Technische Universität München

Essays on Supply Chain Inventories under Uncertainty

Dipl.-Math. oec. Univ. Christian Bohner

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Rainer Kolisch
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Stefan Minner
	2.	Prof. Stephen C. Graves, Ph.D. Massachusetts Institute of Technology Cambridge, MA 02139, USA

Die Dissertation wurde am 28. April 2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15. Juni 2017 angenommen.

To my parents

Acknowledgments

First and foremost, I am very grateful to my supervisor Professor Stefan Minner for his continuous and patient support. His vast knowledge of the state-of-the-art, his deep understanding of methods and problems, and his immense pool of purposeful ideas were very inspiring and motivated me to try to live up to his expectations. Joint research projects, cooperations with the industry, teaching master courses and the supervision of master theses taught me skills that go beyond scientific work. I am also very thankful to Professor Stephen Graves for inviting me to visit MIT from April to August 2016, for coming all the way to Munich to be part of the examination committee, and for our joint work on Chapter 4 of this thesis. I benefited greatly from weekly discussions and learned a lot from his sharp analysis of problems. I also want to thank Professor Rainer Kolisch for being the chairman of the examination committee and for initially arousing my interest in the field by supervising my diploma thesis.

I want to thank all current and former colleagues at the chair of Logistics and Supply Chain Management at TUM for numerous discussions, for their advice, and for joyful lunch breaks and evenings: Yuka Akasaka, Szymon Albinski, Dr. Maximilian Budde, Miray Közen, Sebastian Malicki, Christian Mandl, Thitinan Pholsook, Jun.-Prof. Dr. Anna-Lena Sachs, Dr. Martin Stößlein, Florian Taube and Dariush Tavaghof-Gigloo. My colleague Dr. Pirmin Fontaine deserves special thanks for sharing an office and thoughts on life in general. I also want to thank Evelyn Gemkow for her support, for improving my knowledge about classical music, and for the proofreading of my work. I am also thankful to the German Academic Exchange Service for funding my research visit in Boston and a conference visit in Philadelphia.

Finally, I would like to thank my parents and my close friends for their constant support, patience, critical comments and for providing distraction when I needed it. Very special thanks go to Christine for her patience and support, and for making me smile over and over again.

Abstract

This thesis addresses three problems in the field of inventory management. In all of these problems, we attempt to improve the modeling and managing of real-world uncertainty.

The first problem considers inventory systems that face time-dependent demand. We model demand as non-homogeneous Poisson processes and apply unit-tracking. For single-stage systems, we find a decision rule that yields the optimal order policy. For one-warehouse multi-retailer systems, we characterize the demand distribution at the warehouse, which is no longer Poisson. We determine time-dependent order policies that outperform the repeated application of time-*in*dependent policies significantly. Lastly, we present a simple approximation based on a time-shift.

The second problem focuses on a buyer of multiple products who needs to select suppliers and allocate orders. The suppliers offer quantity and business volume discounts, and they can fail. The buyer needs to find a balance between the benefits of exploiting the discounts and the risk of supplier failure. We present a new mixed-integer linear programming formulation that determines the optimal trade-off between economies of scale and failure risk.

The third problem considers the final order of a product at the end-of-production. The final order needs to satisfy future demand and can be complemented by future product returns. The decision maker determines the size of the final order, the remanufacturing and the disposal of returns. Forecasts for demands and returns are frequently updated. We prove the structure of the optimal policy. By stochastic dynamic programming, we find that forecast evolution has interesting effects on the size of the final order, the remanufacturing and disposal policies. Further, forecast evolution yields significant cost savings; the largest part of these savings is caused by updates of the demand forecasts.

Contents

List of Tables

List of Figures

Acronyms

cdf	cumulative distribution function.
EOP	End-Of-Production.
ETP	Expected Total Penalty costs.
FCFS	First-Come First-Served.
HL	Half Lead time.
MILP	Mixed-Integer Linear Program.
MMFE	Martingale Model of Forecast Evolution.
PC	Purchasing Costs.
pdf	probability density function.
RCLL	Right-Continuous with Left Limits.
SMC	Supplier Management Costs.
TD	Time Decomposition.
TEC	Total Expected Costs.
TPP	Transportation Problem.
UT	Unit-Tracking.

Chapter 1

Introduction

1.1 Motivation

The World Economic Forum's Global Information Technology Report 2016 (Baller et al., 2016) states that, across all industries, the average length of a product lifecycle fell by 24% between 1997 and 2012. This development is mainly due to the rapid technological progress and the customers' fast adaption to new technology. Large competition forces companies to release new versions of their products in shorter intervals. For electronic products, lifecycles of six to twelve months are common (Graves and Willems, 2008). These lifecycles pose new challenges for supply chain management. For instance, they call for a dynamic management of inventories to consider both varying demand structures and supply availabilities.

Today's inventory management is characterized by both new challenges and new opportunities. Besides short product lifecycles, a main challenge is the selection of suppliers, as globalization offers many choices. Suppliers can fail due to external factors such as earthquakes, or internal factors, e.g., bankruptcy. That is, one might choose to have multiple suppliers for the same product, which increases the complexity of the central questions of inventory management, i.e. (a) when to order, (b) how much, and (c) from whom. The age of *big data*, however, offers new potential for inventory management. Technology enables us to track inventories or predict failures of critical parts. A wise use of the available data allows for better forecasts and a reduction of the uncertainty that is tied to our decisions.

Within the field of inventory management, the management of spare parts is very prominent. Most *aftersales services* build on spare parts. A study by the consulting firm Oliver Wyman (2015) found the global profits in the automotive market for 2014 at EUR 273 billion (bn). Aftersales services constituted almost 38% of these global profits: their share was EUR 103 bn. For the German market, the consulting firm Arthur D. Little (2008) found a profit of EUR 6.8 bn from aftersales in 2007, which constitutes as much as 54% of the total profits.

This work is devoted to covering some important challenges of today's inventory/spare parts management. We consider (i) product lifecycles and seasonalities: the time-dependence of demand; (ii) the trade-off between having multiple suppliers and economies of scale; and (iii) the impact of new information on decision-making.

1.2 Problem Statement

This thesis adresses three problems within the area of inventory management. First, an inventory problem with one warehouse and multiple retailers is considered. That is, customer demands (e.g., for spare parts) arise at the retailers who order their stock at the warehouse. The focus of this problem is to determine ordering policies for the retailers and the warehouse. Existing literature mainly assumes customer demand to be independent of time, i.e. to have the same distribution at all times, and determines a time-independent ordering policy. However, this assumption is often not realistic, and demand is rather time-dependent. Although there are previous works that include time-dependence, some important problems remain unsolved. In particular, if an expected time-dependent shape of demand is known, e.g., from an expected lifecycle of a product, the optimal ordering policies are not yet known.

The second problem to be addressed in this thesis deals with a buyer's selection of suppliers. That is, if there are several suppliers that can deliver certain products, the buyer has to decide where to order them. The suppliers usually offer discounts: they might offer quantity discounts per sales volume of a single product, or business volume discounts, i.e. a price reduction depending on the total sales volume for several products. Sourcing from only a few suppliers allows the buyer to take advantage of the discounts and realize economies of scale. However, if a supplier fails and the orders are not satisfied, this can lead to additional costs, a lack of service, or a stop of production. The challenge is to integrate the selection of suppliers and the allocation of orders into one problem, and to account for both the economies of scale and the risk of supplier failure.

The third problem addressed in this thesis considers the final order of a product. At some point, prior to the end of the lifecycle, the production of a part stops. However, the part might still be required to satisfy demand within a specified future time frame. That is, a final order needs to be placed just before production stops in order to ensure the part's future availability. This final order can be complemented by future product returns that might be used to satisfy future demand. This problem has been thoroughly studied in the existing literature. However, none of the studies ask about the value of information: the more information about the future is available, e.g., through improved forecasts, the better are the decisions. That is, improved future information leads to a decrease of uncertainty, and anticipating the lower level of future uncertainty influences today's decisions.

1.3 Structure of the Thesis

The remainder of the thesis is structured as follows. In Chapter 2, we focus on continuous review inventory systems that face non-stationary Poisson demand. We consider single-unit subproblems and track every unit's way through the system until it is matched to a customer demand. For single-stage systems, we find a decision rule that yields the optimal policy. In one-warehouse multi-retailer systems, we find that demands at the warehouse are no longer Poisson distributed. We characterize their arrival time distribution and develop ordering policies. In a numerical study, we find large cost savings for our time-dependent policy. This chapter is based on Bohner and Minner (2017a).

In Chapter 3, we consider a buyer, e.g., a manufacturer, who aims to procure multiple products. In order to do so, she needs to select suppliers and allocate her orders. The suppliers offer quantity and business volume discounts, and they might fail to deliver. We consider both all-units and incremental quantity discounts and introduce a new mixed-integer linear programming formulation that finds optimal solutions. In a numerical study, we discuss the trade-off between economies of scale and failure risk, and show the cost reduction of our exact approach compared to a previously proposed heuristic. This chapter is based on Bohner and Minner (2017b).

In Chapter 4, we consider the final order problem of a service provider who faces customer demands and receives product returns. The returns can be remanufactured or disposed/salvaged. Both demand and return forecasts exist. We consider an evolution of forecasts and prove the structure of the optimal policy. By stochastic dynamic programming, we find that forecast evolution has interesting effects on the size of the final order, the remanufacturing and disposal policies. We show that forecast updates yield large cost savings. These are mainly caused by updates of the demand forecasts; the influence of updating the return forecasts is only small. This chapter is the result of joint work with Stephen C. Graves (Massachusetts Institute of Technology) and Stefan Minner (Technical University of Munich) and based on Bohner et al. (2017).

Chapter 5 concludes the thesis and summarizes the main contributions. We state the limitations of the work and present opportunities for further research.

Chapter 2

Inventory Systems with Non-Homogeneous Poisson Demand

Product lifecycles and demand seasonality are important characteristics of inventory systems. To account for these characteristics, we consider continuous review inventory systems with non-homogeneous Poisson demand. We find a decision rule that yields the optimal policy for single-stage systems. In one-warehouse multi-retailer systems, we find that demands at the warehouse are no longer Poisson distributed. We characterize their arrival time distribution and develop ordering rules. In a numerical study, the resulting policies are compared to the solution from a time decomposition, i.e. the repeated application of a stationary model on a rolling horizon. We obtain large cost savings. Lastly, we suggest an effective alternative for practitioners that is based on a simple time shift.

2.1 Introduction

Lifecycles of products become shorter across markets. Electronic products are the most prominent example, where lifecycles of 6 to 12 months are common (Graves and Willems, 2008). The different phases of the lifecycle cause variation in demand over time and also transfer to the demand for service parts of these products. In the automotive industry, service and parts operations account for 36% of the revenues; in the technology sector even for more than 50% (Guajardo et al., 2015). A study by McKinsey (2013) estimates that the revenue from service parts in the Chinese automotive industry will grow from EUR 20 billion in 2012 to almost EUR 100 billion in 2020. In the U.S., service part sales and after-sales services contribute to the gross domestic product by 8% (Cohen et al., 2006).

Having positive lead times for service parts, e.g., weeks or months, successful inventory management needs to include information on time-varying demand due to lifecycles or seasonalities when setting up ordering policies. This requires dynamic inventory control models.

Modeling time-varying demand as a non-homogeneous Poisson process, as we intend to do, requires knowledge of the time-dependent demand rate, i.e. *how* the demand rate changes, but also *when* it changes. This, in fact, requires a reliable forecast. Undoubtedly, there are problems where such a forecast is hard to obtain and where modeling a time-varying behavior by other means, e.g., Markov-modulated demand, might be more promising. However, there is a variety of problems that have sufficiently accurate demand forecasts. You might think of electronic products, for instance smartphones, where new generations are released at regular intervals. That is, the product lifecycle of the most recent generation resembles the lifecycles of the previous ones. For modeling such lifecycle (or seasonal) demand, non-homogeneous Poisson processes are very suitable.

We will show that non-homogeneous Poisson demand allows us to derive (optimal) time-dependent ordering policies without the use of dynamic programming or any other recursive procedure. This direct approach improves the understanding of the real dynamics of time-dependent ordering in inventory systems.

(i) We present a decision rule that finds the optimal policy of a single-stage inventory system with time-dependent Poisson demand. It is known that the optimal policy is a base-stock policy (see, e.g., Zipkin, 2000). Song and Zipkin (1993) showed that (1) the myopic policy is an upper bound for the optimal policy in general, and (2) the myopic policy is optimal for a non-decreasing demand rate. We consider general demand rates that may increase or decrease. Our rule is based on a decomposition of the problem into single-unit subproblems as proposed by Axsäter (1990) and Muharremoglu and Tsitsiklis (2008). Within each problem, we track the unit's way through the inventory system. The resulting rule is twofold: (a) We decide whether to *actively* order a unit at a certain time. This yields the well-known myopic rule from Zipkin (1989). (b) We decide whether it is preferable to *passively* order the respective unit as a reorder for a future demand instead. If the latter holds, the unit is not ordered actively. That is, part (b) of our rule reduces the base-stock levels of the myopic policy to the ones of the optimal policy.

(ii) We consider a one-warehouse multi-retailer inventory system. Retailers have time-varying base-stock levels that have the shape of step functions. As a consequence, they do not replenish all customer demands one-for-one. That is, they place an extra order at an upward step of the base-stock level, and they refrain from reordering the next demand after a downward step. Thus, the demand that arrives at the warehouse is not Poisson distributed (as it used to be in the stationary problem). We describe this effect and characterize the arrival time distribution of demand at the warehouse.

(iii) We present a decomposed decision rule for obtaining time-dependent base-stock levels for the one-warehouse multi-retailer problem. Note that, even for stationary demand, the optimal policy is not known. Previous approaches require assumptions such as to process the warehouse demands first-come-first-serve or to imply a balance assumption. Our rule for the non-stationary version of the problem is not optimal either. It can be seen as a *two-stage myopic policy* for the one-warehouse multi-retailer problem, and it yields very good solutions. Most notably, for all demand categories, the policy yields large cost savings if compared to a repeated application of the stationary model on a rolling horizon. (In the following, we will call the latter approach *time decomposition*). Secondly, we find a close approximation for the two-stage myopic policy that is based on shifting the solution from time decomposition. As this approximation can be obtained easily, it can be of great value for practitioners.

