structure of an optimal policy, later efforts on blood/perishable inventory systems mainly focus on developing heuristic and/or approximation policies. We next briefly describe a few recent studies on blood/perishable inventory management.

Among the developed policies in the blood/perishable inventory management literature, a base-stock policy (also known as an order-up-to policy), under which the total inventory is replenished up to the same level at each period, is particularly popular due to its simplicity and near-optimal performance (Chazan and Gal (1977); Chen et al. (2014); Cohen (1976); Cooper (2001); Nahmias (1976); Nandakumar and Morton (1993); Zhang et al. (2018)). While a base-stock policy is easy to interpret and implement and hence appealing from a practical perspective, one drawback of such a policy is that the ordering decision is only based on the total inventory level in the system, but not the age distribution, which is typically suboptimal for

a perishable inventory system. Additionally, the above studies primarily focus on iid demand under which the base-stock level is constant for all periods, which can be restrictive in practice where demand is usually nonstationary over time. Motivated by a real problem for platelet inventory management, Haijema et al. (2007) present a new heuristic policy that modifies the base-stock policy in two directions: (a) a double-level base-stock policy, where one level corresponds to the total inventory and the other, to “young” platelets; and (b) different base-stock levels on different days of a week. They show that a double-level base-stock policy is particularly valuable when there are two classes of demand, one can be satisfied by blood products of any age and the other is preferably satisfied by young products. A base-stock policy with different base-stock levels on each day of a week is particularly relevant given the natural weekly periodic demand pattern, and has been considered by other studies and implemented in real-world blood banks (Fontaine et al. 2009).

In addition to base-stock policies and modified base-stock policies, another simple policy that has been studied in the literature is a constant order policy, under which a constant number of units is ordered at each period (Brodheim et al. (1975); Deniz et al. (2010)). Although a constant order policy may seem too simple to perform well, Deniz et al. (2010) show that it can outperform a base-stock policy when there is age-differentiated demand. Another advantage of a constant order policy is that the ordering quantities are stable and predictable over time, which could be preferred by suppliers. Based on our conversations with blood bank managers in US hospitals, we learned that a variant of the constant order policy, which has a constant ordering quantity on each day of a week (where manual adjustments are allowed based on available inventory), has been a popular policy in practice.

While all of the aforementioned studies assume that blood units can be ordered in each period, Zhou et al. (2011) study a platelet inventory management problem where regular orders can be placed every other day, and between two consecutive regular orders the decision-maker has the option to place a costly expedited order. The authors show that with expedited orders, the cost function is not necessarily convex in general. Nevertheless, they prove the existence and uniqueness of an optimal policy and present an algorithm for obtaining the optimal ordering quantities.

A more recent study by Chen et al. (2020) also studies a platelet inventory management problem by considering both regular replenishments and mid-cycle adjustment opportunities. Unlike Zhou et al. (2011), who only consider expedited orders in the mid-cycle, Chen et al. (2020) consider both expedited orders and product returns. Further, they fully characterize the structure of the optimal inventory policy for both regular and mid-cycle adjustment opportunities.

Platelet inventory management has also been studied by Chen et al. (2019), where the authors consider joint decisions for blood collection and platelet inventory management. The authors characterize the structure of the optimal policy. Interestingly, while the optimal platelet production quantity is non-increasing in the platelet inventory level, the optimal blood collection effort may increase with the platelet inventory level. The authors further demonstrate the value of joint decision-making using real-life data.

A common assumption made in the aforementioned studies is that demand in different periods is independent. Moreover, there is a lack of a worst-case performance guarantee for the developed heuristic policies, such as the base-stock policy. More recently, there is a stream of work that studies approximation policies for stochastic inventory systems under nonstationary and correlated demand processes. The high-level idea of these policies is to balance

the marginal costs of under-ordering and over-ordering, and worst-case performance bounds have been developed for these policies. Among this stream of studies, Chao et al. (2015) were the first to study a balancing-type policy for perishable inventory systems. They propose a proportional-balancing policy for nonstationary and correlated demand and show that it has a worst-case performance bound between two and three. Such a policy is later extended to consider fixed ordering costs, capacity constraints, and positive lead time (Chao et al. (2017); Zhang et al. (2016)).

While a balancing-type policy is appealing in that it captures a general demand process and that it admits a worst-case performance guarantee, Zhang et al. (2021b) observe that in blood inventory systems with high shortage penalties, simply balancing the underage and overage costs can lead to under-ordering. Accordingly, they present a truncated-balancing policy and show that it has a worst-case performance guarantee of two when FIFO is an optimal issuing policy. Moreover, it can significantly outperform the existing policies in blood inventory systems with high shortage penalties.

18.2.2 Optimal Inventory Decisions for Multi-Location Systems

While single-location blood and perishable inventory systems have received extensive attention in the existing literature, fewer studies have considered blood and perishable inventory management in multi-location systems. In this section, we first outline the models and key results of two related papers, Prastacos (1981) and Zhang et al. (2021a). They both study the management of blood and perishable inventory in multi-location systems, and they differ in that Prastacos (1981) studies the optimal allocation of blood products among multiple locations (with exogenous supply), whereas Zhang et al. (2021a) study the jointly optimal transshipment and ordering decisions in a two-location system.

18.2.2.1 Optimal allocation

We first outline the model and key results presented in Prastacos (1981) on blood allocation. Consider a regional blood center that periodically produces and allocates fresh blood products across n hospitals in the region. Once produced, a product is distributed to one of the n hospitals. Each hospital faces a stochastic demand, and products are used to satisfy demand in a FIFO manner. If demand exceeds the available inventory at a hospital, unmet demand is satisfied from outside the region or lost (i.e., there is no backlog of demand), at a unit cost p . No transshipment is allowed during the period. At the end of each period, all on-hand inventory ages by one, and products reaching age L are disposed of, at a unit cost w .

The decision faced by the regional center is to determine how to allocate the inventory in the entire region across the n locations to minimize the long-run average expected shortage and outdating costs. Assume that the regional center considers a rotation allocation policy. That is, at the end of each period, after products of age L are disposed of, the remaining inventory at all hospitals is returned to the regional center. Then, the center allocates its freshly produced inventory and all the returned inventory across the n hospitals. The rotation allocation policy can be thought of as transshipments across locations. Since transshipment costs are typically relatively small compared to shortage and outdating costs, no transshipment cost is considered in Prastacos (1981).

Let $Q_t(i)$ denote the quantity of age- i inventory that is available across the entire region at the beginning of period t ; $i = 0, \dots, L-1$. Let $D_{j,t}$ denote the demand at location j at period t ;

$j = 1, \dots, n$; and let D_t denote the total demand across the entire region (i.e., $D_t = \sum_{j=1}^n D_{j,t}$). When the available fresh inventory (i.e., $Q_j(0)$) and the total demand (i.e., D_t) are stationary over time, let EQ and ED denote their expectation, respectively. Let $A_{j,t}(i)$ denote the quantity of age- i inventory that is allocated to location j at the beginning of period t . Then, the decision rule at each period t , denoted by π_t , is a function that maps from the system state $\{Q_t(i), i = 0, \dots, L-1\}$ to all possible allocations. Let $\pi = \{\pi_t, t = 1, 2, \dots\}$ denote an allocation policy. We say that π is stationary when the decision rule π_t is the same for all t . Finally, let $ES(\pi)$ and $EW(\pi)$ denote the long-run average expected shortage cost and outdating cost, respectively, under a stationary policy π . The objective is to find an optimal (stationary) policy π^* that minimizes

$$EC(\pi) := pES(\pi) + wEW(\pi).$$

While it may appear that there is a trade-off between the shortage and outdating costs, Prastacos (1981) shows that an optimal allocation policy π^* simultaneously minimizes the average expected shortage and outdating costs, i.e., $ES(\pi)$ and $EW(\pi)$.

Theorem 18.3 *For any stationary policy π , $EQ + ES(\pi) = ED + EW(\pi)$.*

Since both EQ and ED are constants independent of π , Theorem 18.3 implies that for any two policies π and π' , if $ES(\pi) < ES(\pi')$, we must have $EW(\pi) < EW(\pi')$. Therefore, an immediate corollary of Theorem 18.3 is as follows:

Corollary 18.1 *The optimal policy π^* that minimizes $EC(\pi)$ minimizes both $ES(\pi)$ and $EW(\pi)$. As a result, the optimal policy π^* is independent of p and w .*

Corollary 18.1 has important practical implications. In particular, the exact shortage and outdating costs can be difficult to quantify in practice. Corollary 18.1 states that if a policy minimizes the expected shortage cost, then it simultaneously minimizes the expected outdating cost, which implies that the optimal policy is independent of the cost parameters. This is in contrast to the results for a single-location system where there is a trade-off between shortage and outdating costs. The main reason behind this contrast is that in the allocation problem considered in this section, the supply is considered exogenous, in which case the shortage and outdating costs go hand-in-hand.

In addition to these key structural results, Prastacos (1981) also studies a myopic allocation policy and shows that it has several appealing properties. Given the system state $\{Q_t(i), i = 0, \dots, L-1\}$, a myopic allocation policy that minimizes the expected shortage and outdating costs at period t is defined as follows:

1. Allocate the inventory of age $L-1$ such that $\mathbb{P}(D_{j,t} \leq A_{j,t}(L-1))$ is equal for all $j = 1, \dots, n$;
2. Allocate the inventory of age $L-2$ such that $\mathbb{P}(D_{j,t} \leq A_{j,t}(L-1) + A_{j,t}(L-2))$ is equal for all $j = 1, \dots, n$; and
3. Continue the same process for age $L-3, \dots, 0$. After the allocation of age 0 products, $\mathbb{P}(D_{j,t} \leq A_{j,t}(L-1) + \dots + A_{j,t}(0))$ is equal for all $j = 1, \dots, n$.

Prastacos (1981) further shows that the myopic allocation policy “is not very different” from the optimal allocation policy π^* in the following sense. Let R_π denote the long-run stationary

(random) amount of return to the regional center at each period (i.e., the amount of remaining inventory that is not used at the n locations, after removing the outdated quantities). Intuitively, a smaller R_π implies that the inventory was allocated to the right locations and hence can lead to a lower total cost. Nevertheless, Prastacos (1981) shows that no stationary policy π whose R_π is stochastically smaller than R_M can do better than the myopic policy M , which implies that the myopic policy M “cannot be very different” from the optimal policy. For brevity, we omit the technical details and refer an interested reader to Prastacos (1981) for a full discussion of this result.

18.2.2.2 Optimal transshipment and ordering

One key assumption in Prastacos (1981) is that the remaining inventory at all locations will be returned to the regional center and reallocated at each period. While this can be a reasonable assumption when transshipment across locations is convenient and inexpensive, collecting all of the remaining inventory across the region may be difficult to implement in practice. We next outline the model and key results presented in a recent paper by Zhang et al. (2021a), who study a blood inventory-sharing problem in a two-location system by explicitly modeling transshipment costs. Unlike Prastacos (1981), who focuses on the allocation decision at each period, the two key decisions considered by Zhang et al. (2021a) are (a) the transshipment quantity between the two locations after demand realization at each period and (b) the ordering quantity at each location at the beginning of each period.

Consider two locations (e.g., hospitals) and a single product (e.g., platelets) with an L -period lifetime. Each location faces a stochastic demand. At the beginning of each period, each hospital orders fresh (i.e., age 0) blood products with a unit cost c . The ordered products arrive with zero lead time. After demand realization, products can be transshipped from one location to another with a unit cost r . Products are only transshipped from one location to the other when they are used to meet demand at the latter location during the current period (i.e., products are not transshipped to be stored at the other location), and products at each location are issued to meet demand in a FIFO manner. Unmet demand is satisfied from outside sources, with a unit cost p . At the end of each period, all on-hand inventory ages by one, and products reaching age L are disposed of, with a unit cost w .

More specifically, for $i = 1, \dots, L-1$, $j = 1, 2$, and $t = 1, \dots, T+1$ let $x_{i,t}^j$ denote the inventory level of age i at location j at the beginning of period t , where T denotes the planning horizon. Let $\mathbf{x}_t^j = (x_{1,t}^j, \dots, x_{L-1,t}^j)$. Let z_t^j denote the ordering quantity at location j at period t , and let $y_t^j = \sum_{i=1}^{L-1} x_{i,t}^j + z_t^j$. Let D_t^j denote the (random) demand at location j at period t . Assume that demand across the two locations can be correlated but is independent over time. Finally, let u_t denote the number of units that are transshipped from location 1 to location 2 at period t (where a negative u_t implies transshipment from location 2 to location 1). Since products are only transshipped to meet demand during the current period, we have $-D_t^1 \leq u_t \leq D_t^2$. Then, the single-period cost of the two-location system is as follows:

$$\begin{aligned} G(\mathbf{x}_t^1, \mathbf{x}_t^2, z_t^1, z_t^2, u_t) = & cz_t^1 + cz_t^2 + p(D_t^1 + u_t - y_t^1)^+ \\ & + p(D_t^2 - u_t - y_t^2)^+ + w(x_{L-1,t}^1 - D_t^1 - u_t)^+ \\ & + w(x_{L-1,t}^2 - D_t^2 + u_t)^+ + r(u_t)^+ + r(-u_t)^+. \end{aligned}$$

The objective is to find optimal transshipment and ordering policies that minimize the expected (discounted) total cost. Let $C_t(\mathbf{x}_t^1, \mathbf{x}_t^2)$ denote the optimal expected cost-to-go function at the beginning of period t . Let α denote the discount factor, and let $x_{0,t}^j = z_t^j$, $j = 1, 2$. Then, the optimality equation is as follows:

$$C_t(\mathbf{x}_t^1, \mathbf{x}_t^2) = \min_{z_t^1, z_t^2 \geq 0} \left\{ \mathbb{E} \left[\min_{-D_t^1 \leq u_t \leq D_t^2} \left\{ G(\mathbf{x}_t^1, \mathbf{x}_t^2, z_t^1, z_t^2, u_t) + \alpha C_{t+1}(\mathbf{x}_{t+1}^1, \mathbf{x}_{t+1}^2) \right\} \right] \right\},$$

where $\mathbf{x}_{t+1}^1, \mathbf{x}_{t+1}^2$ are defined as follows: $x_{i,t+1}^1 = \left(x_{i-1,t}^1 - \left(D_t^1 + u_t - \sum_{k=i}^{L-1} x_{k,t}^1 \right)^+ \right)^+$, and $x_{i,t+1}^2 = \left(x_{i-1,t}^2 - \left(D_t^2 - u_t - \sum_{k=i}^{L-1} x_{k,t}^2 \right)^+ \right)^+$, $i = 1, \dots, L-1$. Moreover, let $C_{T+1}(\mathbf{x}_{T+1}^1, \mathbf{x}_{T+1}^2) = -c \left(\sum_{i=1}^{L-1} x_{i,T+1}^1 + \sum_{i=1}^{L-1} x_{i,T+1}^2 \right)$.

As discussed in Section 18.2.1, finding an optimal inventory policy for a blood and perishable inventory system is particularly challenging even for the single-location case. The state space for a two-location problem is further enlarged. Nevertheless, Zhang et al. (2021a) uncover a simple structure for the direction of optimal transshipment, by showing that the direction of transshipment can be easily determined by comparing a scalar at each location. At any period t , let a_t^j denote the age of the oldest product at location j after meeting demand and $a_t^j = -1$ if nothing is left in inventory at location j (assuming that there is no transshipment). Then, the direction of the optimal transshipment quantity u_t^* is characterized in the following theorem.

Theorem 18.4 $u_t^* \geq 0$ if $a_t^1 > a_t^2$; $u_t^* \leq 0$ if $a_t^1 < a_t^2$; and $u_t^* = 0$ if $a_t^1 = a_t^2$.

Theorem 18.4 says that although the state space of the problem is multi-dimensional, it is sufficient to compare the age of the oldest product at each location after meeting demand (assuming that there is no transshipment) to determine the direction of transshipment. In particular, if the age of the oldest product at location 1 is larger than that at location 2, products should be transshipped from location 1 to location 2, and vice versa.

Based on this and other structural results, Zhang et al. (2021a) further show that a myopic transshipment policy, under which transshipment is used either for reducing shortages or for reducing the current-period outdates, is optimal for a two-period lifetime case and serves as a lower bound on the optimal transshipment policy for longer lifetimes. Specifically, let u_t^M denote the myopic transshipment quantity at period t . Then:

Theorem 18.5 *There exists a threshold $\bar{r} \leq w + \beta c$ such that if $r \leq \bar{r}$, then $u_t^M = u_t^*$ for $L=2$ and $|u_t^M| \leq |u_t^*|$ for $L > 2$.*

Theorem 18.5 shows that if the unit transshipment cost is relatively small, the myopic transshipment policy provides a lower bound on the optimal transshipment policy. That is, whenever there is a shortage or outdate in the current period, it is optimal to trigger transshipment. Moreover, when the product has a two-period lifetime, the myopic transshipment policy is exactly the optimal transshipment.

From a practical perspective, a myopic transshipment policy is also appealing because it is simple and easy to interpret. Although the optimal transshipment policy u_t^* may suggest an even larger transshipment quantity than the myopic policy u_t^M , Zhang et al. (2021a) show

via extensive numerical analyses that the myopic policy u_i^M performs close to optimal under a wide range of model parameters. This is mainly because although the transshipment of younger products can help balance the inventory at the two locations and hence reduce future outdates, such transshipment can be largely made up by future transshipment opportunities. As a result, focusing on the transshipment of products with one-day remaining shelf-life can perform extremely well.

We further note that the finding that a myopic transshipment focusing on the transshipment of products with one-day remaining shelf-life performs (close to) optimal parallels the findings of Li et al. (2016), who underscore that the inventory of products with one-day remaining shelf-life plays an important role in determining the optimal ordering quantity in a single-location perishable inventory system where products are issued in a last-in-first-out (LIFO) manner.

Finally, we briefly discuss how the presence of transshipments affects the optimal ordering decisions for blood inventory systems. In particular, while conventional wisdom suggests that inventory pooling typically reduces the optimal inventory level, Zhang et al. (2021a) show that for perishable products, inventory pooling may instead encourage storing more. The key intuition behind this result is that pooling the inventory of the two locations allows the two locations to consume each other's old inventory, which helps reduce the concern about outdating, and hence can result in a higher optimal total inventory level. Additionally, Zhang et al. (2021a) further show that inventory pooling can lead to a strictly higher total inventory even when one location has a deterministic demand. This result implies that inventory pooling may be valuable for hedging against demand uncertainty for blood and perishable products even if demand at one location is deterministic, which is not the case for nonperishable products. Overall, these results highlight that the established findings for nonperishable products do not necessarily carry over to blood and perishable products, and more research is needed to study inventory pooling for blood/perishable products in multi-location systems.

Unlike single-location blood inventory systems, which have been extensively studied in the literature, much fewer studies have studied blood inventory sharing in multi-location systems (see Li et al. (2021); Zhang et al. (2021a) for recent reviews of the literature). In the remainder of this section, we briefly discuss the key findings in a few other papers on blood/perishable inventory allocation and/or transshipment decisions.

Jennings (1973) discusses the value of blood inventory pooling at a regional level, and numerically shows that pooling the inventory across multiple locations reduces both the shortage and outdating rates compared to single-location cases.

Built on the allocation policy presented in Prastacos (1981), Prastacos and Brodheim (1980) present a decision support framework that jointly determines how many blood products to collect at the regional center and how to allocate the (fresh and old) blood products across the region. The proposed framework has been implemented by a 38-hospital region in Long Island, New York.

A regional blood allocation problem is also studied by Kendall and Lee (1980). In particular, they formulate a goal programming model to consider multiple objectives, where the goal constraints include target inventory levels, availability of fresh blood, average age of the inventory, outdating costs, and the costs of collecting blood.

Federgruen et al. (1986) consider a one-period regional blood allocation problem but extend the framework considered in Prastacos (1981) by considering (a) location-dependent shortage and wastage costs, (b) a unit transportation cost from the center to each location, and (c) a positive initial inventory at each location (i.e., products are not necessarily returned to the center

for reallocation). Federgruen et al. (1986) show that under this general framework, the allocation policy presented in Prastacos (1981) is no longer optimal. Federgruen et al. (1986) then present a Lagrangian relaxation approach for solving the proposed problem. Additionally, Federgruen et al. (1986) further extend their model to consider joint inventory allocation and routing decisions for perishable products. We refer the reader to a recent paper by Crama et al. (2018) for more discussion on inventory routing decisions for perishable products.

Perishable inventory allocation/transshipment has also been studied in a recent paper by Li et al. (2021). In particular, they consider a two-location perishable inventory system where the decisions at each period are to determine: (a) how many fresh products to purchase at each location and (b) how many products of each age to allocate between the two locations (without incurring transshipment costs). Moreover, they focus on a setting where products at each location are used to meet demand in a last-in-first-out manner. Interestingly, Li et al. (2021) show that unlike the FIFO case considered in Prastacos (1981) where products of each age should be “evenly” allocated across different locations, when LIFO is used, it may be optimal to “separate” the old and new products (i.e., store old products at one location and new products at the other location).

18.2.3 Optimal Issuing Decisions

In addition to ordering, allocation, and transshipment decisions, another key decision in blood inventory systems is the issuing decision, that is, to determine which products (older vs. fresher) to use to meet demand. In this section, we start by outlining the model and key results in one of the seminal papers on optimal issuing decisions for blood and perishable inventory systems (Pierskalla and Roach (1972)). Then, we briefly describe several recent papers and the findings therein on optimal issuing decisions.

Consider a single location (e.g., a hospital) that manages the inventory of a single product (e.g., RBCs or platelets). As before, let L denote the lifetime of the product. The replenishment of blood is an exogenous process, and the supply can be products of different ages. For $i = 0, \dots, L-1$, let $x_{i,t}$ denote the inventory level of age- i product at the beginning of period t after replenishment. Further, assume that there are different categories of demand, and let $D_{i,t}$ denote the demand for products of age i or less (i.e., only products of age i or less can be used for fulfilling this demand) at period t . After demand realization, the decision is to determine which products to use to satisfy demand. Let $I_{i,t}$ denote the remaining inventory level of age i after products are issued to meet demand at period t . Finally, at the end of each period, all on-hand inventory ages by one and products reaching age L are disposed of.

Pierskalla and Roach (1972) consider both the backlogging and lost sales cases, and define the following three objective functions to study the optimality of a FIFO issuing policy, which uses the oldest possible products from inventory to meet demand:

1. The first objective maximizes the value of all demands that have been satisfied plus the value of the remaining products in inventory. In particular, for $i = 0, \dots, L-1$, let V_i denote the value of an age- i product, where $V_0 \geq V_1 \geq \dots \geq V_{L-1}$. Let $R_0 = 0$ and $I_{i,0} = 0, i = 0, \dots, L-1$. Then, the cumulative value in period t is

$$R_t = R_{t-1} + \sum_{i=0}^L V_i D'_{i,t} + \sum_{i=0}^{L-1} V_i I_{i,t} - \sum_{i=1}^{L-1} V_i I_{i,t-1} = \sum_{s=1}^t \sum_{i=0}^L V_i D'_{i,s} + \sum_{i=0}^{L-1} V_i I_{i,t},$$

where $D'_{i,t}$ denotes the demand for products of age i or less that is met at period t . Then, the first objective is to maximize R_p , where T is the planning horizon. Under this objective, since meeting demand in different classes results in different values, it is without loss of generality to assume that when products of any given age are issued, they are prioritized to meet the highest value demand possible.

2. The second objective minimizes the total number of shortages over the planning horizon.
3. The third objective minimizes the total number of outdates over the planning horizon.

Theorem 18.6 *If unmet demand is backlogged, then FIFO minimizes all three objectives. If unmet demand is lost, then FIFO minimizes the latter two objectives.*

Theorem 18.6 characterizes the optimality of the FIFO issuing policy in a broad set of circumstances. In particular, it says that for both the backlogging and lost sales cases, the FIFO issuing policy simultaneously minimizes the total number of shortages and the total number of outdates. The main reason for the absence of a trade-off between shortages and outdates is that the decision considered here is the issuing decision while the supply is given. In this case, FIFO issues the oldest products first and hence intuitively minimizes the number of outdates. When the total supply is fixed, a smaller number of outdates implies that more units are used to meet demand, which implies a smaller number of shortages. It is also important to note that this result holds even if there are multiple demand classes.

Theorem 18.6 also implies that when unmet demand is backlogged, FIFO also maximizes the total value of the system defined in the first objective. However, this conclusion does not extend to the lost sales case. A counter-example is presented in Pierskalla and Roach (1972) to explain the underlying intuition, which we omit here for brevity. Additionally, we also remark that one assumption made in the first objective is that the value of satisfying a certain class of demand depends only on the demand class, but not on the age of the product used to satisfy demand (as long as it meets the age requirement). As we will discuss in more detail below, if using a younger product to meet a given demand creates more value for the patient than using an older product, simply applying a FIFO issuing policy may no longer be optimal.

Overall, Theorem 18.6 has important practical implications. In particular, FIFO is an issuing policy that is easy to interpret and implement, and hence is appealing from a practical perspective. Theorem 18.6 provides theoretical support for this simple issuing policy by showing that it is optimal under various objective functions.

The optimality of FIFO has also been studied in a number of other studies. For example, Fries (1972) shows that under iid demand, as long as the total inventory level lies within the ordering region characterized in Theorem 18.2, a younger inventory vector results in a lower expected cost, which implies that issuing older products before younger ones (i.e., FIFO) minimizes the expected total cost.

The optimality of FIFO is also revisited in a recent study by Zhang et al. (2021b). In particular, while the condition presented in Fries (1972) is primarily on the demand process (i.e., iid demand) and the condition presented in Pierskalla and Roach (1972) is primarily on the cost parameters (i.e., there are shortage and outdating costs but no holding costs), Zhang et al. (2021b) extend these findings by presenting a more general condition that characterizes the interplay between the demand process and the cost parameters. More specifically, Zhang et al. (2021b) show that FIFO is optimal when demand does not decrease too much (in a stochastic sense) over time; the less the demand decreases, the less the requirement for cost parameters.

Given its optimality in a broad set of circumstances, FIFO has been a popular issuing policy in practice for the management of certain blood products such as platelets (Zhang et al. (2021a)). In particular, platelets have a short shelf-life of approximately three days after testing, transportation, etc. As a result, there is typically no specific requirement on platelet age for transfusion. For these reasons, FIFO has been considered for the management of platelet inventory in many circumstances (van Sambeek Jhj et al. (2021); Zhang et al. (2021a)).

The issuing policy for RBCs, however, can be more complex. In particular, RBCs can be stored for up to 42 days, and the transfusion of younger RBCs has been reported to be associated with improved health outcomes (Abbasi and Hosseinifard (2014); Atkinson et al. (2012); Sarhangian et al. (2018)). In this case, a FIFO issuing policy may not be preferred because issuing the oldest product first can intuitively result in a large age at issuing. Therefore, there is a trade-off between minimizing the shortage and outdated costs and reducing the age at issuing. To this end, several existing studies propose simple-to-use issuing policies that aim at combining the advantages of FIFO and LIFO issuing policies, where FIFO minimizes the shortage and outdated costs whereas LIFO minimizes the age at issuing.

In particular, Atkinson et al. (2012) study an issuing policy that is characterized by a single threshold on product age. For a given unit of demand, if there are products in the inventory that are younger than the threshold, then the patient receives the oldest product (i.e., FIFO) among all the products that are younger than the threshold. If all products in the inventory are older than the threshold, then the patient receives the youngest product (i.e., LIFO) among all the products. More recently, the threshold policy considered by Atkinson et al. (2012) is revisited by Sarhangian et al. (2018), who present an exact analysis to characterize the age distribution of the issued products as well as the number of shortages and outdates under this policy.

In a similar spirit, Abbasi and Hosseinifard (2014) also present an issuing policy that is characterized by a single threshold on product age that partitions the inventory into two groups. However, different from Atkinson et al. (2012), Abbasi and Hosseinifard (2014) propose to consider FIFO within each group but LIFO across the two groups. More specifically, for a given unit of demand, if there are products in inventory that are younger than the threshold, then the patient receives the oldest product (i.e., FIFO) among all the products that are younger than the threshold. If all products in the inventory are older than the threshold, then the patient receives the oldest product (i.e., FIFO) among all the products. In this policy, the spirit of LIFO is reflected in that the younger group is considered before the older group.

Similar to Pierskalla and Roach (1972), Abouee-Mehrizi et al. (2019) consider a multi-class demand problem where each demand class only accepts products with a remaining lifetime longer than a certain threshold. They show that if unmet demand is backlogged and the backlogging costs for demand in different classes are different, the optimal issuing policy is a sequential rationing policy, where demand with a higher priority should be satisfied first, and within each priority level, older products should be used first to meet demand (i.e., FIFO). Moreover, fresher products should be rationed to certain thresholds for future use.

We conclude this section by noting that for the management of blood inventory (especially for that of RBCs), it is also important to track the ABO types of the blood products and to ensure compatibility when an issuing decision is made. We refer an interested reader to Atkinson et al. (2012) and van Sambeek Jhj et al. (2021) for considerations of ABO types in blood issuing decisions.

18.3 INVENTORY MANAGEMENT FOR PHARMACEUTICAL PRODUCTS

In recent years, pharmaceutical supply chains have received increasing attention in the operations management literature (e.g., Adida (2021); Alev et al. (2021); Bandi et al. (2019); King et al. (2019); Kouvelis et al. (2015); Zhao et al. (2012)). Effective inventory management for pharmaceutical products is critical as it helps to maintain a steady supply of life-saving products while minimizing operating costs. Yet, the management of pharmaceutical inventory is particularly challenging due to high market concentration, complex manufacturing processes that are vulnerable to quality problems, supply disruptions, and multi-stakeholder incentive misalignment. As a result, the United States has seen a rapid increase in drug shortages in recent years, and drug shortages have been considered a national healthcare crisis (Fox et al. (2014); Ventola (2011)).

Maintaining the availability of essential medicines has been an even greater challenge for many developing countries. In particular, many medicines have been unaffordable to the majority of the population in the developing world. Because of the low and uncertain demand, local retailers in these countries are often not incentivized to carry a sufficient inventory of many essential drugs (e.g., malaria drugs) to meet patient demand. The low availability of essential medicines has been a major hurdle to improving health outcomes in many low-income countries (Taylor and Xiao, 2014).

In this section, we first discuss strategies for mitigating drug shortages in the United States (or developed countries) by highlighting the multi-stakeholder nature of a pharmaceutical chain and the effect of pharmaceutical contracting on drug availability. Then, we discuss strategies for improving drug availability in developing countries by highlighting the role of government subsidies as well as fund availability.

18.3.1 Mitigating Drug Shortages in the United States

In the United States, generic sterile injectable drugs account for the majority of drug shortages. These drugs are particularly vulnerable to shortages for several reasons. First, the profit margin for generic sterile injectable drugs is often very low. As a result, manufacturers are not incentivized to invest in capacity for these drugs, which in turn leads to a high market concentration and high capacity utilization. Second, the manufacturing processes for generic sterile injectable drugs are complex and vulnerable to quality problems. Once a (severe) violation of the guidelines is detected, the manufacturer needs to halt the production and sometimes even shut down the facility until the FDA approves the resumption of the production. The high utilization of production lines leaves little time for maintenance, and hence can further increase the likelihood of such quality-related disruptions. In this section, we first outline a base model presented by Jia and Zhao (2017) that captures these challenges and trade-offs. We then briefly discuss the key findings in a few other papers on drug shortages.

Consider a drug supply chain that consists of a drug *manufacturer*, a group of *healthcare providers*, a *group purchasing organization* (GPO) that represents the providers, and the *government/payer* (e.g., Medicare) who reimburses the spending. Consider a purchase contract between the manufacturer and the GPO, which specifies two key parameters: the unit wholesale price of the drug p and a failure-to-supply (FTS) penalty f for each unit of shortages. Jia

and Zhao (2017) consider that the GPO is the contract designer. Next, we define the payoff function of each stakeholder under a given contract (p, f) .

We start by defining the payoff function of the GPO. Let D denote the random demand for the drug, which is assumed to be independent of the price p because many drugs that experience shortages are considered medically necessary. Given a purchase contract (p, f) , let $L(p, f)$ denote the percentage of the expected lost sales over the expected demand $\mathbb{E}[D]$. The GPO's objective is to maximize its profit, which primarily comes from a commission fee charged from manufacturers as a percentage of the total sales to the GPO members. Let α denote the percentage of the total sales the GPO charges. Therefore, the GPO's expected profit is

$$\Pi_O(p, f) = \alpha p(1 - L(p, f))\mathbb{E}[D].$$

We next define the payoff function of the drug manufacturer. The manufacturer's decisions include (a) the production capacity decision for a T -period planning horizon, and (b) the inventory decision in each period. Suppose at the beginning of the planning horizon, the manufacturer's initial capacity for the drug of interest is κ_0 . The manufacturer can adjust its production capacity to $\kappa \geq \kappa_0$ with a cost of $G(\kappa - \kappa_0)$. Further, given the capacity level κ , let $U(p, f, \kappa)$ denote the manufacturer's expected profit by following an optimal inventory policy (which is defined below). Then, the manufacturer's problem is to choose the capacity level κ to maximize

$$\Pi_M(p, f) = \max_{\kappa \geq \kappa_0} U(p, f, \kappa) - G(\kappa - \kappa_0).$$

We next define the government's payoff. Given the unit wholesale price p , assume that the government reimburses $r(p)$ to the providers for each unit of demand that is met. When there are lost sales, the government considers a societal cost β_G for each unit of unmet demand. Then, the government's expected cost is as follows:

$$\Pi_G(p, f) = r(p)(1 - L(p, f))\mathbb{E}[D] + \beta_G L(p, f)\mathbb{E}[D].$$

We finally define the health providers' payoff. The health providers purchase the drug from the manufacturer and are reimbursed by the payer at a reimbursement price denoted by $r(p)$. For each unit of demand that is met, the health providers incur a cost of $p - r(p)$. In Jia and Zhao (2017), it is further assumed that the reimbursement function $r(p)$ satisfies $r(p + \delta) - r(p) \geq \delta, \forall \delta > 0$. One example provided by Jia and Zhao (2017) based on a Medicare reimbursement policy is $r(p) = \gamma p$, where $\gamma \geq 1$. On the other hand, when there are lost sales, each unit of lost sales results in a cost of $\beta_p \geq f$ to the providers due to the negative impact of drug shortages (e.g., suboptimal treatment outcomes). Thus, the providers' expected cost is defined as follows:

$$\Pi_P(p, f) = (p - r(p))(1 - L(p, f))\mathbb{E}[D] + (\beta_p - f)L(p, f)\mathbb{E}[D]$$

Finally, it remains to derive the manufacturer's optimal inventory policy and the corresponding shortage rate under a given contract (p, f) and a given capacity κ . Consider a manufacturer that manages the inventory of a single product over a planning horizon of T periods, with a

production capacity of κ at each period. Assume that the production lead time is one period. As discussed earlier, the production of drugs is subject to disruption risks. When there is disruption, the disruption duration can usually be long and uncertain. In these cases, the unmet demand is usually lost. Such lost sales are generally considered as drug shortages and are subject to FTS penalties. On the other hand, when there is no disruption, the unmet demand is usually backlogged and satisfied in a relatively short amount of time. Typically, such a backlog of demand does not lead to the severe drug shortages discussed above.

When the status of production is normal, assume that there is a probability $1 - \theta(\kappa)$ of disruption, which is a function of the capacity κ because a higher utilization of production lines typically implies a higher risk of disruption (due to less time for maintenance, etc.). When the production status is disrupted, let $i \geq 1$ denote the number of periods since disruption. Given that the production has been disrupted for i periods, assume that there is a probability λ_i that the disruption will continue, and a probability $1 - \lambda_i$ that the production status will be back to normal in the next period.

Let D_t be the demand at each period $t \in \{1, \dots, T\}$. Let $U_t(x)$ denote the manufacturer's optimal expected profit-to-go at period t when the production status is normal and the inventory level at the beginning of period t is x . Let $V_t(i, x)$ denote the manufacturer's optimal expected profit-to-go at period t given that the production status has been disrupted for i periods and the inventory level at the beginning of period t is x . Then, the optimality equations for the manufacturer's problem are as follows:

$$\begin{aligned}
 U_t(x) = & \max_{y \in [x, x+\kappa]} \{-c(y-x) + (1-\alpha)p\mathbb{E}[D_t] - h\mathbb{E}[(x-D_t)^+]\} \\
 & + \theta(\kappa)(-b\mathbb{E}[(D_t-x)^+] + \mathbb{E}[U_{t+1}(y-D_t)]) \\
 & + (1-\theta(\kappa))(-(1-\alpha)p + f)\mathbb{E}[(D_t-x)^+] + \mathbb{E}[V_{t+1}(1, (x-D_t)^+)]); \\
 V_t(i, x) = & (1-\alpha)p\mathbb{E}[\min\{x, D_t\}] - h\mathbb{E}[(x-D_t)^+] - s\mathbb{E}[(D_t-x)^+] \\
 & + \lambda_i\mathbb{E}[V_{t+1}(i+1, (x-D_t)^+)] + (1-\lambda_i)\mathbb{E}[U_{t+1}((x-D_t)^+)].
 \end{aligned}$$

That is, when the production status is normal, the manufacturer chooses a production quantity $y-x$, which results in a production cost $c(y-x)$. For each unit of demand that is eventually met, the manufacturer earns a revenue $(1-\alpha)p$. If there is excess inventory after meeting demand, each unit of excess inventory results in a holding cost h . Suppose the production status remains normal by the end of the period. Then, the unmet demand is backlogged with a unit cost b . The net inventory level at the beginning of the next period is given by $y-D_t$. Suppose the production is disrupted during the period. Then, the new batch of the production $y-x$ fails, and all unmet demand is lost, which results in a revenue loss as well as the FTS penalty. The net inventory level at the beginning of the next period is given by $(x-D_t)^+$.

When the production has been disrupted for i periods, the production status either remains disrupted or returns to normal in the next period, with probability λ_i and $1-\lambda_i$, respectively. The optimality equation is defined in a similar way.

At period $T+1$, let $U_{T+1}(x) = c(x)^+ - ((1-\alpha)p + f)(-x)^+$ and $V_{T+1}(i, x) = cx$. That is, at the end of the planning horizon, in addition to recovering the production cost c , the manufacturer

also returns the revenue p and pays the FTS compensation f for all backlogged quantities if the production status is normal (whereas there are no backorders when the production status is disrupted). Let $y_t^*(x)$ denote the optimal inventory level after production at period t . Then, we have:

Lemma 18.1 *The manufacturer's optimal inventory policy is a base-stock policy. Specifically, there exists a base-stock level s_t at each period t such that*

$$y_t^*(x) = \begin{cases} x + \kappa & \text{if } x \leq s_t - \kappa, \\ s_t & \text{if } s_t - \kappa \leq x \leq s_t, \\ x & \text{if } x \geq s_t. \end{cases}$$

With the manufacturer's optimal inventory policy, its cost function $U(p, f, \kappa)$ is well defined as $U_1(0)$. Then, the optimal capacity level that maximizes $\Pi_M(p, f)$ can be obtained accordingly. Further, with the manufacturer's optimal inventory policy and optimal capacity level, the percentage of the expected lost sales over the expected demand $L(p, f)$ can also be evaluated accordingly.

The remaining question is how to find an optimal contract (p, f) that maximizes $\Pi_O(p, f)$. Although the GPO designs the contract to maximize its own profit, the contract may not be accepted by other stakeholders if it leads to a lower payoff than the status quo. To ensure that all stakeholders are appropriately incentivized, the contract should ideally be Pareto-improving. Specifically, let (p_0, f_0) be the current contract. A contract (p, f) is Pareto-improving if it (weakly) improves the payoffs of all stakeholders compared to the current contract (p_0, f_0) . For any given p , let $f_M(p)$ be the FTS penalty level such that $\Pi_M(p, f_M(p)) = \Pi_M(p_0, f_0)$, and let $f_G(p)$ be the FTS penalty level such that $\Pi_G(p, f_G(p)) = \Pi_G(p_0, f_0)$. Then:

Theorem 18.7 *The optimal Pareto-improving contract (p^*, f^*) satisfies $p^* \geq p_0$ and $f^* \geq f_0$. Moreover, we have $f^* = f_M(p^*)$ and $f^* = f_G(p^*)$.*

Theorem 18.7 implies that an optimal Pareto-improving contract simultaneously improves both the wholesale price and the FTS penalty. Moreover, p^* and f^* can be solved by using the two equations $f^* = f_M(p^*)$ and $f^* = f_G(p^*)$, which imply that the payoffs of the manufacturer and the government remain the same as the status quo. On the other hand, the payoffs of the GPO and the health providers may increase, while the percentage of lost sales may decrease. An important implication of this result is that simply increasing the wholesale price, a solution that is advocated by many, may not be effective enough in mitigating drug shortages. Instead, an increase in the wholesale price needs to be paired with an increase in the FTS penalty. Jia and Zhao (2017) further present a calibrated case study to show that using both the pricing and the FTS penalty levers can be substantially more effective than using the pricing lever alone.