The chapter is organized as follows. Section 2.2 reviews the related literature. Section 2.3 introduces the model, Section 2.4 derives results for single-stage systems. In Section 2.5, we consider one-warehouse multi-retailer systems and in Section 2.6, we present a numerical study. Section 2.7 gives conclusions.

2.2 Related Literature

Sherbrooke (1968) introduces the METRIC model that considers (homogeneous) Poisson processes. He assumes continuous review, Poisson demand and one-for-one-replenishments. These assumptions are particularly appropriate for low-demand products with high holding costs, e.g., spare parts. METRIC allows for a decomposition of the system and yields the exact echelon base-stock level at the installation furthest upstream; it also serves as a very powerful and widely used approximation for all downstream installation base-stock levels. Muckstadt (1973) and Graves (1985) provide extensions to METRIC. Axsäter (1990) presents an exact solution procedure for a two-echelon version of the problem on hand with one warehouse and multiple retailers. He assumes first-come first-served (FCFS), which is common in the inventory literature, although a FCFS allocation is not optimal (Axsäter, 2007; Axsäter and Marklund, 2008; Howard and Marklund, 2011). A compact overview of multi-echelon inventory systems is provided by Simchi-Levi and Zhao (2012).

For many real-life problems, stationary models are not suitable. To make up for it, different streams that deal with demand variation have been discussed in the literature. Karlin (1960) introduces a single-stage problem where demands of different periods are independent but vary. The drivers of the demand are often modeled as external factors or *state-of-the-world variables*. Morton and Pentico (1995) derive upper and lower bounds for the optimal policy under a general (non-stationary and independent) demand process and find heuristics that perform very well. Many of the models assume Markov-modulated demand. That is, distribution parameters depend on the state of an external Markov chain. For single-stage systems, Song and Zipkin (1993) consider linear holding and backorder costs and Poisson-distributed demand with a Markov-modulated rate. For linear ordering costs, they show the optimality of state-dependent base-stock policies. Sethi and Cheng (1997) show the optimality of state-dependent (s, S) policies for Markov-modulated demand distributions and constant ordering costs.

For multi-echelon systems, Song and Zipkin (1996) investigate a one-warehouse multi-retailer problem with Markov-modulated Poisson demand. In their model, retailers have constant base-stock policies while the warehouse base-stock policy is state-dependent. Abhyankar and Graves (2001) consider a serial two-stage system with Markov-modulated Poisson demand and two states. They determine the optimal position of an inventory hedge. Iida (2001) considers near-myopic policies for serial systems and derives error bounds. Chen and Song (2001) consider a serial system with Markov-modulated demand and deterministic lead times. They prove that a base-stock policy with state-dependent order-up-to levels is optimal. Muharremoglu and Tsitsiklis (2008) prove the optimality of state-dependent base-stock policies for an extension with stochastic lead times. In these models, the exogeneous Markov chain is assumed to be stationary.

Non-homogeneous Poisson processes provide a different way of modeling without any assumptions on stationarity. The time-varying demand rate $\lambda(t)$ is assumed to be known, e.g. from forecasts or from assumptions on a product lifecycle. The repair of military parts for the U.S. Air Force is one example that motivated the research of Hillestad (1982), who introduces the Dyna-METRIC model where demand follows a non-homogeneous Poisson process. The author formulates an optimization problem that aims at finding the optimal spare part mix for each echelon with respect to specified service levels. The time-varying base-stock levels for the spare parts are found by applying a static approach to every point in time, because "[...] the methodology for cross-time optimization has not been developed." Pourakbar et al. (2012) consider non-homogeneous Poisson demand for a product and determine both its optimal final order quantity and a switching time of policies. Pinçe et al. (2015) consider a single-stage system with $\lambda(t)$ consisting of two constant functions with a downward step between them. For a single-adjustment policy, i.e. a base-stock policy where the parameter may be changed exactly once, they determine the timing of the change and the parameter before and after the change. Shang (2012) considers an N-stage serial problem with general non-stationary demand and derives a heuristic that solves N independent single-stage problems.

Methodologically, Axsäter (1990) introduces a central idea with his *unit-tracking* approach, where the key is to decompose the problem into single units of inventory, to index each single unit and to match it to *its demand*. For instance, under FCFS allocation, unit k serves the k^{th} demand after its ordering, with $k = 1$ being the oldest unit. Muharremoglu and Tsitsiklis (2008) show that, following Axsäter's idea, the problem can be decomposed into identical single-unit single-customer subproblems. Levi et al. (2017) use unit-tracking to determine worst-case performance guarantees for computationally efficient balancing policies in serial inventory systems.

2.3 Model Formulation

The inventory system has one warehouse ($i = 0$) and N retailers ($i \in I := \{1, ..., N\}$). We consider stochastic non-homogeneous Poisson demand at the retailers on an infinite horizon with deterministic demand rates $\lambda_i(t) > 0,\ \forall\, i \in I,\ t \geq 0$. We aim at minimizing the total expected costs. There are no setup costs and the reordering of items follows a continuous review base-stock policy. Orders at the warehouse and the retailers are processed on a FCFS basis and if an order for an item cannot be fulfilled, it is backordered. Backorder and holding costs are linear. Procurement costs per item exist but can be neglected in the model, since the total expected procurement costs are constant. The warehouse orders from a supplier with infinite supply. We use the following notation.

L_i = the constant transportation lead time for an item to arrive at retailer i from the warehouse, $i \in I$,

L_0 = the constant transportation lead time for an item to arrive at the warehouse,

h_i = the linear holding cost per unit and time unit at retailer/warehouse i, $i \in I \cup \{0\}$,

b_i = the linear backorder cost per unit and time unit at retailer i, $i \in I$,

$S_i(t)$ = the *installation* base-stock level at retailer/warehouse i at time t, $i \in I \cup \{0\}$,

$S_0^e(t) = S_0(t) + \sum_{i=1}^{N} S_i(t)$ = the *echelon* base-stock level at the warehouse at time t.

2.4 Single-Echelon Systems

2.4.1 Unit-Tracking

Following Axsäter (1990), the k^{th} oldest unit at the retailer at time t satisfies the k^{th} customer demand after time t. We will denote the k^{th} oldest unit by *unit k*. The following definitions are presented for the general one-warehouse multi-retailer system. For a fixed t, we define

$T_t^{k,i}$ = the arrival time of the k^{th} demand at retailer/warehouse i after time t, $i \in I \cup \{0\}$.

We will use $x \in (0, \infty)$ to measure the time that elapsed since t. We define the *mean value functions*

$$\Lambda_t^i(x) := \int_t^{t+x} \lambda_i(s)ds. \tag{2.1}$$

The probability density function (pdf) and the cumulative distribution function (cdf) of $T_t^{k,i},\ i \in I$, are given by (see Beichelt, 2006)

$$g_t^{k,i}(x) = \frac{(\Lambda_t^i(x))^{k-1}}{\Gamma(k)}\lambda_i(t+x)e^{-\Lambda_t^i(x)}, \quad G_t^{k,i}(x) = 1-\sum_{j=0}^{k-1}\frac{(\Lambda_t^i(x))^j}{\Gamma(j+1)}e^{-\Lambda_t^i(x)} = 1-F_t^{x,i}(k-1). \tag{2.2}$$

$F_t^{x,i}(k-1)$ is the cdf of the non-homogeneous Poisson distribution. That is, it represents the probability that $k-1$ or less demands arrive in $[t, t+x]$ at retailer i. The pdf and cdf for $T_t^{k,0}$ will be derived later. In the special case of a single-echelon system, we have $N = 1$ and use the brief notation $L := L_1$, $S(t) := S_1(t)$, $T_t^k := T_t^{k,1}$, $g_t^k(x) := g_t^{k,1}(x)$, $G_t^k(x) := G_t^{k,1}(x)$. For each t, we aim to find the parameter $S(t)$ that minimizes the expected total costs for the entire system under an $(S(t)-1, S(t))$ policy.

We consider the case that unit k is ordered at time t and arrives at time $t+L$. For the arrival time T_t^k of the k^{th} customer demand after time t, it either holds that $T_t^k \in [0, L)$, i.e. demand k arrives before unit k, or that $T_t^k \in [L, \infty)$, i.e. demand k arrives after unit k. In the first case, shortage costs of $b(L - T_t^k)$ arise; in the second case, we face holding costs of $h(T_t^k - L)$. The expected costs from ordering unit k at time t are given through

$$C_t(k) = \begin{cases} b\displaystyle\int_0^L (L-x)g_t^k(x)\mathrm{d}x + h\int_L^\infty (x-L)g_t^k(x)\mathrm{d}x, & k \in \mathbb{N}\backslash\{0\}, \\ bL, & k = 0. \end{cases} \tag{2.3}$$

The costs for $k = 0$ are derived from a stockless retailer, i.e. $S(t) = 0$, who orders a unit at the very moment of a demand arising for this same unit. This case can, for example, be found in a make-to-order system.

2.4.2 Decision Rule

In this subsection, we present the key idea for finding the optimal base-stock level at time t. Assume that, without loss of generality, the inventory position at time t is 0. Then we have two options: either to order unit 1 or not. If we decide not to order unit 1, we set the base-stock level $S(t)$ to 0, otherwise we repeat the procedure for unit 2 and $S(t)$ is at least 1. To be more general, we assume to know that it is optimal to order unit $k-1$ immediately. That is, $k-1$ is a lower bound for $S(t)$. In order to decide whether to order unit k immediately, we have to consider two cases.

1) We order unit k *actively.*

2) Unit k is ordered *passively*, that is, unit k will be the reorder for a future demand.

In 2), the demand that triggers the reorder could be the first demand after time t; however, it could also be a later one. We will explain this in detail below. Note that there is the possibility that no reorder takes place at all. That is, unit k might never be ordered. In 1), if unit k is ordered actively, we have to decide about the timing of the order. Either 1a) we order unit k immediately and it arrives with lead time L. Or 1b) we postpone the order by an infinitesimally small amount of time Δt and this artificially extends the lead time to $L + \Delta t$. In other words, we observe whether the expected costs for unit k are an increasing or a decreasing function of the lead time. The answer to this question is given by the derivative of $C_t(k)$ with respect to L. Immediate ordering [1a)] dominates postponing [1b)] if the derivative is positive.

$$\frac{\mathrm{d}C_t(k)}{\mathrm{d}L} > 0 \quad \Leftrightarrow \quad b\int_0^L g_t^k(x)dx - h\int_L^\infty g_t^k(x)dx > 0 \quad \Leftrightarrow \quad G_t^k(L) > \frac{h}{b+h}, \ \forall\ k \geq 1. \tag{2.4}$$

For $k = 0$, it holds that $\frac{\mathrm{d}C_t(0)}{\mathrm{d}L} = b > 0$ and $G_t^0(L) = 1$. Note that (2.4) is the myopic rule known from Zipkin (1989). If $G_t^k(L) \leq \frac{h}{b+h}$ [1b)], there is no immediate ordering of unit k and the optimal base-stock level is set to $S(t) = k - 1$. If $G_t^k(L) > \frac{h}{b+h}$ [1a)], we still do not know whether passive ordering of unit k would be preferable to active ordering. (In the special case of a non-decreasing demand rate, the myopic policy is optimal and so is active ordering, see Song and Zipkin 1993.) We consider both the *cost of active ordering* $C_t(k)$ from (2.3) and the *cost of passive ordering* $\hat{C}_t(k)$. For the latter, the timing of the passive order directly results from the demand process. We have

$$\hat{C}_t(k) = \begin{cases} \displaystyle\int_0^\infty g_t^1(s)C_{t+s}(k-1)\mathrm{d}s, & k > 1, \\ bL, & k = 1. \end{cases} \tag{2.5}$$

This formula assumes that $S(t) = k - 1$ and unit k is a reorder for the first demand after time t. If this first demand and the respective reorder take place at time $t + s$, the former unit k is the system's new unit $k - 1$, and the expected costs for the unit are $C_{t+s}(k - 1)$. As the time of the first demand is uncertain, these costs have to be weighted with the respective probability $g_t^1(s)$ that the first demand after time t takes place at time $t+s$. If both (2.4) and $C_t(k) < \hat{C}_t(k)$ hold, then active ordering is optimal for unit k at time t. We order unit k immediately, set the lower bound for $S(t)$ to k and repeat the procedure for unit $k+1$. Otherwise we do not order unit k and we set $S(t) = k-1$.

Due to time-varying base-stock levels, however, passive ordering is generally more complicated. Equation (2.5) considers a case without steps in the base-stock levels. That is, $S(\tau) = k-1,\ \forall \tau > t$, and unit k replaces unit 1. However, for varying base-stock levels, we must take steps in the base-stock levels into account. We consider two cases.

(i) If there is a downward step in the base-stock level, the first demand after the downward step will not be reordered. By that, the inventory position will be adjusted to the new base-stock level. As a consequence, unit k is not necessarily the replacement for unit 1.

(ii) If there is an upward step in the base-stock level, an order will be triggered without there being a demand. That is, unit k does not necessarily replace unit 1 either.

In the following, we only consider a step size of one. In fact, we will prove below that all steps have size one. If the base-stock level has a single *downward* step at time $t+t_1$, the cost of passive ordering is found through

$$\hat{C}_t(k) = \begin{cases} \int_0^{t_1} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s + \int_{t_1}^{\infty} g_t^1(s) \int_0^{\infty} g_{t+s}^1(\tau) C_{t+s+\tau}(k-2)\mathrm{d}\tau\mathrm{d}s, & k > 2, \\ \int_0^{t_1} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s + bL(1-G_t^1(t_1)), & k = 2. \end{cases} \tag{2.6}$$

(We do not consider $k=1$ as this would mean starting from level $k-1=0$; a downward step would make no sense.) If the first demand after time t takes place before the downward step at $t+t_1$, the reordering is analogous to (2.5). If it takes place after the downward step at $t+t_1$, it is not reordered and unit k will be triggered by the second demand after time t.

If the first change in the base-stock level after time t is an *upward* step at time $t+t_1$, the cost of passive ordering is given through

$$\hat{C}_t(k) = \int_0^{t_1} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s + (1-G_t^1(t_1))C_{t+t_1}(k), \qquad k \geq 1. \tag{2.7}$$

The latter formula is independent of future steps in the base-stock level after time $t+t_1$. This is quite intuitive: We have a base-stock level of $k-1$ and consider unit k. The ordering time for this unit will be no later than the time when the base-stock is raised to k, i.e. at time $t+t_1$. This logic still holds if there are other downward and upward steps before the base-stock level is raised to k; in any case, we only need to consider the time frame until the base-stock level is raised to k for the first time after time t (if this ever happens).

Formulas (2.5), (2.6), and (2.7) are the building blocks for determining the cost of passive ordering in cases where the base-stock level has more steps. Extensions to more steps are given in Section 2.A of the appendix.

Note that finding the cost of passive ordering requires knowledge about future upward and downward steps of the base-stock level. They can be obtained by starting the determination of $S(t)$ at the end of the time horizon. Although this looks similar to dynamic programming, it is not: In order to determine $S(t)$, we only make use of the times of upward and downward steps, but do not require a follow-up cost term. The *entire future* is already part of formulas (2.3), (2.5), (2.6), and (2.7).

Proposition 2.1. *Let $\lambda(t)$ be piecewise continuous and bounded. Then the steps of the retailer base-stock levels are always of size 1.*

Proof. See Section 2.B of the appendix. □

Theorem 2.1. *It is optimal to actively order unit k at time t if*

(a) It is optimal to order unit $k-1$ at time t, and

(b) $G_t^k(L) > \frac{h}{b+h}$, and

(c) $C_t(k) < \hat{C}_t(k)$.

Proof. See Section 2.B of the appendix. □

We can show that our inventory problem is decomposable, i.e. it satisfies conditions A1-A6 from Muharremoglu and Tsitsiklis (2008). That is, we can solve single-unit single-customer subproblems instead of the overall inventory problem; this directly yields the following corollary.

Corollary 2.1. *The repeated application of Theorem 1 yields the optimal time-dependent base-stock levels for all t.*

The following Theorem introduces a relation between the stationary and the non-stationary model. It is the basis for a simple alternative that we will derive later.

Theorem 2.2. *Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be a linear function of time, i.e.* $\lambda(t) = mt + d,\ m, d \in \mathbb{R}$*. Then the myopic solution for the base-stock step function $S(t)$ can be found by setting $\tilde{\lambda}(t) := \lambda(t + \frac{L}{2})$ and applying the stationary solution method for every t.*

Proof. The stationary version of $G_t^k(L)$ is given by

$$G^k(L) = 1 - \sum_{j=0}^{k-1} \frac{(\lambda L)^j}{j!} e^{-\lambda L}. \tag{2.8}$$

For linear $\lambda(t) = mt + d$, we find

$$\tilde{\lambda}(t)L = \lambda\Big(t + \frac{L}{2}\Big)L = \Big(m\Big(t + \frac{L}{2}\Big) + d\Big)L = \Big(\frac{m}{2}s^2 + ds\Big)\ \Big|_t^{t+L} = \int_t^{t+L} (ms + d)\ \mathrm{d}s = \Lambda_t(L).$$

That is, the solution of the non-stationary problem for time t equals the solution of the stationary problem with $\lambda \equiv \tilde{\lambda}(t)$. □

The intuition behind this theorem is that the integral of a linear function over a finite interval equals the linear function's value in the middle of the interval multiplied by the interval length. That is, if we consider the interval $[t, t + L]$ for our decision at time t in the non-stationary model, this decision equals the one found from a stationary model with lead time L and the demand rate at time $t + \frac{L}{2}$. Thus, the base-stock level we obtain from the stationary solution for each point in time becomes valuable if it is shifted to the left by a half lead time.

2.4.3 Numerical Results

In order to assess the value of the optimal policy, we compare it to (suboptimal) alternatives. We consider (i) seasonal demand and (ii) demand along the lifecycle of a product. We use

$$\lambda(t) = a \cdot \sin\Big(c\pi\Big(t - \frac{1}{2c}\Big)\Big) + 2 \tag{2.9}$$

as seasonal demand rate, where the shift by $-\frac{1}{2c}$ ensures that the first season starts at time 0, and

$$\hat{\lambda}(t) = \hat{a}t^2 e^{-\hat{c}t} + d \tag{2.10}$$

as lifecycle demand rate. The latter follows the poly-exponential model found by Pourakbar et al. (2012) through fitting spare part demand data of a consumer electronics manufacturer.

In order to assess different shapes of demand variation, we consider higher and lower amplitudes and frequencies in the demand rates. Further, we vary the risk of having excess stocks at the end of the lifecycle, i.e. we vary d in the lifecycle demand rate. These shapes are combined with higher and lower backorder costs and lead times, while holding costs are fixed. The numerical design is given through

$$a \in \{0.3, 1\}, \quad c \in \left\{\frac{1}{3}, \frac{2}{3}\right\}, \quad \hat{a} \in \{1, 2\}, \quad \hat{c} \in \{0.3, 0.6\},$$
$$h = 1, \quad b \in \{5, 15\}, \quad L \in \{0.5, 1\}, \quad d \in \{0.05, 1\}.$$

In total, we have 48 scenarios. For seasonal demand, we investigate season lengths of 3 and 6 months and mean demands that vary within $[1.7, 2.3]$ or $[1, 3]$, respectively. For product lifecycles, the mean demand's peak is in months 4 or 7 and takes values of $1.5 + d$, $3 + d$, $6 + d$, or $12 + d$.

We compare the optimal policy with three other approaches: (i) The myopic policy, i.e. considering active ordering but ignoring passive ordering. (ii) The policy from a time decomposition (TD) that applies the stationary model at each point in time. That is, for a fixed $\bar{t}$, we solve a stationary problem and use the value $\lambda(\bar{t})$ as the stationary input parameter λ for that model. The resulting base-stock level is used as the base-stock level of the TD policy at time $\bar{t}$. (iii) The policy from shifting TD to the left by a half lead time (HL) as suggested by Theorem 2.2. That is, if $\hat{S}(t)$ denotes the base-stock policy of TD, the base-stock policy of HL is defined as

$$\tilde{S}(t) = \hat{S}\Big(t + \frac{L}{2}\Big). \tag{2.11}$$

The relative cost differences between the optimal policy and the myopic, TD and HL policies are denoted by Δ_{MY}, Δ_{TD} and Δ_{HL}, respectively.

Lifecycle demand. We first consider lifecycle demand with a high risk of excess stocks at the end of the lifecycle. For the short lifecycles ($\hat{c} = 0.6$), the demand rates fall below 0.1 after 14 and 16 months, respectively. For the longer lifecycles ($\hat{c} = 0.3$), the demand rates fall below 0.1 after 34 and 37 months. We evaluate the policies for 100 months.

Table 2.1: Results for single-stage system with lifecycle demand.

				$d = 0.05$			$d = 1$		
$\hat{a}$	$\hat{c}$	L	b	Δ_{MY} (%)	Δ_{TD}	Δ_{HL}	Δ_{MY}	Δ_{TD}	Δ_{HL}
1	0.6	0.5	5	16.69	18.67	16.70	0.03	0.07	0.03
			15	10.99	13.03	11.03	0.02	0.07	0.03
		1	5	17.10	22.91	17.10	0.04	0.29	0.04
			15	8.93	13.69	9.10	0.04	0.28	0.04
	0.3	0.5	5	4.15	5.32	4.15	0.03	0.16	0.03
			15	4.18	5.19	4.19	0.02	0.17	0.02
		1	5	5.28	9.04	5.28	0.04	0.91	0.05
			15	2.94	6.73	2.98	0.03	0.92	0.04
2	0.6	0.5	5	15.43	18.84	15.43	0.04	0.16	0.04
			15	10.86	13.73	10.88	0.04	0.16	0.04
		1	5	17.11	24.64	17.18	0.07	0.8	0.09
			15	10.13	17.21	10.26	0.06	0.76	0.06
	0.3	0.5	5	3.34	4.52	3.34	0.03	0.37	0.03
			15	3.45	4.84	3.48	0.03	0.4	0.03
		1	5	4.23	9.52	4.23	0.04	2.11	0.05
			15	6.45	8.85	6.50	0.04	2.50	0.08

Table 2.1 shows that for a high risk of excess stocks ($d = 0.05$), there is a big advantage in using the optimal policy. The reason is that both the myopic policy and the TD policy reduce their inventories too late. That is, by the time they try to reduce their inventories, they are no longer able to sell off the excess stock fast enough due to a lack of demand. In fact, the optimal policy anticipates the risk of excess stocks by considering the cost of passive ordering. For the short lifecycles ($\hat{c} = 0.6$), the optimal policy saves between 8.9% and 17.1% if compared to the myopic policy. If compared to the TD policy, the savings are even between 13% and 24.6%. The long lifecycles ($\hat{c} = 0.3$) leave more time at the end to sell off the remaining parts, and the savings of the optimal policy are smaller. They range between 2.9% and 6.5% if compared to the myopic policy, and between 4.5% and 9.5% if compared to the TD policy. All results of HL are very close to those of the myopic policy.

If we have lifecycle demand with a low risk of excess stocks at the end of the lifecycle ($d = 1$), the results are different. The cost difference between optimal and myopic policy is less than 0.1% in all instances. The cost difference between optimal and TD policy is less than 2.5%. This is due to the ongoing frequent demand after the decrease of the

Table 2.2: Results for single-stage system with seasonal demand.

a	c	L	b	Δ_{MY}	Δ_{TD}	Δ_{HL}
1	$\frac{2}{3}$	0.5	5	2.51	6.87	2.52
			15	2.73	7.39	2.76
		1	5	0.52	17.63	4.34
			15	1.86	15.87	3.72
	$\frac{1}{3}$	0.5	5	0.77	3.08	0.77
			15	0.98	3.13	1.05
		1	5	0.98	7.73	1.11
			15	1.00	8.17	1.45
0.3	$\frac{2}{3}$	0.5	5	0	0	0
			15	0.01	0.44	0.01
		1	5	0.29	2.50	0.55
			15	0.37	2.46	0.37
	$\frac{1}{3}$	0.5	5	0	0	0
			15	0.05	0.25	0.06
		1	5	0.23	1.28	0.28
			15	0.23	1.30	0.23

lifecycle curve: remaining parts in all policies can be sold off rather quickly.

Seasonal demand. In the case of seasonal demand, every decrease of the base-stock level is followed by an increase. That is, excess stocks dwindle fast. Yet the myopic, the TD, and the HL have higher holding costs than the optimal policy in every season. For any combination that includes small ($a = 0.3$) or slow ($c = \frac{1}{3}$) demand fluctuations, the difference between optimal and myopic policy is less than 1%, and HL is close to the myopic policy (see Table 2.2). For large and fast demand fluctuations ($a = 1$, $c = \frac{2}{3}$), the difference between optimal and myopic policy can be slightly higher, i.e. up to 2.7%. In these instances, HL is no longer quite as close and differs from the optimal policy by up to 4.3%. The comparison between the optimal and the TD policy shows similar results as long as demand fluctuations are small and close to stationarity ($a = 0.3$): their cost difference is less than 2.5%. In fact, for two instances a stationary policy is optimal and all approaches find it. For larger fluctuations ($a = 1$), however, the difference between the optimal and the TD policy ranges from 3.1% up to 17.6%.

In conclusion, we observe that (i) the higher the variation in the demand rate, the farther from optimal is TD; (ii) the myopic policy performs well as long as the risk of excess stock is small; (iii) HL is a close approximation for the myopic policy.

2.5 One-Warehouse Multi-Retailer System

We now consider one warehouse and N retailers. We aim at finding base-stock levels $(S_0(t), S_1(t), ..., S_N(t))$. On top of holding and backorder costs at the retailers, there are holding costs at the warehouse. Each unit is shipped from the warehouse to a retailer i, which requires a fixed amount of time L_i. Following van Houtum and Zijm (1991), we assign warehouse holding costs h_0 for the shipping time; these constant costs are assigned to retailer i.

In the following, we present how the decisions at the retailers and at the warehouse can be decomposed. The resulting decision rule is based on the cost of active ordering, both at the retailers and at the warehouse. We do not claim the optimality of this decision rule. In fact, for obtaining optimal solutions, the cost of passive ordering cannot be excluded. However, including the latter does not allow for decomposition anymore. Despite not being optimal, the rule (a) yields large improvements if compared to the solution from time decomposition, as we will show in a numerical study; (b) can be considered a close approximation of the optimal policy in many cases. These include

(i) *Seasonal demand.* In Section 2.4.3, we observed that, in the majority of instances, the costs of applying the myopic policy to seasonal demand deviate less than 1% from the optimal solution.

(ii) *Lifecycle demand with small risk of excess stocks.* In Section 2.4.3, we made an observation similar to (i) for lifecycle demand with $d = 1$. Here, too, the costs of applying the myopic policy only show a small deviation from the optimal policy.

(iii) *Lifecycle demand on a finite horizon.* If we consider a lifecycle on a finite horizon, e.g., if the demand rate approaches zero, and the stock at the end of the horizon is not decision-relevant, the myopic policy is close to optimal with the same arguments as in Section 2.4.3. (As the stock at the end of the horizon might be disposed, salvaged, and so on, this is an approximation. However, there will no longer be any risk of keeping excess stock for an infinite future.)

(iv) *Non-decreasing demand.* Song and Zipkin (1993) showed (for single-stage systems) that the myopic policy is optimal if the demand rate is non-decreasing.

Thus, a two-stage myopic policy can be a precious tool for time-dependent inventory management. In the following, we develop the theory for the one-warehouse multi-retailer problem.