The shortages of generic sterile injectable drugs in the United States have also been studied by Kim and Morton (2015). Unlike Jia and Zhao (2017), who focuses on one monopoly manufacturer, Kim and Morton (2015) consider competition between two manufacturers, where the two manufacturers produce a single product, and their products are perfect substitutes. Each manufacturer's regular production capacity is subject to random disruption, and the two manufacturers compete by setting their spare capacity levels which are used during the disruption period of either manufacturer. Kim and Morton (2015) find several interesting results:

(a) drug availability is particularly sensitive to price when the disruption probability is high; (b) when the price is high, drug availability may decrease when the disruption probability decreases; and (c) without competition, it is optimal to disallow a temporary price increase during a shortage; with competition, however, allowing a temporary price increase during a shortage may increase drug availability.

More recently, Hotkar and Gupta (2021) investigate the drug shortage problem by focusing on the effect of the FDA's mandate that all manufacturers are required to report any disruptions that can potentially cause drug shortages. Specifically, they consider a setting with two competing manufacturers, where each manufacturer chooses a reliable (and more expensive) capacity and an unreliable (and cheaper) capacity. The unreliable capacity of each firm is subject to a probability of disruption. Hotkar and Gupta (2021) show that both manufacturers can benefit from information sharing since it enables better utilization of their production capacity. However, they also show that such information sharing can lead to a lower level of reliable capacity investment, which can then lead to a lower supply and more shortages. They further propose a tax credit approach as a mitigation strategy to encourage the investment of reliable capacity.

A few recent papers have investigated the problem of drug shortages empirically. For example, Yurukoglu et al. (2017) empirically study the effect of a Medicare Part B reimbursement policy change on drug shortages. They show that after the Medicare Part B policy change, drugs that were impacted more by the policy change experienced more shortages. More recently, Lee et al. (2021) empirically studies the role of the FDA's mandate on disruption information sharing and shows that the mandate helps reduce drug shortages, but its effectiveness depends on the level of competition in the market.

Finally, several recent studies also report the implementation of inventory management tools in the pharmaceutical industry with significant organizational benefits. For example, Liu et al. (2013) report the implementation of a spreadsheet model for the management of inventory in pharmacy chain stores. Similarly, Zhang et al. (2014) report that Kroger Co. has implemented a simulation-optimization approach for optimizing their pharmacy inventory, which led to significant reductions in inventory and costs.

18.3.2 Improving Drug Availability in Developing Countries

As discussed earlier, due to the low margin and the uncertain demand for many pharmaceutical products in the developing world, local retailers are often not incentivized to keep sufficient inventory levels to meet patient demand. In this context, government subsidies play an important role in improving drug availability. In this section, we first outline a base model presented by Taylor and Xiao (2014). Then, we briefly discuss the key findings in a few other papers that examine drug availability in developing countries.

Consider a retailer that sells a single product (drug) to consumers in an infinite horizon $t = 1, 2, \dots$. In each period t , the drug demand depends on the market condition M_t , which is independent and identically distributed. At the beginning of each period, given the initial inventory level x , the retailer chooses an ordering quantity $y - x$ to bring the inventory level up to y . After the realization of $M_t = m$, the retailer sets the price of the drug p , which results in a market demand of $md(p)$, where $d(p)$ is continuous, twice differentiable, and strictly decreasing in p . The additional technical assumptions on $d(p)$ presented in Taylor and Xiao (2014) are omitted here for brevity. Examples of $d(p)$ that satisfy all assumptions include $d(p) = a - bp^k$ ($a, b > 0, k \geq 1$), $d(p) = (a - bp)^k$ ($a, b, k > 0$), and $d(p) = a - be^{kp}$ ($a, b, k > 0$).

Consider a donor that aims to maximize consumption subject to a budget constraint. The donor considers two types of subsidies: subsidizing the retailer by a for each unit of product the retailer purchases from its supplier (i.e., purchase subsidy), and subsidizing the retailer by s for each unit of product the retailer sells to consumers (i.e., sales subsidy). Then, given the (a, s) pair, the retailer's problem is as follows:

$$V(x) = \max_{y \geq x} \{- (c - a)(y - x) + \mathbb{E}[\max_{p \geq 0} \{ (p + s) \min\{y, md(p)\} + \alpha V(y - \min\{y, md(p)\}) \}]\},$$

where c denotes the unit ordering cost, and α denotes the discount factor.

Lemma 18.2 *The retailer's optimal inventory policy is a base-stock policy. Moreover, the optimal order-up-to level y^* and the corresponding optimal price given the realization of market condition $p^*(m, y^*)$ can be obtained by solving*

$$\max_{y \geq 0} \{- (c - a)y + \mathbb{E}[\max_{p \geq 0} \{ (p + s) \min\{y, md(p)\} + \alpha(c - a)(y - \min\{y, md(p)\}) \}]\}.$$

Lemma 18.2 says that a base-stock policy is optimal for the retailer, and the optimal base-stock level y^* and the optimal price $p^*(m, y^*)$ can be obtained in a myopic manner. With this characterization of the retailer's decisions, we next define the donor's problem. The donor chooses the subsidy levels a and s to maximize the expected per-period sales, subject to a constraint that the expected per-period subsidy payment does not exceed a budget level B . The donor's problem is formulated as follows:

$$\begin{aligned} \max_{a, s \geq 0} & \mathbb{E}[\min\{y^*, md(p^*(m, y^*))\}] \\ \text{s.t. } & (a + s) \mathbb{E}[\min\{y^*, md(p^*(m, y^*))\}] \leq B. \end{aligned}$$

Taylor and Xiao (2014) show that the purchase subsidy is more effective in boosting the retailer's inventory level, while the sales subsidy is more effective in reducing the retailer's price. Based on these results, they further show when the market condition is strong, the purchase subsidy is more effective in increasing the number of sales; otherwise (i.e., when the market condition is weak), the sales subsidy is more effective. Moreover, Taylor and Xiao (2014) prove that after taking an expectation across all market conditions, the strong-market-condition scenarios dominate (intuitively, the room for increasing sales under weak market conditions is limited), and therefore the purchase subsidy is generally more effective than the sales subsidy, as stated in the following theorem.

Theorem 18.8 *It is optimal for the donor to only provide purchase subsidy, i.e., $s^* = 0$.*

Taylor and Xiao (2014) further comment that this finding is consistent with the practice of several existing subsidy programs which only provide purchase subsidies. A practical reason why the purchase subsidy can be preferred is that it often has lower administrative costs than

the sales subsidy. Taylor and Xiao (2014) provide further justification for choosing a purchase subsidy by proving its superior effectiveness in increasing the number of sales.

Taylor and Xiao (2014) also consider several extensions to show the robustness of their findings. They show that the optimality of a purchase subsidy continues to hold when there are multiple heterogeneous retailers or when the retailer has to set the price before observing the market conditions. Further, they extend their model to consider the perishability of the products, where it is assumed that a fixed fraction of the unsold products will perish in the next period. In this case, the purchase subsidy continues to be optimal if the fraction that perishes is not too high or if the donor's budget level is not too high. Otherwise (i.e., if both the fraction that perishes and the budget level are high), it can be optimal to provide both purchase and sales subsidies.

The problem of using subsidies to improve the availability of healthcare or other essential products in the developing world has been examined by a number of other studies (e.g., Berenguer et al. (2017); Chick et al. (2008); Cohen et al. (2016); Levi et al. (2017); Martin et al. (2020)). However, these papers typically focus on incentivizing the manufacturer(s) to increase their production capacity in a single-period setting, and do not explicitly study the inventory decisions.

The procurement of pharmaceutical products in the developing world often relies on external funding, which is typically highly variable and unpredictable. Natarajan and Swaminathan (2014) take into account such uncertainty in funding for inventory management in the global health sector. Specifically, they formulate a multi-period inventory management problem under funding constraints where there is uncertainty in the funding amount and timing. In this setting, they show that the optimal inventory policy is a base-stock policy. Moreover, the optimal base-stock level is independent of the available funding. Additionally, Natarajan and Swaminathan (2014) also show that receiving funding early is critical, and less but timely funding can be preferred over delayed full funding.

Drug stock-outs and funding disbursement have also been examined by Gallien et al. (2017). Specifically, using publicly available data for the Global Fund's grants, they build a discrete-event simulation model to study the joint effect of drug procurement and fund disbursement processes on the availability of essential drugs in the Global Fund's recipient countries in Africa. A key model component in Gallien et al. (2017) is the Global Fund's performance-based funding policy, under which fund disbursements to recipients are based on past grant performance. They show that both the high unpredictability of fund disbursements and the high grant performance monitoring frequency are important drivers of stock-out risks. They also compare several interventions and show that shifting some fund disbursements upfront to reduce the uncertainty in fund disbursement timing is particularly effective in reducing drug stock-outs.

18.4 CONCLUSION AND FUTURE DIRECTIONS

In this chapter, we discussed a few inventory management problems for blood products and pharmaceuticals by focusing on the perishability of blood products and the multi-stakeholder nature of a pharmaceutical chain. We conclude this chapter by presenting a few interesting questions for future research. First, while single-location blood inventory systems have been studied extensively in the literature, blood inventory sharing in a multi-location system has

received much less attention. As highlighted in Zhang et al. (2021a), the existing findings on inventory sharing for nonperishable products do not necessarily carry over to blood/perishable products, and more research is needed for the latter. Second, the vast majority of the existing studies on blood and perishable inventory systems focus on identifying the optimal inventory policies in centralized systems. In practice, however, blood inventory systems often involve multiple stakeholders, such as the suppliers and different hospitals. Therefore, it would be interesting to study the management of blood and perishable inventory in decentralized systems. Third, drug shortages have been a serious concern both in the United States and worldwide especially in the past two decades. However, few studies have looked into this problem from an operations perspective. Therefore, it would also be interesting to study the effectiveness of different operational strategies for mitigating drug shortages. Finally, the demand for personal protective equipment (PPE), such as masks and gloves, has dramatically increased during the COVID-19 pandemic. This dramatic increase in demand together with disrupted supply chains resulted in substantial shortages of PPEs. Accordingly, the management of PPE inventory during a pandemic is another interesting topic for future research.

REFERENCES

- Abbasi, B., & Hosseinifard, S. Z. (2014). On the issuing policies for perishable items such as red blood cells and platelets in blood service. *Decision Sciences*, 45(5), 995–1020.
- Abouee-Mehrizi, H., Baron, O., Berman, O., & Chen, R. D. (2019). Managing perishable inventory systems with multiple priority classes. *Production and Operations Management*, 28(9), 2305–2322.
- Adida, E. (2021). Outcome-based pricing for new pharmaceuticals via rebates. *Management Science*, 67(2), 892–913.
- Alev, I., Atasu, A., Toktay, L. B., & Zhang, C. (2021). Extended producer responsibility for pharmaceuticals. *Manufacturing & Service Operations Management* [Forthcoming].
- American Red Cross (2021). US blood supply facts. <https://www.redcrossblood.org/donate-blood/how-to-donate/how-blood-donations-help/blood-needs-blood-supply.html>
- Atkinson, M. P., Fontaine, M. J., Goodnough, L. T., & Wein, L. M. (2012). A novel allocation strategy for blood transfusions: Investigating the tradeoff between the age and availability of transfused blood. *Transfusion*, 52(1), 108–117.
- Ayer, T., Zhang, C., Zeng, C., White, C. C., & Joseph, V. R. (2019). Analysis and improvement of blood collection operations – 2017 M&SOM practice-based research competition. *Manufacturing & Service Operations Management*, 21(1), 29–46.
- Bandi, C., Han, E., & Nohadani, O. (2019). Sustainable inventory with robust periodic-affine policies and application to medical supply chains. *Management Science*, 65(10), 4636–4655.
- Beliën, J., & Forcé, H. (2012). Supply chain management of blood products: A literature review. *European Journal of Operational Research*, 217(1), 1–16.
- Berenguer, G., Feng, Q., Shanthikumar, J. G., & Xu, L. (2017). The effects of subsidies on increasing consumption through for-profit and not-for-profit newsvendors. *Production and Operations Management*, 26(6), 1191–1206.
- Brodheim, E., Derman, C., & Prastacos, G. (1975). On the evaluation of a class of inventory policies for perishable products such as blood. *Management Science*, 21(11), 1320–1325.
- Centers for Medicare & Medicaid Services. (2021). NHE fact sheet. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet>
- Chao, X., Gong, X., Shi, C., Yang, C., Zhang, H., & Zhou, S. X. (2017). Approximation algorithms for capacitated perishable inventory systems with positive lead times. *Management Science*, 64(11), 5038–5061.

- Chao, X., Gong, X., Shi, C., & Zhang, H. (2015). Approximation algorithms for perishable inventory systems. *Operations Research*, 63(3), 585–601.
- Chazan, D., & Gal, S. (1977). A Markovian model for a perishable product inventory. *Management Science*, 23(5), 512–521.
- Chen, K., Song, J. S., Shang, J., & Xiao, T. (2020). Managing hospital platelet inventory with mid-cycle expedited replenishments and returns. *Production and Operations Management*, 31(5), 2015–2037.
- Chen, S., Li, Y., & Zhou, W. (2019). Joint decisions for blood collection and platelet inventory control. *Production and Operations Management*, 28(7), 1674–1691.
- Chen, X., Pang, Z., & Pan, L. (2014). Coordinating inventory control and pricing strategies for perishable products. *Operations Research*, 62(2), 284–300.
- Chick, S. E., Mamani, H., & Simchi-Levi, D. (2008). Supply chain coordination and influenza vaccination. *Operations Research*, 56(6), 1493–1506.
- Cohen, M. (1976). Analysis of single critical number ordering policies for perishable inventories. *Operations Research*, 24(4), 726–741.
- Cohen, M. C., Lobel, R., & Perakis, G. (2016). The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Science*, 62(5), 1235–1258.
- Cooper, W. (2001). Pathwise properties and performance bounds for a perishable inventory system. *Operations Research*, 49(3), 455–466.
- Crama, Y., Rezaei, M., Savelsbergh, M., & Woensel, T. V. (2018). Stochastic inventory routing for perishable products. *Transportation Science*, 52(3), 526–546.
- Deniz, B., Karaesmen, I., & Scheller-Wolf, A. (2010). Managing perishables with substitution: Inventory issuance and replenishment heuristics. *Manufacturing & Service Operations Management*, 12(2), 319–329.
- Federgruen, A., Prastacos, G., & Zipkin, P. H. (1986). An allocation and distribution model for perishable products. *Operations Research*, 34(1), 75–82.
- Fontaine, M. J., Chung, Y. T., Rogers, W. M., Sussmann, H. D., Quach, P., Galel, S. A., Goodnough, L. T., & Erhun, F. (2009). Improving platelet supply chains through collaborations between blood centers and transfusion services. *Transfusion*, 9(10), 2040–2047.
- Fox, E. R., Sweet, B. V., & Jensen, V. (2014). Drug shortages: A complex health care crisis. *Mayo Clinic Proceedings*, 89(3), 361–373.
- Fries, B. (1972). Optimal ordering policy for a perishable commodity with fixed lifetime [Ph.D. Thesis]. Cornell University.
- Fries, B. (1975). Optimal ordering policy for a perishable commodity with fixed lifetime. *Operations Research*, 23(1), 46–61.
- Gallien, J., Rashkova, I., Atun, R., & Yadav, P. (2017). National drug stockout risks and the global fund disbursement process for procurement. *Production and Operations Management*, 26(6), 997–1014.
- Haijema, R., van der Wal, J., & van Dijk, N. M. (2007). Blood platelet production: Optimization by dynamic programming and simulation. *Computers and Operations Research*, 34(3), 760–779.
- Hotkar, P., & Gupta, D. (2021). The strategic role of disruption information sharing on the supply of sterile injectable drugs [Working paper].
- Jennings, J. B. (1973). Blood bank inventory control. *Management Science*, 19(6), 637–645.
- Jia, J., & Zhao, H. (2017). Mitigating the us drug shortages through pareto-improving contracts. *Production and Operations Management*, 26(8), 1463–1480.
- Jones, J. M., Sapiano, M. R. P., Savinkina, A. A., Haass, K. A., Baker, M. L., Henry, R. A., Berger, J. J., & Basavaraju, S. V. (2020). Slowing decline in blood collection and transfusion in the united states–2017. *Transfusion*, 60, S1–S9.
- Karaesmen, I. Z., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. In Karl G. Kempf, Pınar Keskinocak, and Reha Uzsoy (Eds.), *Planning production and inventories in the extended enterprise* (pp. 393–436).
- Kendall, K. E., & Lee, S. M. (1980). Formulating blood rotation policies with multiple objectives. *Management Science*, 26(11), 1145–1157.
- Kim, S. H., & Morton, F. S. (2015). A model of generic drug shortages: Supply disruptions, demand substitution, and price control [Working paper].

- King, G. J., Chao, X., & Duenyas, I. (2019). Who benefits when prescription drug manufacturers offer copay coupons? *Management Science*, 65(8), 3758–3775.
- Kouvelis, P., Xiao, Y., & Yang, N. (2015). Pbm competition in pharmaceutical supply chain: Formulary design and drug pricing. *Manufacturing & Service Operations Management*, 17(4), 511–526.
- Lee, J., Lee, H. S., Shin, H., & Krishnan, V. (2021). Alleviating drug shortages: The role of mandated reporting induced operational transparency. *Management Science*, 67(4), 2326–2339.
- Levi, R., Perakis, G., & Romero, G. (2017). On the effectiveness of uniform subsidies in increasing market consumption. *Management Science*, 63(1), 40–57.
- Li, Q., & Yu, P. (2014). Multimodularity and its applications in three stochastic dynamic inventory problems. *Manufacturing & Service Operations Management*, 16(3), 455–463.
- Li, Q., Yu, P., & Du, L. (2021). Separation of perishable inventories in offline retailing through transshipment. *Operations Research* [Forthcoming].
- Li, Q., Yu, P., & Wu, X. (2016). Managing perishable inventories in retailing: Replenishment, clearance sales, and segregation. *Operations Research*, 64(6), 1270–1284.
- Liu, Q., Zhang, X., Liu, Y., & Lin, L. (2013). Spreadsheet inventory simulation and optimization models and their application in a national pharmacy chain. *INFORMS Transactions on Education*, 14(1), 13–25.
- Martin, P., Gupta, D., & Natarajan, K. V. (2020). Vaccine procurement contracts for developing countries. *Production and Operations Management*, 29(11), 2601–2620.
- Memorial Sloan Kettering Cancer Center (2021). Covid-19 has caused a national blood shortage. People with cancer need your help. Tech. rep. Retrieved July 2, 2021, from <https://www.mskcc.org/news/covid-19-has-caused-national-blood-shortage-people-cancer-need-your-help>.
- Nahmias, S. (1975). Optimal ordering policies for perishable inventory-II. *Operations Research*, 23(4), 735–749.
- Nahmias, S. (1976). Myopic approximations for the perishable inventory problem. *Management Science*, 22(9), 1002–1008.
- Nahmias, S. (2011). *Perishable Inventory Systems*. Springer.
- Nandakumar, P., & Morton, T. (1993). Near myopic heuristics for the fixed-life perishability problem. *Management Science*, 39(12), 1490–1498.
- Natarajan, K. V., & Swaminathan, J. M. (2014). Inventory management in humanitarian operations: Impact of amount, schedule, and uncertainty in funding. *Manufacturing & Service Operations Management*, 16(4), 595–603.
- Pierskalla, W. P. (2005). Supply chain management of blood banks. In Margaret L. Brandeau, François Sainfort, and William P. Pierskalla (Eds.), *Operations Research and Health Care* (pp. 103–145). Springer.
- Pierskalla, W., & Roach, C. (1972). Optimal issuing policies for perishable inventory. *Management Science*, 18(11), 603–614.
- Prastacos, G. P. (1981). Allocation of a perishable product inventory. *Operations Research*, 29(1), 95–107.
- Prastacos, G. P., & Brodheim, E. (1980). PBDS: A decision support system for regional blood management. *Management Science*, 26(5), 451–463.
- Sarhangian, V., Abouee-Mehrizi, H., Baron, O., & Berman, O. (2018). Threshold-based allocation policies for inventory management of red blood cells. *Manufacturing & Service Operations Management*, 20(2), 347–362.
- Taylor, T. A., & Xiao, W. (2014). Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science*, 60(10), 2461–2477.
- van Sambeeck, J. H. J., van Brummelen, S. P. J., van Dijk, N. M., & Janssen, M. P. (2021). Optimal blood issuing by comprehensive matching. *European Journal of Operational Research* [Forthcoming].
- Ventola, C. L. (2011). The drug shortage crisis in the united states: Causes, impact, and management strategies. *Pharmacy and Therapeutics*, 36(11), 740.
- Yurukoglu, A., Liebman, E., & Ridley, D. B. (2017). The role of government reimbursement in drug shortages. *American Economic Journal: Economic Policy*, 9(2), 348–382.
- Zhang, C., Ayer, T., & White, C. C. (2021a). Inventory sharing for perishable products: Application to platelet inventory management in hospital blood banks [Working paper].

- Zhang, C., Ayer, T., & White, C. C. (2021b). Truncated balancing policy for perishable inventory management: Combating large shortage penalties [Working paper].
- Zhang, H., Chao, X., & Shi, C. (2018). Perishable inventory systems: Convexity results for base-stock policies and learning algorithms under censored demand. *Operations Research*, 66(5), 1276–1286.
- Zhang, H., Shi, C., & Chao, X. (2016). Approximation algorithms for perishable inventory systems with setup costs. *Operations Research*, 64(2), 432–440.
- Zhang, X., Meiser, D., Liu, Y., Bonner, B., & Lin, L. (2014). Kroger uses simulation-optimization to improve pharmacy inventory management. *Interfaces*, 44(1), 70–84.
- Zhao, H., Xiong, C., Gavirneni, S., & Fein, A. (2012). Fee-for-service contracts in pharmaceutical distribution supply chains: Design, analysis, and management. *Manufacturing & Service Operations Management*, 14(4), 685–699.
- Zhou, D., Leung, L., & Pierskalla, W. (2011). Inventory management of platelets in hospitals: Optimal inventory policy for perishable products with regular and optional expedited replenishments. *Manufacturing & Service Operations Management*, 13(4), 420–438.

19. Spare parts inventory planning

Rob Basten and Geert-Jan van Houtum

19.1 INTRODUCTION

In this chapter, we aim to give insights into the essence of spare parts inventory models and into recent developments in the research on these models. There are two key reasons why spare parts inventory control is quite different from most other inventory control, for example, in retail. The first key reason is that spare parts are kept in stock to support capital goods, i.e., machines or systems that are used by manufacturers to produce their end-products and by service organizations to deliver their services. The users of capital goods are often faced with disruptions of their primary processes if a capital good is down during an unplanned time interval. Therefore they require *high system availability*, i.e., a high fraction of time that a system is available for production. This means that requirements for spare parts are set at the level of capital goods, which directly leads to multi-item inventory models instead of single-item models.

To explain the second key reason, we first need to look at the maintenance of critical capital goods in a manufacturing or service organization. This maintenance may be executed by its own maintenance department, the Original Equipment Manufacturer (OEM), or a third maintenance party. Traditionally, the maintenance itself consists of planned *preventive maintenance* and unplanned *corrective maintenance*. Preventive maintenance is applied to components that are expected to fail after some time (or a certain amount of usage hours). These components are then replaced by spare parts, which may be sent from a central warehouse. Corrective maintenance is needed when a capital good fails unexpectedly. Corrective maintenance starts with diagnosing why the failure happened and next the actual repair can be executed. If for the repair a spare part is needed to replace a broken component, then the spare part has to be delivered very quickly in general. To enable such fast deliveries, spare parts are kept in stock in local warehouses. However, generally capital goods consist of many different components and it would be far too expensive to keep spare parts of all different components in stock at each local warehouse. What is typically done in spare parts networks is allowing for a requested part at a local warehouse to be supplied by a neighboring local warehouse if the first local warehouse is out of stock. We call this a *lateral (trans)shipment*. By allowing such lateral shipments, local warehouses in the same geographical region operate as if they form one large, virtual warehouse. Another property in spare parts operations is that a part is supplied by an *emergency shipment* from a (more) central warehouse if none of the local warehouses in the neighborhood can deliver a requested part. The presence of lateral and emergency shipments forms the second reason that spare parts inventory control is different from most other inventory control.

Notice that spare parts management is very important from an economic point of view. According to the Aberdeen Group (2005), spare parts sales and services in the United States accounted for 8% of the annual GDP in 2005 and this will be similar in many other developed

economies. This is a key reason why the theory on multi-item inventory models for spare parts in networks consisting of central and local warehouses is rich. It started more than 50 years ago with the seminal paper on the well-known *METRIC model* by Sherbrooke (1968). In this model, multiple items are distinguished that are kept in stock in a two-echelon system consisting of one central warehouse and multiple local warehouses. Demands at the local warehouses occur according to Poisson processes, and basestock policies are used for the inventory control. This means that whenever a part has been demanded at a warehouse, the warehouse immediately orders a new one or orders a repair of the broken part. Demands that cannot be directly met from stock are backordered. The objective is to minimize the inventory holding costs subject to a constraint on the mean total number of backordered demands. The latter measure is directly related to the average system availability for the total installed base of capital goods and hence is called a *system-oriented service measure*. The total number of backordered demands is simply the sum of the mean numbers of backordered demands per item and hence connects a system-related service measure to an underlying service measure for individual items. This connection is key for spare parts inventory models. Sherbrooke developed an efficient approximate approach, called the *METRIC approach*, for the evaluation of a given basestock policy and a greedy heuristic to find a close-to-optimal basestock policy in an efficient way.

The work of Sherbrooke has been extended to networks with more than two-echelon levels, to problems with different system-oriented service measures and constraints, and to systems where parts are replaced at different levels in the system structure and where lower-level parts are needed for the repair of higher-level parts. Further, more accurate approximate evaluation methods, exact evaluation approaches, and alternatives and extensions for the greedy heuristic have been developed.

There has also been extensive research on spare parts networks with lateral and emergency shipments, i.e., networks where unmet demand at a local warehouse is not backordered (as in the METRIC-type models) but satisfied via a shipment from another location or another emergency action. For such networks, more complex evaluation approaches are needed. In this chapter, we focus on these multi-item inventory models with lateral and emergency shipments.

For an overview of the literature on multi-item spare parts inventory models and applications of these models in practice, we refer to the books of Sherbrooke (2004), Muckstadt (2005), and Van Houtum and Kranenburg (2015), and the survey papers of Basten and Van Houtum (2014) and Hu et al. (2018). Section 2 of Basten and Van Houtum (2014) also nicely explains in which settings in practice the various types of models are typically used.

Some of the more recent research on spare parts inventory models explores the possible benefits of 3D printing, sensor technology, and the Internet of Things (IoT). Three-dimensional printing can be used to print spare parts at the moment that they are needed. With advanced printing technology, it is possible for some parts to print them at the same quality as regular parts and then these printed parts can be used to replace the failed part permanently. With general-purpose printers, which are cheaper and can be used for a wide range of parts, generally parts of a lower quality are printed. These parts can be used to replace failed parts temporarily until a regular spare part is available. In general, it is interesting to investigate how 3D printing can relieve the pressure to have many spare parts available in many local warehouses. Sensor technology and IoT can be used to get predictions on upcoming failures. These predictions can be used to replace degrading components just before they fail (*predictive maintenance*) or to

proactively move spare parts to locations close to the systems that have upcoming failures so that the failing parts can be replaced quickly when the failures occur.

In the remainder of this chapter, we address the following topics. In Section 19.2, we present a basic, single-location, multi-item inventory model where emergency shipments are used in case of stockouts. In Section 19.3, we describe multi-item, multi-location models with lateral and emergency shipments. Next, we discuss how 3D printing and failure predictions can be used to get a more efficient spare parts supply in Sections 19.4 and 19.5, respectively. Finally, we conclude in Section 19.6.

19.2 SINGLE-LOCATION, MULTI-ITEM MODEL WITH EMERGENCY SHIPMENTS

In this section, we first formulate the model and the corresponding optimization problem (see Section 19.2.1) and after that we describe a greedy heuristic that generally gives a close-to-optimal solution (see Section 19.2.2). The presented model stems from Section 2.9 of Van Houtum and Kranenburg (2015) and is an extension of the basic single-location, multi-item model with backordering of Sherbrooke (2004).

19.2.1 Problem Formulation

Consider a local warehouse where multiple spare parts are kept in stock to serve an installed base of technical systems of the same type. The systems consist of multiple *critical* components for which corrective maintenance is applied. Components are classified as critical when the system cannot function well without that component. When a critical component fails in a given system, this leads to a *demand* for a spare part. The failed part is replaced by a spare part from the local warehouse if it is available, or by a spare part that is sent via an emergency shipment from another location (e.g., a central warehouse); i.e., we have *repair by replacement*. The failed part is returned to the local warehouse and is immediately sent for repair. We assume that all critical components are repairable. (The model remains the same when a subset of components would be consumable and new parts are ordered one-by-one at suppliers.)

We refer to the critical components as Stock-Keeping Units (SKUs). The set of SKUs is denoted by I , and the number of SKUs is denoted by $|I|$ ($\in \mathbb{N} := \{1, 2, \dots\}$). For notational convenience, the SKUs are assumed to be numbered $i = 1, 2, \dots, |I|$. We assume an infinite time horizon $[0, \infty)$. For each SKU $i \in I$, demand occurs according to a Poisson process with a constant rate λ_i (≥ 0). The rate λ_i denotes the demand rate for all systems together. We assume that the demand processes for different SKUs are mutually independent. The total demand rate for all SKUs together is denoted by $\Lambda = \sum_{i \in I} \lambda_i$ and we assume that $\Lambda > 0$. A demand is fulfilled immediately if possible; otherwise, an emergency procedure is followed to make a ready-for-use spare part available as soon as possible. We assume that such a part is sent from another location via a fast transportation mode. The average time for an emergency shipment for SKU i is L_i^{em} (≥ 0). Each demand is accompanied by the return of a failed part, and the failed part is immediately sent into repair. The time that a failed part is in repair is called the *repair lead time*, which consists of the actual repair time, but also administrative time and delays because of resources that are not immediately available. Repair lead times of parts of different SKUs are assumed to be independent and repair lead times of parts of the same SKU are assumed

to be independent and identically distributed (i.i.d.). The mean repair lead time for SKU i is denoted by L_i (>0).

In case an emergency procedure is applied to satisfy a demand, we assume that the failed part is not sent into repair by the local warehouse, but it is sent to the other location that provided the ready-for-use part. It is then easily verified that the inventory position of SKU i , defined as the physical stock plus parts in repair, is constant. This constant amount is denoted by s_i ($\in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$). Instead of saying that each failed part is immediately sent into repair, we may also say that for each SKU i the stock is controlled by a continuous-review *basestock policy*, with *basestock level* s_i , which is a decision variable. Let $\mathbf{s} = (s_1, \dots, s_{|I|})$ denote the base-stock policy for the whole set of SKUs.

Because of the application of emergency shipments, the number of parts in the repair pipeline of SKU i is bounded from above by s_i . The behavior of the number of parts in repair of SKU i is as the number of jobs in an Erlang loss system, i.e., an $M|G|c|c$ queue with $c=s_i$ parallel servers, arrival rate λ_i , and mean service time L_i . (To see this parallel, familiarity with basic queueing theory is required; see, e.g., Kulkarni (2020).) The fill rate $\beta_i(s_i)$ is the fraction of arriving demands for SKU i that is fulfilled immediately upon request. Because the demands arrive according to a Poisson process, this fill rate is equal to the fraction of time that there is at least one part in stock. This is equal to the fraction of time that at least one server is free in the corresponding Erlang loss system. The latter probability is equal to 1 minus the fraction of time that all servers are occupied, i.e., 1 minus the Erlang loss probability. Hence,

$$\beta_i(s_i) = 1 - L(s_i, \rho_i),$$

where $\rho_i := \lambda_i L_i$, and

$$L(c, \rho) = \frac{\frac{1}{c!} \rho^c}{\sum_{x=0}^c \frac{1}{x!} \rho^x}$$

is the Erlang loss probability of an Erlang loss system with $c \in \mathbb{N}$ servers and offered load $\rho > 0$ (by convention, $L(0, \rho) = 1$ for all $\rho > 0$).

For each SKU $i \in I$, let $W_i(s_i)$ denote the mean waiting time, which is equal to

$$W_i(s_i) = (1 - \beta_i(s_i)) L_i^{\text{em}}.$$

Let $W(\mathbf{s})$ be the mean waiting time for an arbitrary demand. This waiting time is called the *aggregate mean waiting time* and is equal to a weighted sum of the mean waiting times per SKU:

$$W(\mathbf{s}) = \sum_{i \in I} \frac{\lambda_i}{\Lambda} W_i(s_i).$$

Notice that the aggregate mean waiting time $W(\mathbf{s})$ is directly related to the system availability. We assume that the aggregate mean waiting time is required to be at most W^{obj} .

We distinguish the following types of costs. First of all, we have an inventory holding cost rate h_i (> 0) per unit of SKU i . These costs are charged for both the on-hand stock and the parts in the repair pipeline. Hence they are equal to $h_i s_i$ for SKU i . Further, we have costs for repairs of failed parts and for emergency shipments. For each demand for an SKU, we have either a repair organized by the local warehouse or by another location in case of an emergency shipment. We assume that repairs at the local warehouse and other locations are equally expensive and hence the total repair costs per time unit are constant and thus not relevant for the optimization of the basestock levels. For each emergency shipment of SKU i , we have costs c_i^{em} (≥ 0). This covers the fast transport costs and other logistics costs to get the part to the failed system in a fast way. For each SKU i , the average costs per time unit for emergency shipments are equal to $\lambda_i(1 - \beta_i(s_i))c_i^{\text{em}}$. Hence, the average costs for SKU i are equal to

$$C_i(s_i) := h_i s_i + \lambda_i(1 - \beta_i(s_i))c_i^{\text{em}}, \quad (19.1)$$

and the total average costs are equal to $C(\mathbf{s}) = \sum_{i \in I} C_i(s_i)$.

The objective is to minimize the total average costs subject to the aggregate mean waiting time constraint:

$$\begin{aligned} (P) \quad & \min \quad C(\mathbf{s}) \\ & \text{subject to} \quad W(\mathbf{s}) \leq W^{\text{obj}}, \\ & \mathbf{s} \in \mathcal{S}, \end{aligned}$$

where $\mathcal{S} = \{\mathbf{s} = (s_1, \dots, s_{|I|}) \mid s_i \in \mathbb{N}_0, \forall i \in I\}$.

The assumptions that are made in the model of this section are common assumptions in the spare parts literature and they are justified in many applications in practice. Nevertheless, they should always be checked when applying this model.

There is one assumption that deserves special attention: The assumption that the demand rate for each SKU i is constant. This demand rate is the total demand rate for SKU i for the whole installed base of technical systems that are supported. Obviously, the technical systems that are temporarily down because they have to wait for a delivery of a spare part (of SKU i or any other SKU) by an emergency shipment or because of the repair itself do not contribute to the demand rate for SKU i at that moment. When the downtimes are short and/or when the number of unavailable systems is always a small fraction of the total installed base, then it is reasonable to assume a constant demand rate. This assumption implies that the whole steady-state behavior of an SKU is independent of the behavior of other SKUs and this simplifies the analysis enormously.

19.2.2 Optimization

Problem (P) is an integer programming problem with a nonlinear objective function and a nonlinear constraint. For this problem, no exact solution method is known that solves the

problem within a reasonable computation time. Therefore, we present a greedy algorithm that in general generates close-to-optimal solutions for Problem (P) for problem instances with sufficiently many SKUs. The required computation time of this algorithm is relatively low. This greedy algorithm has a theoretical foundation, which we explain as well. This foundation holds for a closely related optimization problem with two objectives. Let us consider Problem (Q) with two objectives, minimization of the total average costs $C(\mathbf{s})$ and minimization of the aggregate mean waiting time $W(\mathbf{s})$:

$$\begin{aligned}
 \text{(Q)} \quad & \min \quad C(\mathbf{s}) \\
 & \min \quad W(\mathbf{s}) \\
 & \text{subject to} \quad \mathbf{s} \in \mathcal{S}.
 \end{aligned}$$

This problem is a *multi-objective optimization problem* (or, more specifically, a bi-objective optimization problem). For this problem, we can derive so-called *efficient solutions*, i.e., solutions that are such that there are no solutions that do better on both objectives. To be precise: A solution $\mathbf{s} \in \mathcal{S}$ is efficient for Problem (Q) if and only if there is no other solution $\mathbf{s}' \in \mathcal{S}$ with $C(\mathbf{s}') \leq C(\mathbf{s})$ and $W(\mathbf{s}') \leq W(\mathbf{s})$, and strict inequality for at least one of these inequalities. Alternatively stated, a solution $\mathbf{s} \in \mathcal{S}$ is efficient for Problem (Q) if and only if $C(\mathbf{s}') > C(\mathbf{s})$, or $W(\mathbf{s}') > W(\mathbf{s})$, or $(C(\mathbf{s}'), W(\mathbf{s}')) = (C(\mathbf{s}), W(\mathbf{s}))$ for all $\mathbf{s}' \in \mathcal{S}$. Let $\mathcal{E}^*$ denote the set of all efficient solutions for Problem (Q). Then the points $(C(\mathbf{s}), W(\mathbf{s}))$, $\mathbf{s} \in \mathcal{E}^*$, constitute an *efficient frontier* for the total average costs vs. aggregate mean waiting time. From this efficient frontier, we can obtain directly an optimal solution for Problem (P): The solution $\mathbf{s} \in \mathcal{E}^*$ for which $W(\mathbf{s})$ is closest to, but not exceeding W^{obj} , is optimal for Problem (P).

To analyze Problem (Q), we need the following definition for functions on a discrete domain.

Definition 19.1 Let $f(x)$ be a function on $\mathbb{Z}$, and $x_0 \in \mathbb{Z}$.

(i) $f(x)$ is decreasing for $x \geq x_0$ if

$$\Delta f(x) = f(x+1) - f(x) \leq 0, \quad x \geq x_0;$$

(ii) $f(x)$ is convex for $x \geq x_0$ if

$$\Delta^2 f(x) = \Delta f(x+1) - \Delta f(x) \geq 0, \quad x \geq x_0.$$

Notice that $\Delta f(x+1) - \Delta f(x) = f(x+2) - 2f(x+1) + f(x)$, $x \in \mathbb{Z}$. The definitions for strictly decreasing and strictly convex are obtained by replacing the inequality signs with strict inequality signs. The definitions for (strictly) increasing and (strictly) concave are obtained by turning the (strict) inequality signs around.

Let us now look at the behavior of the cost functions $C_i(s_i)$ and waiting time functions $W_i(s_i)$. Karush (1957) has shown that the Erlang loss probability is convex and decreasing

as a function of the number of servers (see Remark 2 in Kranenburg and Van Houtum (2007) for how to circumvent a small error in the proof of Karush (1957)). This implies that $\beta_i(s_i)$ is concave and increasing on its whole domain. As a result:

- For each $i \in I$, $W_i(s_i)$ is decreasing and convex on its whole domain.
- For each $i \in I$, the first term in Equation (19.1) for $C_i(s_i)$ is strictly increasing and linear and the second term is decreasing and convex. This implies that $C_i(s_i)$ is convex on its whole domain.

Let

$$s_{i,\min} := \text{the first } s_i \geq 0 \text{ for which } C_i(s_i) < C_i(s_i + 1).$$

Then, obviously, for Problem (P) and its corresponding multi-objective optimization Problem (Q), all solutions with $s_i < s_{i,\min}$ for some $i \in I$, can be excluded (replacing s_i by $s_{i,\min}$ in this solution would give a better or equally good solution).

Because of the properties of the functions $C_i(s_i)$ and $W_i(s_i)$, we can prove that a set of efficient solutions can be generated by a greedy algorithm. A first efficient solution $\mathbf{s} = (s_1, \dots, s_{|I|})$ is obtained by setting $s_i = s_{i,\min}$ for each SKU $i \in I$. This solution is efficient because it has the lowest possible total average costs. Next, for each SKU i , we compute the decrease in $W(\mathbf{s})$ relative to the increase in $C(\mathbf{s})$ when s_i would be increased by one unit. The decrease in $W(\mathbf{s})$ is equal to $-\Delta_i W(\mathbf{s}) = -(\lambda_i / \Lambda) \Delta W_i(s_i)$, while the increase in $C(\mathbf{s})$ is equal to $\Delta_i C(\mathbf{s}) = \Delta C_i(s_i)$. Let $\Gamma_i := -(\lambda_i \Delta W_i(s_i)) / (\Lambda \Delta C_i(s_i))$ be their ratio. This Γ_i denotes how much $W(\mathbf{s})$ is decreased per unit increase in $C(\mathbf{s})$ when s_i is increased by one unit. The SKU i with the highest value for Γ_i gives the “biggest bang for the buck” and is increased to obtain a new solution (ties may be broken with equal probabilities, but other rules may be used as well). This new solution can be proved to be efficient as well (see Lemma 19.1 below) and is added to a set of efficient solutions. The generation of efficient solutions is continued until a given aggregate mean waiting time or total average costs has been reached, or until some other stop criterium is met. The formal procedure is described in Algorithm 19.1, where $\mathbf{e}_k$ is an $|I|$ -dimensional unit row-vector, i.e., a vector with a 1 at position k and zeros at all other positions.