2.5.1 Retailer Decision Rule

We have the expected costs at retailer i

$$C_t^i(k) = \begin{cases} b_i \int_0^{L_i} (L_i - x) g_t^{k,i}(x) dx + h_i \int_{L_i}^{\infty} (x - L_i) g_t^{k,i}(x) dx + h_0 L_i, & k \in \mathbb{N} \backslash \{0\}, \\ (b_i + h_0) L_i, & k = 0. \end{cases} \tag{2.12}$$

The decision rule for active ordering at retailer i is

$$\frac{\mathrm{d}C_t^i(k)}{\mathrm{d}L_i} > 0 \;\Leftrightarrow\; b_i \int_0^{L_i} g_t^{k,i}(x) dx - h_i \int_{L_i}^{\infty} g_t^{k,i}(x) dx + h_0 > 0 \;\Leftrightarrow G_t^{k,i}(L_i) > \frac{h_i - h_0}{b_i + h_i}, \quad \forall\, k \geq 1. \tag{2.13}$$

That is, the smaller the difference between retailer holding costs h_i and warehouse holding costs h_0, the more stock is kept at retailer i. However, there is an upper bound to the amount of stock kept at retailer i at time t. It is denoted by $\bar{S}_i(t)$, $i \in \{1, ..., N\}$, and determined such that

$$G_t^{\bar{S}_i(t),i}(L_0 + L_i) > \frac{h_i}{b_i + h_i}, \quad G_t^{\bar{S}_i(t)+1,i}(L_0 + L_i) \leq \frac{h_i}{b_i + h_i}. \tag{2.14}$$

In (2.14), we assume that there is no warehouse between retailer i and the supplier. That is, the lead time for retailer i increases to $L_0 + L_i$ and we consider retailer i as a single-stage system. $\bar{S}_i(t)$ denotes the base-stock level at time t in this system. (2.14) is a constraint for rule (2.13); it becomes binding if h_i and h_0 are close.

2.5.2 Warehouse Decision Rule

Following Axsäter (1990), we use FCFS allocation. The retailer base-stock levels $S_i(t)$ are obtained from equation (2.13) and fixed for all t. We now extend the unit-tracking approach to the warehouse in order to find the time-dependent parameters for an $(S_0(t) - 1, S_0(t))$ policy.

Figure 2.1 illustrates the unit's way through a serial system with $N = 1$. At time t,

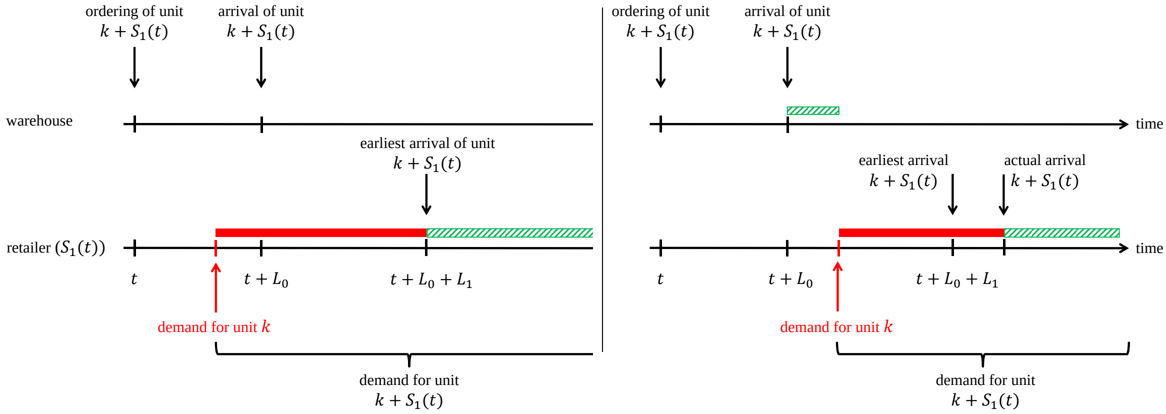


Figure 2.1: Unit-tracking in a serial two-echelon system.

the warehouse must decide to either order unit k (= *system unit:* $k+\sum_{i=1}^{N} S_i(t)$) or not. If the warehouse orders unit k at time t, unit k arrives at the warehouse at time $t+L_0$. As soon as the k^{th} demand at the warehouse after time t occurs, at time $T_t^{k,0}$, and the k^{th} unit is available at the warehouse, it is shipped to the retailer who ordered it.

If $T_t^{k,0} < L_0$, i.e. on the left side of Figure 2.1, the unit is shipped to the retailer immediately after its arrival at time $t+L_0$. Otherwise, i.e. on the right side of Figure 2.1, the unit remains at the warehouse at a holding cost of h_0 until $T_t^{k,0}$. While the unit is being shipped from the warehouse to the retailer, holding costs of h_0 are charged. In contrast to a single-stage system, the backordering time of a unit can now exceed L_i. This happens if the actual demand for the unit at retailer i arrives even before the unit arrives at the warehouse, i.e. both $T_t^{k,0} = s < L_0$ and $T_{t+s}^{S_i(t+s),i} < L_0 - s$.

For $N > 1$, we not only observe the time when the order for unit k arrives, but we also need to know from which retailer. We therefore derive the distribution for the arrival time $T_t^{k,0}$ of the k^{th} demand at the warehouse as

$$G_t^{k,0}(x) = \sum_{i=1}^{N} G_t^{k,0,i}(x), \text{ with } G_t^{k,0,i}(x) := \mathbb{P}\big(\{T_t^{k,0} \leq x\} \cap \{\text{unit } k \text{ is sent to retailer } i\}\big).$$

With the *partial distribution functions* $G_t^{k,0,i}$, the expected costs for unit k are

$$\begin{aligned} C_t^0(k) &= h_0 \int_{L_0}^{\infty} (s - L_0) g_t^{k,0}(s) \mathrm{d}s \\ &+ \sum_{i=1}^{N} \Bigg(\int_0^{L_0} g_t^{k,0,i}(s) \Big(b_i \int_0^{L_0+L_i-s} g_{t+s}^{S_i(t+s),i}(x)(L_0 + L_i - s - x)\mathrm{d}x \\ &\quad + h_i \int_{L_0+L_i-s}^{\infty} g_{t+s}^{S_i(t+s),i}(x)(s + x - L_0 - L_i)\mathrm{d}x \Big) \mathrm{d}s \\ &\quad + \int_{L_0}^{\infty} g_t^{k,0,i}(s) \Big(b_i \int_0^{L_i} g_{t+s}^{S_i(t+s),i}(x)(L_i - x)\mathrm{d}x + h_i \int_{L_i}^{\infty} g_{t+s}^{S_i(t+s),i}(x)(x - L_i)\mathrm{d}x \Big) \mathrm{d}s \Bigg). \end{aligned} \tag{2.15}$$

The term in the first line of (2.15) yields the expected holding costs at the warehouse and the term in the last line yields the expected holding and backorder costs at the retailers in scenario $T_t^{k,0} > L_0$ (right side of Figure 2.1). The terms in the second and third lines in (2.15) yield the expected holding and backorder costs at the retailers in scenario $T_t^{k,0} < L_0$ (left side of Figure 2.1), where no warehouse holding costs arise but backordering at the warehouse takes place.

For an evaluation of the holding and backorder costs at the retailers in both scenarios, let $t+s$ be the time of the k^{th} warehouse demand after time t that was caused by retailer i. That is, warehouse unit k is then shipped to retailer i at time $t+s$. The holding and backorder costs for warehouse unit k at retailer i depend (analogously to the single-stage system) on the arrival time of the $S_i(t+s)^{th}$ customer demand at retailer i after time $t+s$, given by $T_{t+s}^{S_i(t+s),i}$.

Analogously to the retailer decision rule, we decide whether to order unit k immediately or to postpone the ordering. That is, we consider the derivative of (2.15) w.r.t. lead time L_0.

$$\begin{aligned} \frac{\mathrm{d}C_t^0(k)}{\mathrm{d}L_0} =& - h_0(1 - G_t^{k,0}(L_0)) \\ &+ \sum_{i=1}^{N} \Big((b_i + h_i) \int_0^{L_0} g_t^{k,0,i}(s) G_{t+s}^{S_i(t+s),i}(L_0 + L_i - s)\mathrm{d}s - h_i G_t^{k,0,i}(L_0) \Big). \end{aligned} \tag{2.16}$$

Proposition 2.2. *The installation order-up-to level $S_0(t)$ at time t is found by*

$$\frac{\mathrm{d}C_t^0(S_0(t))}{\mathrm{d}L_0} > 0, \quad \frac{\mathrm{d}C_t^0(S_0(t)+1)}{\mathrm{d}L_0} \leq 0, \quad \textit{with} \quad \frac{\mathrm{d}C_t^0}{\mathrm{d}L_0} \quad \textit{from (2.16)}. \tag{2.17}$$

Proof. We show the existence and uniqueness of the base-stock level $S_0(t)$. The existence follows from

$$\lim_{k\to\infty} G_t^{k,0,i}(L_0) = 0 \quad \Rightarrow \quad \lim_{k\to\infty} \frac{\mathrm{d}C_t^0(k)}{\mathrm{d}L_0} = -h_0 < 0. \tag{2.18}$$

For the uniqueness, we show that $\frac{\mathrm{d}C_t^0(k)}{\mathrm{d}L_0}$ is monotonically decreasing in k. Therefore, for arbitrary t, we set

$$t_i^*(t) := \arg\min_{s\in[0,L_0]} G_{t+s}^{S_i(t+s),i}(L_0 + L_i - s). \tag{2.19}$$

For $t_i^*(t)$, it holds that

$$t_i^*(t) \leq L_0 \quad \Rightarrow \quad L_0 + L_i - t_i^*(t) \geq L_i \quad \Rightarrow \quad G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_0 + L_i - t_i^*(t)) \geq G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_i),$$

and since the retailer condition holds for $S_i(t + t_i^*(t))$, i.e.

$$\begin{aligned} & G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_i) > \frac{h_i - h_0}{b_i + h_i}, \qquad \text{and } G_t^{k,0,i} \text{ is decreasing in } k, \\ \text{i.e.} \quad & G_t^{k,0,i}(L_0) - G_t^{k+1,0,i}(L_0) \geq 0, \qquad \text{it follows} \\ & G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_0 + L_i - t_i^*(t)) > \frac{h_i - h_0}{b_i + h_i} \\ & \Rightarrow (b_i + h_i) G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_0 + L_i - t_i^*(t))(G_t^{k,0,i}(L_0) - G_t^{k+1,0,i}(L_0)) \\ & > (h_i - h_0)(G_t^{k,0,i}(L_0) - G_t^{k+1,0,i}(L_0)). \end{aligned} \tag{2.20}$$

Further, again using (2.19), it holds that

$$\begin{aligned} & (b_i + h_i) \int_0^{L_0} (g_t^{k,0,i}(s) - g_t^{k+1,0,i}(s)) G_{t+s}^{S_i(t+s),i}(L_0 + L_i - s)\mathrm{d}s \\ & \geq (b_i + h_i) G_{t+t_i^*(t)}^{S_i(t+t_i^*(t)),i}(L_0 + L_i - t_i^*(t))(G_t^{k,0,i}(L_0) - G_t^{k+1,0,i}(L_0)). \end{aligned} \tag{2.21}$$

Summing over all i, from (2.20) and (2.21) together, it follows that

$$\begin{aligned}&\sum_{i=1}^{N}(b_i+h_i)\int_0^{L_0}(g_t^{k,0,i}(s)-g_t^{k+1,0,i}(s))G_{t+s}^{S_i(t+s),i}(L_0+L_i-s)\mathrm{d}s\\&>\sum_{i=1}^{N}(h_i-h_0)(G_t^{k,0,i}(L_0)-G_t^{k+1,0,i}(L_0))\Leftrightarrow\frac{\mathrm{d}C_t^0(k)}{\mathrm{d}L_0}>\frac{\mathrm{d}C_t^0(k+1)}{\mathrm{d}L_0}.\ \square\end{aligned}\tag{2.22}$$

For a serial system with $N=1$, we know that (i) if $S_1(t+s)=S_1(t)$, the number of retailer orders (k) in interval $[t,t+s]$ equals the number of customer demands; (ii) if $S_1(t+s)>S_1(t)$, i.e. we have more upward than downward steps in $[t,t+s]$, there were more retailer orders than customer demands ($T_t^{k,0}<T_t^{k,1}$); (iii) if $S_1(t+s)<S_1(t)$, there were more customer demands than retailer orders ($T_t^{k,0}>T_t^{k,1}$). Note that, although the retailer might not replenish one-for-one, it still holds that system unit $k+S_1(t)$ at time t serves the $k+S_1(t)^{th}$ customer demand after time t.

2.5.3 Arrival Time Distribution at the Warehouse in a Serial System

Each step of $S_i(t)$ is a point of discontinuity. We denote the corresponding times by $\tilde{t}_j^i$, being defined as the j^{th} step of S_i after time t at retailer i. That is, j depends on t. It holds that $j\in\{1,...,M_i^t\}$ where $M_i^t\in\mathbb{N}_0$ is the total number of steps of S_i on (t,∞). We set $\tilde{t}^i_{M_i^t+1}:=\infty$. The functions $S_i(t)$ are *right-continuous with left limits (RCLL)*, i.e.

$$\begin{aligned}&(a)\ \lim_{x\to\tilde{t}_j^i,x<\tilde{t}_j^i}S_i(x)\ \text{ and }\ \lim_{x\to\tilde{t}_j^i,x>\tilde{t}_j^i}S_i(x)\ \text{ exist, and}\\&(b)\ \lim_{x\to\tilde{t}_j^i,x>\tilde{t}_j^i}S_i(x)=S_i(\tilde{t}_j^i).\end{aligned}$$

We define for step j after time t in S_i

$$\delta_j^i:=\lim_{x\to\tilde{t}_j^i,x<\tilde{t}_j^i}S_i(x)-S_i(\tilde{t}_j^i),\qquad\Delta_j^i:=\sum_{k=1}^{j}\delta_k^i,\ \Delta_0^i:=0.$$

δ_j^i gives the direction of the step at $\tilde{t}_j^i$, i.e. upward ($\delta_j^i=-1$) or downward ($\delta_j^i=1$). Δ_j^i represents the cumulative first j steps since time t, which are positive if there have been more downward steps than upward steps. We now define $t_j^i:=\tilde{t}_j^i-t$ measuring the time between t and the j^{th} step of S_i after t, and set $t_0^i := 0$.

As this subsection considers the special case of a serial system, we omit the superscript 1 and set $\tilde{t}_j := \tilde{t}_j^1,\ t_j := t_j^1,\ \Delta_j := \Delta_j^1$ for ease of notation. At time t, we determine the distribution of the arrival time of the k^{th} demand at the warehouse after time t, $T_t^{k,0}$. First, we present auxiliary functions that help to define the warehouse arrival time cdf piecewise between the individual steps of $S_1(t)$.

Definition 2.1 (Auxiliary Functions)**.** *For* $t \in \mathbb{R}_+,\ j \in \{1, ..., M_1^t\}$*, we define*

$$U_0^k(x) := \mathbb{P}(T_t^{k,1} \leq x), \qquad x \geq 0, \tag{2.23}$$

$$U_j^k(x) := \mathbb{P}\big(T_t^{k+\Delta_j,1} \leq x, T_t^{k+\Delta_{m-1},1} > t_m,\ \forall\ m \in \{1, ..., j\}\big), \qquad x \geq t_j. \tag{2.24}$$

We consider intervals that are bounded by two consecutive steps of the retailer's base-stock level. For any interval j (i.e. between the j^{th} and the $(j+1)^{th}$ step), the k^{th} unit at the warehouse at time t is triggered by the $(k+\Delta_j)^{th}$ customer demand after time t. However, this happens *only if* this same k^{th} warehouse unit has not been requested before. That is, $T_t^{k+\Delta_{m-1},1} > t_m$ needs to hold for all prior intervals; otherwise unit k would have been triggered by the $(k+\Delta_{m-1})^{th}$ demand in an earlier interval. Therefore, the probability that unit k is requested in interval $(t_j, x]$ is given by the joint probability $U_j^k(x)$ from Definition 2.1. The overall probability that unit k is requested prior to time x is consequently obtained by summing the joint probabilities over the disjoint intervals up to x.

Proposition 2.3. *For* $x \in [t_j, t_{j+1}),\ j \in \{0, 1, ..., M_1^t\}$*, the cdf of the arrival time of the* k^{th} *demand at the warehouse after time* t*,* $G_t^{k,0}(x) := \mathbb{P}(T_t^{k,0} \leq x)$*, is given through*

$$G_t^{k,0}(x) = \sum_{m=0}^{j-1} U_m^k(t_{m+1}) + U_j^k(x). \tag{2.25}$$

Proof. If the k^{th} customer demand takes place before time $\tilde{t}_1$, it triggers the k^{th} demand at the warehouse. Otherwise, if $T_t^{k,1} \geq t_1$, then $G_t^{k,0}(x)$ is found by the following algorithm that starts either in (i) or in (ii), depending on the first step at time $\tilde{t}_1$.

(i) The step function decreases at time $\tilde{t}_1$ ($\tilde{t}_j$). If $T_t^{k+1,1} < t_2$ ($T_t^{k+\Delta_j,1} < t_{j+1}$), the k^{th} demand at the warehouse is triggered by the $(k+1)^{th}$ ($(k+\Delta_j)^{th}$) customer demand. If $T_t^{k+1,1} \geq t_2$ ($T_t^{k+\Delta_j,1} \geq t_{j+1}$) and the step function decreases at time

$\tilde{t}_2$ ($\tilde{t}_{j+1}$), again apply (i) with $j := j + 1$. If the step function increases at time $\tilde{t}_2$ ($\tilde{t}_{j+1}$), apply (ii) with $j := j + 1$.

(ii) The step function increases at time $\tilde{t}_1$ ($\tilde{t}_j$). If $T_t^{k-1,1} \geq t_1$ ($T_t^{k+\Delta_j,1} \geq t_j$) and $T_t^{k-1,1} < t_2$ ($T_t^{k+\Delta_j,1} < t_{j+1}$), the k^{th} demand at the warehouse is triggered by the $(k-1)^{th}$ ($(k+\Delta_j)^{th}$) customer demand. If the $(k-1)^{th}$ ($(k+\Delta_j)^{th}$) customer demand takes place before time $\tilde{t}_1$ ($\tilde{t}_j$), the k^{th} warehouse demand arises at time $\tilde{t}_1$ ($\tilde{t}_j$) without a customer demand since the retailer has to raise the inventory position. If $T_t^{k-1,1} \geq t_2$ ($T_t^{k+\Delta_j,1} \geq t_{j+1}$) and the step function decreases at time $\tilde{t}_2$ ($\tilde{t}_{j+1}$), apply (i) with $j := j + 1$. If the step function increases at time $\tilde{t}_2$ ($\tilde{t}_{j+1}$), again apply (ii) with $j := j + 1$. □

Figure 2.2 depicts the different starting points of the algorithm. The dashed horizontal lines represent the possible course of $S_1(t)$ after $\tilde{t}_2$. If $\Delta_j < \Delta_{j-1}$, i.e. $S_1(t)$ has an upward step at $\tilde{t}_j$, then $U_j^k(t_j) > 0$. In particular, $G_t^{k,0}(x)$, too, has a point of discontinuity at $x = t_j$, i.e. an upward step. This means that $G_t^{k,0}(x)$ is only RCLL at t_j and the corresponding pdf has a positive point mass whenever $S_1(t)$ has an upward step. If $\Delta_j > \Delta_{j-1}$, i.e. $S_1(t)$ has a downward step at $\tilde{t}_j$, then $U_j^k(t_j) = 0$, i.e. $G_t^{k,0}(x)$ is continuous at $x = t_j$. The pdf $g_t^{k,0}(x)$ is obtained as the derivative of $G_t^{k,0}(x)$ wherever $G_t^{k,0}(x)$ is differentiable. If $G_t^{k,0}(x)$ is not differentiable, i.e. for times of upward steps $x = t_j$, the positive point mass of the pdf is given through

$$g_t^{k,0}(t_j) = U_j^k(t_j) = G_t^{k,0}(t_j) - \lim_{x \to t_j, x < t_j} G_t^{k,0}(x). \tag{2.26}$$

At each upward step, the retailer's inventory position is adjusted to the new base-stock level. If (a) the inventory position equaled the prior base-stock level before the step,

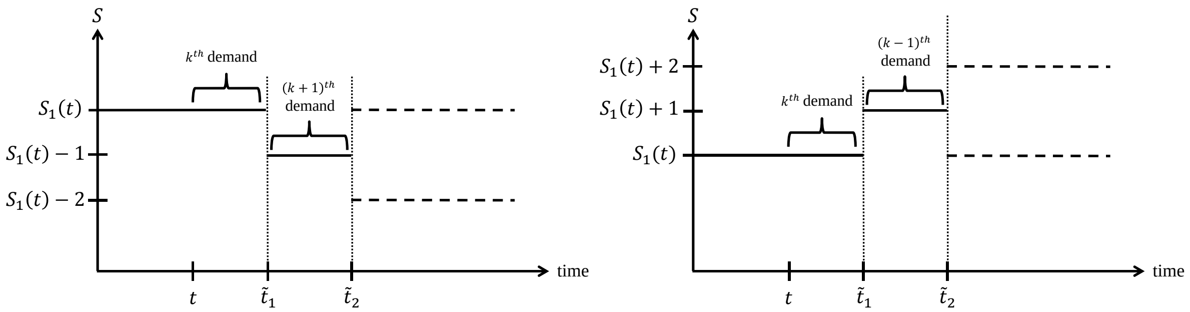


Figure 2.2: Decreasing and increasing step functions.

then one unit is ordered. If (b) the inventory position exceeded the prior base-stock level before the step, there is no instant ordering. The consequences are identical in both cases.

Corollary 2.2. *If the number of upward steps in S_1 on $[t, \tilde{t}_j]$ is greater than or equal to the number of downward steps, i.e. $\Delta_j \leq 0$, the distribution function $G_t^{k,0}(x)$ simplifies to*

$$G_t^{k,0}(x) = \mathbb{P}(T_t^{k+\Delta_j,1} \leq x), \;\; x \in [t_j, t_{j+1}]. \tag{2.27}$$

Proof. Follows directly from Proposition 2.3 and Lemma 2C.1, which is presented and proved in Section 2.C of the appendix. □

Corollary 2.3. *If S_1 is monotonically increasing on $[t, \tilde{t}_j]$, the distribution function $G_t^{k,0}(x)$ simplifies to*

$$G_t^{k,0}(x) = \mathbb{P}(T_t^{k-j,1} \leq x), \;\; x \in [t_j, t_{j+1}]. \tag{2.28}$$

2.5.4 Arrival Time Distribution at the Warehouse in a Multi-Retailer System

If we have an inventory system with one warehouse and N retailers, applying the decision rule (2.17) requires knowledge of the partial distribution functions $G_t^{k,0,i}(x)$ as defined in Section 2.5.2. The aim of this section is to define these partial distribution functions. Having N retailers requires combinatorial case distinctions. Therefore, we first define sets $\mathcal{M}_{t,i}^{k,j}$ that yield a partition of the probability space. Next, we use these sets to define auxiliary functions $U_{t,i}^k$ (similar to Section 2.5.3). These auxiliary functions finally allow for the determination of $G_t^{k,0,i}(x)$.

Having multiple retailers, we use $i \in I$ to indicate the partial distribution function $G_t^{k,0,i}(x)$ that we want to determine; that is, the k^{th} warehouse unit is shipped to retailer i. As this distribution requires knowledge about steps of all retailers, we also use $n \in I$ (in addition to $i \in I$) to account for all retailers (including retailer i).

Before the first step in a base-stock level of any retailer takes place, i.e. at time $x \in \bigcap_{n=1}^{N} [0, t_1^n)$, there exist k events at each retailer i that can trigger the k^{th} warehouse demand, namely any of the demands $1, 2, ..., k$. We find N^{k-1} disjoint sequences that constitute the event that unit k will be ordered by retailer i in time interval $[t, t + x]$.

If the arrival $T_t^{k,0}$ lies in $\bigcap_{n=1}^{N}[0, t_1^n)$, it can be triggered by a total number of $k \cdot N$ different causes, each being a customer demand. Let one index $m_n \in \{0, ..., M_n^t\}$ be fixed for each retailer n and let $t_{m_n}^n$ denote the time of the m_n^{th} step at retailer n after time t. Then, more generally, if the interval $\bigcap_{n=1}^{N}[t_{m_n}^n, t_{m_n+1}^n) = [\max_{n \in I} t_{m_n}^n, \min_{n \in I} t_{m_n+1}^n)$ is non-empty, $T_t^{k,0}$ may occur in this interval. That is, it can be triggered (a) by the $k \cdot N$ customer demands as above; (b) by the $\bar{k}^{th}$ customer demand at a retailer with $\bar{k} > k$ if there were downward steps at this retailer; (c) at the beginning of the interval and independent of a customer demand, as a consequence of an upward step at the particular retailer $\bar{n}$ with $t_{m_{\bar{n}}}^{\bar{n}} = \max_{n \in I} t_{m_n}^n$. That is, $T_t^{k,0}$ also depends on the sequence of upward and downward steps, i.e. on $\delta_{m_n}^n$, rather than only on the cumulative steps $\Delta_{m_n}^n$. We define the event that the k^{th} warehouse demand is triggered by the j^{th} customer demand after time t at retailer i and that it occurs in $[\max_{n \in I} t_{m_n}^n, x]$ by

$$\mathcal{M}_{t,i}^{k,j}\left(\Delta_{m_1}^1, ..., \Delta_{m_N}^N, \delta_{m_i}^i; x\right) = \left\{T_t^{k,0} \geq \max_{n \in I} t_{m_n}^n, T_t^{k,0} = T_t^{j,i} \leq x\right\}. \tag{2.29}$$

If it holds that $i = \bar{n}$ from (c), and the step at time $t_{m_i}^i$ is upward ($\delta_{m_i}^i < 0$), $T_t^{k,0}$ can be triggered by this upward step; we denote this as the 0^{th} demand at retailer i, although there could have been real demands at retailer i before. Further, $T_t^{k,0}$ can be triggered by any of the demands $1, ..., k + \Delta_{m_i}^i$ at retailer i after time t. That is, j in equation (2.29) ranges in set $J_i(x)$ with

$$J_i(x) = \begin{cases} \{0, 1, ..., k + \Delta_{m_i}^i\}, & \text{if } i = \bar{n},\ \delta_{m_i}^i < 0, \\ \{1, ..., k + \Delta_{m_i}^i\}, & \text{otherwise.} \end{cases}$$

Obviously, $x \in \bigcap_{n=1}^{N}[t_{m_n}^n, t_{m_n+1}^n)$ for exactly one combination $(m_1, ..., m_N)$. As our focus is not on the particular demand j that triggers the k^{th} warehouse demand, we will now focus on the union

$$\bigcup_{j \in J_i(x)} \mathcal{M}_{t,i}^{k,j}(\Delta_{m_1}^1, ..., \Delta_{m_N}^N, \delta_{m_i}^i; x). \tag{2.30}$$

As the events are disjoint, the probability mass is given by the sum of the respective probabilities. Similar to Section 2.5.3, we denote the resulting probabilities as *auxiliary functions* that will later be used for defining the arrival time cdf at the warehouse.

Definition 2.2 (Auxiliary Functions). *From the disjoint sets* $(\mathcal{M}_{t,i}^{k,j})$, *we define the* auxiliary functions *as*

$$U^k_{t,i,m_1,\dots,m_N}(x) = \begin{cases} \sum\limits_{j\in J_i(x)} \mathbb{P}\big(\mathcal{M}^{k,j}_{t,i}(\Delta^1_{m_1},\dots,\Delta^N_{m_N},\delta^i_{m_i};x)\big), & \text{if } x \in \bigcap_{n=1}^N [t^n_{m_n}, t^n_{m_n+1}), \\ \lim\limits_{\substack{x\to\bar{t}\\ x<\bar{t}}} \sum\limits_{j\in J_i(x)} \mathbb{P}\big(\mathcal{M}^{k,j}_{t,i}(\Delta^1_{m_1},\dots,\Delta^N_{m_N},\delta^i_{m_i};x)\big), & \text{if } x = \bar{t} := \min\limits_{n\in I} \ t^n_{m_n+1}, \\ 0, & \text{if } x \notin \bigcap_{n=1}^N [t^n_{m_n}, t^n_{m_n+1}]. \end{cases} \tag{2.31}$$

If there is an upward step at the upper bound of the intersection and $x = \min\limits_{n\in I} \ t^n_{m_n+1}$, its positive probability mass is assigned to the adjacent interval. This is stated by the second line in equation (2.31). Next, we determine the partial distribution functions $G^{k,0,i}_t(x)$. We sum over the probability masses of all non-empty intersections, i.e. our auxiliary functions, of the intervals to the left of x.