Algorithm 19.1 (Greedy algorithm)

Step 1 $s_i := s_{i,\min}$ for all $i \in I$, and $\mathbf{s} = (s_{1,\min}, \dots, s_{|I|,\min})$;
 $\mathcal{E} := \{\mathbf{s}\}$;
 Compute $C(\mathbf{s})$ and $W(\mathbf{s})$.
 Step 2 $\Gamma_i := -(\lambda_i \Delta W_i(s_i)) / (\Lambda \Delta C_i(s_i))$ for all $i \in I$;
 $k := \arg \max\{\Gamma_i : i \in I\}$;
 $\mathbf{s} := \mathbf{s} + \mathbf{e}_k$;
 $\mathcal{E} := \mathcal{E} \cup \{\mathbf{s}\}$.
 Step 3 Compute $C(\mathbf{s})$ and $W(\mathbf{s})$;
 If “stop criterium”, then stop, else go to Step 2.

In the following lemma, it is formally stated that Algorithm 19.1 generates efficient solutions for Problem (Q). The proof of this lemma follows from Section 8 of Fox (1966).

Lemma 19.1 *At the termination of Algorithm 19.1, the set $\mathcal{E}$ consists of efficient solutions for Problem (Q).*

The greedy algorithm generates the set $\mathcal{E} = \{s^0, s^1, s^2, \dots\}$ of efficient solutions for Problem (Q), where $W(s^0) > W(s^1) > W(s^2) > \dots$ and $C(s^0) < C(s^1) < C(s^2) < \dots$. In general, the greedy algorithm does *not* generate all efficient solutions, but it does generate many efficient solutions (notice that generating all efficient solutions for a multi-objective optimization problem is difficult, see, e.g., Ruzika and Wiecek (2005)). So, formally, the set $\mathcal{E}$ is a subset of the set $\mathcal{E}^*$ with all efficient solutions. For Problem (P) with a given target W^{obj} , one can easily obtain a feasible solution from the subset $\mathcal{E}$ generated by the greedy algorithm. One just takes the first solution $s^l \in \mathcal{E}$ with $W(s^l) \leq W^{obj}$. This solution is optimal if and only if there is no solution $s \in \mathcal{E}^*$ with $W(s^l) < W(s) \leq W^{obj}$. In general, the solution s^l will be close to optimal if $W(s^l)$ is close to W^{obj} . For real-life problems, we often have many SKUs and hence we get small differences between the aggregate mean waiting times of subsequent solutions of $\mathcal{E}$ and the obtained solution for Problem (P) will be close to optimal.

This completes the analysis of the basic single-location, multi-item model. This analysis can be extended easily to problems with extra features, e.g., ordering in batches and condemnation, and alternative system-oriented service measures that are used in practical applications, e.g., system availability; see Chapter 2 of Van Houtum and Kranenburg (2015). Finally, we would like to note that single-location models can be directly applicable to problems in practice, e.g., for optimizing the local stock owned by and held at a single factory site. Further, they may serve as building blocks for planning concepts for complex networks, e.g., for optimizing the stock at the central warehouse of a network.

19.3 MODELS FOR NETWORKS WITH LATERAL AND EMERGENCY SHIPMENTS

In this section, we study in detail an extension of the model of Section 19.2 to multiple local warehouses that apply lateral shipments if such a shipment is faster than an emergency shipment in case of a stockout. The model and the corresponding optimization problem are presented in Section 19.3.1. Next, we describe a fast and accurate approximate evaluation method for a given basestock policy and we extend the greedy heuristic to find a close-to-optimal policy, see Sections 19.3.2 and 19.3.3, respectively. Finally, in Section 19.3.4, we show how a good policy can be found for a network of central and local warehouses, and we discuss the application of the spare parts model of this section to inventories/products other than spare parts.

19.3.1 Single-Echelon, Multi-Location Model

Let N denote a (non-empty) set of local warehouses (LWs), numbered $n = 1, \dots, |N|$. Each LW supports a number of technical systems. The systems supported at the various LWs are all similar. They consist of critical components, which we denote as SKUs. Let I denote the (non-empty) set of SKUs; they are numbered as $1, \dots, |I|$. Demands for each SKU $i \in I$ and LW $n \in N$ are assumed to occur according to a Poisson process with a constant rate $\lambda_{i,n}$ (≥ 0).

We define Λ_n as the total demand rate for LW n , i.e., $\Lambda_n := \sum_{i \in I} \lambda_{i,n}$, $n \in N$, and we assume that $\Lambda_n > 0$, $n \in N$. The target aggregate mean waiting time for LW $n \in N$ is denoted by $W_n^{\text{obj}} (> 0)$.

If LW n is out of stock for SKU i at the moment that a demand arrives for that SKU, then it may try to obtain the part by means of a lateral shipment from one of the other LWs. The corresponding transportation time for this lateral shipment from LW $m \in N$, $n \neq m$, to LW n is $L_{n,m}^{\text{lat}} (\geq 0)$ and the corresponding cost is $c_{n,m}^{\text{lat}} (\geq 0)$. It is also possible to provide the required part by an emergency shipment from the central warehouse. We assume that this is always possible, which is equivalent to assuming that the central warehouse has infinite stock. The corresponding transportation time is $L_n^{\text{em}} (\geq 0)$ and the corresponding cost is $c_n^{\text{em}} (\geq 0)$. The emergency shipment cost models the extra cost of sending a part from the central warehouse to a LW in comparison to a replenishment (i.e., a shipment at a standard speed to replenish the stock). The transportation times for lateral and emergency shipments (these are often in terms of hours/days) are generally short compared to the replenishment lead times (which are often in terms of weeks/months). When a demand is satisfied by a lateral and emergency shipment, we assume that this demand is immediately coupled to a part in the other LW or the central warehouse and that part may be sent directly to the technical system that is waiting for it. We assume that the costs and times for lateral and emergency shipments are mainly determined by the distances, and therefore they are assumed to be SKU-independent (it would be straightforward to make these times and costs SKU-dependent).

We often observe in practice that, in case of a stockout at an LW n , other LWs are checked in increasing order of their distance to LW n and that only LWs are checked that have a lateral shipment time that is lower than the emergency shipment time. We, therefore, make the assumption that a selection of other LWs is checked in a pre-specified order. For each LW n , this is described by a permutation $\sigma(n) = (\sigma_1(n), \sigma_2(n), \dots, \sigma_{p(n)}(n))$, where $p(n)$ is the number of other LWs that are checked. LW $\sigma_1(n)$ is the first other LW that is checked. If this LW is also out of stock, then LW $\sigma_2(n)$ is checked. This procedure is continued until LW $\sigma_{p(n)}(n)$ has been checked. If this last LW is also out of stock, then an emergency shipment is applied. Given how $\sigma(n)$ is constructed, it holds that

$$L_{n,\sigma_1(n)}^{\text{lat}} \leq L_{n,\sigma_2(n)}^{\text{lat}} \leq \dots \leq L_{n,\sigma_{p(n)}(n)}^{\text{lat}} \leq L_n^{\text{em}},$$

and $L_{n,m}^{\text{lat}} \geq L_n^{\text{em}}$ for all m that do not occur in $\sigma(n)$. Recall that we assumed that costs and times for lateral and emergency shipments are mainly determined by the distances. Hence, it is likely that it also holds that

$$c_{n,\sigma_1(n)}^{\text{lat}} \leq c_{n,\sigma_2(n)}^{\text{lat}} \leq \dots \leq c_{n,\sigma_{p(n)}(n)}^{\text{lat}} \leq c_n^{\text{em}},$$

and $c_{n,m}^{\text{lat}} \geq c_n^{\text{em}}$ for all m that do not occur in $\sigma(n)$.

In Figure 19.1, an example of a network with three LWs is given. LW 1 first checks LW 2 for a lateral shipment in case of a stockout situation. Next, an emergency shipment is applied if also LW 2 is out of stock. It does not check LW 3 because that would be less fast and more expensive than satisfying the demand by an emergency shipment. LW 2 first checks LW 1 and then LW 3 for a lateral shipment in case of a stockout; if necessary, an emergency shipment is applied. LW 3 checks first LW 2 for a lateral shipment, and next an emergency shipment is

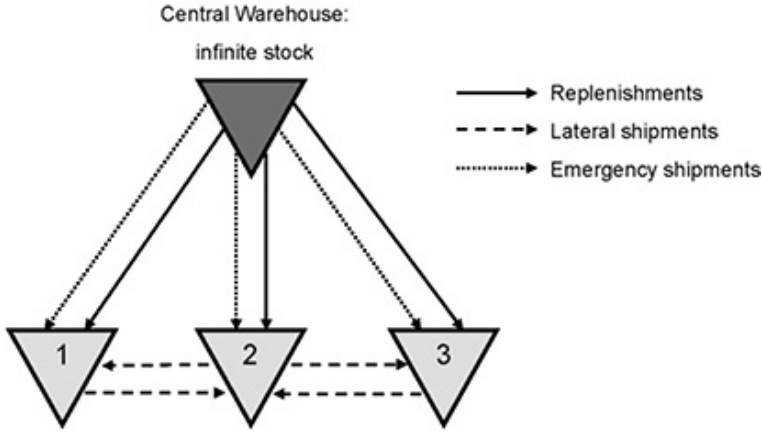


Figure 19.1 A network with three LWs

applied if LW 2 is out of stock. LW 1 is not checked for a lateral shipment because that would be slower and more expensive than satisfying the demand by an emergency shipment.

The stock in all local warehouses is controlled by a basestock policy. The basestock level for SKU $i \in I$ in LW $n \in N$ is denoted by $s_{i,n}$ ($\in \mathbb{N}_0$). Let $\mathbf{s}_i := (s_{i,1}, \dots, s_{i,|N|})$, $i \in I$, denote the vector of basestock levels for SKU i , and let a basestock policy for the whole system be denoted by

$$\mathbf{S} = \begin{pmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,|N|} \\ s_{2,1} & s_{2,2} & \dots & s_{2,|N|} \\ \vdots & \vdots & \ddots & \vdots \\ s_{|I|,1} & s_{|I|,2} & \dots & s_{|I|,|N|} \end{pmatrix}$$

Once a part in a local warehouse is used to satisfy a demand, immediately a new part is requested from the central warehouse. This part will be delivered after a replenishment lead time with mean L_n . Replenishment lead times for the same SKU and LW are assumed to be independent and identically distributed (i.i.d.) and they are independent of replenishment lead times for other SKUs and/or LWs. Notice that generally the replenishment lead times are much larger than the times for emergency and lateral shipments.

The costs for emergency shipments have been defined as the extra costs in comparison to a replenishment. Hence, the costs for the replenishments form a constant factor, and therefore they are excluded from our model. The cost of holding one part of SKU $i \in I$ in stock for one time unit is h_i . We assume that inventory holding costs are also incurred for parts in replenishment.

With respect to the fulfillment of a demand for SKU $i \in I$ at LW $n \in N$, we introduce the following notation:

- $\beta_{i,n}(\mathbf{s}_i)$: for the (item) fill rate, i.e., the fraction of demand for SKU i at LW n that is fulfilled immediately upon request, i.e., from the stock at LW n itself;

- $\alpha_{i,n,m}(\mathbf{s}_i)$, $m \in N$, $m \neq n$: for the fraction of demand for SKU i at LW n that is fulfilled from LW m by means of a lateral shipment (by definition, $\alpha_{i,n,m}(\mathbf{s}_i) = 0$ if m does not occur in $\sigma(n)$);
- $\theta_{i,n}(\mathbf{s}_i)$: for the fraction of demand for SKU i at LW n that is fulfilled from the central warehouse by an emergency shipment.

Notice that for each SKU $i \in I$ at each LW $n \in N$, it holds that

$$\beta_{i,n}(\mathbf{s}_i) + \theta_{i,n}(\mathbf{s}_i) + \sum_{m \in N, m \neq n} \alpha_{i,n,m}(\mathbf{s}_i) = 1. \quad (19.2)$$

In Section 19.3.2, we show how all fractions for SKU i can be determined via an approximate evaluation procedure for a given basestock policy $\mathbf{s}_i$ for SKU $i \in I$. Based on these fractions, we can calculate the total costs and the aggregate mean waiting times at each of the LWs.

Let $W_{i,n}(\mathbf{s}_i)$, $i \in I$, $n \in N$, denote the mean waiting time if SKU i is requested at LW n , under a given basestock policy $\mathbf{s}_i$ for SKU i . Then $W_{i,n}(\mathbf{s}_i)$ can be calculated as follows:

$$W_{i,n}(\mathbf{s}_i) = \sum_{m \in N, m \neq n} L_{n,m}^{\text{lat}} \alpha_{i,n,m}(\mathbf{s}_i) + L_n^{\text{em}} \theta_{i,n}(\mathbf{s}_i).$$

For LW $n \in N$, the aggregate mean waiting time equals

$$W_n(\mathbf{S}) = \sum_{i \in I} \frac{\lambda_{i,n}}{\Lambda_n} W_{i,n}(\mathbf{s}_i).$$

For SKU $i \in I$, the average costs per time unit are given by

$$C_i(\mathbf{s}_i) = \sum_{n \in N} h_i s_{i,n} + \sum_{n \in N} \lambda_{i,n} \left(\sum_{m \in N, m \neq n} c_{n,m}^{\text{lat}} \alpha_{i,n,m}(\mathbf{s}_i) + c_n^{\text{em}} \theta_{i,n}(\mathbf{s}_i) \right),$$

where the first term denotes the inventory holding costs per time unit and the second term the costs for lateral and emergency shipments. The total average costs for all SKUs together are equal to $C(\mathbf{S}) = \sum_{i \in I} C_i(\mathbf{s}_i)$.

The objective is to minimize the total average costs under the condition that the target aggregate mean waiting times are met:

$$\begin{aligned} (P') \quad & \min \quad C(\mathbf{S}) \\ & \text{subject to} \quad W_n(\mathbf{S}) \leq W_n^{\text{obj}}, \quad n \in N, \\ & \mathbf{S} \in \mathcal{S}, \end{aligned}$$

where $\mathcal{S} = \{\mathbf{S} = (s_{i,n})_{i \in I, n \in N} \mid s_{i,n} \in \mathbb{N}_0, \forall i \in I, \forall n \in N\}$.

19.3.2 Approximate Evaluation Procedure

In this subsection, we describe an approximate approach for the fractions $\beta_{i,n}(\mathbf{s}_i)$, $\alpha_{i,n,m}(\mathbf{s}_i)$, and $\theta_{i,n}(\mathbf{s}_i)$. We follow the approach of Reijnen et al. (2009) and describe this approach for the example network of Figure 19.1; see Reijnen et al. (2009) for the description for a general network (the approach of Reijnen et al. (2009) is an extension of the approach described in Chapter 5 of Van Houtum and Kranenburg (2015)).

For the example network of Figure 19.1, it holds that $|N| = 3$, $\sigma(1) = (2)$, $\sigma(2) = (1, 3)$, $\sigma(3) = (2)$, and for the rest the general notation of Section 19.3.1 is followed. The evaluation can be executed per SKU i . We want to calculate the following fractions:

$$\begin{aligned} &\beta_{i,1}(\mathbf{s}_i), \beta_{i,2}(\mathbf{s}_i), \beta_{i,3}(\mathbf{s}_i), \\ &\alpha_{i,1,2}(\mathbf{s}_i), \alpha_{i,2,1}(\mathbf{s}_i), \alpha_{i,2,3}(\mathbf{s}_i), \alpha_{i,3,2}(\mathbf{s}_i), \\ &\theta_{i,1}(\mathbf{s}_i), \theta_{i,2}(\mathbf{s}_i), \theta_{i,3}(\mathbf{s}_i). \end{aligned}$$

Recall that the transportation times and costs for lateral and emergency shipments were (or can be) taken into account when the permutations $\sigma(n)$ are defined; the above fractions depend on the permutations $\sigma(n)$ and hence indirectly also on the lateral/emergency shipment times and costs.

The main idea of the approximate approach is that we look at each LW n as an individual, decoupled LW that receives demands via its own Poisson demand stream with rate $\lambda_{i,n}$ and demands for lateral shipments from other LWs. The demands for lateral shipments placed by LW n at LW $m \neq n$ are modeled as a so-called *overflow stream* with rate $\lambda_{i,n,m}$. We pretend that each overflow stream is a Poisson stream. Then the total demand process at each LW n is also a Poisson process. The corresponding rate is denoted by $\lambda_{i,n}^{\text{tot}}$. It holds that

$$\lambda_{i,1}^{\text{tot}} = \lambda_{i,1} + \lambda_{i,2,1}, \lambda_{i,2}^{\text{tot}} = \lambda_{i,2} + \lambda_{i,1,2} + \lambda_{i,3,2}, \lambda_{i,3}^{\text{tot}} = \lambda_{i,3} + \lambda_{i,2,3}. \quad (19.3)$$

Let us now look at LW n as an individual, decoupled LW. The number of parts in the replenishment pipeline of LW n behaves as the number of jobs in an Erlang loss system with s_i servers and offered load $\lambda_{i,n}^{\text{tot}} L_n$ (similarly as in Section 19.2.1). Hence, the fraction of demand that is satisfied from stock equals (for both the own demand stream and the incoming overflow demand streams)

$$\beta_{i,n}(\mathbf{s}_i) = 1 - L(s_i, \lambda_{i,n}^{\text{tot}} L_n), \quad n = 1, 2, 3. \quad (19.4)$$

For the own demand stream with rate $\lambda_{i,n}$ at LW n , a fraction $1 - \beta_{i,n}(\mathbf{s}_i)$ is not satisfied and flows over to LW $\sigma_1(n)$. For LW 2, we have a second overflow demand stream consisting of demands that are not satisfied by LW $\sigma_1(2)$. The rates of the overflow streams are equal to

$$\begin{aligned} \lambda_{i,1,2} &= \lambda_{i,1}(1 - \beta_{i,1}(\mathbf{s}_i)), \lambda_{i,2,1} = \lambda_{i,2}(1 - \beta_{i,2}(\mathbf{s}_i)), \\ \lambda_{i,2,3} &= \lambda_{i,2,1}(1 - \beta_{i,1}(\mathbf{s}_i)), \lambda_{i,3,2} = \lambda_{i,3}(1 - \beta_{i,3}(\mathbf{s}_i)). \end{aligned} \quad (19.5)$$

For the overflow stream from LW 1 to LW 2, a fraction $\beta_{i,2}(\mathbf{s}_i)$ is satisfied by LW 2. This implies that $\alpha_{i,1,2}(\mathbf{s}_i)$ is equal to $(1 - \beta_{i,1}(\mathbf{s}_i))\beta_{i,2}(\mathbf{s}_i)$, and similar expressions are obtained for the other lateral shipment probabilities. We obtain:

$$\begin{aligned}\alpha_{i,1,2}(\mathbf{s}_i) &= (1 - \beta_{i,1}(\mathbf{s}_i))\beta_{i,2}(\mathbf{s}_i), \\ \alpha_{i,2,1}(\mathbf{s}_i) &= (1 - \beta_{i,2}(\mathbf{s}_i))\beta_{i,1}(\mathbf{s}_i), \\ \alpha_{i,2,3}(\mathbf{s}_i) &= (1 - \beta_{i,2}(\mathbf{s}_i))(1 - \beta_{i,1}(\mathbf{s}_i))\beta_{i,3}(\mathbf{s}_i), \\ \alpha_{i,3,2}(\mathbf{s}_i) &= (1 - \beta_{i,3}(\mathbf{s}_i))\beta_{i,2}(\mathbf{s}_i).\end{aligned}\tag{19.6}$$

Finally, in Equation (19.2), it holds that

$$\begin{aligned}\theta_{i,1}(\mathbf{s}_i) &= 1 - \beta_{i,1}(\mathbf{s}_i) - \alpha_{i,1,2}(\mathbf{s}_i), \\ \theta_{i,2}(\mathbf{s}_i) &= 1 - \beta_{i,2}(\mathbf{s}_i) - \alpha_{i,2,1}(\mathbf{s}_i) - \alpha_{i,2,3}(\mathbf{s}_i), \\ \theta_{i,3}(\mathbf{s}_i) &= 1 - \beta_{i,3}(\mathbf{s}_i) - \alpha_{i,3,2}(\mathbf{s}_i).\end{aligned}\tag{19.7}$$

Equations (19.3) to (19.7) include 17 equations in total with 17 unknown variables. We solve them as follows. We first solve the Equations (19.3)–(19.5) iteratively. We start with zero rates for the overflow streams (i.e., all $\lambda_{i,n,m}$ are set equal to 0) and then we repeatedly calculate the variables $\lambda_{i,n}^{\text{tot}}$ via Equation (19.3), $\beta_{i,n}(\mathbf{s}_i)$ via Equation (19.4), and $\lambda_{i,n,m}$ via Equation (19.5). We do this until the values for these variables have converged. Finally, we obtain the variables $\alpha_{i,n,m}(\mathbf{s}_i)$ via Equation (19.6) and $\theta_{i,n}(\mathbf{s}_i)$ via Equation (19.7).

The approximate approach makes two approximate steps. It assumes that the overflow streams are Poisson distributed and that the behavior of the inventory at each LW is independent of the behavior at other LWs (see the expressions for the fractions $\alpha_{i,n,m}(\mathbf{s}_i)$). Both steps lead to only limited inaccuracies for problem instances with sufficiently high fill rates $\beta_{i,n}(\mathbf{s}_i)$ (the overflow streams are thin streams in that case; hence they are close to Poisson and lead to only a weak coupling between the LWs). Notice that such instances are typically the relevant instances when solving Problem (P), because the target aggregate mean waiting times are usually very low.

The approximate approach has been tested extensively by Reijnen et al. (2009). It converges for all problem instances, and the calculation times are low (less than a millisecond per single-item problem instance with up to 14 LWs). The resulting approximations for the fractions $\beta_{i,n}(\mathbf{s}_i)$, $\alpha_{i,n,m}(\mathbf{s}_i)$, and $\theta_{i,n}(\mathbf{s}_i)$ are generally good (the absolute errors are of the order 0.01–0.02 and at most 0.03 for instances with fill rates of at least 0.80).

19.3.3 Greedy Heuristic

For the model of Section 19.2, we considered the closely related Problem (Q) with two objectives, and we formulated a greedy algorithm to obtain efficient solutions for the aggregate mean waiting time and total costs. For the problem in this section, we have a constraint on the aggregate mean waiting time at each LW, which means that we cannot consider a related

problem with only two objectives. Hence, we apply the greedy logic differently. As before, in the first stage of the greedy heuristic, we search for a solution $\hat{\mathbf{S}}$ that minimizes the total costs $C(\mathbf{S})$. Per SKU i , we follow a greedy logic to obtain a close-to-optimal solution $\hat{\mathbf{s}}_i$ that minimizes $C_i(\mathbf{s}_i)$: We start with $\mathbf{s}_i = (0, \dots, 0)$ and increase in each step the $s_{i,n}$ that gives the largest decrease in $C_i(\mathbf{s}_i)$. The solutions $\hat{\mathbf{s}}_i$, $i \in I$, form a close-to-optimal solution $\hat{\mathbf{S}}$ for the minimization of $C(\mathbf{S})$.

For the second stage of the greedy heuristic, we define a distance function $d(\mathbf{S})$:

$$d(\mathbf{S}) := \sum_{n \in N} (W_n(\mathbf{S}) - W_n^{\text{obj}})^+, \quad \mathbf{S} \in \mathcal{S},$$

with $x^+ := \max\{0, x\}$ for all $x \in \mathbb{R}$. This function denotes the distance to the set of feasible solutions. If $d(\hat{\mathbf{S}}) = 0$, then $\hat{\mathbf{S}}$ is feasible and will be close to optimal for Problem (P'). In that case, no further steps are needed. If $d(\hat{\mathbf{S}}) > 0$, then we calculate

$$\Delta_{i,n} d(\hat{\mathbf{S}}) = d(\hat{\mathbf{S}} + \mathbf{E}_{i,n}) - d(\hat{\mathbf{S}}), \quad i \in I, n \in N,$$

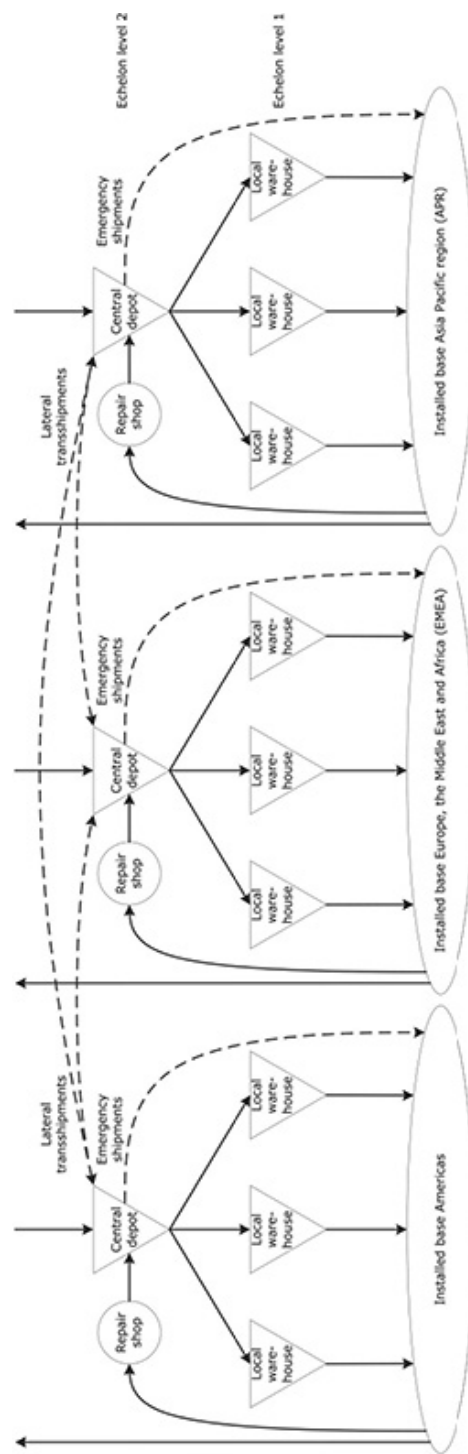
$$\Delta_{i,n} C(\hat{\mathbf{S}}) = C(\hat{\mathbf{S}} + \mathbf{E}_{i,n}) - C(\hat{\mathbf{S}}), \quad i \in I, n \in N,$$

where $\hat{\mathbf{S}} + \mathbf{E}_{i,n}$ is the same solution as $\hat{\mathbf{S}}$ but with $s_{i,n}$ increased by one unit ($\mathbf{E}_{i,n}$ is a matrix with a 1 at position (i, n) and zeros at all other positions). The combination (i, n) with the highest value for $\Gamma_{i,n} = -\Delta_{i,n} d(\hat{\mathbf{S}}) / \Delta_{i,n} C(\hat{\mathbf{S}})$ gives the strongest decrease in distance per unit increase in costs. Hence, $s_{i,n}$ for this combination (i, n) is increased by 1 unit, and similarly further steps can be executed. This is continued until a feasible solution is obtained.

The greedy heuristic as formulated here has been tested by Wong et al. (2005) and performs well, both in terms of quality of the generated solution and computation time.

19.3.4 Application in Other Settings

The models and optimization methods that we have introduced can also be used for more complicated networks in practice, with central and local warehouses. When we look at spare parts networks in practice, we see many variations. For networks managed by OEMs, we often observe the archetypal network of Figure 19.2, where the world is divided into three regions (Americas; Europe, Middle-East, Africa (EMEA); and Asia-Pacific) with a central depot per region and many LWs that are replenished from that central depot. The LWs per region may be divided into multiple groups where only the LWs in the same group support each other via lateral shipments. For such networks, a possible approach is to first calculate basestock levels for the central depots based on a multi-item model for them with given targets for the mean replenishment lead times toward the LWs. Next, per group of LWs in each of the regions, the model of Section 19.3.1 can be used to calculate the basestock levels for all SKUs and all LWs of that group. This gives an approach that works for systems of real-life size. Van Aspert (2014) used this approach to develop and implement a planning concept for the service network of ASML, a supplier of photolithography equipment for the semiconductor industry. He applied his planning method to a real-life problem instance with 3,462 SKUs, one central depot, and 35 LWs. The computation time for this problem instance was 12 minutes and this method gave a significant improvement in comparison to the previous planning concept.



Source: Basten and van Houtum (2014).

Figure 19.2 Archetypical network of an OEM who maintains many of its sold systems

Obviously, other approaches are also possible to get a planning method for a whole network. To the best of our knowledge, in the literature, no other planning concepts are available that work for large real-life networks like the network at ASML. Nevertheless, several methods have been studied that do work for smaller problems; see Basten and Van Houtum (2014) and Paterson et al. (2011). Notice that for multi-echelon networks without emergency and lateral shipments (i.e., METRIC-type models), solutions methods for real-life size problems are available and they have been implemented in commercial software packages (see Section 7 of Basten and Van Houtum (2014)).

The logic that we use for spare parts is also relevant for other scarce resources with low-demand rates and the requirement of fast deliveries, such as low-demand items sold by e-tailers (see, e.g., Acimovic and Graves (2015)) and rental goods like library books and tools (see, e.g., Van der Heide et al. (2018)). The general challenge for spare parts networks is that, on the one hand, we want to keep parts in stock at close distance from the places where they are needed (so that fast deliveries are possible), and on the other hand we want to keep them at stock at central places because only at that level the demand streams are sufficiently thick, i.e., we can benefit from the so-called pooling effect. By the use of lateral and emergency shipments, we are able to keep spare parts in stock at many local stock points, but the total stock is operated as if we have one large stock point. This logic is also useful in the aforementioned settings.

19.4 3D PRINTING

A recent development is the 3D printing of spare parts, more formally known as additive manufacturing (AM). In this section, we first give a brief introduction to the technique of 3D printing and its use for spare parts supply. We next give a broad review of the literature on 3D printing of spare parts, after which we discuss the cases in which existing models can be used. We finally discuss several cases for which new models have been developed or need to be developed.

The technique of AM has been around for decades, but it has been used mainly for prototyping. 3D printing of spare parts becomes economically interesting, now that some early patents have expired, the technique is improving, and the industry is digitalizing. Still, 3D printing can, for now, only be used for components consisting of one or sometimes a few different materials, but not for electronic components, for example. The key advantage of 3D printing is that it is fast. Printing a plastic component may take less than a day, while printing a metal part can be done in a few days. Often, some post-processing is required after the printing. Both printing speeds and post-processing requirements are improving, making printing even faster. Due to its speed, the main promise of 3D printing is that it may reduce the number of emergency shipments that need to be performed while simultaneously reducing inventory levels. Emergency shipments may be costly and are often performed via air, which is an environmentally unfriendly way of transportation. Three-dimensional printing may be another way to deal with spare parts shortages that can be less costly and more environmentally friendly. However, currently, it is often not possible to print parts at high speed and at the same quality as their conventional counterparts, especially in the case of metal parts. Therefore, 3D printing may be used in various ways, requiring different models.

For an overview of the *recent literature* on the 3D printing of spare parts, see the literature review by Westerweel et al. (2021). There has been empirical research on determining which

spare parts are economically interesting to print Knofius et al. (2016), Heinen and Hoberg (2019) or how companies may share their printing capacity Hedenstierna et al. (2019). There has further been analytical research on models that include the option of 3D printing (Song and Zhang (2020); Westerweel et al. (2021)), investigating when to switch from conventional manufacturing to 3D printing (Westerweel et al. (2018)), or investigating how to arrange IP licensing such that users of spare parts may be allowed to print spare parts themselves (Zhang et al. (2022)).

We can sometimes use the *existing models* for spare parts inventory control, if 3D printed parts have the same quality as regular parts. For example, if printing is faster and less costly than performing an emergency shipment, then 3D printing can replace the emergency option and we can simply use the lead time and costs of the 3D printing option as the lead time and costs of the emergency option in the models of Sections 19.2 and 19.3. As a second example, if printing is faster and/or less costly than the regular supply option, it could replace that option, or act as a second source of regular supply. In the latter case, we would have a *dual supply model*, on which there exists extensive literature (see Chapter 9 in this book by Van Mieghem and Xin (2022)). Knofius et al. (2021) go one step further and consider a dual supply model in which the 3D printing option leads to lower-quality parts.

In cases where 3D printing is used locally, it might be relevant to take into account that the *3D printing capacity* is limited. While in spare parts inventory models it is typically assumed that the repair or supply capacity is infinite, this is not the case if only one or a few printers are available locally. Taking queueing effects into account, like Song and Zhang (2020) do, would then be necessary, complicating models considerably. However, in most cases the printers would not be used only to print spare parts, in which case an agreed lead time that does not depend on the demand for spare parts may be realistic.

It may be possible to print parts relatively fast using a general-purpose printer that prints lower-quality parts. In that case, printed parts may be used as a *temporary replacement*. Westerweel et al. (2021) study this problem in a setting with periodic review and a so-called order cycle, meaning that regular replenishments arrive at predetermined, equidistant moments in time. In their case study, they consider the peacekeeping mission of the Royal Netherlands Army in Mali. Weekly, parts may be supplied from the central warehouse in the Netherlands to replenish the local stock point. If there is a shortage on any day (period) in between, an expensive emergency shipment may be performed or a part with lower quality may be printed. When a replenishment arrives, the printed parts that have been installed are replaced by regular parts. They find that even if the parts have significantly lower quality (failure rates are ten times as high), there can be huge cost savings and inventory reductions of 47% and 66%, respectively. Based on the sample of parts that they have investigated, they estimate that 10–20% of spare parts can technically be printed.

Using lower-quality printed parts as temporary replacements and next to the emergency option in the single-location model of Section 19.2 is far from straightforward. The reason is that using the emergency option leads to a lost sale for the local warehouse, i.e., if the emergency option is used the part will not need to be supplied later by the local warehouse, while using the 3D printing option would not take away the demand for the local warehouse. It may then also not be straightforward to decide when to replace the temporary replacement with a regular part. One option would be to use the first incoming part, but if a demand arrives soon afterwards, there is another shortage requiring a temporary replacement. So, extra research is needed to find good policies for how to deal with temporary replacements. Such research is

also needed for the models discussed in Section 19.3 when they are enriched with the option of temporary replacements.

We expect that much more research will happen on the topic of 3D printing of spare parts, both on high-quality and lower-quality parts, and on analytical inventory control models as well as on the other topics that we mentioned at the beginning of this section. This should help companies in practice to understand when and how to use 3D printing.

19.5 EXPLOITING FAILURE PREDICTIONS

With the industry digitalizing and sensor technology developing, the IoT allows for more information on the use of equipment and the degradation and failure of components to be obtained, analyzed, and used at a low cost. This can be used, first, to improve periodic maintenance intervals. For example, if the supplier of equipment does not know exactly how often the equipment is used, it may recommend performing preventive replacements yearly. However, if usage can be monitored, the supplier may recommend performing the preventive replacement after a certain number of running hours. Second, the condition of the component may be monitored and that knowledge may be used to perform preventive replacements shortly before a failure may occur, i.e., to implement *predictive maintenance*. If there is no or limited understanding of the failure modes that may occur in a component, then data from many systems can be combined and analyzed using simple statistical methods or more advanced data mining or artificial intelligence techniques. Having knowledge of the underlying failure modes may improve results considerably, especially also because there is often limited failure data available: If equipment is relatively new, only few components will have failed. If some components have failed more often, they are typically upgraded, rendering the collected failure data useless. So, in the case of high-tech equipment with a limited installed base size, it may only be possible to base failure predictions purely on data later in the life cycle of the equipment. Earlier in the life cycle, true understanding of the underlying failure behavior is necessary.

An advantage of periodic maintenance is that the replacement time can be predicted well in advance: If replacements are performed yearly, the exact timing is given for the next few years, and if replacements are performed every so many running hours, an accurate prediction of the replacement moment is possible well in advance. The advantage of predictive maintenance is that knowledge of a *specific* component is used to determine its replacement moment, instead of knowledge on the *average* component, implying that much less remaining useful life is thrown away on average, while the probability of acting too late is also smaller. The predictions of failures and replacement moments can be used in the spare parts supply. This holds even if the failure predictions are not used to preventively replace components, but only to ensure that all required resources (spare parts, service engineers, and tooling) are available at the equipment at the moment the failure occurs. This is especially the case for expensive equipment that is used 24/7, e.g., lithography systems in front-end wafer fabs. In that case, users of the equipment do not want to shut down equipment preventively, since every hour of downtime is costly, no matter whether it is planned or not.

Failure predictions lead to Advance Demand Information (ADI; see Chapter 13 in this book; Kadiyala et al. (2022)). If the ADI arises only a limited time before the replacement is required, then for spare parts inventory control, the demand is very similar to a demand resulting from corrective maintenance. However, if the ADI for all demands for spare parts of SKU

$i \in I$ is perfect and arises l_i time units before the actual demand, while the replenishment lead time is deterministic, then Hariharan and Zipkin (1995) show that we may replace the original replenishment lead time L_i by $L_i - l_i$ in the inventory control models of Sections 19.2 and 19.3. If $l_i \geq L_i$, this implies that spare parts can be ordered just-in-time, leading to zero inventory holding and penalty costs. In practice, not all failures can be predicted, and sometimes failures are predicted that do not actually occur. The ratio of predicted failures that really occur (true positives) divided by the number of failures that occurs (true positives plus false negatives), is denoted by the recall, while the percentage of predicted failures that really occur (true positives) divided by the total number of predicted failures (true positives plus false positives), is denoted by the precision. If the recall q is less than 1, while the precision is 1 and $l_i \geq L_i$, then the spare parts for the predicted failures can be supplied just-in-time. For the remaining, unpredicted failures, spare parts can be stocked using the models from Sections 19.2 and 19.3, but with a lower demand rate $(1 - q)\lambda_i$. If the precision is lower than 1 or the timing is imperfect, things become more complicated. Topan et al. (2018) propose an inventory control model in which recall, precision, and timing may be imperfect.

19.6 CONCLUDING REMARKS

This chapter should have made clear that spare parts inventory control in practice is not easy. The amount of money that is spent on spare parts divided by the number of spare parts used, is very high: Demand rates may be low, spare parts may be expensive, and customers require high system uptimes, implying that spare parts should be stocked close to customers at many different locations, and implying that in case of a stockout, action needs to be taken. Emergency shipments from a central warehouse and lateral shipments from other local warehouses are options to consider in case of a stockout; deciding between them at a tactical level is difficult, and at an operational level it may be even more difficult: A lateral shipment from one location to another may lead to problems at the former location if the stock level is already low there. With the rise of Industry 4.0, we get more options that we may use through 3D printing, and we have more information to act upon through sensors and the IoT. This also allows us to proactively perform emergency or lateral shipments, again giving more options to consider. To manage all the information and to make good decisions, service control towers will need to be developed. All available data from sensors and information on the supply chain status come together in such a control tower. Decision support systems can use techniques from artificial intelligence and operations research to decide how to solve or prevent problems, not only concerning the spare parts supply, but starting from predictive maintenance decisions and including decisions on sending or repositioning service engineers and expensive tooling. In some cases, these decisions may be made automatically, but in the foreseeable future, we expect that most decisions will be validated and confirmed by human decision makers.

To ensure that all information from the installed base is actually available in the service control tower, the service provider needs to have access to the sensor data. Next to the technical challenges that should be overcome, this requires the right agreements between customers and service providers. Since we also see a trend that users want to focus on their core business, and delivering services often has higher margins than selling products, we see that increasingly, service providers offer service contracts or even offer the use of the product instead of the product itself: servitization (see, e.g., Cohen et al. (2006); Guajardo and Cohen (2018)).

Ensuring that service level agreements are met over a finite time horizon requires different models than the infinite horizon models that we discussed in this chapter. Lamghari-Idrissi et al. (2021) give an example of the kind of models that may be required and the mathematical complications this may lead to.

Spare parts inventory control models have been developed for more than half a century. Because of new developments, plenty of new, practically relevant research challenges have emerged. We look forward to many new studies on these challenges.