Proposition 2.4. *For $x \in \bigcap_{n=1}^N [t^n_{p_n}, t^n_{p_n+1}) \neq \emptyset$ with $p_n \in \{0,\dots,M^t_n\}$, $\forall\ n \in I$, we find the partial distribution functions as*

$$G^{k,0,i}_t(x) = \sum_{m_1=0}^{p_1} \cdots \sum_{m_N=0}^{p_N} U^k_{t,i,m_1,\dots,m_N}\Big(\min\big\{\min_{n\in I}\{t^n_{m_n+1}\}, x\big\}\Big). \tag{2.32}$$

Proof. The above derivation proves the proposition. □

Since our decision rule only requires knowledge of $G^{k,0,i}_t(x)$ for $x \in [0, L_0]$, we do not need to determine these functions for $x \in (L_0, \infty)$. This is why, for our purposes, there are many cases where only very few (0, 1, or 2) steps of the retailer step functions need to be considered. We present explicit expressions of $G^{k,0,i}_t(x)$ for the two-retailer case in Section 2.D of the appendix.

2.5.5 Cost Evaluation

For the evaluation of expected costs, we consider a finite horizon. Having a one-warehouse multi-retailer problem with stationary λ, the expected average costs of a policy $(S_0, S_1, \dots, S_N)$ can be found in Axsäter (2006). However, if λ is time-dependent, this requires adjustment. The total expected costs of a policy $(S_0(t), S_1(t), \dots, S_N(t))$

over time $[0, T]$ are given through

$$TEC = \int_0^T \Big(h_0^e \mathbb{E}[IL_0^{e+}(t)] + \sum_{i=1}^{N} \big(h_i^e \mathbb{E}[IL_i^+(t)] + b_i \mathbb{E}[IL_i^-(t)] \big) \Big) \mathrm{d}t, \tag{2.33}$$

with *echelon* holding costs $h_i^e := h_i - h_0$, $h_0^e := h_0$, *echelon* warehouse inventory level $IL_0^e(t)$, and retailer inventory level $IL_i(t)$. For brevity, we use $x^+ = \max\{x, 0\}$, and $x^- = -\min\{x, 0\}$ in (2.33). In a non-stationary setting, inventory positions are stochastic processes. As long as there is no decrease in $S(t)$, it holds that $IP(t) = S(t)$, since we can always order up to $S(t)$. However, as soon as $S(t)$ decreases, we might have excess stocks (or *overstocks*) $OV(t)$ that follow a stochastic process. Therefore, the inventory position at time t is

$$IP(t) = S(t) + OV(t) \geq S(t).$$

In order to determine $\mathbb{E}[IL_0^{e+}(t)]$, $\mathbb{E}[IL_i^+(t)]$, and $\mathbb{E}[IL_i^-(t)]$, we require the distributions of the inventory positions of the warehouse and the retailers (i.e. the distributions of the overstocks), and the distributions of retailer backorders at the warehouse. We present the formulas for the expected values and derive the respective distributions in Section 2.E of the appendix.

2.6 Numerical Study

2.6.1 Results for Serial Systems

We consider the demand rates $\lambda(t)$ and $\hat{\lambda}(t)$ from Section 2.4.3. We consider a finite horizon for lifecycle demand and set $d = 0$ in $\hat{\lambda}(t)$. We vary the parameters as in Section 2.4.3, along with h_0 and L_0. This leads to a full factorial design with

$$a \in \{0.3, 1\}, \quad c \in \left\{\frac{1}{3}, \frac{2}{3}\right\}, \quad \hat{a} \in \{1, 2\}, \quad \hat{c} \in \{0.3, 0.6\},$$
$$h_0 \in \{0.3, 0.7\}, \quad h_1 = 1, \quad b \in \{5, 15\}, \quad L_0 \in \{0.5, 1\}, \quad L_1 \in \{0.5, 1\}.$$

We have 64 different scenarios each for seasonal and lifecycle demand.

In order to assess the value of our policy from unit-tracking (UT), we compare it to two alternative policies. The first alternative is TD that applies the two-stage stationary

Table 2.3: Results for serial system with seasonal demand.

				$a=1,\ c=\frac{1}{3}$		$a=0.3,\ c=\frac{1}{3}$		$a=1,\ c=\frac{2}{3}$		$a=0.3,\ c=\frac{2}{3}$	
L_0	L_1	h_0	b	$\hat{\Delta}$ (%)	$\tilde{\Delta}$ (%)	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$
0.5	0.5	0.3	5	2.32	0.00	0.34	0.08	8.26	0.41	0.46	0.05
			15	3.02	0.41	0.17	0.05	10.53	1.29	0.70	0.05
		0.7	5	2.81	0.03	0.40	0.05	10.80	3.14	0.10	0.05
			15	3.54	0.17	0.53	0.03	10.24	0.23	0.71	0.00
	1	0.3	5	7.34	0.00	0.40	0.19	21.93	3.23	3.49	2.40
			15	8.81	0.03	0.77	0.00	27.12	1.88	1.61	0.03
		0.7	5	7.32	0.68	0.52	0.00	15.35	1.79	2.16	0.26
			15	9.07	0.11	1.75	1.15	23.52	2.23	4.62	3.80
1	0.5	0.3	5	6.69	0.19	0.30	0.01	18.55	1.63	0.00	0.10
			15	8.25	0.11	0.25	0.02	26.56	3.37	1.44	0.00
		0.7	5	7.30	0.40	0.59	0.01	17.91	4.18	1.66	0.27
			15	9.32	0.10	0.79	0.31	26.54	3.38	4.33	3.21
	1	0.3	5	12.99	0.10	0.77	0.03	27.17	5.56	1.97	0.02
			15	18.05	1.02	1.18	0.05	40.32	8.06	3.52	1.84
		0.7	5	13.06	0.77	0.80	0.22	21.26	7.84	3.98	4.61
			15	19.35	1.01	0.93	0.01	32.40	9.74	0.13	4.04

model at each point of time t with $\lambda \equiv \lambda(t)$. The second alternative is HL that shifts the base-stock levels from TD to the left. That is, we shift the base-stock levels from TD by $\frac{L_1}{2}$ and $\frac{L_0+L_1}{2}$, respectively. Denoting the solution from TD by $(\hat{S}_1, \hat{S}_0^e)$, the base-stock policy of HL is formally given by

$$\tilde{S}_1(t) = \hat{S}_1\Big(t + \frac{L_1}{2}\Big), \quad \tilde{S}_0^e(t) = \hat{S}_0^e\Big(t + \frac{L_1 + L_0}{2}\Big). \tag{2.34}$$

For seasonal demand, we evaluate the expected costs over one season. For lifecycle demand, we evaluate the expected costs from 0 to T, where T is the time when the three policies first match at the end of the lifecycle. Tables 2.3 and 2.4 show the results. $\hat{\Delta}$ and $\tilde{\Delta}$ give the relative differences in expected costs between $(\hat{S}_1, \hat{S}_0^e), (\tilde{S}_1, \tilde{S}_0^e)$ and UT, respectively.

Comparing UT to TD, we see that the TD solution performs significantly worse with up to 41.86% in additional costs. The average cost difference over the 64 cases of lifecycle demand is 13.66%; over the 64 cases of seasonal demand, it is 8.11%.

Looking more closely at TD's expected costs, their high value can be explained by the fact that base-stock levels are increased too late and cause a lot of backordering. Consequently, in *most* of the 128 cases, especially for cases with lifecycle demand, the TD solution differs more significantly from UT if backorder costs are 15 than in the identical situations when backorder costs are only 5.

Considering the sensitivities of the cost differences with respect to lead times, the relative cost differences increase in both L_0 and L_1. This is in line with our observation on backorder costs: The longer the lead times, the larger the lateness of the ordering of TD.

In the case of seasonal demand (Table 2.3), the cost differences between TD and UT decrease in the amplitude of $\lambda(t)$, i.e. the smaller the changes in the demand rate ($a = 0.3$ vs. $a = 1$), the smaller the advantage of UT. This is a rather intuitive result; in particular, if demand is stationary, both UT and TD yield the optimal policy. Further, the cost differences between TD and UT decrease if the season length increases (season length 3 months ($c = \frac{2}{3}$) vs. 6 months ($c = \frac{1}{3}$)).

In the case of lifecycle demand (Table 2.4), the differences between TD and UT increase with the shortening of the product lifecycle ($\hat{c} = 0.6$ yields a shorter lifecycle than $\hat{c} = 0.3$).

HL performs very well. For 59 out of the 64 cases with lifecycle demand, the deviation of HL from UT is less than 1%. A similar observation can be made for seasonal demand with $c = \frac{1}{3}$, i.e. a season length of six months, where, in 29 of the 32 cases, HL deviates less than 1% from UT. However, for $c = \frac{2}{3}$, i.e. a season length of three months, the results deteriorate. We know that HL produces the same results as UT at the retailer for a linear demand rate (Theorem 2.2). The reason for its worse performance in the case of shorter periods is the increased curvature that moves us further away from linearity. Nonetheless, HL offers us an alternative that (i) yields very good results and (ii) is easy to implement as it is based on the stationary model.

2.6.2 Results for One-Warehouse Two-Retailer Systems

We now study the case of $N = 2$ identical retailers both of whom face seasonal or lifecycle demand and apply the full factorial design from Section 2.6.1. Having identical lead times for both retailers, $L_1 = L_2$, we again define the alternative policies TD and

Table 2.4: Results for serial system with lifecycle demand.

				$\hat{a}=2,\ \hat{c}=0.6$		$\hat{a}=1,\ \hat{c}=0.6$		$\hat{a}=2,\ \hat{c}=0.3$		$\hat{a}=1,\ \hat{c}=0.3$	
L_0	L_1	h_0	b	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$
0.5	0.5	0.3	5	5.49	0.04	3.20	0.02	3.91	0.06	2.26	0.02
			15	10.01	0.02	5.56	0.09	6.31	0.10	3.83	0.05
		0.7	5	5.07	0.08	1.90	0.19	3.34	0.00	1.88	0.05
			15	12.75	1.07	5.65	0.06	6.20	0.05	4.11	0.02
	1	0.3	5	18.65	0.44	10.03	0.28	9.38	0.78	8.22	0.15
			15	36.34	0.27	22.62	0.43	23.73	0.79	14.67	0.32
		0.7	5	14.80	0.50	8.08	2.15	5.10	0.28	6.34	0.20
			15	31.01	0.36	23.38	1.62	14.35	0.40	13.03	0.07
1	0.5	0.3	5	15.49	0.15	9.36	0.02	9.28	0.15	6.74	0.08
			15	24.15	0.13	16.14	0.34	13.86	0.18	12.26	0.11
		0.7	5	15.25	0.31	8.77	0.26	7.08	0.09	6.98	0.10
			15	26.20	1.95	20.82	0.25	3.83	0.24	11.51	0.04
	1	0.3	5	25.81	0.68	20.99	0.21	4.79	0.60	12.44	0.28
			15	41.12	0.56	38.15	0.25	16.43	0.50	20.72	0.47
		0.7	5	19.21	0.73	19.94	0.92	3.37	0.14	9.59	0.30
			15	32.17	0.83	41.86	1.94	2.75	0.07	15.99	0.19

Table 2.5: Results for one-warehouse two-retailer system with lifecycle demand.

				$\hat{a}=2,\ \hat{c}=0.6$		$\hat{a}=1,\ \hat{c}=0.6$		$\hat{a}=2,\ \hat{c}=0.3$		$\hat{a}=1,\ \hat{c}=0.3$	
L_0	L_1	h_0	b	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$
0.5	0.5	0.3	5	7.66	0.05	4.05	0.07	6.80	0.27	4.02	0.51
			15	13.86	0.48	8.50	0.05	12.47	0.27	7.43	0.85
		0.7	5	5.44	0.05	2.35	0.46	4.92	0.74	2.16	0.20
			15	15.01	0.08	8.73	0.27	10.33	0.38	4.98	0.38
	1	0.3	5	19.98	0.05	11.06	0.00	18.95	0.00	10.16	0.09
			15	40.90	1.14	23.60	0.50	34.78	0.31	19.59	0.25
		0.7	5	16.28	0.19	5.38	0.33	12.25	0.12	7.55	0.41
			15	34.33	0.35	21.17	0.15	25.86	0.27	16.38	0.31
1	0.5	0.3	5	25.44	1.77	12.89	0.88	24.56	0.28	13.82	0.16
			15	45.28	2.27	23.07	0.28	44.92	1.77	26.00	0.64
		0.7	5	19.72	1.13	11.04	0.96	14.15	0.15	10.91	0.97
			15	41.52	2.78	26.35	1.43	29.36	0.68	20.60	0.64
	1	0.3	5	38.24	0.14	23.00	0.21	32.10	0.43	24.10	0.39
			15	71.35	1.70	43.67	0.30	60.04	0.16	43.87	0.14
		0.7	5	30.06	1.13	19.81	0.51	17.07	0.54	14.92	0.84
			15	57.23	2.51	41.02	1.57	36.27	0.60	29.38	0.05

Table 2.6: Results for one-warehouse two-retailer system with seasonal demand.