REFERENCES

- AberdeenGroup. (2005). *The service parts management solution selection report, SPM strategy and technology selection handbook*. AberdeenGroup.
- Acimovic, J., & Graves, S. C. (2015). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing and Service Operations Management*, 17, 34–51.
- Basten, R. J. I., & van Houtum, G. J. (2014). System-oriented inventory models for spare parts. *Surveys in Operations Research and Management Science*, 19(1), 34–55.
- Cohen, M. A., Agrawal, N., & Agrawal, V. (2006). Winning in the aftermarket. *Harvard Business Review*, 84(5), 129–138.
- Fox, B. (1966). Discrete optimization via marginal analysis. *Management Science*, 13(3), 210–216.
- Guajardo, J. A., & Cohen, M. A. (2018). Service differentiation and operating segments: A framework and an application to after-sales services. *Manufacturing and Service Operations Management*, 20(3), 440–454.
- Harihara, R., & Zipkin, P. (1995). Customer-order information leadtimes, and inventories. *Management Science*, 41(10), 1599–1607.
- Hedenstierna, C. P. T., Disney, S. M., Eysers, D. R., Holmström, J., Syntetos, A. A., & Wang, X. (2019). Economies of collaboration in build-to-model operations. *Journal of Operations Management*, 65(8), 753–773.
- Heinen, J. J., & Hoberg, K. (2019). Assessing the potential of additive manufacturing for the provision of spare parts. *Journal of Operations Management*, 65(8), 810–826.
- Hu, Q., Boylan, J. E., Chen, H., & Labib, A. (2018). Or in spare parts management: A review. *European Journal of Operational Research*, 266(2), 395–414.
- Kadiyala, B., Lee, H., & Özer, O. (2023). Information and incentives in inventory management. In Jing-Sheng Jeannette Song (Ed.) *Research Handbook on Inventory Management*. Edward Elgar.
- Karush, W. (1957). A queueing model for an inventory problem. *Operations Research*, 5(5), 693–703.
- Knofius, N., Van der Heijden, M. C., Sleptchenko, A., & Zijm, W. H. M. (2021). Improving effectiveness of spare parts supply by additive manufacturing as dual sourcing option. *OR Spectrum*, 43(1), 189–221.
- Knofius, N., Van der Heijden, M. C., & Zijm, W. H. M. (2016). Selecting parts for additive manufacturing in service logistics. *Journal of Manufacturing Technology Management*, 27(7), 915–931.
- Kranenburg, A. A., & van Houtum, G. J. (2007). Cost optimization in the $(s-1, s)$ lost sales inventory model with multiple demand classes. *OR Letters*, 35(4), 493–502.
- Kulkarni, V. G. (2020). *Modeling and analysis of stochastic systems* (3rd ed.). CRC Press.
- Lamghari-Idrissi, D., Basten, R. J. I., & van Houtum, G. J. (2021). Reducing risks in spare parts service contracts with a long downtime constraint. *IIE Transactions*, 53(10), 1067–1080.
- Muckstadt, J. A. (2005). *Analysis and algorithms for service parts supply chains*. Springer Series in Operations Research & Financial Engineering. Springer.
- Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2), 125–136.
- Reijnen, I. C., Tan, T., & van Houtum, G. J. (2009). *Inventory planning for spare parts networks with delivery time requirements*. Beta Research School, Eindhoven University of Technology. Working paper 280.
- Ruzika, S., & Wiecek, M. M. (2005). Approximation methods in multiobjective programming. *Journal of Optimization Theory and Applications*, 126(3), 473–501.

- Sherbrooke, C. C. (2004). *Optimal inventory modeling of systems: Multi-echelon techniques*. International Series in Operations Research and Management Science. Kluwer.
- Sherbrooke, C. C. M. (1968). A multi-echelon technique for recoverable item control. *Operations Research*, 16(1), 122–141.
- Song, J. S., & Zhang, Y. (2020). Stock or print? Impact of 3-D printing on spare parts logistics. *Management Science*, 66(9), 3799–4358.
- Topan, E., Tan, T., van Houtum, G. J., & Dekker, R. (2018). Using imperfect advance demand information in lost-sales inventory systems with the option of returning inventory. *IIE Transactions*, 50(3), 246–264.
- Van Aspert, M. (2014). Design of an integrated global warehouse and field stock planning concept for spare parts. Ph.D. Thesis, Eindhoven University of Technology. PDEng thesis.
- Van der Heide, G., Van Foreest, N. D., & Roodbergen, K. J. (2018). Optimizing stock levels for rental systems with a support warehouse and partial backordering. *European Journal of Operational Research*, 265(1), 107–118.
- van Houtum, G. J., & Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints*. International Series in Operations Research and Management Science. Springer.
- Van Mieghem, J. A., & Xin, L. (2023). Dual-sourcing, dual-mode dynamic stochastic inventory models. In Jing-Sheng Jeannette Song (Ed.) *Research Handbook on Inventory Management*. Edward Elgar.
- Westerweel, B., Basten, R., den Boer, J., & van Houtum, G. J. (2021). Printing spare parts at remote locations: Fulfilling the promise of additive manufacturing. *Production and Operations Management*, 30(6), 1615–1632.
- Westerweel, B., Basten, R. J. I., & van Houtum, G. J. (2018). Traditional or additive manufacturing? Assessing component design options through lifecycle cost analysis. *European Journal of Operational Research*, 70(2), 570–585.
- Wong, H., van Houtum, G. J., Cattrysse, D., & Van Oudheusden, D. (2005). Simple efficient heuristics for multi-item multi-location spare parts systems with lateral transshipments and waiting time constraints. *Journal of the Operational Research Society*, 56(12), 1419–1430.
- Zhang, Y., Westerweel, B., Basten, R. J. I., & Song, J. S. (2022). Distributed 3D printing of spare parts via IP licensing. *Manufacturing and Service Operations Management*, 24(5), 2685–2702.

20. Retail inventory systems

Stefan Minner and Anna-Lena Sachs

20.1 INTRODUCTION

Retailing is an important application domain of inventory management. Many retail-specific topics require extension and adjustments of traditional inventory models and solution algorithms to provide competitive decision support (Fisher, 2009). Retail operations considerably increase the problem complexity of inventory management (Mou et al., 2018). Contrary to standard inventory control models,

- 1) Retail practice requires the simultaneous consideration of multiple products defining an assortment and sharing available space and often being ordered, transported and handled jointly.
- 2) Multiple stores need to be considered simultaneously, which are supplied from a single or multiple central warehouses to cross-dock incoming supply and share transportation resources in delivery tours. Therefore, the problem inhibits a multi-echelon structure, which can be further complicated by allocation problems and potential interaction through transshipments.
- 3) In recent e-commerce strategies, retail inventory management problems need to consider multiple sales channels, i.e., a direct online sales channel and a traditional store sales channel.
- 4) Demands are typically hard to be modeled exogenously as they depend on endogenous factors such as decisions about promotions, advertisements, and pricing and exogenous factors such as weather and competitor decisions. Therefore, building data-driven demand models and integrating rather than separating forecasting and inventory decisions are core for performance, in particular, in low-margin retail environments.
- 5) Demand is mostly stochastic in stationary retail, as retailers do not know in advance how many customers are going to shop at their stores and which items they will buy. In online retailing (see Chapter 21), this can be different if customers place their orders in advance or have booked a recurring delivery as it is often the case with groceries. Furthermore, demand is often non-stationary, due to demand patterns varying over time. Demand variations occur depending on the time of the day, day of the week, beginning or end of the month, holidays, and seasons (Ehrenthal et al., 2014). The type of demand variation is usually store-specific, e.g., a store in a busy city location faces a different demand pattern during the day compared to a store in a rural area from the same retail chain.
- 6) If demand cannot be met, a stockout situation occurs. Part of the demand is then shifted to a substitute product, whereas the remaining part results in a lost sale (Corsten & Gruen, 2003). Frequent out-of-stock situations result in customer dissatisfaction and eventually, customers might decide to buy elsewhere in the long run (Andersen et al., 2006). Taking out-of-stock situations into account is important in any type of retail situation,

as otherwise sales could be mistaken for demand observations which would result in an underestimation of demand in the future, and thus, too low service levels.

For inventory management practices to be successful, i.e., by better matching supply and demand, decision makers should consider these characteristics that are prevalent in retailing.

Many different standard software providers like SAP offer tailored retail inventory management solutions and automated store ordering (ASO) systems are a valuable resource to support retailers with their replenishment decisions. However, ASO systems usually apply standard inventory policies or heuristics and do not distinguish between perishables and non-perishables (Van Donselaar et al., 2006). The ASO systems then recommend order quantities every order cycle (e.g., once per day) and the store managers can overwrite the recommendation. Van Donselaar et al. (2010) found that store managers might have different incentives to overwrite ASO recommendations by considering in-store handling costs and stimulating sales through higher service levels. ASO systems can also leverage additional data by sharing information between supply chain partners. As a result, replenishment decisions can be coordinated more effectively, which leads to a reduction in food waste (up to 20% according to Kiil et al., 2018).

In the following sections, we discuss the different aspects that a decision-maker should take into account when dealing with inventory decisions in retailing. In Section 20.2, we outline the data that is required as input for inventory decisions and the key performance indicators that should be considered when setting an objective for inventory optimization. We then present several strategic and tactical decisions in Section 20.3 that determine the environment in which a retailer can make inventory decisions. In Section 20.4, we describe inventory models that address different settings and decisions a retailer might face. We conclude this chapter in the last section and provide an outlook on future research areas.

20.2 DATA AND KEY PERFORMANCE INDICATORS

Before making any inventory decisions, the retailer collects data from various sources, e.g., through point-of-sale scanner systems and its ERP system, and identifies the characteristics of the particular retail setting to be taken into account. This analysis should guide the retailer when selecting appropriate inventory models. The data required to make inventory decisions consists of information on demand, lead time, costs and revenues, service levels, and inventory records.

20.2.1 Demand Characteristics

We describe several demand-related aspects, such as classification and segmentation, demand distribution and forecasting, unobservable lost sales and stockout-based substitution.

20.2.1.1 Classification and segmentation

Classification and segmentation are important instruments to assign suitable inventory modeling and control techniques to different items. ABC and XYZ classification belong to the standard repertoire of every inventory management solution (Silver et al., 2016) not only for distinguishing fast and slow-moving consumer goods, but also for selecting different demand

models with regard to demand patterns, demand predictability, choice of service level types and values and the selection of appropriate inventory control rules. While for A-category products, a few determine the largest turnaround value, many C-category products only show low sales volumes and values and therefore contribute less. As a widely used rule of thumb, the A-category makes up for 80% of the annual consumption value of products while containing only 20% of the items; the B-category with 30% of the items representing 15% of the annual consumption value, whereas the C-category only represents 5% of the consumption value with 50% of the items included. As a consequence, different data and planning efforts should be spent on the different categories. Similarly, X-category products have regular demand and can be better forecasted than Z-category items with many zero and erratic demands. As a rule of thumb here, the X-category could include all items with a coefficient of variation of demand less than 0.5 whereas the Z-category could include those items with many zero demands and a coefficient of variation larger than one. Svoboda and Minner (2021) summarize the relevant literature and present a machine-learning-based decision tree approach for classifying items and suggesting demand distribution, service level type and size and inventory control policy for each class. However, a remaining challenge is the robustness of such a classification over time. In particular, for seasonal and promotional products, the categorization might outdate quickly within a dynamically changing assortment.

20.2.1.2 Demand distribution

We first develop an understanding of demand patterns and characteristics before choosing a suitable inventory model. The distribution with the best fit may vary by product, store, or even time in retailing. The normal or the Poisson distribution are often used in inventory models (Agrawal & Smith, 1996). The normal distribution is appropriate when demands are large and less variable, while the Poisson assumption is more appropriate for small quantities (e.g., bakery items), where rounding would result in comparatively large errors (Schulte & Sachs, 2020). Agrawal and Smith (1996) find that the negative binomial distribution provides a generally better fit to their data than either the normal or Poisson, as it combines the advantages of both distributions. The negative binomial distribution has the ability to capture a potentially large variability of demand (e.g., due to weather, promotions) while accounting for the discrete nature of the demand. Theoretically, the negative binomial distribution results from Poisson demands with an unknown Gamma-distributed mean or a compound Poisson demand with a logarithmic transaction size distribution (Johnston et al., 2003). In some settings, it can be difficult to find one distribution that fits to all products and/or stores, and a non-parametric approach avoids having to impose some kind of distribution that may not fit to all (Beutel & Minner, 2012).

20.2.1.3 Forecasting

In a recent review, Fildes et al. (2022) distinguish between different levels of aggregation for forecasting in retailing. We will focus on what they call “product-level demand forecasting” as this is the most commonly used level for the inventory decisions we will discuss in Section 20.4. For forecasting different aggregation level demands within the M5-competition on Walmart data, see Spiliotis et al. (2021). Product-level forecasting implies that a large number of forecasts are required, as retailers often carry tens of thousands of different products in each store. To enable inventory decisions at the store level, demand forecasts for each product and store are required, which means that the number of forecasts is a multiple of the number

of products and the number of stores. As a result, there is usually not one forecasting method that fits to all product-store combinations and clustering can be helpful (Boylan et al., 2014). The high granularity of the data can also pose additional challenges if the data contains many zeros and is thus intermittent (Kolassa, 2016). The retailer has to decide whether time-series or causal forecasting methods are more suitable. This decision depends on the data available, the structures within the data, the decisions the store manager has to make and the customer demand behavior. For a general introduction to time-series and causal forecasting methods, we refer to Ord et al. (2017).

When considering time-series elements, either for choosing a time-series method or including explanatory variables in a causal method, the retailer should check for seasonality (e.g., the month of the year, beginning/end of month effects, weekdays) and trend. Furthermore, calendar events such as holidays like Christmas, sports events like the WorldCup, or local festivals, often have an important effect on demand. The effect might occur for all products or only a few and might even be in opposing directions depending on the category. For example, demand for beer and barbecue products probably increases during the World Cup. Other external factors, such as price and promotions, are well-known to have an effect on demand. Not only the current selling price of a product may be relevant, but also price differences (absolute or relative) to previous periods or to other products in the same category. In some settings, e.g., if the retailer applies an every day low price strategy, there are no price promotions and the sales price stays the same. In this case, and if there are not other important explanatory variables, time-series methods provide better forecasts than methods that take more factors into account (Kolassa, 2016). Demand often also depends on the weather. E.g., demand for ice cream increases when it is warm and sunny. However, the difficulty here is that the weather is also not known in advance and only the weather forecast could be used as explanatory variable, but not the actual weather. The longer the forecast horizon, the more difficult it gets to obtain a reliable weather forecast.

20.2.1.4 Unobservable lost sales

An important aspect of retailing is that demand exceeding supply is often unobservable. In this case, we have unobservable lost sales and need to estimate demand based on sales. Ignoring unobservable lost sales would underestimate demand in the future and hence, order too little. There are different approaches for estimating unobservable lost sales. Parametric approaches assume that demand follows a theoretical demand distribution. From knowing the properties of a distribution and the sales observed, we can infer the unobservable lost sales.

When choosing an estimation method for unobservable lost sales, several aspects need to be considered:

1. How granular is the sales data and does it contain information that is helpful to determine stockout times?
2. Can the inventory level be determined accurately?
3. Does a theoretical demand distribution provide a good fit for the data?

The first question is important to identify whether and when a stockout occurred. Let us assume that a retailer collects hourly sales data, as it is a commonly used feature of point-of-sale scanner systems. On days where the retailer observes sales until the store closes, the retailer assumes that the observed sales are complete and thus serve as a good estimate of

actual demand. On days when sales stop at some point during the day, and more demand could be expected according to the previous sales history, the last time of a sale is an indicator of the stockout time. Jain et al. (2015) show the importance of considering the timing of stockouts versus only the event of a stockout occurrence.

The answer to the second question can help to identify stockouts if no information on timing is available or to refine the estimate if both types of data can be obtained. Stockouts can be identified by comparing available inventory and sales. If accumulated sales equal available inventory, the retailer has sold all inventory and a stockout occurs. However, retailers do not always have accurate inventory records, in which case, this comparison might not be possible. In particular, if a product has a shelf-life longer than the replenishment intervals, inventory of different ages are mixed on the shelves and some items expire or do not look good for sale anymore, in which case, inventory records would still show available inventory, but the customers would not perceive it as such and do not buy it.

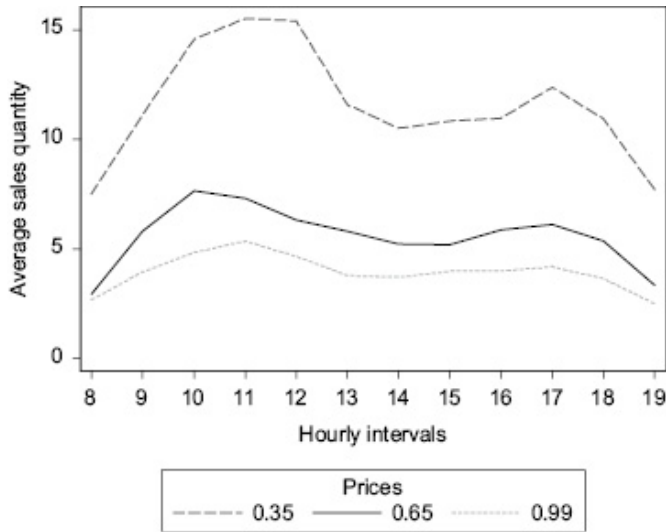
The third question helps to select a parametric estimation method, in particular, if stockout timing is not available and therefore, non-parametric estimation techniques would not provide suitable estimates. Which theoretical distribution fits best often varies between stores and products, and potentially other characteristics (such as weekdays, seasons, etc.). This makes it more challenging to find a distribution that fits to all products and stores. There are different approaches that consider the properties of the respective distribution, e.g., Wecker (1978) or Bell (1981) for the normal distribution.

The non-parametric approach by Lau and Lau (1996) works well with point-of-sale (POS) scanner data and does not require fitting a demand distribution. The approach is also quite intuitive by establishing sales patterns from periods when no stockout occurred and then transferring this information to periods with stockouts. Often, the sales data is collected on an hourly basis from POS scanner systems so that the approach establishes a sales pattern over each day. In Figure 20.1, we show sales patterns for butterhead lettuce in a store that opens at 8 am and closes at 8 pm. The product was sold for three different prices over time and we can see from the figure that the overall pattern is similar for all three prices, but shifted upwards when the price decreases compared to the highest price observed during that period which was €0.99.

20.2.1.5 Stockout-based substitution

If a product stocks out, this may have an effect on other products. Depending on whether suitable substitutes exist, a portion of the unfulfilled demand may be shifted to a substitute product. As a result, the demand for the substitute is higher than it would normally be had the other product still been available. The difficulty in estimating substitution rates is that the substitution decision of a customer is unobservable and we can only observe the change in demand for the substitute. A customer not finding their first choice is unlikely to ask the store manager about substitutes but rather make the decision themselves. Therefore, when estimating substitution rates, it is important to select appropriate products that might serve as potential substitutes.

Karabati et al. (2009) suggest a state-space-based approach that requires only point-of-sale scanner data and does not make any assumptions about the inventory policy used or the distribution of customer demand. The data is categorized by whether an item is available or stocks out during a POS interval (e.g., one hour). They minimize the squared error from fitting the parameters for substitution to the observed sales data. A function describes each possible



Source: Sachs and Minner (2014).

Figure 20.1 Sales patterns for different prices

combination of available and unavailable products. Each item may stockout and may be a substitute for another product.

The approach by Anupindi et al. (1998) follows a similar logic by categorizing the observations according to which products are available and considering the timing of stockouts. Their approach assumes that demand is Poisson and the arrival rates are estimated using Maximum Likelihood. When collaborating with a large European retail chain, we estimated the substitution rates for butterhead and iceberg lettuce using the approach by Anupindi et al. (1998). The Poisson distribution provides a good fit for both products. The two products seem reasonably similar to be mutual substitutes. As a result, we obtained a substitution rate from butterhead to iceberg lettuce of 70%. This means that if a customer wanted to buy a butterhead lettuce as their first choice, they would then buy an iceberg lettuce with a substitution rate of 70% instead. Interestingly, the opposite case, where the iceberg lettuce is out of stock only yields a substitution rate of 27%. Before showing the results, we asked several managers what their estimate would be and received answers very close to our estimates. The managers explained that even though both lettuces were used for the same purposes (e.g., salads and sandwiches), an iceberg lettuce lasts longer than a butterhead lettuce. Therefore, a customer planning to consume an iceberg lettuce over a longer period of time might not find that a butterhead lettuce is a suitable substitute.

20.2.2 Lead Times, Costs, and Prices

Lead times are a core data figure for the replenishment of items in inventory theory and we refer to Part I (Fundamentals – Theory and Methodologies) of the book for different modeling options and challenges. Although many inventory models assume constant or random

lead times, in retail these might depend on external factors such as the country of origin for fresh fruits and therefore can benefit from state-of-the-world-based modeling (Zipkin, 2000).

Many well-known inventory structures and policy analyses assume rather simple cost structures, i.e., linear ordering, holding and stockout penalty costs. However, in retail, both the ordering and the demand fulfillment process typically show more complex cost structures with dependencies across products (Krisnadewi & Soewarno, 2019). Further, handling efforts determine an important cost when ordering, transporting, and storing units (Curşeu et al., 2009) and ordering costs occur per order, per ordered product and per unit ordered. Ordering cost can exhibit both convex and concave structures modeling either economies of scale and discounts (Munson & Rosenblatt, 1998) on the one hand side or increasing costs when certain quantities and inventory levels are exceeded. For complex ordering cost structures related to transportation, we refer to Sections 20.4.2 and 20.4.3 of this chapter. Further, holding cost rates might be inventory level dependent. Finally, stockout penalty costs can occur per stockout, per unit out of stock or per unit and unit of time being out of stock (Silver et al., 2016).

In particular in retail, the interface between marketing and operations is important where typically, price setting is seen as a marketing decision and inventory replenishment as an operations decision. However, both decisions interact and it has been shown, that integrated rather than sequential decision-making offers substantial improvements or even renders unprofitable products profitable. Prices might not only affect the demand level, but also its demand volatility. For joint pricing and inventory replenishment in the context of the Economic Order Quantity (EOQ), we refer to Eliashberg and Steinberg (1993), for summaries of price-setting newsvendor models to Petruzzzi and Dada (1999) and DeYong (2020), and for approaches at the interface of revenue and inventory management to Chen and Simchi-Levi (2012).

20.2.3 Service Levels, On-Shelf Availability, Inventory Record Inaccuracy, Tracking, and Tracing

An important measure to control for customer satisfaction in inventory models is service level. Due to the difficulty to measure stockout costs in many applications, non-stockout probability and fill-rate are frequently used performance measures (Silver et al., 2016). The non-stockout measure is empirically determined by the ratio of periods without stockout and the number of periods involved in the evaluation horizon. The (per stock-keeping unit) fill-rate is determined by the fraction of demand satisfied from stock and the total demand over the evaluation period. In particular, the concept of on-shelf availability has received considerable attention in the retail literature (Trautrimis et al., 2009). Besides the different types of service levels incorporating stockout occasions and magnitudes, measures can be defined by item, by assortment category (in particular group service levels for substitutable products) and also with regard to different times of the day and planning horizons (Minner & Transchel, 2010).

Digitalization of inventory management requires an accurate inventory record tracking to support decision automation. However, for various reasons (DeHoratius & Raman, 2008), inventory records in particular in bricks-and-mortar stores are often inaccurate and deteriorate performance when not being incorporated in determining inventory positions and render automation difficult. Therefore, technological solutions for continuously monitoring inventory levels and statistical approaches to infer potentially inaccurate records from sales anomalies have been suggested to improve data quality (Rekik et al., 2019). For perishable products with units of different best-before dates on the shelf, inventory depletion (e.g., FIFO and LIFO as two extremes)

has a considerable impact on sales and outdated inventory (Minner & Transchel, 2017), however, tracking the inventory per age category is a challenge (Broekmeulen & Van Donselaar, 2019).

Additional important performance indicators are relevant in online retailing, such as order split rate, fulfillment speed and order fulfillment service level. We refer to Chapter 21 in this book on online retailing inventory management.

20.3 STRATEGIC AND TACTICAL DECISIONS

The retailer's medium- and long-term decisions determine which options the retailer has available when it comes to placing orders according to the inventory models presented in Section 20.4. A very important aspect is the assortment, which determines for which products orders will be placed. The replenishment pattern controls when and how often orders are placed and the case-pack size of the quantities that can be ordered.

20.3.1 Assortment Planning

The products held in an assortment have important implications for all other inventory decisions in the retail context. The assortment that is carried by a store can be determined at a global level, where all stores carry the same assortment, or vary by store depending on customer preferences. Either strategy has its own advantages, with the former being easier to administer if all stores require the same decisions, and the latter being more flexible to target different customer groups. The pioneering work by Pentico (1974) addresses the assortment planning problem with stochastic demand.

The assortment available to a customer also influences the demand observed for all the products in the assortment. While carrying a larger assortment means that the customer can choose from a larger variety (which potentially increases demand), the retailer also incurs higher inventory costs due to having more products in stock. Van Ryzin and Mahajan (1999) analyze this trade-off between inventory costs and product variety benefits. More product variety can result in cannibalization effects, where some demand for a product is shifted to another one that was added to the assortment (assortment-based substitution). The model of Van Ryzin and Mahajan (1999) has resulted in many extensions, such as including pricing (Maddah & Bish, 2007) or varying the assortment over the selling period (Topaloglu, 2013). Another interesting approach can be to learn from customer purchases about their preferences. In the assortment planning problem studied by Sauré and Zeevi (2013), the retailer varies the products that are available on a limited shelf space. As a result, the retailer identifies the best set of products that should be part of the assortment.

The assortment planning problem is closely linked to the inventory decision problem with stockout-based substitution. Depending on which products are available as part of the assortment, stockout-based substitution can take place only for these products. Similarly, if customers perceive products as close substitutes with high substitution rates, the retail manager should consider whether carrying close substitutes in the assortment is the best use of space or whether other products, for which no substitute exists, should be carried instead. Smith and Agrawal (2000) solve the joint inventory and assortment problem to determine which products to carry in the assortment and the inventory level for each product. Further constraints such as shelf space can be considered as part of their approach. This model is extended by Mahajan

and Van Ryzin (2001), who allow for dynamic consumer substitution. In this case, there is no restriction on the number of substitution attempts and the substitution rates depend on the products available. They show that a retailer should stock more of popular items and less of unpopular items compared to a newsvendor without substitution. This finding results from two opposing effects: with substitution, demand for a product that serves as a substitute increases which should lead to higher order quantities. In contrast, the underage cost decreases as understocking may result in sales of a substitute rather than a lost sale. Depending on the popularity of a product, one of the two effects dominates the other one.

Motivated by the decision problem the retail company Albert Heijn faces in the Netherlands, Kök and Fisher (2007) present an approach to determine the demand and substitution parameters and then iteratively maximize expected profit while considering shelf space. Shelf space is reflected by the number of facings available. They achieve significant improvements compared to the previous assortment at Albert Heijn by adjusting the space allocated to subcategories, changing the products in the assortment and the considering the effect of case packs on the facings allocated.

20.3.2 Replenishment Patterns

Before making any inventory decisions, the retailer has to decide how often the inventory should be reviewed and when orders should be placed. It is rare that inventory is reviewed continuously in retailing. The store manager rather assesses the inventory available on the shelf (and potentially in the backroom) in regular time intervals, such as daily or weekly. The replenishment pattern is usually fixed for at least the next three to six months, if not longer, as a regular pattern not only facilitates inventory planning, but also other operational decisions such as workforce scheduling and planning transportation routes (Holzapfel et al., 2016).

Taube and Minner (2018) present a data-driven optimization approach and a case study to set replenishment patterns in retail. Kuhn and Sternbeck (2013) provide in-depth insights into the replenishment patterns employed at 28 European grocery retailers based on semi-structured interviews. Most companies use different replenishment patterns depending on the type of product. For example, fresh produce is frequently replenished from nearby suppliers to ensure short lead times and freshness. Most products are replenished at least once a week, with the majority being replenished three to six times per week. The retailers usually determine the delivery frequency first and then set the replenishment pattern, such as specific weekdays on which an order is placed. They also distinguish between product groups and stores. The chosen frequency and pattern then affect all parties involved in the supply chain.

20.3.3 Case Packs

Inventory decisions are often restricted to case-pack sizes, i.e., they have to be a multiple of a given case-pack size. The case pack thus determines the minimum order quantity and consequently affects the frequency with which an order is placed and in-store inventory levels (Wensing et al., 2018). The case-pack size can also affect the shelf space required for a product in a store as companies often employ a “packout” rule, where the shelf space allocated to a product is such that all items in a case pack fit on the shelf, or a “pack-and-a-half” rule (Eroglu et al., 2011).

The smaller the case-pack size, the more flexible the inventory decisions become. Ketzenberg et al. (2002) show the benefits of breaking case packs into individual products at the warehouse, which means that the stores can carry a broader assortment in the same space or require less space for a given assortment. However, this comes at higher in-store handling costs, which account for the largest share of operational logistics costs for non-perishable products in the retail supply chain according to Van Zelst et al. (2009). They show that increasing case-pack sizes results in large efficiency gains due to a faster shelf-stacking process.

Case-pack sizes typically vary between perishable products with smaller case-pack sizes and non-perishable products with larger case-pack sizes (Van Donselaar et al., 2006). Eroglu et al. (2011) show the effects of case-pack size, shelf space and consumer demand on stock-outs as well as moderation effects between the different aspects. They consider a case with backroom storage, so if there is not enough space available on the shelf, the remaining items can be moved to the backroom. However, backroom storage comes with additional challenges, such as ensuring a reliable backroom-to-shelf replenishment process. The store managers must monitor inventory in two locations (backroom and shelf), but benefit from having additional storage capacity, which should be taken into account when making a new order decision (Eroglu et al., 2013).

Wensing et al. (2018) optimize case-pack sizes for a large European retail chain, which results in potential cost savings of more than 20%. They compare two different policies, one with the same case-pack size for all stores, and the other one with individual case-pack sizes per store. While the latter is much more complex in the overall logistics operations, the former already captures the majority of the savings and is preferred by the retail chain. The choice of a case-pack size not only affects the retailer's operations along the whole supply chain, but also its supplier's logistics. Therefore, case-pack sizes are usually determined as part of the tactical planning for the next six to twelve months (Hübner et al., 2013). Teulings and Van der Vlist (2001) propose the use of mixed-loads in multi-product environments, i.e., to mix several products in given quantities on a load unit (e.g., pallet) to increase order quantity flexibility, but retain transportation efficiency.

Recent innovations in warehousing have opened up new opportunities to leverage the efficiencies of large case-pack sizes, but allowing for smaller order quantities at the customer level or store level by breaking case packs into individual products for storage. Using robots that transport inventory on pods from storage to replenishment stations, the KIVA system described in Chapter 21 allows for reacting more flexibly to changing customer demand and adapting the warehouse design accordingly.

20.4 OPERATIONAL DECISIONS

After having identified the data available and the conditions under which an inventory decision is to be made, we present core inventory models that address the characteristics prevalent in retailing. In addition to the more standard and sequential predict-then-optimize models, we present several data-driven models in this chapter, where prediction and optimization are integrated and which emerged more recently in the literature. We describe the data-driven newsvendor model with unobservable lost sales in Section 20.4.1 as it considers external factors, integrates forecasting and inventory optimization, and considers unobservable lost sales

due to stockout situations. In Section 20.4.2, we consider joint replenishment problems, where the ordering pattern is determined. The inventory routing problem covered in Section 20.4.3 combines inventory management and transportation and Section 20.4.4 discusses several approaches to the multi-echelon inventory problem most common in retailing, which is a one-warehouse multi-retailer system. We conclude this chapter with a discussion of dynamic pricing in Section 20.4.5, which has become increasingly popular in retailing and strongly affects inventory management.

20.4.1 Data-Driven Newsvendor with Unobservable Lost Sales

In the standard newsvendor model, the decision-maker determines the order quantity for the future assuming that the demand distribution is known. Any inventory at the end of the day is discarded. The standard newsvendor is a special case of the models described in Chapters 1 and 2.

Store managers often face a similar problem for perishable products, but the demand distribution is not known and demand might depend on external factors such as price and weekdays, which can help to make a better order decision. Any demand not satisfied from stock is lost. Beutel and Minner (2012) suggest a data-driven newsvendor model, where the optimal order quantity is determined as a function of the external variables using linear programming. For example, if the price is low, demand is high and the retailer should order more than on a day with a high price. This order function is fit to the demand observations taking the underage and overage costs into account. If the underage costs increase, the coefficients of the function change so that fewer lost sales occur. Elmachetoub and Grigas (2022) highlight the importance of incorporating the decision-maker's objective and constraints when designing prediction models and measuring their performance using the same parameters as in the objective. Ban and Rudin (2019) solve the data-driven newsvendor model using empirical risk minimization and an algorithm based on kernel-weights optimization. They derive bounds on the out-of-sample cost based on in-sample information, and highlight the importance of controlling for potential overfitting issues in data-driven optimization. An extension to non-linear problems is proposed by Huber et al. (2019) using different machine-learning approaches and quantile regression. They show that data-driven approaches often outperform sequential decision-making, but only if there is enough data available to make a reliable estimation.

In the following, we consider the practical situation where the retailer's historical observations only consist of sales information, not demand, if there was a stockout, from Sachs and Minner (2014). The retailer collects historical sales data from POS scanner systems for a newsvendor product and demand depends on external variables such as price, weather forecast and weekdays. The retailer's objective is to minimize leftover inventory and shortage penalty costs, as in the standard newsvendor problem. The data required for this approach consists of historical sales observations D_i for each period i , the values of the external factors in this period X_{ji} , where m external factors are considered with $j = 1, \dots, m$. Consequently, we can write the target inventory level B_i as product sum of coefficients B_j and external factors X_{ji} :

$$B_i = \sum_{j=0}^m \beta_j X_{ji}. \quad (20.1)$$

Note that the sum starts at $j=0$, as the function includes a constant β_0 and hence, $X_{0i} = 1$. The difficulty in this setting is that historical demand was not fully observable if there was a

stockout. For this purpose, the model estimates unobservable lost sales using a sales pattern based on previous full demand observations based on Lau and Lau (1996).

To establish the sales patterns, each period i (e.g., day) is divided into $t = 1, \dots, T$ discrete time intervals (e.g., hours), for which sales h_{ti} are recorded. The cumulative sales H_{ti} correspond to demand D_i at the end of each period if there was no stockout. In this case, $H_{Ti} \equiv D_i$ and $\frac{H_T}{D} = 1$. We denote the average cumulative demand calculated across all periods without stockouts as $\bar{H}_t$, and hourly ($\bar{h}_t$) and daily demand ($\bar{D}$), respectively.

The retailer categorizes the historical data into full ($F = \{1, \dots, g\}$) and censored ($C = \{g+1, \dots, N\}$) demand observations. Each demand observation in $i = 1, \dots, N$ is either assigned to F or C , depending on whether demand was fully observable or a stockout occurred. In the latter case, the time of the stockout is recorded as $t = e_i$. If this information is not available, it can be approximated by observing if a product's sales stop at some point during the day, when more sales could normally be expected.

We can then work backward to obtain the fraction of demand that took place in the previous time interval with

$$\frac{\bar{H}_{T-1}}{\bar{D}} = 1 - \frac{\bar{h}_T}{\bar{D}} \quad \text{with } \bar{h}_t = \frac{\sum_{i=1}^g h_{ti}}{g} \quad (20.2)$$

and then continue by obtaining the ratios for the remaining time intervals as

$$\frac{\bar{H}_t}{\bar{D}} = \frac{\bar{H}_{t+1}}{\bar{D}} - \frac{\bar{H}_{t+1}}{\bar{D}} \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}} = \frac{\bar{H}_{t+1}}{\bar{D}} \left(1 - \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}} \right) = \frac{1}{E_t}. \quad (20.3)$$

Consequently, it follows that

$$E_t = \frac{E_{t+1}}{\left(1 - \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}} \right)} \quad \text{with } E_T = 1. \quad (20.4)$$

For the periods with censored demand observations, we can obtain demand up to the point in time when the stockout occurred in e_i . By multiplying the cumulative sales that were observed before the stockout occurred with the estimator, we can obtain an estimate of demand. Note that a stockout typically occurs during a time interval, so we take the average of the two estimators E_{e_i} and E_{e_i-1} and obtain $\hat{D}_i$ as an estimate of demand as

$$\hat{D}_i = \frac{H_{e_i}(E_{e_i-1} + E_{e_i})}{2} \quad \forall i \in C. \quad (20.5)$$

We can then formulate a mixed-integer linear programming model to determine the coefficients of the inventory function. The objective function in Equation (20.6) balances the trade-off between ordering too much or too little. For every unit of leftover inventory y_i , a holding cost h is incurred. A shortage cost v penalizes for every unit of demand D_i that exceeds sales s_i for complete demand observations in $i = 1, \dots, g$. In the case of censored demand observations, demand is approximated by its estimate in Equation (20.5).

$$\min C = \sum_{i=1}^N h y_i + \sum_{i=1}^g v(D_i - s_i) + \sum_{i=g+1}^N v \left(\frac{H_{e_i}(E_{e_i-1} + E_{e_i})}{2} - s_i \right) \quad (20.6)$$

s.t.

$$s_i \leq D_i \quad \forall i \in F \quad (20.7)$$

$$s_i \leq \frac{H_{e_i}(E_{e_i} + E_{e_i-1})}{2} \quad \forall i \in C \quad (20.8)$$

$$s_i \leq \sum_{j=0}^m \beta_j X_{ji} \quad \forall i \in F \cup C \quad (20.9)$$

$$y_i \geq \sum_{j=0}^m \beta_j X_{ji} - D_i \quad \forall i \in F \quad (20.10)$$

$$y_i \geq \sum_{j=0}^m \beta_j X_{ji} - \frac{H_{e_i}(E_{e_i} + E_{e_i-1})}{2} \quad \forall i \in C \quad (20.11)$$

$$s_i, y_i \geq 0 \quad \forall i \in F \cup C \quad (20.12)$$

$$\beta_j \in \mathfrak{R} \quad \forall j = 0, \dots, m \quad (20.13)$$

Constraints in Equation (20.7) for full and Equation (20.8) for censored demand observations ensure that sales cannot exceed demand. The constraint in Equation (20.9) limits sales to the amount of inventory that is available as a function of the external factors X_{ji} . The constraint in Equation (20.10) determines the leftover inventory y_i , which is discarded at the end of the day. The leftover inventory can be calculated as the difference between the inventory level and demand. Similarly, the constraint in Equation (20.11) determines the leftover inventory for estimated demand in the presence of stockouts. Note that while there was a stockout observed, there might not be a stockout in the data-driven model if the model had set a higher order quantity than what was observed in the historical data. Equations (20.12) and (20.13) define the decision variables, where $\mathfrak{R}$ is a real number.

20.4.2 Joint Replenishments

As mentioned in Section 20.3, it is not optimal to replenish every product every period. Implicitly, the choice of order frequency determines the review period of a product. Such a restriction typically only slightly increases inventory and stockout costs, but considerably reduces operational costs for ordering, transportation, picking, and shelf stacking. Taube and Minner (2018) develop a two-stage data-driven mixed-integer linear programming approach that determines ordering patterns at the first stage and ordering execution at the lower level

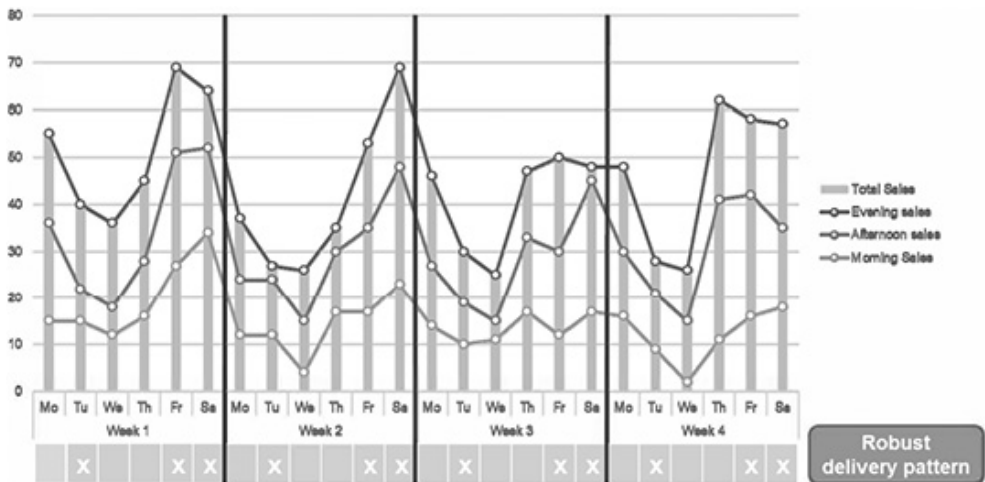
under non-stationary demand patterns using real data (see Figure 20.2). For larger problems, they suggest a genetic algorithm.

Another type of decision-support model that allows for more realistic assumptions in inbound delivery coordination is the joint replenishment problem. Aksoy and Erenguc (1988) review deterministic and stochastic models for joint replenishment problems. The following model assumes discrete time periods $t = 1, \dots, T$ within a finite planning horizon of length T . Multiple products $k = 1, \dots, K$ with dynamic demands d_{kt} without permitting backordering are considered. Inventories are subject to holding costs h_k for product k per unit per unit of time. The fixed cost structure includes a major setup cost A independent of the number of products included in the replenishment and minor setup costs A_k for each product replenished in a period but independent of the order quantity. The major setup cost addresses the replenishment, e.g., truck delivery, whereas the minor setup costs account for handling and processing per product. The solution of the model coordinates the inbound logistics across products, i.e., which products to replenish together and at what frequency. The following mixed-integer linear program supports such joint and coordinated replenishments. Decision variables are the order quantity of product k in period t , q_{kt} , the binary indicator γ_t if there is any (major setup) order in period t , binary indicators u_{kt} if there is an order (minor setup) for product k in period t , and inventory levels y_{kt} of product k at the end of period t . Initial inventories y_{k0} are given. Then, the optimization problem is

$$\min \sum_{t=1}^T (A\gamma_t + \sum_{k=1}^K (h_k y_{kt} + A_k u_{kt})) \quad (20.14)$$

s.t.

$$y_{kt} = y_{k,t-1} + q_{kt} - d_{kt} \quad t = 1, \dots, T, k = 1, \dots, K \quad (20.15)$$



Source: Taube and Minner (2018).

Figure 20.2 Delivery patterns for different prices

$$q_{kt} \leq u_{kt}M \quad t = 1, \dots, T, k = 1, \dots, K \quad (20.16)$$

$$u_{kt} \leq \gamma_t \quad t = 1, \dots, T, k = 1, \dots, K \quad (20.17)$$

$$q_{kt} \geq 0, y_{kt} \geq 0, u_{kt} \in \{0, 1\}, \gamma_t \in \{0, 1\} \quad t = 1, \dots, T, k = 1, \dots, K \quad (20.18)$$

The objective function in Equation (20.14) minimizes the sum of major and minor ordering costs and inventory holding costs for all products k and periods t . Inventory balances in Equation (20.15) enforce that the final inventory is equal to the initial inventory plus added order quantities minus demanded units. Constraints in Equations (20.16) and (20.17) represent logical constraints ensuring that an order quantity of a product can only be positive if the corresponding indicator is equal to one and the product-specific indicator itself can only be one if the major indicator is one. The parameter M is set to a sufficiently large value (e.g., the sum of all demands) to not restrict any order quantity q_{kt} .

This mixed-integer-linear programming formulation can be solved by standard solvers. To do so, it might be advantageous to use a different model formulation, see Narayanan and Robinson (2006). While this basic formulation assumes dynamic but deterministic and therefore known demand, these assumptions might not be realistic and require extension, in particular for retail inbound logistics. Other extensions with a joint warehouse space constraint are presented by Minner (2009) and by Minner and Silver (2005).

20.4.3 Inventory Routing Problems

Inventory routing combines the two fundamental problems in logistics, inventory management, and transportation. The basic dynamic multi-period single-product lot-sizing model is combined with the vehicle routing problem. For a literature review and introduction, see Coelho et al. (2014) and Bertazzi and Speranza (2012). The following model formulation combines the two traditional mixed-integer linear programming models for lot-sizing and vehicle routing. All deliveries to customers $i = 1, \dots, n$ originate from a single central depot $i=0$. Depot and customers are located at nodes $i = 0, \dots, n$. Customer demands d_{it} for periods $t = 1, \dots, T$ need to be satisfied, i.e., backorders are not permitted. Transportation between nodes i and j causes distance-dependent transportation costs c_{ij} and (homogenous) trucks have a limited capacity of W . Inventories at customer i at the end of a period are subject to holding costs h_i per unit and unit of time. The decision variables are the binary delivery to customer i in period t , γ_{it} , the delivery quantities q_{it} to customer i in period t , the inventory levels y_{it} of customer i in period t , and binary routing variables x_{ijt} if a truck goes from customer i to j in period t . The variable u_{it} defines the remaining capacity of a truck after supplying customer i in period t , the use of which additionally serves the purpose of avoiding short cycles. The optimization model is

$$\min \sum_{t=1}^T \left(\sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ijt} + \sum_{i=1}^n h_i y_{it} \right) \quad (20.19)$$

s.t.

$$y_{it} = y_{i,t-1} + q_{it} - d_{it} \quad i = 1, \dots, n; t = 1, \dots, T \quad (20.20)$$

$$\sum_{i=0}^n x_{ijt} = \gamma_{jt} \quad j = 1, \dots, n; t = 1, \dots, T \quad (20.21)$$

$$\sum_{j=0}^n x_{ijt} = \gamma_{it} \quad i = 1, \dots, n; t = 1, \dots, T \quad (20.22)$$

$$q_{it} \leq M\gamma_{it} \quad i = 1, \dots, n; t = 1, \dots, T \quad (20.23)$$

$$u_{0t} = W \quad t = 1, \dots, T \quad (20.24)$$

$$u_{jt} \leq u_{it} - q_{it} + (1 - x_{ijt})M \quad t = 1, \dots, T, i, j = 0, \dots, n; i \neq j \quad (20.25)$$

$$x_{ijt} \in \{0, 1\}, u_{it} \geq 0 \quad i, j = 0, \dots, n, i \neq j, t = 1, \dots, T \quad (20.26)$$

$$q_{it} \geq 0, y_{it} \geq 0, \gamma_{it} \in \{0, 1\} \quad i = 1, \dots, n; t = 1, \dots, T \quad (20.27)$$

The objective function in Equation (20.19) minimizes the sum of transportation and inventory holding costs. The constraints for every period t represent inventory balances in Equation (20.20), truck arrival and departure in Equations (20.21) and (20.22) at locations that require delivery during that period, and logical constraints limiting supply quantities to those days that are scheduled for delivery in Equation (20.23). Loading capacity constraints and the avoidance of sub-tours are achieved through Equations (20.24) and (20.25).

As for the vehicle routing problem, several extensions are possible to this model, i.e., time windows and forbidden days, the combination of pickup and delivery when multiple suppliers deliver to multiple plants, etc. Turan et al. (2017) present an approach using a variable neighborhood search for a perishable (newsvendor-type) product with an option for resupplying stock once during the sales day. The inbound coordination problem is the combined routing, delivery timing and resupply quantity allocation problem. Inventory routing integrates inventory and transportation decisions instead of solving the replenishment and delivery problems sequentially. Delivery costs are interdependent for all locations with an order in the same period, the inventory routing problem is related to the joint replenishment problem as multiple stores being delivered in a tour originating from the same lot-size share the setup cost. Due to the problem's complexity, real-world instances are typically solved by heuristic approaches, e.g., Malicki and Minner (2021) develop a joint transportation and inventory savings-based heuristic for a multi-location problem under uncertain demand and service level constraints.

20.4.4 One-Warehouse Multi-Retailer Systems

The multi-echelon, one-warehouse multiple-retailer problem is one of the traditional and most fundamental stochastic inventory management models (for a general typology and review, see De Kok et al., 2018). Extending on single-echelon stochastic inventory problems, the (two-stage) model not only determines the right level of inventory, but also the placement of it (Simpson, 1958; Clark & Scarf, 1960). Coordination of replenishments (depot effect) and risk

pooling (portfolio effect) favor the centralization of inventories in the warehouse (Federgruen, 1993). However, this problem is still one of those where the optimal policy is unknown and only of a simple form under additional assumptions or approximations (the so-called balance assumption). It also adds an additional decision to the problem, the allocation (rationing) of stock in case the central warehouse has insufficient inventory to serve all the retailers as desired.

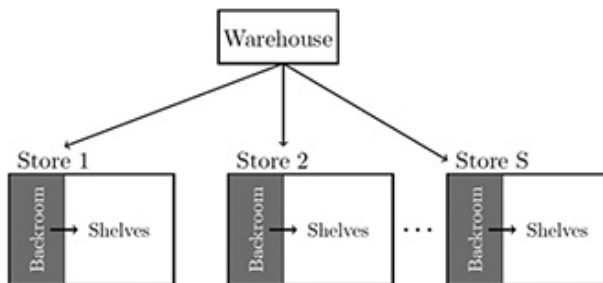
In the following, we formally present the periodic review one-warehouse multi-retailer inventory problem (see Figure 20.3). Assume a single warehouse ($i=0$) and $n=1, \dots, N$ retailers. Each retailer faces random demand following a theoretical distribution f_n with expected value μ_n and standard deviation σ_n (or more generally: covariance matrix) that is fully backlogged (or lost) if not satisfied. The warehouse replenishes inventory from an uncapacitated external supplier at a unit cost c (and potentially a fixed cost A_0) that is delivered at a constant lead time of L_0 periods. The retailers order from the warehouse and if sufficient stock can be shipped, orders arrive after a constant lead time L_i (including the review period). Inventories at all locations $i=0, \dots, N$ are subject to holding costs h_i and stockouts at the retailers are subject to a penalty cost per unit and unit of time p_i . In case the warehouse does not have sufficient inventory, the available stock is allocated to the retailers according to some rationing policy.

Then, the optimal inventory decisions are determined from the following stochastic dynamic program (see Inderfurth, 1994). The variables $x = x_0, \dots, x_n$ define the (echelon) inventory positions (stock on hand plus outstanding orders minus backorders at the beginning of a period before ordering and $y = (y_0, \dots, y_n)$ the target inventory levels after ordering. The value function $g_n(x)$ defines the minimum cost for an n -period problem and initial echelon inventory positions x_i . The functional equation is

$$g_n(x) = \min_y \sum_{i=0}^n (K_i(y_i - x_i) + C_i(y_i)) + \int_D g_{n-1}(y - D) f(D) dD \quad (20.28)$$

$$s.t. y_i \geq x_i, i = 1, \dots, n \quad (20.29)$$

$$\sum_{i=1}^n y_i \leq x_0 \quad (20.30)$$



Source: Taube and Minner (2018).

Figure 20.3 One warehouse, multi-retailer system

with K_i defining the ordering cost function, C_i being the single period expected holding and penalty cost function and $f(D)$ the multivariate demand distribution of all retailers and $f_{L_i}(d)$ being the density of cumulative demand d over L_i periods.

$$C_i(y_i) = h_i \int_0^{y_i} (y_i - d) f_{L_i}(d) dd + p_i \int_{y_i}^{\infty} (d - y_i) f_{L_i}(d) dd \quad (20.31)$$

Clark and Scarf (1960) and Federgruen and Zipkin (1984) have shown that under the so-called balance assumption, an echelon order-up-to policy is optimal, i.e., depot and retailer raise their (echelon) inventory positions to S_i . Based on this basic model, several enhancements and extensions have been developed, see Federgruen (1993) and Inderfurth (1994) for reviews. For extensions with particular importance to retail inventory management, we refer to perishable inventory management (Karaesmen et al., 2011), demand correlation (Erkip et al., 1990), and transshipments (Paterson et al., 2011).

20.4.5 Dynamic Pricing and Assortments

While price is typically an exogenous variable in many inventory models, the close relationship between price, demand, and inventories can also be leveraged by retailers using dynamic pricing. With dynamic pricing, the retailer adjusts the price over time. Dynamic pricing has become increasingly popular in retailing over the past two decades. Several factors have contributed to this development, such as more data being collected electronically and better data processing capabilities to learn about the price-demand relationship, electronic price tags allowing to change prices more easily and decision-support tools to optimize pricing decisions (Elmaghraby & Keskinocak, 2003; Den Boer, 2015; Chen & Chen, 2015).

Dynamic pricing is attractive from an inventory management perspective since it allows the retailer to sell leftover inventory that would otherwise be discarded. For example, grocery retailers sell products close to the expiration date at a discounted price. Fashion retailers hold end-of-season clearance sales at the end of a selling season. Caro and Gallien (2012) provide a very detailed description of the clearance pricing process at Zara before they implemented a model to determine a clearance price strategy. They show that the actual problem is much more complex than textbook solutions, as the clearance pricing decisions affect a large assortment and not individual products. They solve the multi-product price optimization problem by clustering products into categories, for which clearance prices are determined. This makes it easier to handle in-store operations for different prices as different products that are sold at the same price can be grouped together. Also, this strategy is easier to communicate with customers, which is another important consideration in pricing.

The example of Zara's decision problem highlights the importance of considering not only the problem from an inventory perspective, but also taking the customer into account. Elmaghraby and Keskinocak (2003) name three main characteristics of a retailer's market environment that determine the type of dynamic pricing problem a retailer needs to solve: Strategic customer behavior, dependence of demand over time and inventory replenishment.

Customers who behave strategically consider potential future end-of-season clearance pricing before making a purchase decision during the season. If they anticipate that the item will be sold at a lower price in the future, they might decide that savings from a lower price

outweigh the benefit of receiving the item earlier. Aviv et al. (2019) analyze under which conditions a retailer should apply a dynamic strategy considering that customers might behave strategically and delay the purchase decision. They also find an interesting effect: the “active learning effect”, which describes that the retailer might be able to learn less about customer demand if customers delay their purchase decisions, and as a result, underestimate the market size.

Demand is often dependent on the demand in previous periods. If a customer buys a product when it is sold at a lower price and stores the items at home until they are consumed, the demand in the future will be lower. Whether demand is dependent or not is often influenced by the lifetime of a product. While durable products can easily be kept at home for a longer period of time and demand is thus dependent, the demand for perishable products is often independent over time (Elmaghraby & Keskinocak, 2003).

Another possibility to react more flexibly to customer demand and take inventory constraints into account, is to customize the assortment. While this is usually not feasible in traditional brick-and-mortar stores, it can be applied in online retailing where retailers have collected a lot of information on customer preferences and where it is technically possible to vary the assortment shown to different customers. If customers have heterogeneous preferences, it can be beneficial for the retailer to target the product offering to each customer. Based on a dataset from an online retailer, Golrezaei et al. (2014) show that customizing the assortment in real-time can yield over 10% improvements in revenues compared to offering the same assortment to every customer. By showing only a subset of the assortment to arriving customers, the retailer can reserve products with low inventories for those customers with a strong preference for these products (Bernstein et al., 2015). Rusmevichientong et al. (2020) suggest a policy for the dynamic assortment optimization problem for reusable products. Combining dynamic assortments and pricing, Ma and Simchi-Levi (2020) suggest an algorithm to decide both on the assortment offered to an arriving customer and on the price at which items, such as airline tickets, will be sold.

20.5 SUMMARY AND FUTURE RESEARCH

The material presented in this chapter provides an overview of data and decision-support models for retail inventory management. These are important prerequisites for further digitalization, technology use, and automation in retail inventory management. Although many inventory management models and algorithms are available, we currently observe a further integration of data and decision-making rather than treating forecasting and decision-making sequentially. In an ever faster and dynamically changing environment such as retail, learning about model parameters and inventory control policy structures for execution will receive increasing importance. In particular, the multi-product feature of retail deserves more integrated and data-driven work. However, integration of traditionally independently treated decisions increases model complexity and therefore, requires new algorithmic approaches for solving large-scale problems under uncertainty, decomposition and coordination schemes. Despite automation, certain tasks might still be relegated to or require approval by human decision makers. Therefore, more behavioral operations work in retail going beyond simple newsvendor decision-making and an integration with empirical data is required.

REFERENCES

- Agrawal, N., & Smith, S. A. (1996). Estimating negative binomial demand for retail inventory management with unobservable lost sales. *Naval Research Logistics*, 43(6), 839–861. [https://doi.org/10.1002/\(SICI\)1520-6750\(199609\)43:6<839::AID-NAV4>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1520-6750(199609)43:6<839::AID-NAV4>3.0.CO;2-5)
- Aksoy, Y., & Erenguc, S. S. (1988). Multi-item inventory models with co-ordinated replenishments: A survey. *Journal of Operations & Production Management*, 8(1), 63–73. <https://doi.org/10.1108/eb054814>
- Andersen, E. T., Fitzsimons, G. J., & Simester, D. (2006). Measuring and mitigating the costs of stockouts. *Management Science*, 52(11), 1751–1763. <https://doi.org/10.1287/mnsc.1060.0577>
- Anupindi, R., Dada, M., & Gupta, S. (1998). Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science*, 17(4), 406–423. <https://doi.org/10.1287/mksc.17.4.406>
- Aviv, Y., Wei, M. M., & Zhang, F. (2019). Responsive pricing of fashion products: The effects of demand learning and strategic consumer behavior. *Management Science*, 65(7), 2982–3000. <https://doi.org/10.1287/mnsc.2018.3114>
- Ban, G.-Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1), 90–108. <https://doi.org/10.1287/opre.2018.1757>
- Bell, P. C. (1981). Adaptive sales forecasting with many stockouts. *Journal of the Operational Research Society*, 32(10), 865–873. <https://doi.org/10.1057/jors.1981.180>
- Bernstein, F., Kök, A. G., & Xie, L. (2015). Dynamic assortment customization with limited inventories. *Manufacturing and Service Operations Management*, 17(4), 538–553. <https://doi.org/10.1287/msom.2015.0544>
- Bertazzi, L., & Speranza, M. (2012). Inventory routing problems: An introduction. *EURO Journal on Transportation and Logistics*, 1(4), 307–326. <https://doi.org/10.1007/s13676-012-0016-7>
- Beutel, A. L., & Minner, S. (2012). Safety stock planning under causal demand forecasting. *International Journal of Production Economics*, 140(2), 637–645. <https://doi.org/10.1016/j.ijpe.2011.04.017>
- Boylan, J. E., Chen, H., Mohammadipour, M., & Syntetos, A. (2014). Formation of seasonal groups and application of seasonal indices. *Journal of the Operational Research Society*, 65(2), 227–241. <https://doi.org/10.1057/jors.2012.126>
- Broekmeulen, R. A., & Van Donselaar, K. H. (2019). Quantifying the potential to improve on food waste, freshness and sales for perishables in supermarkets. *International Journal of Production Economics*, 209(3), 265–273. <https://doi.org/10.1016/j.ijpe.2017.10.003>
- Caro, F., & Gallien, J. (2012). Clearance pricing optimization for a fast-fashion retailer. *Operations Research*, 60(6), 1404–1422. <https://doi.org/10.1287/opre.1120.1102>
- Chen, M., & Chen, Z.-L. (2015). Recent developments in dynamic pricing research: Multiple products, competition, and limited demand information. *Production and Operations Management*, 24(5), 704–731. <https://doi.org/10.1111/poms.12295>
- Chen, X., & Simchi-Levi, D. (2012). Pricing and inventory management. In Ö. Özer & R. Phillips (Eds.), *The Oxford handbook of pricing management*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199543175.013.0030>
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490. <https://doi.org/10.1287/mnsc.6.4.475>
- Coelho, L. C., Cordeau, J.-F., & Laporte, G. (2014). Thirty years of inventory routing. *Transportation Science*, 48(1), 1–19. <https://doi.org/10.1287/trsc.2013.0472>
- Corsten, D., & Gruen, T. (2003). Desperately seeking shelf availability: An examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail and Distribution Management*, 31(12), 605–617. <https://doi.org/10.1108/09590550310507731>
- Curşeu, A., Van Woensel, T., Fransoo, J., Van Donselaar, K., & Broekmeulen, R. (2009). Modelling handling operations in grocery retail stores: An empirical analysis. *Journal of the Operational Research Society*, 60(2), 200–214. <https://doi.org/10.1057/palgrave.jors.2602553>
- De Kok, T., Grob, C., Laumanns, M., Minner, S., Rambau, J., & Schade, K. (2018). A typology and literature review on supply chain inventory management. *European Journal of Operational Research*, 269(3), 955–983. <https://doi.org/10.1016/j.ejor.2018.02.04>
- DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: An empirical analysis. *Management Science*, 54(4), 627–641. <https://doi.org/10.1287/mnsc.1070.0789>

- Den Boer, A. V. (2015). Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys in Operations Research and Management Science*, 20(1), 1–18. <https://doi.org/10.1016/j.sorms.2015.03.001>
- DeYong, G. D. (2020). The price-setting newsvendor: Review and extensions. *International Journal of Production Research*, 58(6), 1776–1804. <https://doi.org/10.1080/00207543.2019.1671624>
- Ehrenthal, J. C., Honhon, D., & Van Woensel, T. (2014). Demand seasonality in retail inventory management. *European Journal of Operational Research*, 238(2), 527–539. <https://doi.org/10.1016/j.ejor.2014.03.030>
- Eliashberg, J., & Steinberg, R. (1993). Marketing-production joint decision-making. In *Marketing, Handbooks in Operations Research and Management Science* (Vol. 5, pp. 827–880). Elsevier. [https://doi.org/10.1016/S0927-0507\(05\)80041-6](https://doi.org/10.1016/S0927-0507(05)80041-6)
- Elmachtoub, A. N., & Grigas, P. (2022). Smart “predict, then optimize”. *Management Science* 68(1), 9–26. <https://doi.org/10.1287/mnsc.2020.3922>
- Elmaghraby, W., & Keskinocak, P. (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10), 1287–1309. <https://doi.org/10.1287/mnsc.49.10.1287.17315>
- Erkip, N., Hausman, W. H., & Nahmias, S. (1990). Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands. *Management Science*, 36(3), 381–392. <https://doi.org/10.1287/mnsc.36.3.381>
- Eroglu, C., Williams, B. D., & Waller, M. A. (2011). Consumer-driven retail operations: The moderating effects of consumer demand and case pack quantity. *International Journal of Physical Distribution and Logistics Management*, 41(5), 420–434. <https://doi.org/10.1108/09600031111138808>
- Eroglu, C., Williams, B. D., & Waller, M. A. (2013). The backroom effect in retail operations. *Production and Operations Management*, 22(4), 915–923. <https://doi.org/10.1111/j.1937-5956.2012.01393.x>
- Federgruen, A. (1993). Centralized planning models for multi-echelon inventory systems under uncertainty. In *Logistics of Production and Inventory, Handbooks in Operations Research and Management Science* (Vol. 4, pp. 133–173). Elsevier. [https://doi.org/10.1016/S0927-0507\(05\)80183-5](https://doi.org/10.1016/S0927-0507(05)80183-5)
- Federgruen, A., & Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), 818–836. <https://doi.org/10.1287/opre.32.4.818>
- Fildes, R., Ma, S., & Kolassa, S. (2022). Retail forecasting: Research and practice. *International Journal of Forecasting* 38(4), 1283–1318. <https://doi.org/10.1016/j.ijforecast.2019.06.004>
- Fisher, M. (2009). Rocket science retailing: The 2006 Philip McCord Morse Lecture. *Operations Research*, 57(3), 527–540. <https://doi.org/10.1287/opre.1090.0704>
- Golrezaei, N., Nazerzadeh, H., & Rusmevichientong, P. (2014). Real-time optimization of personalized assortments. *Management Science*, 60(6), 1532–1551. <https://doi.org/10.1287/mnsc.2014.1939>
- Holzapfel, A., Hübner, A., Kuhn, H., & Sternbeck, M. G. (2016). Delivery pattern and transportation planning in grocery retailing. *European Journal of Operational Research*, 252(1), 54–68. <https://doi.org/10.1016/j.ejor.2015.12.036>
- Huber, J., Müller, S., Fleischmann, M., & Stuckenschmidt, H. (2019). A data-driven newsvendor problem: From data to decision. *European Journal of Operational Research*, 278(3), 904–915. <https://doi.org/10.1016/j.ejor.2019.04.043>
- Hübner, A. H., Kuhn, H., & Sternbeck, M. G. (2013). Demand and supply chain planning in grocery retail: An operations planning framework. *International Journal of Retail and Distribution Management*, 41(7), 512–530. <https://doi.org/10.1108/IJRDM-05-2013-0104>
- Inderfurth, K. (1994). Safety stocks in multistage divergent inventory systems: A survey. *International Journal of Production Economics*, 35(1), 321–329. [https://doi.org/10.1016/0925-5273\(94\)90098-1](https://doi.org/10.1016/0925-5273(94)90098-1)
- Jain, A., Rudi, N., & Wang, T. (2015). Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need. *Operations Research*, 63(1), 134–150. <https://doi.org/10.1287/opre.2014.1326>
- Johnston, F., Boylan, J., & Shale, E. (2003). An examination of the size of orders from customers, their characterisation and the implications for inventory control of slow moving items. *Journal of the Operational Research Society*, 54(8), 833–837. <https://doi.org/10.1057/palgrave.jors.2601586>
- Karabati, S., Tan, B., & Öztürk, Ö. C. (2009). A method for estimating stock-out-based substitution rates by using point-of-sale data. *IIE Transactions*, 41(5), 408–420. <https://doi.org/10.1080/07408170802512578>

- Karaesmen, I. Z., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. In K. G. Kempf, P. Keskinocak, & R. Uzsoy (Eds.), *Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook* (Vol. 151). Springer; *International Series in Operations Research & Management Science*, 151, 393–436.
- Ketzenberg, M., Metters, R., & Vargas, V. (2002). Quantifying the benefits of breaking bulk in retail operations. *International Journal of Production Economics*, 80(3), 249–263. [https://doi.org/10.1016/S0925-5273\(02\)00258-X](https://doi.org/10.1016/S0925-5273(02)00258-X)
- Kiil, K., Dreyer, H. C., Hvolby, H. H., & Chabada, L. (2018). Sustainable food supply chains: The impact of automatic replenishment in grocery stores. *Production Planning and Control*, 29(2), 106–116. <https://doi.org/10.1080/09537287.2017.1384077>
- Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, 55(6), 1001–1021. <https://doi.org/10.1287/opre.1070.0409>
- Kolassa, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. *International Journal of Forecasting*, 32(3), 788–803. <https://doi.org/10.1016/j.ijforecast.2015.12.004>
- Krisnadewi, K. A., & Soewarno, N. (2019). Competitiveness and cost behaviour: Evidence from the retail industry. *Journal of Applied Accounting Research*, 21(1), 125–141. <https://doi.org/10.1108/JAAR-08-2018-0120>
- Kuhn, H., & Sternbeck, M. G. (2013). Integrative retail logistics: An exploratory study. *Operations Management Research*, 6(1–2), 2–18. <https://doi.org/10.1007/s12063-012-0075-9>
- Lau, H. S., & Lau, A. H. L. (1996). Estimating the demand distributions of single-period items having frequent stockouts. *European Journal of Operational Research*, 92(2), 254–265. [https://doi.org/10.1016/0377-2217\(95\)00134-4](https://doi.org/10.1016/0377-2217(95)00134-4)
- Ma, W., & Simchi-Levi, D. (2020). Algorithms for online matching, assortment, and pricing with tight weight-dependent competitive ratios. *Operations Research*, 68(6), 1787–1803. <https://doi.org/10.1287/opre.2019.1957>
- Maddah, B., & Bish, E. (2007). Joint pricing, assortment, and inventory decisions for a retailer's product line. *Naval Research Logistics*, 54(3), 315–330. <https://doi.org/10.1002/nav.20209>
- Mahajan, S., & Van Ryzin, G. (2001). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, 49(3), 334–351. <https://doi.org/10.1287/opre.49.3.334.11210>
- Malicki, S., & Minner, S. (2021). Cyclic inventory routing with dynamic safety stocks under recurring non-stationary interdependent demands. *Computers and Operations Research*, 131, 105247. <https://doi.org/10.1016/j.cor.2021.105247>
- Minner, S. (2009). A comparison of simple heuristics for multi-product dynamic demand lot-sizing with limited warehouse capacity. *International Journal of Production Economics*, 118(1), 305–310. <https://doi.org/10.1016/j.ijpe.2008.08.034>
- Minner, S., & Silver, E. (2005). Multi-product batch replenishment strategies under stochastic demand and a joint capacity constraint. *IIE Transactions on Scheduling and Logistics*, 37(5), 469–479. <https://doi.org/10.1080/07408170590918254>
- Minner, S., & Transchel, S. (2010). Periodic review inventory control for perishable products under service level constraints. *OR Spectrum*, 32(4), 979–996. <https://doi.org/10.1007/s00291-010-0196-1>
- Minner, S., & Transchel, S. (2017). Order variability in perishable product supply chains. *European Journal of Operational Research*, 260(1), 93–107. <https://doi.org/10.1016/j.ejor.2016.12.016>
- Mou, S., Robb, D. J., & DeHoratius, N. (2018). Retail store operations: Literature review and research directions. *European Journal of Operational Research*, 265(2), 399–422. <https://doi.org/10.1016/j.ejor.2017.07.003>
- Munson, C., & Rosenblatt, M. (1998). Theories and realities of quantity discounts: An exploratory study. *Production and Operations Management*, 7(4), 352–369. <https://doi.org/10.1111/j.1937-5956.1998.tb00129.x>
- Narayanan, A., & Robinson, E. P. (2006). More on 'models and algorithms for the dynamic-demand joint replenishment problem'. *International Journal of Production Research*, 44(2), 383–397. <https://doi.org/10.1080/002075405000270562>
- Ord, K., Fildes, R. A., & Kourentzes, N. (2017). *Principles of business forecasting*. Wessex Press Publishing Co.

- Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2), 125–136. <https://doi.org/10.1016/j.ejor.2010.05.048>
- Pentico, D. W. (1974). The assortment problem with probabilistic demands. *Management Science*, 21(3), 286–290. <https://doi.org/10.1287/mnsc.21.3.286>
- Petruzzi, N. C., & Dada, M. (1999). Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2), 183–194. <https://doi.org/10.1287/opre.47.2.183>
- Rekik, Y., Syntetos, A., & Glock, C. (2019). Modeling (and learning from) inventory inaccuracies in e-retailing/B2B contexts. *Decision Sciences*, 50(6), 1184–1223. <https://doi.org/10.1111/deci.12367>
- Rusmevichientong, P., Sumida, M., & Topaloglu, H. (2020). Dynamic assortment optimization for reusable products with random usage durations. *Management Science*, 66(7), 2820–2844. <https://doi.org/10.1287/mnsc.2019.3346>
- Sachs, A.-L., & Minner, S. (2014). The data-driven newsvendor with censored demand observations. *International Journal of Production Economics*, 149, 28–36. <https://doi.org/10.1016/j.ijpe.2013.04.039>
- Sauré, D., & Zeevi, A. (2013). Optimal dynamic assortment planning with demand learning. *Manufacturing and Service Operations Management*, 15(3), 387–404. <https://doi.org/10.1287/msom.2013.0429>
- Schulte, B., & Sachs, A. L. (2020). The price-setting newsvendor with Poisson demand. *European Journal of Operational Research*, 283(1), 125–137. <https://doi.org/10.1016/j.ejor.2019.10.039>
- Silver, E. A., Pyke, D. F., & Thomas, D. J. (2016). *Inventory and production management in supply chains* (4th ed.). CRC Press.
- Simpson, K. (1958). In-process inventories. *Operations Research*, 6(6), 863–873. <https://doi.org/10.1287/opre.6.6.863>
- Smith, S. A., & Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, 48(1), 50–64. <https://doi.org/10.1287/opre.48.1.50.12443>
- Spiliotis, E., Makridakis, S., Kaltsounis, A., & Assimakopoulos, V. (2021). Product sales probabilistic forecasting: An empirical evaluation using the m5 competition data. *International Journal of Production Economics*, 240, 108237. <https://doi.org/10.1016/j.ijpe.2021.108237>
- Svoboda, J., & Minner, S. (2021). Tailoring inventory classification to industry applications: The benefits of understandable machine learning. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2021.1959078>
- Taube, F., & Minner, S. (2018). Data-driven assignment of delivery patterns with handling effort considerations in retail. *Computers and Operations Research*, 100(12), 379–393. <https://doi.org/10.1016/j.cor.2017.08.004>
- Teulings, M., & Van der Vlist, P. (2001). Managing the supply chain with standard mixed loads. *International Journal of Physical Distribution and Logistics Management*, 31(3), 169–186. <https://doi.org/10.1108/09600030110389442>
- Topaloglu, H. (2013). Joint stocking and product offer decisions under the multinomial logit model. *Production and Operations Management*, 22(5), 1182–1199. <https://doi.org/10.1111/j.1937-5956.2012.01423.x>
- Trautrim, A., Grant, D., Fernie, J., & Harrison, T. (2009). Optimizing on-shelf availability for customer service and profit. *Journal of Business Logistics*, 30(2), 231–247. <https://doi.org/10.1002/j.2158-1592.2009.tb00122.x>
- Turan, B., Minner, S., & Hartl, R. (2017). A VNS approach to multi-location inventory redistribution with vehicle routing. *Computers and Operations Research*, 78(2), 526–536. <https://doi.org/10.1016/j.cor.2016.02.018>
- Van Donselaar, K., Van Woensel, T., Broekmeulen, R., & Fransoo, J. (2006). Inventory control of perishables in supermarkets. *International Journal of Production Economics*, 104(2), 462–472. <https://doi.org/10.1016/j.ijpe.2004.10.019>
- Van Donselaar, K. H., Gaur, V., Van Woensel, T., Broekmeulen, R. A., & Fransoo, J. C. (2010). Ordering behavior in retail stores and implications for automated replenishment. *Management Science*, 56(5), 766–784. <https://doi.org/10.1287/mnsc.1090.1141>

- Van Ryzin, G., & Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, 45(11), 1496–1509. <https://doi.org/10.1287/mnsc.45.11.1496>
- Van Zelst, S., Van Donselaar, K., Van Woensel, T., Broekmeulen, R., & Fransoo, J. (2009). Logistics drivers for shelf stacking in grocery retail stores: Potential for efficiency improvement. *International Journal of Production Economics*, 121(2), 620–632. <https://doi.org/10.1016/j.ijpe.2006.06.010>
- Wecker, W. E. (1978). Predicting demand from sales data in the presence of stockouts. *Management Science*, 24(10), 1043–1054. <https://doi.org/10.1287/mnsc.24.10.1043>
- Wensing, T., Sternbeck, M. G., & Kuhn, H. (2018). Optimizing case-pack sizes in the bricks-and-mortar retail trade. *OR Spectrum*, 40(4), 913–944. <https://doi.org/10.1007/s00291-018-0515-5>
- Zipkin, P. (2000). *Foundations of inventory management*. Irwin Publishing.

21. Online retailing inventory management

Mengxin Wang and Zuo-Jun Max Shen

21.1 OVERVIEW

Online retail inventory management focuses on enabling fast order fulfillment at a low cost. A United Parcel Service (UPS) survey showed that 63% of customers value delivery speed when choosing a product, and 27% consider the same-day delivery option when choosing an online retailer (Jehl 2020). E-retailers aim to cut delivery times to attract more customers. Amazon, for example, offers free same-day or 24-hour delivery for Prime members. JD.com, which services over 99% of the Chinese population across great distances, a “211 program”, which promises same-day delivery for orders placed before 11 a.m. and next-day delivery before 3 p.m. for orders placed between 11 a.m. and 11 p.m.

Traditional brick-and-mortar retailers hold a small number of store-keeping units (SKUs) in a small number of stores. Each store serves the customers in a nearby region. Stores can regularly replenish large volumes of stock of each SKU. On the contrary, online retailing players hold millions of SKUs and have to fulfill numerous small orders from vast regions. Online retail players have a large number of fulfillment centers (FCs) covering vast areas, which is a strategy taken to ensure order delivery speed. For e-retailers to be a go-to store, all products need to be listed online. However, it is impossible to have all SKUs in each FC. The efficient management of all SKUs in the fulfillment network becomes a key factor for quick order fulfillment.

These differences lead to challenging inventory management problems in online retailing. In this chapter, we provide an overview of online retail inventory management, examining both literature and industry practices. We will focus specifically on two critical inventory management problems in e-retailing: inventory placement and order picking. To provide a detailed analysis of these issues, we will examine the practices of two major e-retailing players, Amazon and JD.com, and discuss some representative works in this area.

Inventory placement determines which fulfillment center holds inventory for which SKUs. This is a tactical-level problem that arises before making any replenishment or transshipment decision. The large volume of SKUs and fast delivery requirement introduce unprecedented challenges in determining the placement of inventory. Order picking is the process where individual items are picked from an FC to fulfill customers’ orders. It is an important warehouse control problem at an operational level. The advent of new warehouse technologies for e-retailing introduces new questions for order picking: How can a company efficiently manage millions of SKUs in a warehouse? How can a company efficiently pick a large volume of small-sized orders in a warehouse? What type of warehouse infrastructure and order-picking policy is suitable for different e-retail practices?

Both inventory placement and order picking play a critical role in order fulfillment efficiency. With high service requirements on product availability and timely delivery, online retailing is generating new challenges while looking for innovative approaches to solve these problems.

In the following section, we provide a brief review of the standard inventory management literature, linking these studies to the practice of the two major e-retailing players. We organize our overview of the literature into three parts: the replenishment problem, the transshipment problem, and warehouse operations. Following this, in Section 21.2, we delve into the inventory placement problem, and then we discuss the order-picking problem in Section 21.3.

21.1.1 Replenishment

Inventory management aims to minimize the costs incurred in an inventory system, including holding, fixed, purchase, and stock-out costs. Stock-out may occur because both customer demand and vendor lead time are highly random and difficult to forecast in practice. In addition, products may be required sooner than suppliers can provide them. A mismatch of inventory and demand causes stock-out and leads to customer dissatisfaction and churn. This then leads to the problem of replenishing the inventory.

We first discuss single-echelon replenishment problems, where the replenishment decisions are only related to a single stage in a supply chain. With stochastic demand and vendor lead time, managing inventory leads to the setting of two inventory levels: cycle stock and safety stock. Cycle stock, or working inventory, is the inventory intended to meet an anticipated demand. Safety stock is the inventory needed to face forecast inaccuracies or variability in demand and vendor lead time while limiting the incurred stock-out cost.

The separation of inventory between safety stock and cycle has been used to design classic replenishment policies such as the newsvendor model, periodic review inventory (s, S) policy, and continuous review inventory (r, Q) policy (Snyder and Shen 2011). The newsvendor problem considers a single-period setting in which the remaining inventory or stock-out cannot be carried over to the next period at the end of the replenishment period. The newsvendor policy specifies an order-up-to level that balances the stock-out and inventory holding costs.

The periodic review policy checks an inventory position and replenishes it at discrete times in a multi-period setting. The (s, S) policy specifies the *reorder point* s and *order-up-to level* S . If an inventory position drops below the reorder point at the review time, the inventory position is replenished up to S . In a continuous review system, the inventory position is monitored continuously and can be ordered at any time. The widely adopted (r, Q) policy decides the ordering quantity Q , where an order of size Q is placed every time the inventory position reaches the reorder point r .

In practice, e-retailers may have several layers of warehousing in their fulfillment networks. For example, JD.com has a fulfillment network with regional distribution centers (RDCs) and forward distribution centers (FDCs). The RDCs are large facilities in remote areas that hold the majority of inventory of all types of SKUs and replenish the FDCs. FDCs are small facilities that fulfill orders close to customer locations. Amazon is also expanding into a multi-layered fulfillment network, including fulfillment centers (for everyday products) and non-sortable fulfillment centers (for larger items), smaller fulfillment centers catering to same-day orders, and distribution centers that supply products to downstream fulfillment centers. The fulfillment centers in one layer (or echelon) serve as suppliers to the other layers (Amazon 2021). This is called a multi-echelon system. The multi-layered network structure allows more FCs to be located closer to where customers live. This plays a critical role in fast delivery. Meanwhile, it entails more complex replenishment decisions.

In a multi-echelon system, the inventory levels of various SKUs must be determined at each level of the supply chain. A multi-echelon system can assume different topologies, such as serial, divergent, or convergent. A divergent system consists of a single central node and several successors. For instance, JD.com's fulfillment network has several FDCs for each RDC, whereas each FDC has a unique parent RDC. A convergent system has one end node with several predecessors. Here, a serial chain is a special case that falls under both previous categories: each node in the network has a single successor and single predecessor.

The theory of a multi-echelon supply chain was first studied by Clark and Scarf (1960), who analyzed the basic model of a serial system. The main result was that the optimal base-stock levels in a chain could be computed sequentially, starting at the end stage and working upstream. Each step was computed by minimizing a one-dimensional convex function. The optimality was shown with a simplified proof to hold for serial and convergent systems, as well as continuous time models (Chen and Zheng 1994). A generalized model of a serial supply chain under the assumption of Markov-modulated demand was proposed by Chen and Song (2001).

The exact method proposed by Clark and Scarf (1960) was based on dynamic programming techniques and required discretization of stock levels, which limited the computational efficiency. Alternative methods have been developed in subsequent research to find exact or approximate solutions. Federgruen and Zipkin (1984) extended the work of Clark and Scarf (1960) by building on the model of a stochastic dynamic program with a finite horizon to solve the problem of an infinite horizon. A simple approximate solution using a weighted average of the upper and lower bounds of the optimal base-stock level was proposed by Shang and Song (2003) and yielded a remarkably accurate empirical solution.

Other studies have extended the work of Clark and Scarf (1960) to convergent and divergent systems. A convergent system was shown to be equivalent to a serial system (Rosling 1989; Langenhoff and Zijm 1990). However, a divergent system was considerably more complicated. In this case, a major difficulty arose from the fact that a given stage must decide the allocation of inventory to its successors when facing insufficient inventory. Therefore, an allocation decision had to be made jointly along with a replenishment decision. The simplest type of divergent system is a single-depot multi-warehouse system. The projection algorithm is one of the best-known algorithms that solves the exact solution for a single-depot multi-warehouse system (Graves 1985 and Axsäter 1990). Heuristic methods have been proposed for more general divergent type systems (see, e.g., Graves 1985; Sherbrooke 1968; Gallego et al. 2007).

21.1.2 Transshipment

Dealing with many suppliers and customers implies a high variability in demand and vendor lead time. It is impossible to guarantee in-stock status for every order in a single FC with a fixed inventory profile. Therefore, e-retailers regularly move inventory within their fulfillment network to better coordinate with customer demand. Replenishment has solved the problem of choosing the inventory level to maintain in each layer of the network. Transshipment deals with the problem of managing inventory levels between facilities in a delivery network.

Transshipment is an additional step that simplifies inventory rebalancing, helps with order fulfillment, and adds more robustness to the replenishment policy. With transshipment, the overall inventory level can be reduced while the service level remains unchanged or is improved (Herer et al. 2002). Replenishment in a multi-echelon network can be regarded

as vertical transshipment from upstream to downstream stages. In addition to vertical transshipment, a more flexible system allows lateral transshipment, that is, the movement of items within the same echelon of a supply chain.