				$a=1,\ c=\frac{1}{3}$		$a=0.3,\ c=\frac{1}{3}$		$a=1,\ c=\frac{2}{3}$		$a=0.3,\ c=\frac{2}{3}$	
L_0	L_1	h_0	b	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$	$\hat{\Delta}$	$\tilde{\Delta}$
0.5	0.5	0.3	5	4.04	0.50	0.12	0.04	12.97	0.11	0.39	0.05
			15	3.36	0.70	0.23	0.46	14.51	0.03	0.48	0.24
		0.7	5	8.48	2.10	3.12	2.59	16.04	4.84	4.31	2.86
			15	5.81	0.37	2.28	2.05	16.51	0.57	12.47	11.93
	1	0.3	5	9.32	0.97	1.27	0.13	21.86	3.51	3.34	1.15
			15	11.88	1.40	1.17	0.16	30.79	3.20	3.26	0.64
		0.7	5	10.15	0.95	3.94	2.94	17.60	2.66	4.68	0.26
			15	13.06	0.47	2.58	1.38	27.19	3.15	4.01	1.48
1	0.5	0.3	5	8.04	0.11	0.37	0.10	7.89	3.02	2.01	1.29
			15	7.16	1.21	0.32	0.06	8.16	0.02	1.55	0.12
		0.7	5	15.85	9.82	10.74	9.96	18.82	12.90	5.04	4.54
			15	10.31	1.95	3.45	3.03	12.05	6.84	3.08	3.48
	1	0.3	5	11.86	0.52	2.26	0.99	7.17	9.66	3.08	2.47
			15	12.62	2.22	0.48	0.33	9.84	8.47	1.57	0.36
		0.7	5	13.84	7.57	8.30	7.98	5.89	12.61	8.62	10.58
			15	16.81	5.73	7.17	6.05	8.13	11.93	6.37	6.72

HL by base-stock parameters $(\hat{S}_0^e, \hat{S}_1, \hat{S}_2)$ and $(\tilde{S}_0^e, \tilde{S}_1, \tilde{S}_2)$ with

$$\tilde{S}_i(t) = \hat{S}_i\Big(t + \frac{L_i}{2}\Big),\ i \in \{1,2\}, \quad \tilde{S}_0^e(t) = \hat{S}_0^e\Big(t + \frac{L_1 + L_0}{2}\Big). \tag{2.35}$$

Looking at the results for lifecycle demand in Table 2.5, we make identical observations as in the serial case. The cost differences between TD and UT increase in the backorder costs and in both lead times L_0 and L_1, and the shorther the lifecycle, the more advantageous is UT. HL again performs very well in most of the cases.

Considering seasonal demand in Table 2.6, the cost differences between TD and UT, too, increase for most cases in both lead times L_0 and L_1. However, this does not hold for the case with large demand variation and a season length of 3 months ($a = 1,\ c = \frac{2}{3}$) when the lead time L_0 changes from 0.5 to 1. In this case, cost differences decrease, which can be explained by the fact that, in the one-warehouse two-retailer system with lead time $L_0 = 1$, the TD solution builds up much more system stock at certain points in time than in the case with lead time $L_0 = 0.5$ or the serial system with lead time

$L_0 = 1$. This high stock level, however, now results in comparably lower costs due to the short season length of 3 months. The left-over stocks from one season reduce the backordering of the subsequent one.

A further consequence of the high stock levels of TD is the sensitivity to warehouse holding costs h_0. As opposed to the serial system, the relative cost differences between TD and UT in Table 2.6 now increase with the warehouse holding costs h_0.

2.7 Conclusion

This chapter makes a contribution by offering a thorough analysis of this type of problem with large practical application potential in spare parts. We believe that our approach to analyze inventory systems is important with regard to further applications. Our results include: (i) We found a decision rule that yields the optimal policy of a single-stage inventory system with non-homogeneous Poisson demand. (ii) We found that the arrival of demands at the warehouse is influenced by the ordering policies of the retailers, and no longer Poisson distributed: we characterized the actual arrival time distributions. (iii) We derived a myopic decision rule for time-dependent base-stock levels in one-warehouse multi-retailer systems and found large improvements over a time decomposition solution. (iv) We found a simple and effective alternative: the repeated application of the stationary model is promising if handled properly; that is, the solution needs to be shifted by a half lead time. As finding stationary solutions is simple, this is of particular interest for practice.

Possibilities for future research include the extension of the demand structure to non-homogeneous compound Poisson demand or the optimization of batch-ordering policies with time-dependent review periods.

2.A Cost of Passive Ordering

If the first step is an upward step, the cost of passive ordering is found by (2.7) for all cases. Therefore, it remains to consider cases where the first step is downward. We present formulas for two steps after time t, where the second step is either upward or downward. The extensions to more than two steps are analogous and omitted. We further only consider the case $k \geq 3$; for $k < 3$, we only require to plug-in bL instead of C whenever it is necessary. In the case of two downward steps, we find the cost of passive ordering as

$$\begin{aligned}\hat{C}_t(k) = &\int_0^{t_1} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s \\ &+ \int_{t_1}^{t_2} g_t^1(s)\Bigg(\int_0^{t_2-s} g_{t+s}^1(\tau) C_{t+s+\tau}(k-2)\mathrm{d}\tau \\ &\qquad + \int_{t_2-s}^{\infty} g_{t+s}^1(\tau)\Big(\int_0^{\infty} g_{t+s+\tau}^1(x) C_{t+s+\tau+x}(k-3)\mathrm{d}x\Big)\mathrm{d}\tau\Bigg)\mathrm{d}s \\ &+ \int_{t_2}^{\infty} g_t^1(s)\Big(\int_0^{\infty} g_{t+s}^2(\tau) C_{t+s+\tau}(k-3)\mathrm{d}\tau\Big)\mathrm{d}s.\end{aligned} \tag{2.36}$$

That is, if the first demand after time t arrives prior to t_1, it is immediately reordered (line 1). If it arrives after t_1, it is not reordered. If, in the latter scenario, the second demand arrives prior to t_2, it is immediately reordered (line 2). If it arrives after t_2, it is not reordered. In the latter case, unit k will be triggered by the third demand after time t (line 3). If the first demand arrives after time t_2, unit k will also be triggered by the third demand after time t (line 4).

If the first step is downward and the second is upward, we find the cost of passive ordering as

$$\begin{aligned}\hat{C}_t(k) = &\int_0^{t_1} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s \\ &+ \int_{t_1}^{t_2} g_t^1(s)\Bigg(\int_0^{t_2-s} g_{t+s}^1(\tau) C_{t+s+\tau}(k-2)\mathrm{d}\tau \\ &\qquad + \big(1 - G_{t+s}^1(t_2 - s)\big) C_{t+t_2}(k-1)\Bigg)\mathrm{d}s \\ &+ \int_{t_2}^{\infty} g_t^1(s) C_{t+s}(k-1)\mathrm{d}s.\end{aligned} \tag{2.37}$$

In (2.37), if the first demand after time t arises at time $s \in [t, t+t_1) \cup [t+t_2, \infty)$, unit k's expected costs are $C_{t+s}(k-1)$ as in (2.5). If the first demand takes place between $t+t_1$ and $t+t_2$, unit k is either a reorder for unit 2 and triggered by the second demand after time t, if this demand takes place before $t+t_2$. Or it is ordered at time $t+t_2$ in order to adjust the inventory position to the increased base-stock level.

2.B Proofs

2.B.1 Proof of Theorem 2.1

Following Muharremoglu and Tsitsiklis (2008), we can optimize the total expected cost of the system by optimizing the expected cost of every single unit in the system. We consider the decision at time t; we assume to know the optimal base-stock level $S(\tau)$ for $\tau > t$. We divide time into infinitessimally small units Δt. Thus, $S(t+\Delta t)$ denotes the first optimal base-stock level after time t. In all following cases and subcases, there is always the opportunity (i) to order unit k immediately (and actively), or (ii) not to do so. While the total expected costs for immediate ordering of unit k remain the same throughout all cases ($C_t(k)$), the costs of the alternative vary. In the following, we consider unit k at time t and present all that can happen to this particular unit if it is not ordered immediately: what can trigger its order; when it is ordered; the mode of its ordering (active/passive).

We distinguish the two cases: (a) $k \leq S(t+\Delta t)$, and (b) $k = S(t+\Delta t)+1$. (There is no need to consider $k > S(t+\Delta t)$ since the step size is always one.) Further, we consider the two subcases for (b): (1) the first step in S after time t is upward, and (2) the first step in S after time t is downward (or there is no step at all). For (a), these subcases are not relevant. In the following, we present the distinctions for unit k.

(a) If $k \leq S(t+\Delta t)$, the alternatives to ordering unit k immediately are:

(i) It is ordered (*passively*) as the substitute for unit 1, i.e. when the first demand after time t takes place, if this happens prior to $t+\Delta t$ (the time of the first upward step if k $\big(\leq S(t+\Delta t)\big)$ is not ordered at t). The probability of this event converges to zero as $\Delta t \to 0$.

(ii) It is ordered (*actively*) at time $t+\Delta t$ to raise the base-stock level from $k-1$ to $S(t+\Delta t) > k-1$. That is, the decision is whether to order unit k (*actively*) at time t or (*actively*) at time $t+\Delta t$. This decision is obtained by the decision rule for active ordering $G_t^k(L) > \frac{h}{b+h}$.

(b) If $k = S(t+\Delta t)+1$, the alternatives for ordering unit k immediately are:

(1&2) It is ordered (*passively*) as the substitute for unit 1, i.e. when the first demand after time t takes place, if this happens prior to the first step after time t.

(1) It is ordered (*actively*) to raise the base-stock level at the time of the first upward step, if there is no demand before that time.

(2.1) It is ordered (*passively*) as the substitute for unit 2, if (i) the first demand after time t takes place after the first downward step, and (ii) the first and the second demand after time t take place between the first downward step and the subsequent step.

(2.2) It is ordered (*passively*) as the substitute for unit 1, if the first demand after time t takes place after both (i) the first downward step and (ii) the subsequent upward step.

(2.3) It is ordered (*actively*) at the time of the second step after time t, if (i) the first step is a downward step, (ii) the second step is an upward step, and (iii) there was no demand prior to the first step and exactly one demand between the first and the second step.

(2.4) It is ordered (*passively*) as the substitute for unit 3, if (i) the first two steps are downward steps, (ii) the first demand after time t takes place after the first downward step, and (iii) the second demand after time t takes place after the second downward step.

The remaining case distinctions (for cases with more than two steps and demand taking place in the proper intervals) are similar and not given. The distinctions form a partition of the probability space and cover all possibilities for unit k. This includes that it may never be ordered actively; however, it can always be ordered passively if sufficiently many demands take place. Being actively or passively ordered constitutes all possibilities for unit k. Therefore, the rule in Theorem 2.1 yields the optimal decision. □

2.B.2 Proof of Proposition 2.1

Lemma 2B.1. *If $\lambda(x)$ is bounded and piecewise continuous, then*

(a) $\Lambda(x)$ is continuous,

(b) $\Lambda_t(x) = \Lambda(t+x) - \Lambda(t)$ is continuous in t,

(c) $G_t^k(x)$ is continuous in t.

Proof. (a) follows directly from the definition of $\Lambda(x)$, (b) follows from the fact that sums of continuous functions are continuous and (c) follows from (b) as composition of continuous functions. □

Lemma 2B.1 allows us to prove Proposition 2.1.

Proposition 2.1. *Let $\lambda(t)$ be piecewise continuous and bounded. Then the steps of the retailer base-stock levels are always of size 1.*

Proof. There are two cases: (i) Active ordering is prefered over passive ordering for all k, i.e. $C_t(k) < \hat{C}_t(k)$ for all k with $G_t^k(L) > \frac{h}{b+h}$. (ii) Passive ordering is prefered for $k \geq \bar{k}$, i.e. there exists $\bar{k}$ with $G_t^{\bar{k}}(L) > \frac{h}{b+h}$ but $C_t(\bar{k}) \geq \hat{C}_t(\bar{k})$. We define

$$G^k(t) := G_t^{k,i}(L_1) - \frac{h_i - h_0}{b_i + h_i}.$$

We consider (i) and prove the proposition for downward steps. Let $G^k(t) > 0$, $G^k(t+\Delta) < 0$ and w.l.o.g. $G^{k+1}(x) < 0$, $\forall x \in [t, t+\Delta]$. For interarrival times of a Poisson process it always holds $G^{k-1}(t) > G^k(t)$ and with Lemma 2B.1, i.e. the continuity of

$G^k(t)$, and the intermediate value theorem $\exists\, t^* \in [t, t+\Delta] : G^k(t^*) = 0$. Then on $[t, t^*)$ we find $S = k$. With $G^{k-1}(t^*) > G^k(t^*) = 0$ it follows $S(t^*) = k-1$. The proof for upward steps is analogous. (ii) is proved similarly; we only give an outline. If (ii) holds, the additional holding costs from *actively instead of passively* ordering unit $\bar{k}$ exceed the additional backorder costs from *passively instead of actively* ordering unit $\bar{k}$. The additional backorder costs for passive instead of active ordering are decreasing in k; the additional holding costs of active instead of passive ordering are increasing in k. That is, $\hat{C}_t(k) - C_t(k)$ is decreasing in k. The remainder follows with the arguments on continuity from (i). □

2.C Arrival Time Distributions in Serial Systems

2.C.1 Simplifications for Upward Steps

Lemma 2C.1. *If an upward step in $S_1(t)$ follows a downward step, say at times $\tilde{t}_{j-1}$ and $\tilde{t}_j$, then*

$$U^k_{j-2}(t_{j-1}) + U^k_{j-1}(t_j) + U^k_j(x) = U^k_{j-2}(x),\ \ x \in [t_j, t_{j+1}]. \tag{2.38}$$

More generally, whenever $\Delta_k = \Delta_i$, $k < i$ and $\Delta_{k+1}, ..., \Delta_{i-1}$ are larger than Δ_k, the same simplification follows.