In the literature, two main types of lateral transshipment have been studied: proactive transshipment and reactive transshipment. Proactive transshipment balances inventory in a fulfillment network periodically in advance of customer orders. Proactive transshipment exists from a simple single-period single-transshipment with no network inventory replenishment model to a more complex model with multi-period, multi-transshipment, and a network. Reactive transshipment balances inventory in a fulfillment network in response to customer orders. Reactive transshipment has been studied in the context of both periodic review and continuous review models and can be further classified into single and multi-echelon systems. We refer readers to a comprehensive review by Paterson et al. (2011) on lateral transshipment for more details.

The problem of balancing inventory between facilities to minimize stock-out has been extensively studied in the literature (see, e.g., Allen 1958 and Allen 1961), where varying numbers of items, echelons, facilities, and costs have been analyzed. In addition to stock-out, *order split* is becoming a new concern in e-retailing. As mentioned before, e-retailers sell millions of SKUs of various types. An FC cannot hold all the SKUs. However, e-retail orders usually contain multiple SKUs of different types. An order split occurs when an order cannot be fulfilled by a single FC and multiple FCs are assigned to fulfill the order. Order splits increase the handling, packaging, and delivery costs and delays delivery when an order is routed to distant FCs. As a result, the inventory should be carefully allocated and transshipped in the fulfillment network to minimize order splits. There is a growing interest in reactive transshipment because it reduces both the stock-out and order-split rates.

21.1.3 Warehouse Operations

Online retailers have FCs to receive and store inventory and pick items according to customer orders. Replenishment and transshipment deal with the tactical-level decisions of how many and which SKUs to stock up at each FC and how to move the SKUs in the fulfillment network. At the operational level, FCs perform a large amount of receiving, storage, and order-picking operations on a daily basis. Therefore, order-fulfillment speed depends on efficient warehouse control in FCs.

Receiving, shipping, storage, and order picking are the four major components of warehouse operations. Receiving and shipping manage the inbound and outbound flows of a warehouse. For example, truck scheduling, truck-order assignment, and loading/unloading scheduling are all receiving and shipping decisions. The storage decision determines the assignment of replenished items to different storage locations in an FC. Order picking is the process of retrieving items from storage areas and sorting them into individual orders for shipment. Among all the warehouse operations, order picking is the most labor intensive and capital intensive. Order picking incurs more than 50% of warehouse operational costs (De Koster et al. 2007).

Warehouse operation performance hinges on warehouse layout and technology. Conventional warehouses use a multi-parallel aisle layout where storage racks are arranged in parallel; aisles between the racks are used by human workers to store and pick items to and from the racks. The major disadvantage of the picker-to-part system is the unproductive picker walking time, which is multiple orders of magnitude more significant than the picking time.

Conventional picker-to-parts systems have difficulty fitting e-retailing with large assortments, small orders, and rapid delivery requirements (Boysen et al. 2019). Currently, an increasing number of e-retailing warehouse systems are automated. The *automated storage/retrieval system* (AS/RS) is a widely adopted automated warehouse system. It was introduced in the 1950s and has been widely used in real distribution and production environments since the 1990s (Roodbergen and Vis 2009). The AS/RS adopts the multi-parallel aisle layout, with automated machines installed on the racks for picking and storage operations. JD.com, for example, has a large network of automated warehouses with AS/RS in China (Cao 2020).

The AS/RS provides a highly automated solution for efficiently handling large volumes of orders and SKUs. The downsides of an AS/RS include a long design cycle, high investment cost, and inflexibility. The efficiency of AS/RS hinges on a properly designed facility infrastructure and automated control techniques that are capital intensive. The deployment of an AS/RS can take several years. Once the physical infrastructure has been built, it is difficult to modify it to cope with changing demand. E-retailers need accurate long-term demand forecasts to ensure that the system design will meet future demand. In addition, economies of scale play a role; the warehouse scale and transaction volume must be large enough to recover the high capital investment cost.

Amazon Robotics LLC, previously Kiva Systems LLC, came up with an innovative solution: the *robotic mobile fulfillment system* (RMFS). The RMFS utilizes robots to carry storage units called “pods” from an inventory area to human operators who are at picking/replenishment stations. The robots are efficiently managed to save the unproductive walking time of conventional picker-to-part systems. An RMFS has the advantage of rapid deployment, flexibility, and expandability. In addition, warehouse managers do not need to accurately specify the long-term demand for system design. An RMFS can be expanded easily by adding more robots; the system layout can also be easily modified by changing the locations of the pods and stations. When facing increasing or decreasing demand, new robot units can be purchased or withdrawn to cope with changes. Despite the upsides, the success of an RMFS largely depends on the efficient management of the robots. This introduces new challenges for order-picking control. In this chapter, we focus on the AS/RS and RMFS in terms of order-picking operations, compare these two types of systems in e-retailing, and identify potential research directions.

21.2 INVENTORY PLACEMENT

In the previous section, we have provided an overview of the inventory management literature and connected it to the practices of the e-retailing industry. In this section, we discuss a specific inventory management problem in the e-retailing industry that is critical for order-fulfillment efficiency: the inventory placement problem.

Shipping speed is a key differentiator between e-retailers. The main players in e-retailing invest heavily in ever-faster delivery to attract more customers. In a traditional supply chain, there are a limited number of SKUs; the warehouses hold a fixed set of SKUs being replenished and transshipped on a regular basis. In e-commerce, however, there are millions of SKUs, and an FC may only hold a small subset of them. Meanwhile, the assortment in an FC must vary with fluctuating demand. As a result, allocating different SKUs to FCs has become a vital problem for e-retailers. Optimizing the placement of SKUs is essential for e-retailers to offer reliable and fast delivery.

The policy for selecting which SKU to store in which distribution center (DC) depends on the delivery network of the e-retailer. Various types of networks exist, here depending on the company's business scope, scale, and geographical constraints. One of the most common delivery networks is the lateral fulfillment network. In this type of network, customer nodes and FC nodes are fully connected: an FC in the network can fulfill any customer order, and a customer order can be fulfilled by any FC. Amazon, for example, operated a lateral fulfillment network in its EU network (Merriam 2007).

This type of network maximizes the in-stock rate: that is, the number of orders that can be fulfilled by an FC. Another advantage of such networks is their high robustness to block off delivery routes. The disadvantage of such a network is the delivery cost. Having a delivery arc between every DC and customer node increases the number of routes quadratically with the number of FCs. Economies of scale can be achieved for any given delivery arc. However, because this network increases the number of arcs quadratically in the number of FCs, a significant order volume is needed to sustain such a network. This type of network structure restricts the total number of DCs, which increases the average distance between DCs and customers and decreases the delivery speed.

JD.com, on the other hand, has a tree-shaped delivery network. JD.com owns two types of DCs: RDCs are large storage facilities located in remote areas far from most customer locations, and FDCs are smaller storage facilities located closer to end customers. In this type of network, each FDC has a unique parent RDC for inventory replenishment, and each customer node relies on a unique FDC for delivery. Unlike the lateral network, the number of routes grows linearly with the number of FDCs and customer nodes, maintaining a low shipment cost. This type of network structure allows the e-tailer to build a large number of FDCs close to the customer locations, thus shortening the delivery time. The downside of such a network is a lower in-stock rate. FDCs can only hold a limited inventory of distinct SKUs. Consequently, many orders cannot be fulfilled by the parent FDC. When this occurs, JD.com routes the order to the parent RDC, thus increasing the delivery cost and time. This type of multi-layered network has been gaining much attention in recent years because it is able to successfully meet the needs of same-day delivery. Amazon is also building a multi-layered network, including fulfillment centers (for everyday products) and non-sortable fulfillment centers (for larger items), smaller fulfillment centers catering to same-day orders, and distribution centers that supply products to downstream fulfillment centers (Amazon 2021).

In the following, we discuss the inventory placement problem in these two types of networks using representative models. In Section 21.2.1, we present how JD.com allocates the inventory to its RDC-FDC network to minimize order split. In Section 21.2.2, we present an inventory placement model for the lateral fulfillment network and a case study based on Amazon's fulfillment network.

21.2.1 JD.com's Allocation to Minimize Order Split

JD.com focuses on same-day delivery because delivery speed has been reported as a major consideration for e-shopping customers (Jehl 2020). JD.com proposes the "211 program" for their same-day delivery service. The 211 program promises same-day delivery for orders placed before 11 a.m. and next-day delivery before 3 p.m. for orders placed between 11 a.m. and 11 p.m.

As explained above, JD.com's delivery network limits the number of delivery routes at the cost of a lower in-stock rate. In JD.com's delivery network, an FDC can only hold a limited number of SKUs. In addition, each customer node relies on a single FDC for delivery. Consequently, it is essential for the e-retailer to optimize SKU allocation to the FDC to maximize the in-stock rate.

On JD.com's e-commerce stores, 30% of the orders contain multiple items (Jehl 2020). When one of the items in the order is missing from the DC, the order is split. In JD.com's operations, a split order has one or several items being shipped from the RDC. Because RDCs are often located in remote areas, order splits result in delayed delivery. These split orders cannot be delivered with the service level of the 211 program. Because of these operational constraints, JD.com optimizes its SKU allocation from the RDC to the FDC to minimize order splits. The allocation happens between an RDC/FDC pair.

21.2.1.1 Single-pair allocation

The inventory placement problem between a single RDC/FDC pair was first studied by Jehl (2020), who formulated it as a mixed-integer program (MIP). Suppose that the RDC holds a universal set of SKUs denoted by I . FDC is dedicated to a single customer area. The FDC has a limited capacity k , such that only a subset of the SKUs can be held at the FDC. Once an order in the customer area has been placed, either all items in the order are fulfilled by the FDC or the order is split into the RDC. The objective of the optimization is to maximize the expected number of orders that can be fulfilled by the FDC in the upcoming period, that is, to minimize the expected number of order splits.

Let O be the set of all order types of the FDC. Each order type $o \in O$ represents a combination of items $o \subset I$. The demand for order type o is a random variable D_o with mean μ_{D_o} . The binary variable X_i denotes whether product i must be placed at the FDC. The binary variable Y_o denotes whether order type o can be fulfilled at the FDC. The inventory placement problem is formulated as follows:

$$\max \mathbb{E}[\sum_{o \in O} D_o Y_o] = \sum_{o \in O} \mu_{D_o} Y_o \quad (\text{P1})$$

$$s.t. \sum_{i \in I} X_i \leq k \quad (21.1)$$

$$\begin{aligned} Y_o &\leq X_i, & \forall o \in O, i \in o \\ X_i &\in \{0, 1\}, Y_o \in \{0, 1\}, & \forall i \in I, o \in O \end{aligned} \quad (21.2)$$

The constraint in Equation (21.1) specifies that the total number of SKUs placed at the FDC does not exceed the capacity k . The constraint in Equation (21.2) ensures that an order type can be fulfilled at the FDC only if every item in the order is placed at the FDC. The k -densest subgraph problem can be reduced to a specific instance of this problem. Therefore, this problem is generally NP-hard. Jehl (2020) utilized Lagrangian relaxation to solve this problem. The Lagrangian relaxation can be expressed as follows:

$$\begin{aligned}
& \max \sum_{o \in O} \mu_{D_o} Y_o - \lambda \sum_{i \in I} X_i \\
& s.t. Y_o \leq X_i, \quad \forall o \in O, i \in o \\
& X_i \in \{0,1\}, Y_o \in \{0,1\}, \quad \forall i \in I, o \in O
\end{aligned} \tag{P1-LR}$$

(P1-LR) can be regarded as a selection problem, where each order is a set of benefits μ_{D_o} and each product is an item with cost λ . One can observe that the constraint matrix in (P1-LR) is totally unimodular. Therefore, solving the linear relaxation of (P1-LR) yields an integer solution. Moreover, it was shown by Balinski (1970) that the selection problem can be solved faster by solving a minimum s - t cut problem.

The minimum s - t cut instance for (P1-LR) can be generated as follows: Construct a bipartite graph where, on the left-hand side, there is a set of nodes V_I , where each $v_i \in V_I$ corresponds to a unique SKU $i \in I$. On the right-hand side of this graph, there is a set of nodes V_O such that each $v_o \in V_O$ corresponds to a unique order type $o \in O$. Let $G = (V = \{s\} \cup V_I \cup V_O \cup \{t\}, A)$, where V is the set of nodes and A is the set of arcs. s and t represent the source and sink nodes, respectively. For each order type $o \in O$, there is an arc of infinite capacity from v_i to v_o for all $i \in I$ such that $i \in o$. In addition, there exists an arc from the source node s to every SKU node $v_i \in V_I$ with capacity λ , which is the Lagrangian multiplier in (P1-LR). Furthermore, for each order type $o \in O$, there is an arc of capacity μ_{D_o} to the sink node t . Figure 21.1 illustrates the concept of this construction. An s - t cut in this graph is a partition of the node set V into two subsets S and $T = V \setminus S$, such that $s \in S$ and $t \in T$. The capacity of the s - t cut $C(S, T)$ is defined as the sum of the capacities of the arcs linking a node in S to a node in T . Given a minimum s - t cut solution (S^*, T^*) where $S^* = \operatorname{argmin}_S \{ |S| : (S, T) \text{ is a minimum cut in } G \}$, we can construct an optimal solution for (P1-LR) by setting $X_i = 1[v_i \in T^*]$ for all $i \in I$ and $Y_o = 1[v_o \in T^*]$ for all $o \in O$.

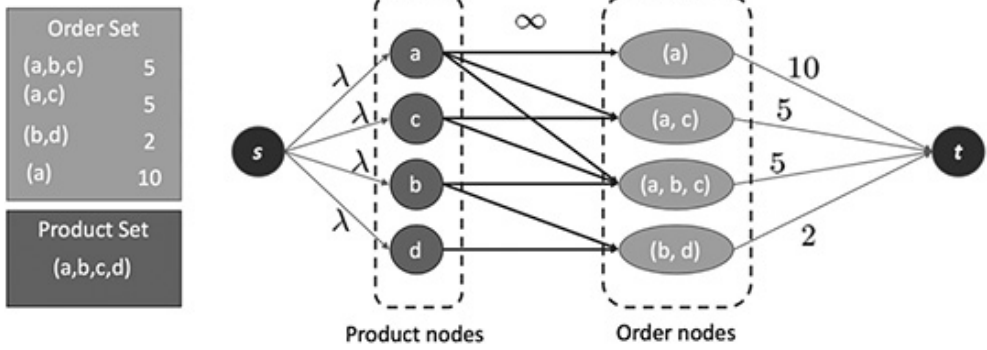
Jehl (2020) also showed that ranging the value of the Lagrangian multiplier λ can yield a list of nested assortments of products that were optimal for specific capacity constraints. This provided an upper bound for the optimal value given any FDC capacity, which served as a guideline for JD.com to decide the FDC capacity.

(P1) requires that each customer area have a dedicated FDC and that customer orders be fulfilled by this FDC or its parent RDC. In the next section, we discuss a model where flexibility is allowed in the fulfillment network so that the customer orders can be fulfilled by more than one FDC.

21.2.1.2 Flexibility in fulfillment

In the previous section, the inventory placement problem was between a single RDC/FDC pair, where each customer area had a dedicated FDC. Jehl (2020) further considered allowing flexibility in a fulfillment network. Flexibility in fulfillment allows a group of FDCs to cover a larger set of orders.

Consider an RDC with set J of the child FDCs. The RDC holds every SKU in set I . Each FDC $j \in J$ is dedicated to a single customer area. The number of SKUs an FDC can hold is upper bound by its physical capacity and daily replenishment time consumption. Let k_j denote the maximum number of SKUs that can be held at FDC j . When an order is placed at FDC j , it is either fulfilled by the FDC or routed to other warehouses, including the RDC and other



Source: Jehl (2020).

Figure 21.1 An example of a graph corresponding to a set of orders

FDCs in J . If an order is routed to the RDC, it is modeled as a complete order split, with a penalty of 1. If an order placed at FDC j is routed to another FDC j' , it incurs, by definition, an order-split penalty $p_{j,j'} \in [0,1]$. $p_{j,j} = 0$. The goal is to optimize the assortment $\mathcal{S}$ placed at each FDC $j \in J$ such that the total order-split penalty is minimized.

Let O be the set of all order types of an FDC. Each order type $o \in O$ represents a combination of items $o \subset I$. The demand of order type o at FDC j is a random variable D_o^j with mean $\mu(D_o^j)$. Let the binary variable X_i^j denote whether product i must be placed at FDC j . The binary variable Y_o^j denotes whether the order type o can be fulfilled at FDC j . Let W_o^j denote the unit reward from the flexible shipments of order type o placed at FDC j . The flexible inventory placement problem in an RDC-FDC network is formulated as follows:

$$\max \sum_{j \in J, o \in O} \mu(D_o^j) W_o^j \quad (\text{f-P1})$$

$$s.t. W_o^j = \max_{j' \in J} \{Y_o^{j'} (1 - p_{j',j})\}, \quad \forall o \in O, j \in J \quad (21.3)$$

$$\sum_{i \in I} X_i^j \leq k_j, \quad \forall j \in J, \quad (21.4)$$

$$\begin{aligned} Y_o^j &\leq X_i^j, & \forall o \in O, i \in o, j \in J \\ X_i^j &\in \{0,1\}, & \forall i \in I, j \in J \end{aligned} \quad (21.5)$$

The constraint in Equation (21.4) specifies that the total number of SKUs placed at FDC j does not exceed capacity k_j . The constraint in Equation (21.5) ensures that an order type can be fulfilled at FDC j only if every item of the order is placed at the FDC. The constraint in Equation (21.3) ensures that, among all FCs that can fulfill order type o , the one with the lowest penalty

is selected to finally ship the order. The objective is to maximize the total expected reward from flexible fulfillment.

In contrast to the (P1) for a single RDC/FDC pair, (f-P1) cannot be solved using the same Lagrangian relaxation framework. However, Jehl (2020) showed that the closed chain structure found in Jordan and Graves (1995) achieves more than half of the fully connected structures empirically. Under the closed chain structure, each FDC had a unique neighbor FDC; the orders placed at the FDC could be fulfilled by itself or by the neighboring FDC. Let $n(j)$ denote the unique neighbor FDC for each FDC j ; with the closed chain structure, the (f-P1) then becomes the following:

$$\begin{aligned}
 & \max \sum_{j \in J, o \in O} p_j \mu(D_o^j) Y_o^j + (1 - p_j) \mu(D_o^j) (1 - Z_o^{j, n(j)}) \\
 & \text{s.t. } Z_o^{j, n(j)} + Y_o^j + Y_o^{n(j)} \geq 1 \quad \forall o \in O, j \in J \\
 & \sum_{i \in I} X_i^j \leq k_j \quad \forall j \in J \\
 & Y_o^j \leq X_i^j, \quad \forall o \in O, i \in o, j \in J \\
 & X_i^j, Y_o^j \in \{0, 1\}, \quad \forall o \in O, i \in I, j \in J
 \end{aligned}$$

p_j denotes the penalty of fulfilling an order at FDC j from its neighboring FDC. $Z_o^{j, n(j)}$ is a binary variable indicating that the order type o placed at FDC j was fulfilled by the FDC itself or by the neighbor FDC $n(j)$. Under this closed chain structure, (f-P1) can be solved via a similar Lagrangian relaxation framework as the single-pair allocation problem. In particular, by relaxing the capacity constraints $\sum_{i \in I} X_i^j \leq k_j$, the resulting problem can be efficiently solved by utilizing the minimum s - t cut method when there is an even number of FDCs in the network.

21.2.2 Inventory Placement in a Lateral Fulfillment Network

The problem of allocating inventory in a lateral fulfillment network has been long recognized in the industry. In 2006, Amazon EU extended its single-node FC to a lateral FC network. Merriam (2007) described how Amazon EU has considered the inventory placement problem in their fulfillment network. As stated by Merriam (2007), the same order assignment policy in Amazon US is directly applicable to the EU network, but the inventory placement logic is not. The main reason for this is that, at the time, the US fulfillment network was set up such that most orders could be fulfilled locally; this was not the case for the EU network, where an order could, by chance, be fulfilled by any of the FCs in the network. Therefore, deciding what product line to hold at each FC and how much inventory to allocate in an FC became a critical problem. Amazon EU formulates and solves the inventory allocation problem given a fixed assortment in each FC as a linear program, minimizing the total transportation and labor costs subject to FC storage capacity, labor capacity, and supplier delivery constraints. This model considers fixed assortment allocation while ignoring the additional transportation cost of split

orders. Therefore, Merriam (2007) proposed a short-term and a long-term plan for optimizing assortment holding at each FC in the EU network.

A general inventory placement model based on the lateral fulfillment network structure is proposed by Chen and Graves (2021). In their model, each item is indexed by $i \in I$, and each FC is represented by $n \in N$. The entire service region is divided into a set R of sub-regions from which customer demand originates. An inventory placement plan is determined at the beginning of a certain planning period (e.g., one week), which is specified by the following decision variables:

- $\mathbf{x} = \{x_{nr}^i \in [0,1] : i \in I, n \in N, r \in R\}$, the fraction of region r , item i demand to be shipped from FC n , and
- $\mathbf{y} = \{y_n^i \in \{0,1\} : i \in I, n \in N\}$, where $y_n^i = 1$ indicates that item i is placed at FC n and $y_n^i = 0$ otherwise.

Let b_n denote the throughput capacity of FC n , which is the number of units that it can receive and ship during the entire planning horizon. Let c_{nr}^i denote the unit variable cost (e.g., the outbound shipping cost, inbound shipping cost, and the handling cost at the FC) of item i from FC n to region r . Let f_n^i denote the fixed cost for holding item i in FC n . Let d_r^i denote the demand for item i originated from region r . The inventory placement problem is formulated as follows:

$$\max_{\mathbf{x}, \mathbf{y}} \sum_{i \in I} \sum_{n \in N} \sum_{r \in R} c_{nr}^i d_r^i x_{nr}^i + \sum_{i \in I} \sum_{n \in N} f_n^i y_n^i \quad (\text{P2})$$

$$\text{s.t. } \sum_{n \in N} x_{nr}^i = 1, \quad \forall r \in R, \forall i \in I, \quad (21.6)$$

$$\sum_{i \in I} \sum_{r \in R} d_r^i x_{nr}^i \leq b_n, \quad \forall n \in N, \quad (21.7)$$

$$\sum_{r \in R} d_r^i x_{nr}^i \leq \left(\sum_{r \in R} d_r^i \right) y_n^i, \quad \forall n \in N, \forall i \in I, \quad (21.8)$$

$$x_{nr}^i \geq 0, y_n^i \in \{0,1\}, \quad \forall n \in N, \forall r \in R, \forall i \in I.$$

Equation (21.6) ensures that all the demands are satisfied. Equation (21.7) ensures that the total demand assigned to FC n does not exceed its throughput capacity. Equation (21.8) ensures the demand for item i can be shipped from FC n but only if FC n holds item i .

(P2) is a computationally challenging problem, even with state-of-the-art commercial solvers. Chen (2017) reported an example that reflects the computational burden to solve an inventory placement problem with a practical size: for a problem instance with $|N| = 88$ FCs, $|R| = 99$ regions, and $|I| = 1,000$ items, the Gurobi optimizer (version 7.0.1) on a personal laptop reaches an optimality gap of around 6–7% after five hours; in addition, there are not many improvements in the gap after an additional five hours. Chen and Graves (2021) proposed a large-scale optimization framework that works in an “aggregate–optimize–disaggregate”

fashion. In particular, the items are first grouped into clusters. Then, the cluster-level problem is solved via column generation. Finally, the solution is disaggregated into item-level decisions.

A large-scale synthetic numerical experiment was conducted based on Amazon's fulfillment network. The synthetic numerical experiment used data of Amazon's fulfillment centers in the United States (MWPVL International Inc. 2015), where there were $|N| = 88$ FCs as of December 2015. $|R| = 98$ regions were generated according to the 2010 US Census data (United States Census Bureau 2010). The FC locations and customer regions are depicted in Figure 21.2. $|I| = 1,000,000$ items were generated independently and uniformly at random. The numerical results showed that the algorithm framework found near-optimal solutions. In addition, it outperformed the sequential placement heuristic, which represented the status quo practice.

21.3 ORDER PICKING

In the previous section, we discussed the inventory placement problem, which is a tactical-level decision for online retailing. In this section, we focus on the operational-level decision in warehouses.

Order picking is widely recognized as the most labor-intensive and capital-intensive warehouse operation. More than 50% of warehouse operational costs are incurred by the order-picking process (De Koster et al. 2007). E-commerce platforms manage an enormous number of SKUs in large warehouses. Meanwhile, they receive a large number of orders that need to be delivered in a short time. In this context, it is becoming even more critical and challenging for e-commerce platforms to efficiently handle the order-picking process.



Figure 21.2 The Amazon FC locations and customer regions in Chen and Graves (2021)

The conventional order-picking process involves the batching, sequencing, routing, and sorting of picking requests. The batching problem involves partitioning a set of placed orders into batches, such that each batch is picked and sorted together during a specific time window. The batching decisions include both the size of the batch and assignment of orders to each batch. Batching reduces the total length of the pick tours at the expense of the longer fulfillment time of individual orders. Sequencing and routing determine the best sequence and route for the completion of picking requests. Sorting is required when multiple orders are batched together. Because items from multiple orders are picked together, they need to be sorted back into individual orders for packaging and shipping.

There are two types of sorting: sort-while-picking and sort-after-picking. Sort-while-picking complicates the picking process. Sort-after-picking requires a separate downstream sorting process. In practice, sorting time is often not the bottleneck of order picking in an automated warehouse system. Therefore, simple heuristic methods could work, and few studies have focused on this problem (Gu et al. 2007). We focus our attention on order batching, sequencing, and routing problems.

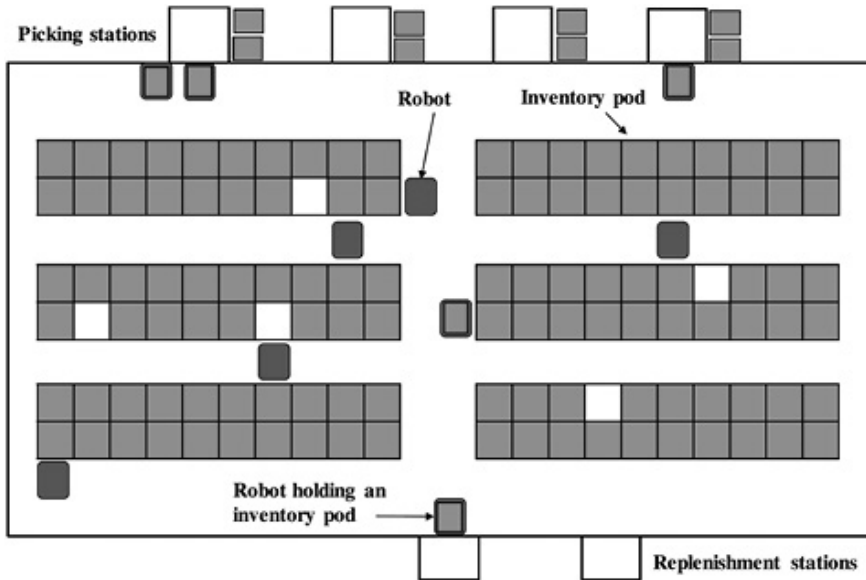
The choice of order-picking policy depends on the physical infrastructure of the warehouse and on storage technologies. As mentioned in Section 21.1.3, we focus on the two typical automated warehouse systems adopted by the major players in the e-commerce era: the AS/RS and RMFS. We provide a detailed description of these two systems and discuss the order-picking process for each option. We give a brief review of the literature on the order-picking problem for the AS/RS. Then, we discuss the order-picking problem for RMFS in terms of the literature and a recent industry practice of JD.com.

21.3.1 Automated Warehouses

In this section, we give a detailed description of the AS/RS and RMFS. The conventional AS/RS contains four components: racks, aisles, cranes, and I/O points. The racks are physical locations arranged in parallel for storage; the cranes are automated machines that perform picking and storage operations on the racks; the aisles are the empty spaces between the racks for the cranes to move; and the I/O points are where input loads are picked up for storage and where the picking orders are to be dropped off. There are different types of AS/RSs, here depending on the actual configuration. Interested readers are directed to a thorough overview of the various types of AS/RS in Roodbergen and Vis (2009). A typical AS/RS has fixed racks, each of which has one crane that cannot travel across racks; each crane can carry a unit load. This is called the *unit-load AS/RS*.

An RMFS is an innovative system with a very different structure from conventional warehouses. A typical RMFS includes four components: 1) inventory *pods* to store items replenished at the warehouse, 2) a robot with a lifting mechanism to lift the pods off the ground and carry them around the warehouse, 3) picking stations where the pods stop and allow workers to pick the items from the pod to fulfill the orders, and 4) replenishment stations where the pods stop and allow workers to store the replenished items onto the pods. Typically, an RMFS is arranged in a grid layout, where storage zones are in the middle and the pick/replenishment stations are at the periphery. Figure 21.3 provides an illustration of an RMFS.

The two major players have adopted these two typical warehouse systems, and both have used them successfully. Amazon has been rapidly expanding its use of an RMFS since its purchase of Kiva in 2012. According to a 2020 report, it utilized 200,000 robots in fulfillment centers (Edwards 2020). JD.com has a large network of automated warehouses in China, with



Source: Kim et al. (2020).

Figure 21.3 Illustration of an RMFS (Kim et al. 2020)

a hybrid of AS/RS and RMFS. Opening the first automated warehouse in 2014, JD.com gradually deployed 28 automated warehouses in China by May 2020 to greatly improve the delivery service level for 200 cities (Cao 2020).

21.3.2 Order Picking in an AS/RS

In this section, we discuss the order-picking problem in an AS/RS. An AS/RS operates with the conventional multi-parallel aisle layout, where automated cranes are installed on the storage racks. The orders are picked up and dropped off by the cranes at I/O points. Throughout this section, we use the terminology AS/RS, specifically the unit-load AS/RS, unless otherwise noted. The order-picking problem in an AS/RS is a thoroughly studied area. Roodbergen and Vis (2009) and Gu et al. (2007) provided an in-depth review and discussions on this issue. In particular, order picking in an AS/RS can be separated into two parts: order batching and order sequencing and routing.

Order batching determines how to divide arriving orders into batches for pickup. It is a computationally challenging problem. Given fixed batch sizes, the order-batching problem is a variant of the vehicle-routing problem (VRP), where the orders represent pickup locations on the rack to be assigned to crane routes. Each order contains multiple SKUs that represent multiple pick locations. VRP-based heuristic methods have been proposed for the order-batching problem of an AS/RS. There are two major types of heuristics. A seed algorithm starts with a random order as a “seed” and adds orders to the seed according to a route-closeness criterion (Elsayed 1981; Elsayed and Stern 1983; Hwang and Lee 1988). A saving heuristic starts by separating all the orders into a single route and iteratively merges the routes according to the

route-closeness criterion (Hwang et al. 1988; Elsayed and Unal 1989). A few studies have used mixed-integer programming techniques to determine the optimal order-batching solution (Armstrong et al. 1979; Gademann and van de Velde 2005).

Order sequencing determines the sequence and routes needed to perform the picking requests. In an AS/RS, there are two types of order sequencing cycle patterns: a single command cycle containing a single picking operation and a dual command cycle that occurs when a storage request and picking request are paired together in a single tour. The dual command cycle has the advantage of shorter total travel time. It was shown in Eben-Chaime and Pliskin (1996, 1997) that a hybrid mode, where a dual command cycle is performed whenever possible and a single command cycle is performed otherwise, can achieve a more stable system with fewer cranes. In certain warehouse situations, the arrival times of the pick and storage requests do not overlap. In this case, a single command cycle is performed. The literature on order sequencing for AS/RSs has focused primarily on the dual command sequencing problem.

There are two typical sequencing rules: block and dynamic sequencing. In block sequencing, a subset of the most urgent picking requests is selected as a static “block”; one block is sequenced and completed at a time. In dynamic sequencing, the entire request list is re-sequenced each time a new picking request is released. The storage operation determines the distribution of the SKUs in the warehouse and affects the retrieval routes. Therefore, the difficulty of the sequencing problem largely depends on the storage operation. The static sequencing problem can be formulated as a transportation problem (van den Berg and Gademann 1999) or an assignment problem (Lee and Schaefer 1997) with dedicated storage; it is NP-hard in general to find the optimal pick sequence with a randomized or class-based storage rule. Nearest-neighbor heuristics have been proposed for solving such problems (Han et al. 1987; Eynan and Rosenblatt 1993).

Block sequencing has the advantages of transparency and simplicity but may not guarantee good performance in a highly non-deterministic environment. Dynamic sequencing provides a more adaptive solution in an uncertain environment. Dynamic sequencing is mainly based on the reoptimization of the static sequencing problem when a new order request arrives. The nearest-neighbor heuristic rule can also be applied to dynamic control. Simulation studies show that the nearest-neighbor method outperforms the first-come, first-served and shortest-leg heuristic (Han et al. 1987), providing more savings when combined with class-based storage (Eynan and Rosenblatt 1993). In recent years, artificial intelligence-based methods have also been applied to find adaptive sequencing solutions in dynamic environments (e.g., Wang and Yih 1997; Seidmann 1988).

21.3.3 Order Picking in an RMFS

An AS/RS offers an automatic method for efficiently picking orders. The downside of an AS/RS is the high investment cost, long design cycle, and inflexibility of the system infrastructure. Online retailers must fulfill a large number of small orders across a million types of SKUs within a short time period. The order-batching process in an AS/RS can be time-consuming without a large order volume, a properly designed warehouse infrastructure, and an automatic control policy. These factors limit the use of an AS/RS for e-retailing.

The RMFS has emerged as a potential solution to these challenges. The order-picking procedure of an RMFS differs from that of conventional picker-to-parts systems and an AS/RS in

several aspects. We describe the order-picking procedure in an RMFS in terms of *order assignment, sequencing, and routing* as follows:

- *Order Assignment*: the orders in an RMFS are assigned to pick stations in real time. Typically, the arriving orders are backlogged. An order from the backlog is assigned to a pick station whenever there is an empty spot. It is rarely the case that multiple pick stations complete orders and have empty slots at the same time. The elimination of order batching is considered a major advantage of an RMFS over the conventional multi-parallel aisle system (Wurman et al. 2008). The elimination of batching saves unproductive waiting and processing time for aggregating orders and sorting back to individual orders. Thus, the first step in the order-picking procedure becomes order assignment without any explicit batching.
- *Sequencing and Routing*: the sequence of orders to be completed at a pick station is essentially an order assignment problem. In addition, new sequencing and routing problems arise in an RMFS: i) pod selection, that is, which pod to assign a pick station, ii) pod storage, that is, where to return the pods in the storage zone after a pick operation, iii) robot assignment, that is, which robot carries the pod movement tasks, and iv) robot path planning.

These new decision problems are highly complex and interdependent. It is difficult to find a jointly optimal solution or even an optimal solution for a single decision. In practice, decisions are made independently in sequence, relying on heuristic decision rules. Amazon Robotics, which was previously the Kiva system, utilizes heuristic rules for the order-picking process (Wurman et al. 2008). For order assignment, the Kiva system assigns similar orders to the same pick station so that multiple items can be picked from the pod in a single tour. Pod selection is based on the distance and number of needed products available on candidate pods. For pod storage, the time to free a robot and time to deliver the next order is considered. Pods with a higher utilization rate are placed closer to a pick station. The utilization of heuristics in implementation shows great flexibility for system control. Field systems have also verified these heuristics in various configurations and operational environments.

Rule-based solutions have also been reported in the literature. Merschformann et al. (2019) studied multiple decision rules for the order assignment, pod selection, and pod storage problems of an RMFS. They used simulations to compare the performances of different combinations of decision rules, which they called the *rule configurations* (RCs). They found that an RC's successful performance hinges on high pile-on and short travel distances. Among the order-picking problems, order assignment affected the overall performance the most. They also reported that the Pod-Match rule, which selects the pick order that best matches pods heading to a pick station, performed the best for order assignment. Zou et al. (2017) used semi-open queueing networks to analyze the policy for assigning pods to select stations. They showed that their proposed handling speed-based assignment rule significantly outperformed the randomized policy when the service time of pickers had a large variance. A velocity-based rule for pod storage, in which popular items were placed closer to pick stations, was proposed by Yuan (2016), who used a fluid model and numerical experiments to show that the velocity-based rule led to a significant improvement in randomized pod storage under the randomized replenishment policy. Yuan (2016) also showed that class-based pod storage policies with two or three classes can achieve most of the benefits from an idealized velocity-based policy.

A limited amount of literature has investigated the order-picking problem in an RMFS using mathematical modeling and optimization techniques. Weidinger et al. (2018) studied the pod storage problem for an RMFS by formulating it as a mixed-integer program given a known pod visit schedule at the pick stations on the immediate planning horizon. Boysen et al. (2017) studied the joint order assignment and sequencing of a single pick station and proposed heuristic decomposition procedures that iteratively solved the order sequencing and rack sequencing subproblem, respectively. One of the successful applications of mathematical optimization techniques for order picking in an RMFS is demonstrated by JD.com's intelligent warehouse.

21.3.3.1 JD.com's order-picking optimization for an RMFS

JD.com utilizes mathematical optimization techniques for the order-picking process in their RMFS warehouses. In particular, real-time dispatching decisions among robots, pods, and pick stations are made by solving large-scale integer programs in seconds, which has helped the company decrease its fulfillment expense ratio to a world-leading level of 6.5% (Qin et al. 2021).

JD.com models the joint pod selection and robot assignment problem in RMFS as an online tripartite network flow model (Qin et al. 2021). In particular, the system is characterized by the following sets and parameters:

- $I = I_a \cup I_b$ is the set of the robots indexed by i , where I_a is the set of idle robots and I_b is the set of occupied robots.
- $J = J_a \cup J_b$ is the set of all sides of the pods indexed by j , where J_a is the set of pod sides on the idle pods and J_b is the set of pod sides on the moving pods that are being carried by a robot.
- T is the set of all pods, as indexed by t . Given a pod t , one knows the index of the sides of the pod; and given each side index, one knows the index of the pod that the side belongs to.
- For each pod $t \in T$, let J_t denote the set of pod sides on pod t . JD.com's warehouse consists of one-sided and double-sided pods. Thus, $|J| \leq 2|T|$.
- For each moving robot $i \in I_b$, let t_i denote the pod that it is carrying. Let J_{t_i} denote the indices (index) of the pod sides on this pod.
- K is the set of pick stations indexed by k .
- S is the set of all SKU types, as indexed by s .
- $\{c_{ij}^1\}_{i \in I, j \in J}$ and $\{c_{jk}^2\}_{j \in J, k \in K}$: the travel-distance parameters. c_{ij}^1 denotes the travel distance between the location of robot i and the location of pod side j . c_{jk}^2 denotes the travel distance between the location of pod side j and the pick station k .
- $\{O_{ks}\}_{k \in K, s \in S}$: the number of SKU type s required at pick station k .
- $\{q_{js}\}_{j \in J, s \in S}$: the number of SKU type s on pod side j .
- $\{B_k\}_{k \in K}$: the number of available berths at pick station k ; that is, at most B_k pods can be parked at station k .

The company wants to make the following decisions:

- $\mathbf{x} = \{x_{ij} \in \{0, 1\} : i \in I, j \in J\}$: which robot to assign to which pod side. $x_{ij} = 1$ means that robot i is assigned to pod side j and $x_{ij} = 0$ otherwise.

- $y = \{y_{jk} \in \{0,1\} : j \in J, k \in K\}$: which pod side to bring to which pick station. $y_{jk} = 1$ means that a pod side j is assigned to pick station k and $y_{jk} = 0$ otherwise.
- $z = \{z_{ks} \in \mathbb{Z}_+ : k \in K, s \in S\}$: how much demand of type s is unsatisfied at pick station k .

The goal of the company is to minimize the travel distance cost of robots while fulfilling as many of the orders as possible. This is characterized by the following objective function, which is a weighted sum of the total cost from the travel distance of the robots and the unsatisfied demand at the pick stations:

$$\min \alpha_1 \sum_{i \in I} \sum_{j \in J} c_{ij}^1 x_{ij} + \alpha_2 \sum_{j \in J} \sum_{k \in K} c_{jk}^2 y_{jk} + \alpha_3 \sum_{k \in K} \sum_{s \in S} z_{ks}$$

α_1 , α_2 , and α_3 are the weights assigned to the total cost of moving robots to pod sides, moving pod sides to pick stations, and unsatisfied demand, respectively. The assignment decisions are subject to the following constraints:

$$\sum_{j \in J} x_{ij} \leq 1, \quad \forall i \in I \quad (21.9)$$

$$\sum_{i \in I} x_{ij} \leq 1, \quad \forall j \in J \quad (21.10)$$

$$\sum_{k \in K} y_{jk} \leq 1, \quad \forall j \in J \quad (21.11)$$

$$\sum_{j \in J} y_{jk} \leq B_k, \quad \forall k \in K \quad (21.12)$$

$$\sum_{i \in I} x_{ij} \geq \sum_{k \in K} y_{jk}, \quad \forall j \in J \quad (21.13)$$

$$\sum_{j \in J} y_{jk} \cdot q_{js} \geq O_{ks} - z_{ks}, \quad \forall k \in K, s \in S \quad (21.14)$$

$$x_{ij} = 0, \quad \forall i \in I_a, j \in J_b \quad (21.15)$$

$$x_{ij} = 0, \quad \forall i \in I_b, j \notin J_{t_i} \quad (21.16)$$

$$\sum_{i \in I} \sum_{j \in J_t} x_{ij} \leq 1, \quad \forall t \in T \quad (21.17)$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in I, j \in J, \quad (21.18)$$

$$y_{jk} \in \{0,1\}, \quad \forall j \in J, k \in K, \quad (21.19)$$

$$z_{ks} \in \mathbb{Z}_{+}, \quad \forall k \in K, s \in S. \quad (21.20)$$

Equation (21.9) requires that each robot can be assigned to at most one pod side. Equation (21.10) requires that at most one robot can be assigned to each pod side. Equation (21.11) requires that each pod side can be assigned to at most one pick station. Equation (21.12) ensures that the number of pod sides assigned at a workstation is no more than the number of available berths at the workstation. Equation (21.13) ensures that a robot is assigned to a pod side if and only if the pod side is assigned to a pick station. Equation (21.14) characterizes the unsatisfied demand at each station given the pod sides assigned to it.

Equation (21.15) requires that an idle robot cannot be assigned to a moving pod. Equation (21.16) requires that an occupied robot can only be assigned to the pod sides of the pod that it is currently carrying. Equation (21.17) ensures that only one side of a multi-sided pod can be assigned to a robot.

There are $|I| \times |J| + |J| \times |K| + |K| \times |S|$ decision variables. This is a very large-scale optimization problem for a typical warehouse: $|I| = 250$ robots, $|T| = 1,800$ pods ($|J| = 3300$ pod sides), and $|K| = 50$ workstations, with $|S| = 2,000$ types of SKUs, which results in $O(10^6)$ variables and constraints. However, this tripartite network flow problem needs to be solved on the fly, which takes around three seconds per operating period.

To tackle this challenge, Qin et al. (2021) proposed a Lagrangian relaxation algorithm that can solve this problem efficiently and provide solutions of a good quality. The algorithm works by relaxing Equation (21.13), which is the only constraint that links the decision variables $\mathbf{x}$ and $\mathbf{y}$. Let $\lambda = \{\lambda_j \geq 0 : j \in J\}$ denote the Lagrangian multiplier for Equation (21.13). The Lagrangian relaxation problem can be written as follows:

$$\begin{aligned} P(\lambda) &= \min_{\mathbf{x}, \mathbf{y}, \mathbf{z}} \alpha_1 \sum_{i \in I} \sum_{j \in J} c_{ij}^1 x_{ij} + \alpha_2 \sum_{j \in J} \sum_{k \in K} c_{jk}^2 y_{jk} + \alpha_3 \sum_{k \in K} \sum_{s \in S} z_{ks} + \sum_{j \in J} \lambda_j \left(\sum_{k \in K} y_{jk} - \sum_{i \in I} x_{ij} \right) \\ &= \min_{\mathbf{x}, \mathbf{y}, \mathbf{z}} \sum_{i \in I} \sum_{j \in J} (\alpha_1 c_{ij}^1 - \lambda_j) x_{ij} + \sum_{j \in J} \sum_{k \in K} (\alpha_2 c_{jk}^2 + \lambda_j) y_{jk} + \alpha_3 \sum_{k \in K} \sum_{s \in S} z_{ks} \end{aligned}$$

subject to Equations (21.9)–(21.12) and (21.14)–(21.18). This relaxed problem can be decomposed into two subproblems:

$$P^1(\lambda) = \min_{\mathbf{x}} \sum_{i \in I} \sum_{j \in J} (\alpha_1 c_{ij}^1 - \lambda_j) x_{ij}$$

subject to Equations (21.9), (21.10), (21.15), (21.16), (21.17), and (21.18),

and

$$P^2(\lambda) = \min_{\mathbf{y}, \mathbf{z}} \sum_{j \in J} \sum_{k \in K} (\alpha_2 c_{jk}^2 + \lambda_j) y_{jk} + \alpha_3 \sum_{k \in K} \sum_{s \in S} z_{ks}$$

subject to Equations (21.11), (21.12), (21.14), (21.19), and (21.20).

In other words, the relaxed tripartite network flow problem is decomposed into two subproblems: the first subproblem $P^1(\lambda)$ matches the pod sides and robots; the second subproblem $P^2(\lambda)$ matches the robots and pick stations. In general, the Lagrangian relaxation method requires searching for the optimal Lagrangian multiplier $\lambda^* = \arg\max_{\lambda \geq 0} P(\lambda)$ to provide a better-quality solution. However, searching for λ^* is time-consuming. Therefore, the dual variable associated with Equation (21.13) of the continuous relaxation is used as an approximation of the optimal Lagrangian multiplier λ^* . The continuous relaxation is solved to update λ in a longer time interval, only when it is necessary. Given a fixed λ , $P^2(\lambda)$ is an integer program with a smaller size than the original problem. The authors solved the continuous relaxation of $P^2(\lambda)$ and round the fractional solution to an integer solution. They added lifting cuts and minimum cover cuts to further tighten the constraints. Given the solution to $P^2(\lambda)$, $P^1(\lambda)$ is solved with an additional constraint $\sum_{i \in I} x_{ij} \geq \sum_{k \in K} y_{jk}^*, \forall j \in J$, which is the relaxed Equation (21.13). This ensures that the resulting solution is feasible. $P^1(\lambda)$ is an unbalanced assignment problem, which is known to be equivalent to its continuous relaxation and can be solved by existing algorithms such as the Hungarian algorithm in polynomial time.

21.4 CONCLUSION AND DISCUSSION

Online retailing brings new challenges and problems to the retail industry and operations management (OM) community. The inventory placement problem and order-picking problem discussed here have recently gained more attention. This chapter does not mean to be a comprehensive review of the related literature. We have carefully selected and discussed several of the most relevant and representative works in the field. We hope that this review will spark further interest and research into this critical area. Although we highlighted some of the primary research on these problems, a large gap persists between the literature and practice.

The inventory placement model in Section 21.2 focused on two specific network structures. An important future direction would be to study inventory placement decisions on other network structures and determine how the network structure affects inventory placement policy. The methodology that we reviewed was based on combinatorics and mixed-integer programming techniques. In practice, it would be of great interest to develop heuristic rules for placing SKUs in different fulfillment networks. Inventory placement is a first-stage decision that affects the replenishment, transshipment, and real-time fulfillment decisions of the platform. Recent research studied real-time fulfillment decisions (i.e., deciding which FC to fulfill which order in real time) given fixed inventory allocation plans (see, e.g., Acimovic and Graves 2015; Xu et al. 2009; Zhao et al. 2020). Another future direction is to incorporate fulfillment dynamics into making the inventory placement decision.

Although order-picking control has been thoroughly studied within the context of the traditional retail industry and conventional warehouse systems, there is a large gap between the literature and real-world applications for e-retailing. The advent of RMFSs has opened a new era for warehouse control. To the best of our knowledge, a theoretical analysis of the warehouse operations and general modeling framework of RMFSs is lacking. We have observed the success of conventional picker-to-parts systems, the AS/RS, and the RMFS in different types of field systems. However, no clear conclusion can be drawn on which system under which circumstances is better for e-retailing. We highlight some features that may determine the efficiency of a warehouse system. For example, an AS/RS can achieve success for large-scale

companies with both capital capacity and a stable market, while an RMFS might be more suitable for small- or medium-sized companies in an expansion phase or in areas with sparse populations. AS/RS seems to be a better fit for large-scale, non-sortable warehouse facilities that handle bulk items; the RMFS seems to fit sortable warehouses and forward distribution centers that handle daily customer orders. We believe in the value of developing general rules for selecting appropriate warehouse control policies for both systems in e-retailing.

The OM community has traditionally utilized a model-based approach to solve real-world problems, based on theoretical assumptions. However, it is challenging to verify whether these assumptions hold in practice or to evaluate how much they deviate from reality, mainly due to the lack of access to real data characterizing system behavior. With the current high level of digitization and the development of machine learning and artificial intelligence techniques, there is an unprecedented amount of data accessible to researchers in all fields. Online retailing problems are highly dynamic and complex, but they naturally generate large volumes of valuable data. Open-source datasets made available by major retailing platforms, such as JD.com (Shen et al. 2020), RiRiShun Logistics (Guo et al. 2021), Cainiao Network (Tianchi 2018), and a large supermarket in Zhao et al. (2020), offer opportunities for data-driven research. We believe that there is a significant opportunity to exploit this data with domain knowledge of inventory management.

Equally important for the field of retail inventory management is the shift from literature-driven research to practice-driven research (Chen et al. 2023). Online retail inventory management systems are increasingly complex and may deviate from the standard inventory management literature, which highlights the importance of OM researchers collaborating with practitioners to identify core issues, lead developments, and build a unified research and practice framework. Recent studies have verified successful implementation through field experiments with Alibaba (Zhang et al. 2019; Feldman et al. 2021; Sun et al. 2021), Amazon (Cui et al. 2019), and JD.com (Qin et al. 2021, Qi et al. 2021). Achieving this goal requires overcoming the challenges of verifying performance improvement, building trust with practitioners, and balancing model accuracy and simplicity. We hope that this chapter will encourage more researchers to take on these challenges and help close the gap between OM research and practical implementation.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Titouan Jehl for the constructive suggestions and support in structuring this chapter.

REFERENCES

- Acimovic, J., & Graves, S. C. (2015). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management*, 17(1), 34–51.
- Allen, S. G. (1958). Redistribution of total stock over several user locations. *Naval Research Logistics Quarterly*, 5(4), 337–345.
- Allen, S. G. (1961). A redistribution model with set-up charge. *Management Science*, 8(1), 99–108.
- Amazon. (2021). The evolution of Amazon's inventory planning system. Retrieved December 18, 2021, from <https://www.amazon.science/latest-news/the-evolution-of-amazons-inventory-planning-system>

- Armstrong, R. D., Cook, W. D., & Saipé, A. L. (1979). Optimal batching in a semi-automated order picking system. *Journal of the Operational Research Society*, 30(8), 711–720.
- Axsäter, S. (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research*, 38(1), 64–69.
- Balinski, M. L. (1970). Notes on a selection problem. *Management Science*, 17(3), 230–231.
- Boysen, N., Briskorn, D., & Emde, S. (2017). Parts-to-picker based order processing in a rack-moving mobile robot environment. *European Journal of Operational Research*, 262(2), 550–562.
- Boysen, N., De Koster, R., & Weidinger, F. (2019). Warehousing in the e-commerce era: A survey. *European Journal of Operational Research*, 277(2), 396–411.
- Cao, L. (2020). JD adds three new Asia No. 1 logistics parks, additional measures to support 618 promotion. Retrieved December 18, 2021, from Jd.com corporate blog; <https://jdcorporateblog.com/jd-adds-three-new-asia-no-1-logistics-parks-additional-measures-to-support-618-promotion/>
- Chen, A. I., & Graves, S. C. (2021). Item aggregation and column generation for online-retail inventory placement. *Manufacturing & Service Operations Management*, 23(5), 1062–1076.
- Chen, A. I. A. (2017). *Large-scale optimization in online-retail inventory management* [Ph.D. Thesis]. Massachusetts Institute of Technology.
- Chen, F., & Song, J. S. (2001). Optimal policies for multi-echelon inventory problems with Markov-modulated demand under the assumption of Markov modulated demand. *Operations Research*, 49(2), 226–234.
- Chen, F., & Zheng, Y. S. (1994). Lower bounds for multi-echelon stochastic inventory systems. *Management Science*, 40(11), 1426–1443.
- Chen, X., Deng, T., Shen, Z. J. M., & Yu, Y. (2023). Mind the gap between research and practice in operations management. *IIE Transactions*, 55(1), 32–42.
- Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475–490.
- Cui, R., Zhang, D., & Bassambooc, A. (2019). Learning from inventory availability information: Evidence from field experiments on Amazon. *Management Science*, 65(3), 1216–1235.
- De Koster, R., Le-Duc, T., & Roodbergen, K. J. (2007). Design and control of warehouse order picking: A literature review. *European Journal of Operational Research*, 182(2), 481–501.
- Eben-Chaïme, M., & Pliskin, N. (1996). An integrative model for automatic warehousing systems. *International Journal of Computer Integrated Manufacturing*, 9(4), 286–292.
- Eben-Chaïme, M., & Pliskin, N. (1997). Operations management of multiple machine automatic warehousing systems. *International Journal of Production Economics*, 51(1–2), 83–98.
- Edwards, D. (2020). Amazon now has 200,000 robots working in its warehouses. Retrieved December 18, 2021, from <https://roboticsandautomationnews.com/2020/01/21/amazon-now-has-200000-robots-working-in-its-warehouses/28840>
- Elsayed, E. A. (1981). Algorithms for optimal material handling in automatic warehousing systems. *International Journal of Production Research*, 19(5), 525–535.
- Elsayed, E. A., & Stern, R. G. (1983). Computerized algorithms for order processing in automated warehousing systems. *International Journal of Production Research*, 21(4), 579–586.
- Elsayed, E. A., & Unal, O. I. (1989). Order batching algorithms and travel-time estimation for automated storage/retrieval systems. *International Journal of Production Research*, 27(7), 1097–1114.
- Eynan, A., & Rosenblatt, M. J. (1993). An interleaving policy in automated storage/retrieval systems. *International Journal of Production Research*, 31(1), 1–18.
- Federgruen, A., & Zipkin, P. H. (1984). Computational issues in an infinite-horizon, multi-echelon inventory model. *Operations Research*, 32(4), 818–836.
- Feldman, J., Zhang, D., Liu, X., & Zhang, N. (2021). Customer choice models versus machine learning: Finding optimal product displays on Alibaba. *Operations Research*, 70(1), 309–328.
- Gademann, N., & van de Velde, S. (2005). Order batching to minimize total travel time in a parallel-aisle warehouse. *IIE Transactions*, 37(1), 63–75.
- Gallego, G., Özer, Ö., & Zipkin, P. (2007). Bounds, heuristics, and approximations for distribution systems. *Operations Research*, 55(3), 503–517.
- Graves, S. C. (1985). A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science*, 31(10), 1247–1256.

- Gu, J., Goetschalckx, M., & McGinnis, L. F. (2007). Research on warehouse operation: A comprehensive review. *European Journal of Operational Research*, 177(1), 1–21.
- Guo, X., Yu, Y., Allon, G., Wang, M., & Zhang, Z. (2021). RiRiShun logistics: Home appliance delivery data for the 2021 Manufacturing & Service Operations Management data-driven research challenge. *Manufacturing & Service Operations Management*.
- Han, M. H., McGinnis, L. F., Shieh, J. S., & White, J. A. (1987). On sequencing retrievals in an automated storage/retrieval system. *IIE Transactions*, 19(1), 56–66.
- Herer, Y. T., Tzur, M., & Yucesan, E. (2002). Transshipments: An emerging inventory recourse to achieve supply chain agility. *International Journal of Production Economics*, 80(3), 201–212.
- Hwang, H., Baek, W., & Lee, M. K. (1988). Clustering algorithms for order picking in an automated storage and retrieval system. *International Journal of Production Research*, 26(2), 189–201.
- Hwang, H., & Lee, M. K. (1988). Order batching algorithms for a man-on-board automated storage and retrieval system. *Engineering Costs and Production Economics*, 13, 285–294.
- Jehl, T. (2020). *Data-driven decision making algorithms for internet platforms* [Ph.D. Thesis]. University of California, Berkeley.
- Jordan, W. C., & Graves, S. C. (1995). Principles on the benefits of manufacturing process flexibility. *Management Science*, 41(4), 577–594.
- Kim, H. J., Pais, C., & Shen, Z. J. M. (2020). Item assignment problem in a robotic mobile fulfillment system. *IEEE Transactions on Automation Science and Engineering*, 17(4), 1854–1867.
- Langenhoff, L. J. G., & Zijm, W. H. M. (1990). An analytical theory of multi-echelon production/distribution systems. *Statistica Neerlandica*, 44(3), 149–174.
- Lee, H. S., & Schaefer, S. K. (1997). Sequencing methods for automated storage and retrieval systems with dedicated storage. *Computers and Industrial Engineering*, 32(2), 351–362.
- Merriam, K. (2007). *Reducing total fulfillment at costs at amazon EU through network design optimization* [Master's Thesis]. Massachusetts Institute of Technology.
- Merschformann, M., Lamballais, T., De Koster, M. B. M., & Suhl, L. (2019). Decision rules for robotic mobile fulfillment systems. *Operations Research Perspectives*, 6, 100–128.
- MWPVL International Inc. (2015). The Amazon fulfillment center and distribution center network in the United States. Retrieved December 18, 2021, from http://www.mwpvl.com/html/amazon_com.html
- Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2), 125–136.
- Qi, M., Shi, Y., Qi, Y., Ma, C., Yuan, R., Wu, D., & Shen, Z. J. (2023). A practical end-to-end inventory management model with deep learning. *Management Science*, 69(2), 759–773.
- Qin, H., Xiao, J., Ge, D., Xin, L., Gao, J., He, S., Hu, H., & Carlsson, J. G. (2021). JD.com: Operations research algorithms drive intelligent warehouse robots to work. *INFORMS Journal on Applied Analytics*, Forthcoming.
- Roodbergen, K. J., & Vis, I. F. (2009). A survey of literature on automated storage and retrieval systems. *European Journal of Operational Research*, 194(2), 343–362.
- Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), 565–579.
- Seidmann, A. (1988). Intelligent control schemes for automated storage and retrieval systems. *International Journal of Production Research*, 26(5), 931–952.
- Shang, K. H., & Song, J. S. (2003). Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science*, 49(5), 618–638.
- Shen, M., Tang, C. S., Wu, D., Yuan, R., & Zhou, W. (2020). JD.com: Transaction-level data for the 2020 MSOM data driven research challenge. *Manufacturing & Service Operations Management*.
- Sherbrooke, C. C. (1968). METRIC: A multi-echelon technique for recoverable item control. *Operations Research*, 16(1), 122–141.
- Snyder, L. V., & Shen, Z. J. M. (2011). *Fundamentals of supply chain theory*. Wiley.
- Sun, J., Zhang, D., Hu, H., & Van Mieghem, J. (2021). Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations. *Management Science*, 68(2), 809–1589.
- Tianchi. (2018). CAINIAO MSOM data-driven research competition. Retrieved from <https://tianchi.aliyun.com/competition/entrance/231623/introduction>

- United States Census Bureau. (2010). 2010 census gazetteer files: Counties. Retrieved December 18, 2021, from <http://www.census.gov/geo/maps-data/data/gazetteer2010.html>
- van den Berg, J., & Gademann, A. (1999). Optimal routing in an automated storage/retrieval system with dedicated storage. *IIE Transactions*, 31(5), 407–415.
- Wang, J. Y., & Yih, Y. (1997). Using neural networks to select a control strategy for automated storage and retrieval systems (AS/RS). *International Journal of Computer Integrated Manufacturing*, 10(6), 487–495.
- Weidinger, F., Boysen, N., & Briskorn, D. (2018). Storage assignment with rack-moving mobile robots in KIVA warehouses. *Transportation Science*, 52(6), 1479–1495.
- Wurman, P. R., D'Andrea, R., & Mountz, M. (2008). Coordinating hundreds of cooperative, autonomous vehicles in warehouses. *AI Magazine*, 29(1), 9–9.
- Xu, P. J., Allgor, R., & Graves, S. C. (2009). Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management*, 11(2), 340–355.
- Yuan, R. (2016). *Velocity-based storage and stowage decisions in a semi-automated fulfillment system* [Ph.D. Thesis]. Massachusetts Institute of Technology.
- Zhang, D., Dai, H., Dong, L., Wu, Q., Guo, L., & Liu, X. (2019). The value of pop-up stores on retailing platforms: Evidence from a field experiment with Alibaba. *Management Science*, 65(11), 5142–5151.
- Zhao, L., Li, L., & Shen, Z. J. M. (2020). Transactional and in-store display data of a large supermarket for data-driven decision-making. *Naval Research Logistics*, 67(8), 617–626.
- Zhao, Y., Wang, X., & Xin, L. (2022). Multi-item online order fulfillment in a two-layer network. *Chicago Booth Research Paper*, 20–41.
- Zou, B., Gong, Y., Xu, X., & Yuan, Z. (2017). Assignment rules in robotic mobile fulfillment systems for online retailers. *International Journal of Production Research*, 55(20), 6175–6192.

Index

- Abbasi, B. 443
Abhyankar, H. S. 107
Abouee-Mehrizi, H. 33, 35, 443
Achy-Brou, A. 125
Acimovic, J. 208
Action Elimination algorithm 361
active packaging 40–42
adaptive inventory management (AIM) algorithm 349
additive demand 289, 292
additive manufacturing (AM) 470
advanced booking discount (ABD) programs 147
aggregate mean waiting time 458
aggregation and decomposition methods 120
Agrawal, N. 201, 478, 483
Agrawal, S. 24, 354
AIM algorithm *see* adaptive inventory management (AIM) algorithm
Akçay, Y. 194, 199
Akkas, A. 43
Aksoy, Y. 489
Allon, G. 172, 187, 265
 α -approximation algorithm 234
Amazon 262, 288, 500, 511
Amazon Robotics LLC 504
Amil, A. 208
An, M. Y. 313
Anand, K. 284n9
Angelus, A. 64, 74, 80, 81, 83, 89, 91, 93–6
anti-multimodularity 29, 34, 35
Anupindi, R. 481
approximate approach 466–7
approximation algorithms 234
 state-of-the-art overview 234–6
approximation balancing policy 220
approximation ratio 234, 236, 257
Archibald, B. 13, 16
Ardestani-Jaafari, A. 147
Arreola-Risa, A. 54
Arrow, K. J. 307
Arts, J. 19
assemble-to-order (ATO) systems 191–2
 assembly design 208–9
 dynamic ATO models 200–207
 with endogenous prices 209
 e-retailing 207–8
 extensions and research directions 207–9
 one-period models 192–200
 online resource allocation with component commonality 207
assembly design, ATO systems 208–9
assembly systems 64–5
 Clark–Scarf model 89–90
 nonstationary assembly systems 91–6
 stationary assembly systems 90–91
 extensions, multi-echelon inventory models 118–19
assembly systems management, applications in 138
 general assembly systems 138–40
 single-product assemble-to-order systems 140–42
assortment planning 483–4
asymmetric demand, CLT 154–6
asymptotically optimal policies, dynamic ATO models 205
asymptotic optimality
 constant-order policies 9–11
 echelon OUT policy 22–3
 order-up-to policies 5–6
Asynchronous Advantage Actor Critic (A3C) algorithm 187, 337
Atalı, A. 54, 264
Atan, Z. 191, 200, 201
Atkins, D. 125
Atkinson, M. P. 443
ATO systems *see* assemble-to-order (ATO) systems
automated storage/retrieval system (AS/RS) 504, 520
 order picking in 513–14
automated store ordering (ASO) systems 477
automated warehouses 512–13
autoregressive and moving average (ARMA) process 106
Aviv, Y. 264, 322, 494
Axsäter, S. 66, 120, 125, 390
Azoury, K. 335

Babich, V. 380
backlogged demand, nonparametric learning with 366–7
backlogging 52, 304n3
backorder cost rate 383–4
backordering 2, 6, 7, 13, 22, 237
backorder penalty 7–8, 33
backorder system 6–8, 13, 15, 22

- Bai, X. 9
- balanced assumption 120
- balanced base-stock (BBS) policy 202–3, 205
- balanced cost 256
- balanced echelon base-stock policy 118
- balancing quantity 239, 242
- balancing-type policy 436
- Balinski, M. L. 507
- Bandi, C. 148
- Ban, G. 317–20, 329, 330
- Ban, G.-Y. 486
- bang-bang structure 298
- base-stock list-price policy 51, 291, 292
- base-stock (BS) policies 13–15, 47, 56, 61, 288, 321, 352, 435, 449, 458
 - with delay 23
- Bassamboo, A. 265
- Basten, R. J. I. 456
- batch-order contract (BOC) 275–7
- batch ordering constraint 132–3
- batch sizes
 - bounds for reorder points with fixed 112–13
 - simple solution for 113–14
- Baye, M. R. 379
- Bayesian persuasion framework 266
- BEBS *see* branching echelon base stock (BEBS)
- Becker-Peth, M. 418–20
- behavioral bullwhip effect
 - isolating 411–12
 - moderating factors and mitigation strategies 412
- behavioral inventory management 400–401
 - censored-demand/unobservable lost sales 413
 - economic order quantity and review periods 413
 - multi-location retailer inventory sharing 414
 - newsvendor problem 401–8
 - price-setting and revenue management 414
 - serial supply chain 408–13
 - supply risk and dual sourcing 413
- behavioral inventory-management research 400, 414–15
 - decision behavior, addressing “behavioral problems” 415–18
 - designing processes and systems with human behavior considerations 420–25
 - predictable irrationality 418–20
- Belavina, E. 398
- Bendavid, I. 397
- Ben-David, S. 331
- Benjaafar, S. 143, 206, 207
- Bensoussan, A. 292
- Ben-Tal, A. 148, 152, 162
- Benzion, U. 404
- Bergen, M. E. 56
- Berling, P. 126, 143
- Bernstein, F. 302
- Bertsimas, D. 147–8, 158, 162, 163
- Besbes, O. 24, 336, 348, 364
- Beutel, A. L. 486
- Bielecki, T. 56
- Bienstock, D. 147
- Bijvank, M. 12, 13, 15–16, 22, 23
- Billington, C. 104
- Bill-of-Materials 192
- bisection-based algorithm 368–9
- blockchain technology 398
- block sequencing 514
- blood products 432
 - optimal inventory decisions for multi-location systems 436–41
 - optimal inventory decisions for single-location systems 432–6
 - optimal issuing decisions 441–3
- BOC *see* batch-order contract (BOC)
- Bollapragada, R. 64
- Bolton, G. E. 404–8, 424, 425
- Boute, R. N. 176
- Boyacı, T. 56
- bracket policy 9
- Bradley, J. R. 57, 186
- branching echelon base stock (BEBS) 61
- Briesch, R. A. 304n7
- Broder, J. 372, 374
- Brodheim, E. 440
- Bu, J. 173
- budget-balanced 282
- bullwhip effect 263
 - experimental examination of 408–11
- Burnetas, A. N. 337
- Buzacott, J. A. 397
- Cachon, G. P. 66, 67, 395, 401–4
- capacitated assembly systems 94
- capacitated dual-sourcing models 175–6
- capacitated inventory systems
 - applications and methods 65–8
 - competition 67–8
 - information 65–6
 - methodological 65
 - single-installation systems 47–57
 - system configurations 57–65
 - Vendor-Managed Inventory 66
- capacitated setting models 321–3
 - algorithms for 324–6
- capacity 46
 - effect of 48–9
 - interactions between information and 65–6
 - limited echelon 133–4
 - limits affect systems 46

- reactive and proactive uses of 55–6
- capacity investment 54–5
- capacity management, demand information
 - asymmetry in 281–3
- capital-dependent base-stock policy 397
- capital structure irrelevance principle 379
- capped base-stock (CBS) policy 12, 174
- Capped Dual-Index (CDI) policy
 - in continuous-time inventory models 181–5
 - in discrete-time inventory model 174–5
- Caro, F. 414, 493
- case packs 484–5
- cash holding cost rate 386
- cash retention decision 386
- Castañeda, J. A. 414
- CDI policy *see* Capped Dual-Index (CDI) policy
- censored demand 24, 333, 334, 347, 413
 - lost-sales system with 337
 - nonparametric learning with 367–9
 - nonparametric learning with lost sales and 367–9
 - periodic-review inventory system with 349–50
- censored demand data 329–30
- centralized multi-echelon inventory management 261
- Central Limit Theorem (CLT) 147–8, 150–52
 - asymmetric demand 154–6
 - symmetric demand 152–3
- Çetinkaya, S. 66
- chained BOMs 199
- Chan, E. W. 49
- Chan, L. 288
- Chang, H. S. 336
- Chao, X. 108, 216, 217, 235, 336, 337, 357–9, 370, 436
- Charikar, M. 310
- Chen, A. I. 208, 510
- Chen, B. 336–8, 357–9, 362, 363, 366–74
- Chen, F. 65, 102, 103, 106, 112, 113, 120, 133, 151, 502
- Chen, H. 292
- Chen, J. 23
- Chen, K. 43, 435
- Chen, K.-Y. 414
- Chen, L. 107, 263–4, 425
- Chen, M. 288, 303
- Chen, Q. 338
- Chen, S. 34–5, 126, 138, 140, 205, 208
- Chen, X. 173, 289, 292, 294–6, 299, 301–3, 372, 373, 396, 397, 434, 482
- Chen, Y. 298
- Chen, Z.-L. 288, 303
- Cheung, W. C. 313, 320–22, 324, 325, 328, 329, 369
- Chod, J. 209, 397
- Chou, M. C. 232
- Chu, Y. L. 235
- Ciarallo, F. W. 51
- clairvoyant optimal policy 355, 358
- Clark, A. 22, 389
- Clark, A. H. 73–5, 90, 93
- Clark, A. J. 58, 64, 66, 67, 102, 105, 106, 120, 125, 225, 262, 328, 493, 502
- Clark–Scarf decomposition 73, 74, 77
 - of objective cost function 74, 77, 80, 84, 97
- Clark–Scarf model 73–5, 108, 389
 - assembly systems 89–96
 - dimensions of 74
 - generalizations to multiple flows of product 75–89
 - open research problems 96–7
 - systems with expediting 75–8
 - systems with lateral flows of product 78–81
 - systems with reverse logistics 81–9
- classical periodic-review inventory control problem 270
- classic inventory model, with remanufacturing 214–16
- classic lost-sales inventory system 334
- classic T -period inventory management problem 147
- clearance pricing 493–4
- clearance sales 2, 31, 32, 37, 38
- CLT *see* Central Limit Theorem (CLT)
- clustering step, stage cost function 113–14
- Cohen, M. A. 201
- collaborative planning, forecasting, and replenishment (CPFR) 263
- collection period 384
- committed policies 128, 129
- common component inventory allocation 191
- component assembly system 64, 91–2
- component inventory replenishment 191
- consecutive lead times, discrete-time model 167–9
- consignment payment timing contracts 395
- constant-factor approximation algorithms 236
- constant-order (CO) policies 8–9
 - asymptotic optimality 9–11
 - combining order-up-to and 11–14
 - optimizing constant-order quantity 9
- consumer-behavior models 292–3, 303
 - reference price effect 293–6
 - strategic consumers 296–8
- continuous demand 348–9
- continuous relaxation 519
- continuous-review base-stock policy 110
- continuous-review models 23–4
- continuous-review system 100, 101
- continuous-time inventory models 176–8

- Capped Dual-Index policy 181–5
- Dual-Index policy 178–80
 - dual sourcing with capacitated sources 185–7
 - heuristic policies for 177
 - Tailored Base-Surge policy 180–81
- continuous time models 56–7
- contracts
 - in dynamic multi-period settings 267
 - menu of 268
- control policy 128
- conventional order-picking process 512
- conventional picker-to-parts systems 504
- conventional warehouses 503
- convexity 48
- coordinated base-stock (CBS) policy 204, 205
- corrective maintenance 455, 457
- cost balancing 238
- cost information asymmetry 277
 - inventory shortage cost information asymmetry 277–9
 - production cost information asymmetry 279–81
- costs 481–2
 - for emergency shipments 464
 - types of 459
- CPLEX 194
- Crama, Y. 441
- critical components 457
- Croson, R. 411–12, 415–17, 425
- cumulative distribution function (CDF) 307, 314
- cycle-based method 24
- cycle stock/working inventory 501
- Cycle-Update Policy 351–2
- cyclic (periodic) inventory systems 51
- cyclic skimming pricing strategy 295

- Dada, M. 482
- dam models 52, 53
- data and key performance indicators, retailing 477
 - demand characteristics 477–81
 - lead times, costs, and prices 481–2
 - service levels, on-shelf availability, inventory record inaccuracy, tracking, and tracing 482–3
- data-driven inventory and pricing models 303
- data-driven inventory control problem 336
- data-driven methods 263
- Data-Driven Multi-product Algorithm (DDM) 356, 357
- data-driven newsvendor, with unobservable lost sales 486–8
- data-driven optimization approach 484
- Davis, A. M. 414, 423, 424

- DDM *see* Data-Driven Multi-product Algorithm (DDM)
- decentralized supply chain 264
- decision behavior, addressing “behavioral problems” 415–18
- decision-maker (DM) 307, 333
- decision-support model 489
- decision support systems 473
- decomposable of degree 2 78
- decomposition method 120
 - heuristics 195–7
- decoupled policies 129, 132, 135
- DeCroix, G. A. 54, 91, 225–7
- de Kok, T. 73
- Delage, E. 147
- delivery costs 491
- demand characteristics
 - classification and segmentation 477–8
 - demand distribution 478
 - forecasting 478–9
 - stockout-based substitution 480–81
 - unobservable lost sales 479–80
- demand distribution 24
- demand information 263–4
- demand information asymmetry 267
 - in capacity management 281–3
 - sharing non-stationary demand information 270–73
 - sharing stationary demand information 267–70
- demand information sharing 264
- demand–price relationship 333
- demand uncertainty 42, 43
- Deniz, B. 435
- Deshpande, V. 57
- de Tejada Cuenca, A. S. 414
- deterministic heuristic 225
- deterministic lead times, serial systems 130–31
 - batch ordering constraint 132–3
 - costly expediting options 131–2
 - limited echelon capacity 133–4
- DeValve, L. 194, 196, 198, 200, 208
- de Véricourt, F. 57
- DeYong, G. D. 482
- differentiated remanufactured and new products 228
 - joint inventory and pricing optimization with 228–30
 - perishable inventory 230–31
- different unit–customer pairs 131
- Ding, X. 335
- direct inventory-order decisions 425–6
- discrete convex optimization, heuristics 199–200
- discrete demand 348
- discrete-time inventory models 165–7
 - Capped Dual-Index policy 174–5

- consecutive lead times 167–9
- dual sourcing with capacitated sources 175–6
- non-consecutive lead times 169
- single-index and dual-index policies 170–72
- single-sourcing lost-sales model 169–70
- Tailored Base-Surge policy 172–4
- distributionally robust optimization 163
- distribution centers (DCs) 505
- distribution-free 310
- distribution-free/black-box model 236
- distribution system, multi-echelon inventory models 119–20
- Doğru, M. K. 199, 203, 205
- Dong, L. 106
- Donohue, K. 411, 412
- doubling trick 361
- Downs, B. 5
- downstream product remanufacturing 227
- Drake, D. 413
- Drakopoulos, K. 265
- drugs
 - availability in developing countries 448–50
 - shortages in United States 444–8
- dual-balancing policy 16, 19, 237–8
- marginal cost accounting 238–9
- policy description 239
- worst-case analysis 239–42
- Dual-Index Dual-Base-Stock policies 170
- dual-index (S, M) policy 219
- Dual-Index (DI) policy
 - in continuous-time inventory models 178–80
 - in discrete-time inventory model 167, 170–72
- dual-mode dynamic stochastic inventory model
 - see* dual-sourcing dynamic stochastic inventory model
- dual-mode equipment-procurement framework 187
- dual sourcing
 - continuous-time model 185–7
 - discrete-time model 175–6
 - supply risk and 413
- dual-sourcing dynamic stochastic inventory model 165
 - continuous-time model 176–87
 - discrete-time model 165–76
 - research directions 187–8
- dual-sourcing problem 336
- Dual-Sourcing Smoothing (DSS) policy 176
- dual-sourcing system, tailored base-surge policy for 362–4
- dual supply model 471
- dynamic ATO models 200–202
 - asymptotically optimal policies 205
 - lost sales model 206–7
 - optimal policies for N- and W-systems 202–5
- dynamic balanced echelon base-stock policy 140
- dynamic contracts 262, 267
- dynamic inventory management problems
 - information asymmetry in 266
 - information in 262
 - in two-level supply chain setting 273
- dynamic learn-and-screen approach 270
- dynamic multi-echelon inventory systems 390
- dynamic pricing and assortments 493–4
- dynamic pricing models 121, 288
- dynamic programming formulation 237
 - multi-echelon serial systems 20–21
 - single stage inventory systems 3–4
- dynamic-programming framework 57
- dynamic robust inventory management 159–62
- dynamic sequencing 514
- dynamic short-term contracting problem 275
- dynamic stochastic inventory system 53
- Eben-Chaïme, M. 514
- echelon base-stock policy 58, 61, 62, 65, 66, 104, 131, 225, 389
- Echelon formulation 388–9
- echelon inventory positions 91
- echelon- j system 104, 105
- echelon net inventory level 388
- echelon OUT policy 21
 - asymptotic optimality 22–3
 - optimizing S 21–2
- echelon policy 104, 122n1
- echelon stocks 73
- e-commerce 44, 504, 506, 511
- e-commerce ecosystem 261–2
- economic order quantity (EOQ) model 413, 482
- economic surplus 293
- effective working capital 385
- efficient frontier 460
- efficient solutions 460
- electronic data interchange (EDI) information 261, 263
- ElHafsi, M. 206, 207
- Eliashberg, J. 288, 482
- Elimination-Based Half-Q-Learning algorithm 337
- Elmachtoub, A. N. 486
- Elmaghraby, W. 288, 493
- emergency shipments 455
 - application in other settings 468–70
 - approximate evaluation procedure 466–7
 - greedy heuristic 467–8
 - models for networks with 462–70
 - optimization 459–62
 - problem formulation 457–9
 - single-echelon, multi-location model 462–5
 - single-location, multi-item model with 457–62

- Empirical Risk Minimization (ERM) problem 318
- end-of-season clearance pricing 493–4
- endogenous prices, ATO systems with 209
- equilibrium policy 67
- Erenguc, S. S. 489
- e-retailers 500, 501, 504
 - shipping speed 504
- e-retailing 207–8
- Eroglu, C. 485
- Ettl, M. 104
- Evans, R. V. 54, 55, 57
- exact optimal policy 102–3
- exogenous lead times 136–8
- expected-profit function, of newsvendor problem 404–5
- expediting
 - into downstream installation 77–8
 - at installation 75–7
 - model with 107–9
 - systems with 75
- explore-first algorithm 344–6
- ex-post* inventory errors 403, 404
- extended echelon base-stock policies 132
- extensions, multi-echelon inventory models 117
 - assembly system 118–19
 - distribution system 119–20
 - local information 120–21
- factor-based model of uncertainty 147
- failure predictions 472–3
- failure-to-supply (FTS) penalty 444, 446
- Federgruen, A. 22, 51, 52, 67, 68, 99, 102, 121, 175, 288, 289, 292, 322, 328, 364, 440, 441, 493, 502
- Feiler, D. 421, 422
- Feng, Q. 282, 291, 292
- Fildes, R. 478
- financed demands 393
- financed inventory 393
- financial constraints, models with 380–81
 - series system 385–90
 - single-stage system 381–5
- financial flows 390–92
- financial holding cost 82
- financial inventory-related cost assessment 393–4
- financial markets
 - imperfect 384
 - perfect 382–4
- finished-goods inventory (FGI) 165–6
- finite-horizon Markov decision process (MDP) 336
- first-come, first-served (FCFS) allocation rule 200, 201, 203
- first-in-first-out (FIFO) rule 29, 30, 34, 42, 44, 299, 350, 433
- first-ordered, first-consumed basis 238
- Fisher, M. L. 56, 66, 484
- fixed costs 242
 - capacitated systems with 53–4
 - randomized cost-balancing policy 242–3
 - worst-case analysis 244–6
- fixed order costs 288, 291
 - backlog system with 372–4
 - series systems with 109–17
 - series systems without 100–109
- Fleischmann, M. 213
- “flow-through” stages 94
- forced backlogging cost 235
- forecast-adjusted base-stock policies 222
- forecasting 478–9
- forward distribution centers (FDCs) 501, 505–9
- Fox, B. 461
- fractional rounding, heuristics 197–9
- free-return system 101–2
- Friend, C. 151
- Fries, B. 432, 434, 442
- Fry, M. J. 66
- Fu, K. 35, 230, 231
- Fu, M. C. 53
- Fukuda, Y. 56, 175
- fulfillment centers (FCs) 500, 503, 505
- Full-Q-Learning algorithm 337
- gain/loss surplus 293
- Gallego, G. 54, 120, 263, 264
- Gallien, J. 450, 493
- Gao, L. 277, 279, 281
- Gao, X. 292
- Gavirneni, S. 66, 263
- general assembly systems 90, 138–40
- “generalized M-system” 206
- Gershwin, S. B. 56
- Ghobbar, A. 151
- Gijsbrechts, J. 175, 187, 337
- Gilbert, S. M. 288
- Glasserman, P. 59, 61, 62
- Global Fund 450
- Glynn, P. W. 57
- Godfrey, G. A. 336
- Goldberg, D. A. 7–10, 173, 191, 205, 236
- Goldschmidt, K. 413
- Golrezaei, N. 494
- Gong, X. 78, 81, 216–17, 292
- Gong, X.-Y. 334, 337
- Gorissen, B. 147
- gradient-based method
 - with increasing cycle lengths 353

- with stationary cycle lengths 353–4
- Graves, S. C. 104, 107, 208, 509, 510
- greedy heuristic 467–8
- Greenleaf, E. 293
- Grigas, P. 486
- grocery retailing 33, 42
- group purchasing organization (GPO) 431, 444, 445
- Gu, J. 513
- Gupta, D. 397, 448
- Gupta, S. M. 213
- Gurnani, H. 413
- Gurobi 194
- Ha, A. Y. 57
- Hadley, G. 134
- Haijema, R. 435
- Halman, N. 237
- Haran, U. 416
- Hardgrave, B. C. 265
- Harihara, R. 473
- healthcare inventory management 431–2
 - for blood products 432–43
 - for pharmaceutical products 444–50
- Heching, A. 51, 121, 288, 289, 292, 364
- Hedenstierna, C. P. T. 471
- Heilman, C. M. 43
- Heinen, J. J. 471
- Hertog, D. 147
- heterogeneous replenishment lead times 191
- heuristic policies 32, 35
 - reverse logistics 87–9
- heuristic procedures and policies 13
 - base-stock policies 13, 15
 - computational analysis 16–18
 - heuristic replenishment policies 15–16
- heuristics, one-period ATO models 194
 - decomposition 195–7
 - discrete convex optimization 199–200
 - fractional rounding 197–9
- Hill, R. 23
- Ho, T.-H. 414
- Hoberg, K. 471
- Hosseiniard, S. Z. 443
- Hotkar, P. 448
- Hu, Q. J. 397, 456
- Hu, X. 52, 56, 66
- Hu, Z. 295–8, 303
- Hua, Z. 171
- Huber, J. 486
- Huh, W. H. 333, 336, 338, 346, 349, 357
- Huh, W. T. 5, 8, 15, 20–22, 24, 61, 62, 65, 173, 291, 336, 353
- hybrid heuristic 225
- Hyndman, K. 423, 424
- Iancu, D. 148
- Iglehart, D. 335
- IID demand *see* independent and identically distributed (IID) demand
- Ilgin, M. A. 213
- imperfect financial markets 384
- incentive compatibility (IC) constraint 274, 275, 280, 284n11
- incentives, in inventory models 266–83
- increasing failure rate (IFR) 313
- incremental cost 41
- independent and identically distributed (IID) demand 101
- independent base-stock (IBS) replenishment policies 200
- Inderfurth, K. 216, 218
- individual rationality (IR) constraint 274, 280
- induced penalty cost function 113, 389
- infinite-horizon discounted problem 169
- infinite-horizon problem 102
- Infinitesimal Perturbation Analysis (IPA) 53, 62
- information, in inventory models 262–3
 - demand information 263–4
 - in dynamic inventory management problems 262
 - inventory information 264–6
- information asymmetry 266
 - cost information asymmetry 277–81
 - demand information asymmetry 267–73
 - demand information asymmetry in capacity management 281–3
 - in dynamic inventory management problems 266
 - inventory information asymmetry 273–7
 - sources of 266
- information design approach 283
- information set 237
- information sharing 263
- integrality gap 198
- inter-temporal demand 297, 298
- inventory control
 - with pricing decisions 327–8
 - on serial systems 328–9
- inventory control model (multiple period setting) 320
 - algorithms for capacitated setting 324–6
 - algorithms for uncapacitated setting 323–4
 - censored demand data 329–30
 - inventory control on serial systems 328–9
 - inventory control with pricing decisions 327–8
 - uncapacitated and capacitated settings models 321–3
- inventory-dependent demand 292
- inventory dynamic models 289, 298
- inventory information 264–6

- inventory information asymmetry 273–7
- inventory management 333
 - digitalization of 482
 - expediting of 75–8
 - learning in 335–6
 - problem 125
 - retailing problem complexity of 476–7
 - standard time-dependent formulation of 47
 - theory and practice of 261
 - tools 448
- inventory model formulation 355
- inventory models
 - equivalence between dam models and 52, 53
 - information in *see* information, in inventory models
- inventory optimization problem 334
- inventory placement 504–5
 - JD.com 505–9
 - in lateral fulfillment network 509–11
- inventory quantities 414
- inventory record inaccuracy 482–3
- inventory replenishment 386
- inventory routing problems 490–91
- inventory shortage cost information asymmetry 277–9
- inventory tracking technologies 265
- IPA *see* Infinitesimal Perturbation Analysis (IPA)
- Iyer, A. V. 56
- Jain, A. 480
- Janakiraman, G. 5, 7, 12, 20–22, 59, 61, 65, 68, 133, 134, 173, 291, 353
- Jasin, S. 208
- JD.com 500–502, 505–6
 - flexibility in fulfillment 507–9
 - order-picking optimization for RMFS 516–19
 - single-pair allocation 506–7
 - tree-shaped delivery network 504
- Jehl, T. 208, 506, 507, 509
- Jennings, J. B. 440
- Jia, J. 231, 444, 445, 447
- Jia, R. 24, 354
- Jiang, Y. 235
- Johansen, S. 12, 15, 16, 23, 24
- joint capacity limit, multiple items with 54, 55
- joint inventory 364–5
 - backlog system with fixed ordering cost 372–4
 - nonparametric learning with backlogged demand 366–7
 - nonparametric learning with lost sales and censored demand 367–9
 - online learning in 337
 - parametric learning with limited price changes 369–72
 - and pricing optimization 228–30
- joint inventory model 121
- joint replenishments 488–90
- Jordan, W. C. 509
- Kadiyala, B. 267, 268, 270, 284n7
- Kalkanci, B. 413
- Kamesam, P. V. 64
- Kaplan, A. 57
- Kaplan, R. S. 134
- Kaplan–Meier estimator 24, 333, 338
- Kapuściński, R. 56, 57, 59–62, 66, 67, 133, 134, 264, 321, 322
- Karabati, S. 480
- Karaesmen, I. Z. 27
- Karlin, S. 4, 5
- Karush, W. 460
- Katehakis, M. N. 337, 397
- Katircioglu, K. 57, 125
- Katok, E. 404–6, 412, 413, 424, 425
- Kendall, K. E. 440
- Kesavan, S. 426
- Keskin, N. B. 336, 337, 366
- Keskinocak, P. 288, 493
- Ketzenberg, M. 485
- Kiesmüller, G. P. 218, 219
- Kim, C. 77
- Kim, S. H. 447
- Kimemia, J. 56
- Kingman, J. 11
- Kiva system 515
- Kleywegt, A. J. 199
- Knofius, N. 471
- Kocabiykoglu, A. 291
- Kök, A. G. 484
- Koukia, C. 23
- Kouvelis, P. 380
- Kranenburg, B. 456, 457
- Kremer, M. 417
- Krishnan, H. 56
- Küçükgül, C. 266
- Kumar, P. R. 56
- k*-unit subproblem 132
- Kunnumkal, S. 336
- Kurtuluş, M. 263
- Kushwaha, T. 426
- Lago, A. 126, 142, 143
- Lagrangian relaxation 506–7
- Lambrecht, M. 53
- Lamghari-Idrissi, D. 474
- laminar set family 199
- laminar structure of tree families 199
- last-in-first-out (LIFO) rule 32, 36, 37, 42, 44, 440, 442, 443
- lateral flows of product, systems with 78–81

- lateral fulfillment network, inventory placement in 509–11
- lateral transshipments 455, 503
 - application in other settings 468–70
 - approximate evaluation procedure 466–7
 - greedy heuristic 467–8
 - models for networks with 462–70
 - single-echelon, multi-location model 462–5
- lattice-dependent ATO policy 206
- lattice-dependent base-stock policy 206
- lattice-dependent rationing policy 206
- Lau, A. H. L. 480
- Lau, H. S. 480
- Law of the Iterated Logarithm (LIL) 148
 - static robust inventory management 157–8
- Lawson, D. G. 74, 75, 77, 81, 108
- lead/lifetime, models with 298–302
- lead times 481–2
- learn-and-screen approach 268, 269
- learning
 - challenges in 353
 - in inventory management 335–6
 - optimal joint inventory and pricing decisions 364–74
 - (s, S) policy for lot-sizing problem 359–61
 - via Bayesian updating 335–6
- learning algorithm 358–9
 - sketch and regret results 360–61
 - TBS policy 363–4
- learning algorithm design
 - and analysis 356–7
 - challenges of 355
- learning-based newsvendor models 317–20
- learning optimal inventory decisions 348
 - learning (s, S) policy for lot-sizing problem 359–61
 - lost-sales model with positive lead time 352–5
 - lower-bound results 348–9
 - multiple products with substitution 357–9
 - multiple products with warehouse-capacity constraint 355–7
 - periodic-review inventory system with censored demand 349–50
 - perishable inventory systems 350–52
 - tailored base-surge policy for dual-sourcing system 362–4
- Lederer, P. J. 57
- Lee, C.-Y. 66
- Lee, H. 264, 265
- Lee, H. L. 104, 106, 263, 264, 411
- Lee, J. 448
- Lee, S. M. 440
- Lei, J. 163
- letter of credit (LC) 387
- Levi, R. 16, 19, 182, 235–7, 248, 309, 311–13, 315, 316, 320–25, 328–30, 333
- Li, L. 57, 397
- Li, M. 425
- Li, Q. 28–30, 32, 34, 36–8, 40–42, 171, 292, 434, 440, 441
- Li, S. 414
- Li, Y. 301, 302
- light-tailed demand distributions 62
- LIL *see* Law of the Iterated Logarithm (LIL)
- limit theorems of probability 147
- linear ordering cost 304n3
- linear program (LP-ERM) 318, 319
- Liu, F. 111, 282
- Liu, M. 336
- Liu, Q. 448
- Liu, S. 217
- Liu, X. 49
- Liu, Y. 30
- Liyanage, L. H. 314, 315
- Lobel, I. 267, 270–73, 280
- local information, multi-echelon inventory models 120–21
- logistics supply chains, reverse logistics in 81–4
- long-run balance condition 119
- long-term contracts 271, 272, 280
- long-term dynamic contract 267
- long-term partnerships 262
- loss aversion 293
- lost sales ATO model 206–7
- lost-sales inventory systems 2
 - multi-echelon serial systems 20–23
 - pointers to related lost-sales models 23–4
 - single stage 3–19
- lost-sales models 232
 - nonparametric learning with 367–9
 - with positive lead time 352–5
- lot-sizing problem, learning (s, S) policy for 359–61
- Lovejoy, W. S. 279, 335
- lower-bounding policy, reverse logistics 86–7
- lower-bound results 348–9
- lower-bound system, echelon- j system 105–9
- lower confidence bound (LCB) method 354–5
- Lu, L. 203–5
- Lu, Y. 194–6, 203, 291, 292
- Luo, W. 384, 385, 388–90
- Lurie, N. H. 404
- Luss, H. 54, 55
- Lutze, H. 277–9
- Ma, W. 208, 494
- MAB *see* multi-armed bandit problems (MAB)
- Mahajan, S. 483, 484

- make-to-order (MTO) model 229
- make-to-stock (MTS) model 229–30
- Malicki, S. 491
- Mamani, H. 147–9, 154, 159, 161, 162
- marginal backlogging cost accounting 250
- marginal cost accounting scheme 238–9
- marginal holding cost accounting approach 238
- Markov-modulated demand (MMD) 78, 106–7, 133
- Markov-modulated formulation 50
- Markov-modulated lead times 134–6
- Martínez-de-Albéniz, V. 126, 142–3
- Martingale Model of Asymmetric Forecast Evolutions (MMAFE) 282
- Martingale model of forecast evolution (MMFE) 106
- Massart inequality 311, 312
- maximum likelihood estimation (MLE) method 338, 340–42
- Maxwell, W. L. 99
- mean residual life (MRL) assumption 6–8
- MEBS policy *see* modified echelon base-stock (MEBS) policy
- memory factor 294
- Merriam, K. 509, 510
- Merschformann, M. 515
- Mersereau, A. J. 263, 264
- METRIC approach 456
- Miller, M. H. 379, 380
- minimization step, stage cost function 113, 114
- Minner, S. 478, 484, 486, 488, 490, 491
- misaligned incentives 261
- misreported private information 261
- mixed-integer-linear programming formulation 490
- mixed-integer linear program (MILP) solvers 194, 198
- MMAFE *see* Martingale Model of Asymmetric Forecast Evolutions (MMAFE)
- MMD *see* Markov-modulated demand (MMD)
- MM theorem 379–81
- modified base-stock policy (MBS) 175, 176, 321–2
- modified dual base-stock policy 175
- modified echelon base-stock (MEBS) policy 59–61, 63, 66
- modified FIFO (MFIFO) rule 203
- Modigliani, F. 379, 380
- Moinzadeh, K. 171, 178, 179, 185
- monotone/monotonic policies 59, 128, 133, 135
- Moritz, B. B. 412
- Morton, F. S. 447
- Morton, T. 4, 15, 16
- Muckstadt, J. A. 49, 59, 61, 68, 99, 133, 134, 456
- Muharremoglu, A. 107, 126, 129, 131–2, 134, 135, 137–8, 140, 141, 336, 348, 364, 390
- multi-armed bandit problems (MAB) 339
- multi-echelon inventory models 99, 121–2
 - configurations for 99, 100
 - extensions 117–21
 - series systems with fixed order costs 109–17
 - series systems without fixed order costs 100–109
- multi-echelon inventory problem
 - with reverse logistics 83
 - with stock disposals 80
- multi-echelon inventory systems 125, 144
- multi-echelon inventory theory 58, 73–4, 96
- multi-echelon serial systems 20, 501–2
 - dynamic programming formulation 20–21
 - echelon OUT policy 21–3
- multi-item fulfillment problem 207–8
- multi-level state-dependent rationing policy 206
- multi-location retailer inventory sharing 414
- multi-location systems, optimal inventory decisions for 436
 - optimal allocation 436–8
 - optimal transshipment and ordering 438–41
- multimodularity 29
- multimodular set 29
- multi-objective optimization problem 460, 461
- multiple products
 - with substitution 357–9
 - with warehouse-capacity constraint 355–7
- multiplicative demand 289
- multi-product inventory system 336
- multi-return/multi-echelon models 222, 397
 - downstream product remanufacturing 227
 - multiple types of returns 222–4
 - random yield 224–5
 - upstream product remanufacturing 226–7
- multi-sourcing inventory models 165
- multi-stage assembly system 64
- multi-stage echelon base-stock policies
 - exact optimal policy 102–3
 - service-level constrained model 104–5
 - single-stage bounds and heuristics 103–4
- multi-stage system
 - (r, q) policies 111–14
 - (s, T) policies 114–17
- multi-state dynamic program 384
- Murray, G. 335
- Myles, J. 208
- myopic policy 15, 16, 366
- myopic transshipment policy 438–9
- Nadar, E. 206–7
- Nadaraya–Watson kernel regression 320
- Nahmias, S. 27, 54, 247, 432, 434

- Narayanan, A. 412
- Nasiry, J. 295, 304n7
- Natarajan, K. V. 163, 450
- nested marginal cost accounting scheme 248–50
- nested marginal holding cost accounting 249
- nested marginal outdating cost accounting 249–50
- net working capital level 388
- network revenue management (NRM) problems 207
- newsvendor decomposition 195
- newsvendor model (single-period setting) 307–8
 - learning-based newsvendor models 317–20
 - sampling-based newsvendor, beyond SAA 314–16
 - sampling-based newsvendor via sample average approximation 308–14
- newsvendor networks 208
 - and beer games 425
- newsvendor problem 401–2
 - external validity of “pull-to-center” effect 406–8
 - identifying PTC effect 402–4
 - robustness of “pull-to-center” effect 404–6
- no-holdback (NHB) allocation rule 203–5
- non-consecutive lead times, discrete-time model 169
- non-identically distributed random variables 151
- non-identical manufacturing lead times 218–20
- non-identical remanufacturing lead times 218–20
- non-linear regression-based approach 320
- non-parametric approach 480
- non-parametric heuristic policies 16
- nonparametric learning
 - with backlogged demand 366–7
 - with lost sales and censored demand 367–9
- non-perishable inventory systems 36, 39
- nonstationary assembly systems 91–6
- nonstationary demand
 - independent but nonidentically distributed demand 105–6
 - Markov-modulated demand 106–7
 - time-series demand model 106
- objective cost function
 - Clark–Scarf decomposition of 74, 77, 80, 97
 - for inventory problem with expediting 78
 - multi-dimensional 73
 - for multi-echelon inventory problem 80
- Ockenfels, A. 425
- offline learning 333
 - online learning *versus* 333–4
- Oh, S. 209, 282
- one-dimensional dynamic program 35
- one-period ATO models 192–4
 - computational challenges for 194
 - discussions on 200
 - heuristics 194–200
- one-period cost function 39, 76, 82, 85, 92
- one-warehouse-multi-retailer (OWMR) system 235–6, 491–3
- on-hand inventory 238, 264–6
- online demand learning, common approaches to 338–9
- online learning
 - in joint inventory and pricing optimization 337
 - MLE for 340–42
 - offline learning *versus* 333–4
 - in pure inventory management 336–7
- online learning algorithms, evaluation of 337–8
- online optimization 147
- online resource allocation, with component commonality 207
- online retailing inventory management 500–501, 510–20
 - inventory placement 504–11
 - order picking 511–19
 - replenishment 501–2
 - transshipment 502–3
 - warehouse operations 503–4
- on-shelf availability 482–3
- open research problems, Clark–Scarf model 96–7
- operational decisions 485–6
 - data-driven newsvendor with unobservable lost sales 486–8
 - dynamic pricing and assortments 493–4
 - inventory routing problems 490–91
 - joint replenishments 488–90
 - one-warehouse multi-retailer systems 491–3
- operations management literature 46
- Operations Management textbooks 44
- optimal allocation 436–8
- optimal allocation policy 34, 35
- optimal base-stock policy 336, 352
- optimal base-stock target 52–3
- optimal clearance sale policies 31
- optimal clearance sales strategy 34
- optimal contract 270–78
- optimal dynamic contracts 280
- optimal inventory decisions
 - for multi-location systems 436–41
 - for single-location systems 432–6
- optimal inventory management 333
- optimal inventory policy, for ATO systems 191
- optimal issuing decisions 441–3
- optimality equation, defined as 39
- optimal joint pricing 296
- optimal long-term contract 273
- optimal ordering quantity 433–5
- optimal order quantity 32

- optimal policies 14, 15, 41, 54, 55, 101
 - for equivalent component system 94
 - for heuristic policies 16–18
 - for N- and W-systems, dynamic ATO models 202–5
 - serial system, single-unit analysis 129–30
 - series system 389–90
 - structure 4, 42, 52
- optimal reorder intervals 114, 116, 117, 122n1
- optimal tailored vase-surge policy 336
- optimal top-down echelon base-stock policy 108
- optimal transshipment 438–41
- optimal value function, structural properties of 35
- optimization 310
 - emergency shipments 459–62
- Ord, K. 479
- order assignment 515
- order batching 513, 514
- ordering 438–41
- order picking process 503, 511–12
 - in AS/RS 513–14
 - automated warehouses 512–13
 - in RMFS 514–19
- order split 503
- order-up-to (OUT) policies 2, 4–5
 - asymptotic optimality 5–6
 - combining constant-order and 11–14
 - MRL assumption 6–8
 - optimizing S 5
- Original Equipment Manufacturer (OEM) 455, 468, 469
- Oroojlooyjadid, A. 320, 336, 412
- outsourcing, single-installation systems 49–51
- outstanding orders 238
- Ovchinnikov, A. 420
- overflow stream 466–7
- overshoot 362
- Özbay, N. 147
- Özer, Ö. 54, 56, 66, 74, 83, 89, 95, 96, 263–5, 277–9, 282
- packaging, types of 40
- Pang, Z. 301
- parametric learning, with limited price changes 369–72
- Parker, R. P. 59–62, 66, 67, 133, 134, 264
- partial consignment payment timing contracts 395–6
- Pasternack, B. A. 67
- Paterson, C. 503
- payment, triggers and processes 392–3
- payment period 384
- payment scheme 393
- payment times, models with 390
- financial inventory-related cost assessment 393–4
- payment triggers and processes 392–3
- physical and financial flows 390–92
- supply chains through payment timing contracts 394–6
- 3PLs role in supply-chain finance 396–7
- payment timing contracts, supply chains through 394–6
- PBC rule *see* priority-based backorder clearing (PBC) rule
- PBMs *see* pharmacy benefits managers (PBMs)
- PB policy *see* proportional-balancing (PB) policy
- Peng, C. 187
- Pentico, D. W. 483
- Perakis, G. 163
- perfect Bayesian equilibrium 275
- perfect financial markets 382–4
- periodic maintenance 472
- periodic-review inventory system, with censored demand 349–50
- periodic-review models 201
- periodic review policy 501
- periodic-review priority (PRP) allocation rule 205
- periodic-review system 100, 101, 110
- perishable inventory systems 27, 230–31, 235, 350
 - control of lifetimes 40–43
 - empirical research 43–4
 - learning algorithm design 351
 - model formulation 350–51
 - models with one class of demand and location 27–32
 - with m shelf life 246–56
 - multiple classes of demand 33–5
 - multiple locations 36–40
 - nested marginal cost accounting scheme 248–50
 - proportional-balancing policy 250–51
 - regret analysis 351–2
 - worst-case analysis 251–6
- persistent *versus* non-persistent 267
- personal protective equipment (PPE) 451
- Petruzzi, N. C. 482
- pharmaceutical products, healthcare inventory management 444
 - drug availability in developing countries 448–50
 - drug shortages in United States 444–8
- pharmacy benefits managers (PBMs) 431
- phased exploration-and-exploitation algorithm 336
- physical flows 390–92
- physical holding cost 82

- Pierskalla, W. 441
 pipeline errors 63–4
 Plambeck, E. L. 205, 209
 Pliskin, N. 514
 point-of-sale (POS) data 261, 268, 480
 pooling effect 470
 Popescu, I. 291, 295, 304n7
 Porteus, E. L. 74, 75, 77, 81, 91, 93, 94, 96, 108
 positive lead time, lost-sales model with 352–5
 POUT-TBS policy 176
 Powell, W. B. 336
 Prabhu, N. U. 52
 Prastacos, G. P. 27, 436–8, 440, 441
 predictable irrationality 418–20
 predictive maintenance 472, 473
 preventive maintenance 455
 price changes, parametric learning with 369–72
 prices 481–2
 price-setting 414
 pricing decisions 364–5
 backlog system with fixed ordering cost 372–4
 inventory control with 327–8
 nonparametric learning with backlogged demand 366–7
 nonparametric learning with lost sales and censored demand 367–9
 online learning in 337
 parametric learning with limited price changes 369–72
 single-installation systems 51
 pricing optimization, online learning in 337
 pricing policy 121
 Prince, J. 379
 priority-based backorder clearing (PBC) rule 205
 proactive transshipment 503
 problematic periods 244
 problem formulation, emergency shipments 457–9
 product demands
 models with one class of 27–32
 multiple classes of 33–5
 production cost information asymmetry 279–81
 production–inventory problem 186
 product-level forecasting 478
 product locations
 models with one class of 27–32
 multiple 36–40
 product-transforming supply chains 73
 reverse logistics in 84–7
 projected inventory level policy 19
 promised lead-time contract 277–8
 proportional-balancing (PB) policy 250–51
 proportional order-up-to (POUT) policy 176
 Protopappa-Sieke, M. 397
 PRP allocation rule *see* periodic-review priority (PRP) allocation rule
 pseudo cost, defined as 359
 pull-to-center (PTC) effect
 external validity of 406–8
 identification of 402–4
 robustness of 404–6
 pure inventory management, online learning in 336–7
 “push-ahead” effect 58, 59
 Qin, H. 303, 320, 327, 328, 518
 quantity-based models 298
 quantity-based pricing problem 292
 quasiconvexity 48
 queueing theory 56
 Radio Frequency Identification (RFID) 265
 Rajagopalan, S. 55
 Ramachandran, K. 303, 414
 Raman, A. 56, 264
 random customer demands 78, 81
 randomized cost-balancing policy 242–3
 random utility model 303n1
 random yield, multi-echelon models 224–5
 rationality, documenting deviations from 425
 reactive capacity 56
 REDO 416–17
 reference price effect 293–6
 regional distribution centers (RDCs) 501, 505–9
 regret 337–8
 regular packaging 40–42
 Reijnen, I. C. 466, 467
 Reiman, M. I. 9, 23, 177, 178, 185, 205
 reinforcement learning 336
 remanufacturing 213
 classic inventory model with 214–16
 Ren, Y. 415–17, 425
 Ren, Z. J. 281, 282
 Renewal Reward Theorem 360
 reorder intervals
 bounds for base-stock levels with fixed 117
 simple solution for 117
 repair lead time 457
 replacement of old inventories (ROI) system 352
 replenishment, online retailing 501–2
 replenishment patterns 484
 reservation-price model 289, 292, 293, 304n4
 responsive pricing 209
 restricted base-stock policy 16
 retail inventory systems 476–7
 data and key performance indicators 477–83
 operational decisions 485–94
 strategic and tactical decisions 483–5

- returned product acquisition and pricing 217–18
- revelation principle 262, 268, 271, 274, 278
- revenue management 414
- reverse logistics, systems with 81
 - heuristic policy 87–9
 - in logistics supply chains 81–4
 - lower-bounding policy 86–7
 - in product-transforming supply chains 84–6
- reverse-order schedule 81, 82
- RFID *see* Radio Frequency Identification (RFID)
- RMFS *see* robotic mobile fulfillment system (RMFS)
- Roach, C. 441
- Robinson, L. W. 134
- robotic mobile fulfillment system (RMFS) 504, 520
 - order picking in 514–19
- robust inventory management 147, 162–3
 - dynamic 159–62
 - literature review 147–8
 - static 148–59
- Roels, G. 163
- Rong, Y. 120
- Roodbergen, K. J. 512, 513
- Rosling, K. 13, 64, 66, 75, 90, 91, 96, 118, 120, 126, 138, 140, 205
- rounding schemes 197–9
- Roundy, R. O. 5, 99, 236, 353
- routing 515
- Rudi, N. 193, 208, 413
- Rudin, C. 317–20, 486
- rule-based solutions 515
- rule configurations (RCs) 515
- Rusmevichientong, P. 24, 336, 346, 349, 357, 372, 374
- Russo, D. 342
- Sachs, A.-L. 486
- safety stock 501
- sample-average-approximation (SAA) approach 24, 199, 324, 338, 343–4
 - sampling-based newsvendor beyond 314–16
 - sampling-based newsvendor via 308–14
- sample-based approximation algorithms, for series systems 122
- sampling-based inventory control model 320, 321
- sampling-based newsvendor
 - beyond SAA 314–16
 - via sample average approximation 308–14
- sampling-based newsvendor problem 310
- sampling-based serial system problem 329
- sampling-based setting 323
- Sarhangian, V. 443
- Sauré, D. 483
- Scarf, H. 22, 58, 64, 66, 67, 73–5, 90, 93, 102, 105, 106, 120, 125, 163, 225, 262, 307, 328, 330, 335, 493, 502
- Scarf, K. 389
- Schäl, M. 173
- Scheller-Wolf, A. 54, 171, 194, 196–8, 208
- Schmidt, C. P. 54, 171, 178, 179, 185
- Schweitzer, M. E. 401–4
- secondary market sales 78
 - inventory states, decisions, and product flows in 79
- second-stage allocation-decision problem 192
- second-stage allocation problem 199
- See, C. 147, 162
- Seifert, R. W. 397
- Selten, R. 425
- sequencing 514, 515
- serial supply chain 408
 - behavioral bullwhip moderating factors and mitigation strategies 412
 - experimental examination of bullwhip effect 408–11
 - isolating “behavioral bullwhip effect” 411–12
- serial systems 58
 - inventory control on 328–9
 - with Markov-modulated demand 126
 - optimality results 58–61
 - single-unit analysis 126
 - extension to general lead times 130
 - optimal policies 129–30
 - policy classification and problem decomposition 128–9
 - problem description 126–8
 - stability and computational time 61–4
- serial systems management, applications in 130
 - with deterministic lead times and additional considerations 130–34
 - with stochastic lead times 134–8
- series systems 385–8
 - Echelon formulation 388–9
 - with fixed order costs 109–10
 - multi-stage (r, q) policies 111–14
 - multi-stage system, (s, T) policies 114–17
 - single-stage (r, q) and (s, T) policies 110–11
 - without fixed order costs 100
 - model with expediting 107–9
 - multi-stage echelon base-stock policies 102–5
 - nonstationary demand 105–7
 - single-stage base-stock policy 100–102
 - optimal policy 389–90
- service-level constrained model 104–5
- service levels 482–3
- SGD method *see* stochastic gradient descent (SGD) method

- shadow dynamic program 323
- Shalev-Shwartz, S. 331
- Shang, K. H. 103, 105–7, 113, 116–17, 120, 121, 122n1, 384, 385, 388–90, 397, 502
- Shanthikumar, J. G. 292, 314, 315
- Shaoxiang, C. 53, 54
- sharing non-stationary demand information 270–73
- sharing stationary demand information 267–70
- Shen, X. 77, 78, 292
- Shen, Z. M. 235
- Sheopuri, A. 2, 169, 170
- Sherbrooke, C. C. 456, 457
- Shi, C. 235, 298, 336, 355–7, 362, 363
- Shmoys, D. B. 310
- shortage cost 277, 279
- shortfall 52
- short-term contracts 273, 275
- short-term dynamic contract 267
- Sierra Trading Post 265
- Silver, E. 335, 490
- Sim, M. 147, 162
- Simchi-Levi, D. 289, 291, 292, 303, 313, 320–22, 324, 325, 328, 329, 331, 334, 337, 372, 373, 482, 494, 520
- Simpson, V. 213–15, 217, 218, 220, 222, 224, 227
- single-echelon, multi-location model 462–5
- single-echelon continuous-review system 143
- single-echelon replenishment problems 501
- single-echelon stochastic inventory problems 491
- Single-Index Dual-Base-Stock policy 167–70, 187
- single-index (S, M) policy 219
- single-index (SI) policy
 - in discrete-time model 170–72
- single-installation systems 47
 - applications 49–52
 - base model 47–8
 - capacitated systems with fixed cost 53–4
 - capacity investment 54–5
 - continuous time models 56–7
 - effect of capacity 48–9
 - multiple items with joint capacity limit 54, 55
 - optimal base-stock target 52–3
 - outsourcing 49–51
 - pricing 51
 - reactive and proactive uses of capacity 55–6
 - uncertain capacity 51–2
- single-location blood inventory systems 440
- single-location systems, optimal inventory decisions for 432–6
- single-pair allocation, JD.com 506–7
- single-product assemble-to-order (ATO) systems 126, 140–42
- single-product joint pricing 289
- single-sourcing lost-sales model 169–70, 185
- single-stage, single-return inventory models 214
 - capacity constraints 216–17
 - classic inventory model with remanufacturing 214–16
 - dependent demands and returns 220–22
 - non-identical manufacturing and remanufacturing lead times 218–20
 - returned product acquisition and pricing 217–18
- single-stage-based heuristics, for multi-stage inventory systems 99
- single-stage base-stock policy 100–102
- single-stage bounds and heuristics 103–4
- single stage inventory systems
 - combining order-up-to and constant-order policies 11–14
 - constant-order policies 8–11
 - dynamic programming formulation 3–4
 - general demand models and balancing policies 13, 19
 - heuristic procedures and policies 13, 15–18
 - optimal policy structure 4
 - order-up-to policies 4–8
 - projected inventory level policy 19
- single-stage system 381–2
 - (r, q) and (s, T) policies 110–11
 - imperfect financial markets 384
 - perfect financial markets 382–4
 - trade credit 384–5
- single-unit analysis 125–6, 144–5
 - applications in managing assembly systems 138–42
 - applications in managing serial systems 130–38
 - related models 142–4
 - toy example on serial system 126–30
- single unit–customer pair 131
- single-unit single-customer subproblem 136, 139, 140
- Sinha, A. 208
- Sinha, S. 148, 157, 162, 163
- SLLN *see* Strong Law of Large Numbers (SLLN)
- Smith, C. E. 337
- Smith, S. A. 478, 483
- Sobel, M. J. 358, 366, 397
- Solyalı, O. 147, 148, 162
- Song, J.-S. 55, 103, 105–7, 111, 125, 134, 178, 179, 185–7, 191, 193–6, 201, 203, 302, 471, 502
- sort-after-picking 512
- sorting process 512
- sort-while-picking 512
- Souza, G. C. 213

- spare parts inventory planning 455–7, 473–4
 - emergency shipments *see* emergency shipments
 - exploiting failure predictions 472–3
 - lateral shipments *see* lateral transshipments
 - 3D printing 470–72
- Speck, C. J. 58
- SPIES 416
- spline approximation-based algorithm 368
- standard dynamic programming approach 238
- standard inventory control approach 143
- standard payment timing contracts 395
- standard theory 400
- standard vector base-stock policy 15
- Stangl, T. 413, 426
- state-dependent base-stock policies 206, 237
- state-dependent echelon base-stock policies 134, 136
- state-space-based approach 480
- static contract 267
 - revelation principle in 271
- static robust inventory management 148–50
 - Central Limit Theorem 150–56
 - choice of uncertainty set 158–9
 - Law of the Iterated Logarithm 157–8
 - Strong Law of Large Numbers 156–7
- stationary assembly systems 90–91
- stationary echelon base-stock policy 329
- statistical method 338
- Steinberg, R. 288, 482
- Sterman, J. D. 408–11
- stochastic concavity in midpoint 291
- stochastic gradient descent (SGD) method 338–9, 346
- stochastic inventory models 335
- stochastic inventory systems
 - approximation algorithms for *see* approximation algorithms
 - approximation results for 236
 - dual-balancing policy 237–42
 - with fixed costs 242–6
 - model formulation 237
- stochastic lead times, serial systems 134
 - exogenous lead times 136–8
 - Markov-modulated lead times 134–6
- stochastic linearity in midpoint 292
- Stock-Keeping Units (SKUs) 457–9
- stockout-based substitution 480–81
- store-keeping units (SKUs) 500, 503–5, 511
- strategic and tactical decisions 483
 - assortment planning 483–4
 - case packs 484–5
 - replenishment patterns 484
- strategic consumers 296–8
- Strong Law of Large Numbers (SLLN) 148, 309
 - static robust inventory management 156–7
- Su, X. 424
- subsidies 449–50
- Sun, J. 174
- supply-chain efficiency 424
- supply-chain finance, 3PLs role in 396–7
- supply-chain management 379
 - inventory in 264
- supply chain networks, complexity of 263
- supply chains
 - matching supply and demand in 288
 - through payment timing contracts 394–6
- supply-line under-weighting 409
- supply state information 279
- Svoboda, J. 478
- Swaminathan, J. M. 404, 450
- Swamy, C. 310
- symmetric demand, CLT 152–3
- system-optimal inventory-control policies 99
- system-oriented service measure 456
- tailored base-surge (TBS) policy
 - in continuous-time inventory models 180–81
 - in discrete-time inventory model 172–4
 - for dual-sourcing system 362–4
- Talluri, K. 290
- Tao, Z. 219–22, 224, 225, 235
- target inventory level 349
- Taube, F. 484, 488
- Taylor, T. A. 444, 448–50
- Tayur, S. R. 52, 59, 61–2, 64, 321–2
- TBS policy *see* Tailored Base-Surge (TBS) policy
- temporary replacement 471
- Teulings, M. 485
- Thaler, R. 419
- Thiele, A. 147, 162, 163
- third-party logistics providers (3PLs) 396–7
- Thompson sampling (TS) method 339, 342
- Thonemann, U. W. 413, 426
- Thorstenson, A. 23
- 3-approximation algorithm 235, 244
- three-component assembly system 139
- 3D printing 470–72
- 3D printing capacity 471
- time-locked sales 266
- time-series demand model 106
- Toktay, B. L. 56
- Tong, J. 390, 395, 416, 417, 421, 422
- Topaloglu, H. 336
- Topan, E. 473
- top-down echelon base-stock policy 77, 81
- Topkis, D. M. 57

- tournament-based algorithm 337
- tracing 482–3
- tracking 482–3
- trade credit 384–5
- traditional brick-and-mortar retailers 500
- transshipment 36–40, 42
 - online retailing 502–3
- trimmed on-hand inventory level 235, 251, 252
- Truong, V.-A. 235, 236
- Tsiros, M. 43
- Tsitsiklis, J. N. 107, 126, 129, 131–2, 134, 135, 137, 140, 390
- TS method *see* Thompson sampling (TS) method
- Tu, Y. 49
- Turan, B. 491
- 2-approximation algorithm 235, 236
- 211 program 505, 506
- two-stage data-driven mixed-integer linear programming approach 488–9
- two-stage serial system 67
- two-stage stochastic programming (SP) formulation 192
- “two-tier base-stock” policy 59, 134
- type-*k* cores 224

- UCB-based algorithms 337
- UCB method *see* upper confidence bound (UCB) method
- uncapacitated setting models 321–3
 - algorithms for 323–4
- uncensored demand observation 270
- uncertain capacity, single-installation systems 51–2
- unit–customer subsystem 129
- United Parcel Service (UPS) survey 500
- United States
 - drug shortages in 444–8
 - spare parts sales and services in 455–6
- unit holding cost 82
- unit-load AS/RS 512
- unit-matching approach 251
- unit regular-ordered 76, 79, 82
- unknown stochastic distribution 163
- unobservable lost sales 413, 479–80
 - data-driven newsvendor with 486–8
- upper-bound system, echelon-*j* system 105–9
- upper confidence bound (UCB) method 339, 346–8
 - advanced line search method with 354–5
- upstream product remanufacturing 226–7

- value function 269
- value of information 262
- value-to-go function 32

- Van Aspert, M. 468
- Van der Vlist, P. 485
- van der Wal, J. 58
- Van Donselaar, K. H. 477
- van Houtum, G. J. 456, 457
- van Jaarsveld, W. 19, 194, 196–8, 208
- Van Mieghem, J. A. 172, 174, 176, 187, 193, 208
- Van Roy, B. 342
- van Ryzin, G. 290, 483, 484
- Van Zelst, S. 485
- Veatch, M. H. 56, 60
- vector base-stock policy 171
- Veeraraghavan, S. 171
- vehicle-routing problem (VRP) 513
- velocity-based rule 515–16
- vendor-managed inventory (VMI) 261, 263, 267–8
 - agreement 268–70
 - frameworks 270
- Villa, S. 414
- virtual valuation 304n4
- Vis, I. F. 23, 512, 513

- Wagner, M. 147, 148, 160, 163
- Walmart 288
- Wan, H. 205
- Wang, L. 397
- Wang, Q. 205
- Wang, T. 56, 78, 81
- Ward, A. R. 205, 209
- warehouse-capacity constraint, multiple products with 355–7
- warehouse operations, online retailing 503–4
- weakly reverse exponential (WRE) 276–7, 284n12
- Wei, W. 66
- weighted average cost of capital (WACC) 383
- weighted mean spread 315
- Wein, L. M. 56, 60
- Wensing, T. 485
- Westerweel, B. 470, 471
- Whang, S. 264
- Whitten, T. M. 134
- Wijngaard, J. 53
- Willems, S. P. 104
- Williams, T. 151
- Wong, H. 468
- worst-case analysis
 - dual-balancing policy 239–42
 - perishable inventory systems 251–6
 - stochastic inventory systems with fixed costs 244–6
- worst-case performance analysis 235
- worst-case performance guarantee 234, 251

- WRE *see* weakly reverse exponential (WRE)
- Wu, D. Y. 412
- Wu, Q. 12
- Wu, S. 298
- Wu, Y. 295
- Xiao, W. 267, 270–73, 280, 444, 448–50
- Xin, L. 8–10, 12, 173, 174, 177, 178, 180–82, 184, 185, 218, 219, 236
- Xu, S. H. 194, 199
- Xue, Z. 302
- Yan, X. 228–30
- Yang, N. 126, 134, 137, 138, 292
- Yano, C. A. 288
- Yu, B. 151
- Yu, M. 56
- Yu, P. 28–30, 34, 171, 434
- Yu, Y. 143, 217, 218
- Yuan, H. 336, 359–61
- Yuan, R. 515
- Yurukoglu, A. 448
- Zalkind, D. 134
- Zeevi, A. 366, 483
- Zenios, S. 271
- Zhang, C. 38–40, 436, 438–40, 442, 451
- Zhang, H. 24, 235, 271, 273, 336, 348–9, 351–4, 364
- Zhang, K. 122, 313, 316, 320, 328, 329
- Zhang, R. 292
- Zhang, R. Q. 397
- Zhang, X. 448
- Zhang, Y. 187, 471
- Zhang, Z. 292
- Zhao, H. 57, 400, 414, 444, 445, 447
- Zhao, Y. 203, 519
- Zheng, S. 292
- Zheng, Y.-S. 99, 102, 103, 112, 113
- Zhou, D. 42, 43, 435
- Zhou, S. X. 108, 116, 117, 122n1, 217–25, 235
- Zhu, W. 64, 94, 95
- Zipkin, P. H. 4, 8, 15, 22, 50, 52, 67, 68, 91, 102, 107, 114, 118, 119, 125, 134, 176, 178, 179, 185, 186, 191, 193, 199, 201, 328, 352, 395, 473, 493, 502