Proof. We only prove this proposition for the case $j = 2$, i.e. the first step in $S_1(t)$ is downward, the second step is upward and $x \in [t_2, t_3]$. For all other cases, this can be done analogously.

$$\begin{aligned}
&U^k_0(t_1) + U^k_1(t_2) + U^k_2(x) = \mathbb{P}(T^k_t \le t_1) + \mathbb{P}(T^{k+1}_t \le t_2, T^k_t > t_1) + \mathbb{P}(T^k_t \le x, T^{k+1}_t > t_2, T^k_t > t_1) \\
&= \mathbb{P}(T^k_t \le t_1) + \mathbb{P}(T^{k+1}_t \le t_2, t_1 < T^k_t \le t_2) + \mathbb{P}(t_1 < T^k_t \le x, T^{k+1}_t > t_2) \\
&= \mathbb{P}(T^k_t \le t_1) + \mathbb{P}(T^{k+1}_t \le t_2, t_1 < T^k_t \le t_2) + \mathbb{P}(t_1 < T^k_t \le x) - \mathbb{P}(t_1 < T^k_t \le x, T^{k+1}_t \le t_2) \\
&= \mathbb{P}(T^k_t \le t_1) + \mathbb{P}(T^{k+1}_t \le t_2, t_1 < T^k_t \le t_2) + \mathbb{P}(t_1 < T^k_t \le x) - \mathbb{P}(T^{k+1}_t \le t_2, t_1 < T^k_t \le t_2) \\
&= \mathbb{P}(T^k_t \le t_1) + \mathbb{P}(t_1 < T^k_t \le x) = \mathbb{P}(T^k_t \le x) = U^k_0(x).
\end{aligned}$$

It has been used that if the $(k+1)^{th}$ demand arrives before time t_2, the k^{th} demand cannot arrive after t_2. □

2.C.2 Representation of Auxiliary Functions

We now derive $U_j^k(x)$. If we have an upward step at $\tilde{t}_1$, we find

$$U_1^k(x) = U_0^{k-1}(x) - U_0^k(t_1), \quad x \geq t_1,$$

and the corresponding pdf

$$u_1^k(x) = \frac{\mathrm{d}}{\mathrm{d}x} U_1^k(x), \text{ if } x > t_1, \quad u_1^k(t_1) = U_0^{k-1}(t_1) - U_0^k(t_1).$$

Note that $T_t^{k,1}$, $k \in \{1, 2, ...\}$, are *not* independent, but $T_t^{k,1}$ and $(T_t^{k+1,1} - T_t^{k,1})$ are independent increments. If we have a downward step at $\tilde{t}_1$, for $x \geq t_1$ we find

$$\begin{aligned} U_1^k(x) &= \mathbb{P}(T_t^{k+1,1} \leq x, T_t^{k,1} > t_1) = \mathbb{P}(T_t^{k,1} + (T_t^{k+1,1} - T_t^{k,1}) \leq x, T_t^{k,1} > t_1) \\ &= \int_{t_1}^{x} \int_0^{x-s_1} g_t^{k,1}(s_1) g_{t+s_1}^{1,1}(s_2) \mathrm{d}s_2 \mathrm{d}s_1, \end{aligned}$$

and the corresponding pdf

$$u_1^k(x) = \int_{t_1}^{x} g_t^{k,1}(s_1) g_{t+s_1}^{1,1}(x - s_1) \mathrm{d}s_1.$$

For $j > 1$, $x > t_j$ and an upward step at $\tilde{t}_j$ we find

$$\begin{aligned} U_j^k(x) &= U_{j-1}^{k-1}(x) - U_{j-1}^k(t_j), \quad x \geq t_1 \\ u_j^k(x) &= \frac{\mathrm{d}}{\mathrm{d}x} U_j^k(x), \text{ if } x > t_1, \quad u_j^k(t_1) = U_{j-1}^{k-1}(t_j) - U_{j-1}^k(t_j). \end{aligned}$$

For $j > 1$, $x > t_j$ and a downward step at $\tilde{t}_j$ we find

$$\begin{aligned} U_j^k(x) &= \int_{t_j}^{x} \int_0^{x-s_1} u_{j-1}^k(s_1) g_{t+s_1}^{1,1}(s_2) \mathrm{d}s_2 \mathrm{d}s_1, \\ u_j^k(x) &= \int_{t_j}^{x} u_{j-1}^k(s_1) g_{t+s_1}^{1,1}(x - s_1) \mathrm{d}s_1. \end{aligned}$$

2.D Arrival Time Distributions in One-Warehouse Two-Retailer Systems

We present special cases for $N = 2$ and denote the two retailers by A and B for ease of notation. We give $G_t^{k,0,A}(x)$ for three explicit cases in terms of joint probabilities:

(a) Neither S_A nor S_B have steps:

$$G_t^{k,0,A}(x) = \sum_{j=1}^{k} \mathbb{P}(T_t^{k,0} = T_{t,A}^{j} \leq x, T_{t,A}^{j} \leq T_{t,B}^{k-j+1}, T_{t,A}^{j} > T_{t,B}^{k-j}),\ \forall\ x \in [0, \infty)$$

(b) S_A has one step at t_1^A, S_B has no step:

$$G_t^{k,0,A}(x) = \begin{cases} \sum_{j=1}^{k} \mathbb{P}(T_t^{k,0} = T_{t,A}^{j} \leq x, T_{t,A}^{j} \leq T_{t,B}^{k-j+1}, T_{t,A}^{j} > T_{t,B}^{k-j}), & x \in [0, t_1^A), \\ \sum_{j=1}^{k} \mathbb{P}(T_t^{k,0} = T_{t,A}^{j} \leq t_1^A, T_{t,A}^{j} \leq T_{t,B}^{k-j+1}, T_{t,A}^{j} > T_{t,B}^{k-j}) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = T_{t,A}^{j+\Delta_1^A} \leq x, T_{t,A}^{j+\Delta_1^A} \leq T_{t,B}^{k-j+1}, & \\ \qquad T_{t,A}^{j+\Delta_1^A} > T_{t,B}^{k-j}, T_{t,A}^{j} > t_1^A, T_{t,A}^{j+\Delta_1^A} > t_1^A\big) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = t_1^A, t_1^A \leq T_{t,B}^{k-j+1}, & \\ \qquad t_1^A > T_{t,B}^{k-j}, T_{t,A}^{j+\Delta_1^A} < t_1^A, T_{t,A}^{j} > t_1^A\big), & x \in [t_1^A, \infty). \end{cases}$$

(c) S_B has one step at t_1^B, S_A has no step:

$$G_t^{k,0,A}(x) = \begin{cases} \sum_{j=1}^{k} \mathbb{P}(T_t^{k,0} = T_{t,A}^{j} \leq x, T_{t,A}^{j} \leq T_{t,B}^{k-j+1}, T_{t,A}^{j} > T_{t,B}^{k-j}), & x \in [0, t_1^B), \\ \sum_{j=1}^{k} \mathbb{P}(T_t^{k,0} = T_{t,A}^{j} \leq t_1^B, T_{t,A}^{j} \leq T_{t,B}^{k-j+1}, T_{t,A}^{j} > T_{t,B}^{k-j}) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = T_{t,A}^{j} \leq x, T_{t,A}^{j} \leq T_{t,B}^{k-j+1+\Delta_1^B}, & \\ \qquad T_{t,A}^{j} > T_{t,B}^{k-j+\Delta_1^B}, T_{t,A}^{j} > t_1^B, T_{t,B}^{k-j} > t_1^B\big) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = T_{t,A}^{j} \leq x, T_{t,A}^{j} \leq T_{t,B}^{k-j+1+\Delta_1^B}, & \\ \qquad T_{t,A}^{j} > T_{t,B}^{k-j}, T_{t,A}^{j} > t_1^B, T_{t,B}^{k-j} \leq t_1^B\big), & x \in [t_1^B, \infty). \end{cases}$$

If we include information on the type of step, upward or downward, the representation simplifies.

Corollary 2.4. *Considering the type of step (upward or downward), we give two examples:*

(a) S_A has a downward step at t_1^A, S_B has no step:

$$G_t^{k,0,A}(x) = \begin{cases} \sum_{j=1}^k \mathbb{P}(T_t^{k,0} = T_{t,A}^j \le x, T_{t,A}^j \le T_{t,B}^{k-j+1}, T_{t,A}^j > T_{t,B}^{k-j}), & x \in [0, t_1^A), \\ \sum_{j=1}^k \mathbb{P}(T_t^{k,0} = T_{t,A}^j \le t_1^A, T_{t,A}^j \le T_{t,B}^{k-j+1}, T_{t,A}^j > T_{t,B}^{k-j}) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = T_{t,A}^{j+1} \le x, T_{t,A}^{j+1} \le T_{t,B}^{k-j+1}, & \\ \qquad T_{t,A}^{j+1} > T_{t,B}^{k-j}, T_{t,A}^j > t_1^A, T_{t,A}^{j+1} > t_1^A\big), & x \in [t_1^A, \infty). \end{cases}$$

(b) S_B has an upward step at t_1^B, S_A has no step:

$$G_t^{k,0,A}(x) = \begin{cases} \sum_{j=1}^k \mathbb{P}(T_t^{k,0} = T_{t,A}^j \le x, T_{t,A}^j \le T_{t,B}^{k-j+1}, T_{t,A}^j > T_{t,B}^{k-j}), & x \in [0, t_1^B), \\ \sum_{j=1}^k \mathbb{P}(T_t^{k,0} = T_{t,A}^j \le t_1^B, T_{t,A}^j \le T_{t,B}^{k-j+1}, T_{t,A}^j > T_{t,B}^{k-j}) & \\ \quad + \mathbb{P}\big(T_t^{k,0} = T_{t,A}^j \le x, T_{t,A}^j \le T_{t,B}^{k-j}, & \\ \qquad T_{t,A}^j > T_{t,B}^{k-j-1}, T_{t,A}^j > t_1^B, T_{t,B}^{k-j} > t_1^B\big), & x \in [t_1^B, \infty). \end{cases}$$

The expressions simplify because for all $j \in \{1, .., k\}$ one summand vanishes in both (a) and (b), but not the same one.

2.E Distributions for the Cost Evaluation

In the following, we derive the required expected values and distributions for determining the total expected cost in (2.33). Let $S_0^e(t)$ denote the echelon base-stock level at the warehouse at time t. We define the maximum number of echelon overstocks in the system and the maximum number of overstocks at retailer i at time t by

$$N_{0,t}^e := \max_{x \in [0,t]} \{S_0^e(x)\} - S_0^e(t), \qquad N_{i,t} := \max_{x \in [0,t]} \{S_i(x)\} - S_i(t).$$

That is, $N^e_{0,t}$ and $N_{i,t}$ represent the differences of the base-stock levels at time t and the highest base-stock levels prior to time t. We have

$$\mathbb{E}[IL_0^{e+}(t)] = \sum_{n=S_0^e(t-L_0)}^{S_0^e(t-L_0)+N^e_{0,t-L_0}} \sum_{j=0}^{n} (n-j)\mathbb{P}(D^0_{t-L_0}(L_0)=j)\mathbb{P}(IP_0^e(t-L_0)=n), \tag{2.39}$$

$$\mathbb{E}[IL_i^{+}(t)] = \sum_{n=S_i(t-L_i)}^{S_i(t-L_i)+N_{i,t-L_i}} \sum_{j=0}^{n} (n-j)\mathbb{P}(D^i_{t-L_i}(L_i)+B_i(t-L_i)=j)\mathbb{P}(IP_i(t-L_i)=n), \tag{2.40}$$

$$\mathbb{E}[IL_i^{-}(t)] = \mathbb{E}[IL_i^{+}(t)] - \mathbb{E}[IL_i(t)]. \tag{2.41}$$

These equations extend the stationary analogues (Axsäter, 2006) to time-dependence. Further, they account for the fact that the inventory position is a random variable (through the first sum and the last probability in both (2.39) and (2.40)). $B_i(t-L_i)$ represents the number of backorders of retailer i at the warehouse at time $t-L_i$. Since $D^i_{t-L_i}(L_i)$ represents demand after time $t-L_i$ and $B_i(t-L_i)$ only depends on demands before time $t-L_i$, these random variables are independent and therefore

$$\mathbb{P}\big(D^i_{t-L_i}(L_i)+B_i(t-L_i)=j\big) = \sum_{k=0}^{j} \mathbb{P}\big(D^i_{t-L_i}(L_i)=k\big)\mathbb{P}\big(B_i(t-L_i)=j-k\big). \tag{2.42}$$

It remains to determine the distributions of the backorders and the inventory position. For the latter, we need to determine the distribution of overstocks.

In order to compute (2.42), retailer i's probability of facing warehouse backorders at time t, $B_i(t)$, remains to be determined. With

$$IL_0(t) = IL_0^e(t) - \sum_{i=1}^{N} S_i(t) \qquad \text{and} \qquad IL_0^e(t) = S_0^e(t-L_0) - D^0_{t-L_0}(L_0)$$

we have

$$\mathbb{P}(IL_0^-(t)=k) = \mathbb{P}(IL_0(t)=-k) = \mathbb{P}\Big(D^0_{t-L_0}(L_0) = S_0^e(t-L_0) - \sum_{i=1}^{N} S_i(t) + k\Big). \tag{2.43}$$

In a stationary setting, where it holds that

$$S_0^e - \sum_{i=1}^{N} S_i = S_0, \qquad \text{we obtain the well-known} \qquad \mathbb{P}(IL_0^- = k) = \mathbb{P}(D_0(L_0) = S_0 + k).$$

Now we find

$$\begin{aligned} \mathbb{P}(B_i(t) = j) = & \sum_{k=j}^{\infty} \mathbb{P}(IL_0^-(t) = k) \binom{k}{j} \left(\frac{\Lambda^i_{t-L_0}(L_0)}{\Lambda^0_{t-L_0}(L_0)} \right)^j \left(\frac{\Lambda^0_{t-L_0}(L_0) - \Lambda^i_{t-L_0}(L_0)}{\Lambda^0_{t-L_0}(L_0)} \right)^{k-j}, \\ & j > 0, \qquad (2.44) \\ \mathbb{P}(B_i(t) = 0) = & 1 - \sum_{j=1}^{\infty} \mathbb{P}(B_i(t) = j). \qquad (2.45) \end{aligned}$$

In the following, we derive the distribution of overstocks, which characterizes the distribution of the inventory position. We distinguish between the echelon inventory position at the warehouse $IP_0^e(t)$, the echelon overstocks at the warehouse $OV_0^e(t)$, the inventory positions at the retailers $IP_i(t)$ and the overstocks at retailer i, $OV_i(t)$. $D_t^0(s-t)$ represents the *system demand* in the time interval $[t, s]$, i.e. $D_t^0(s-t) = \sum_{i=1}^N D_t^i(s-t)$, where $D_t^i(s-t)$ is the demand at retailer i in $[t, s]$. Recall the maximum number of echelon overstocks in the system and the maximum number of overstocks at retailer i at time t

$$N_{0,t}^e := \max_{x \in [0,t]} \{S_0^e(x)\} - S_0^e(t), \qquad N_{i,t} := \max_{x \in [0,t]} \{S_i(x)\} - S_i(t).$$

For $n \in \{1, ..., N_{0,t}^e\}$, we define the time of the last downward step of the echelon warehouse base-stock level from level $S_0^e(t) + n$ before time t as

$$t_0(n) := \sup_{[0,t]} \{x, S_0^e(x) = S_0^e(t) + n\}.$$

Steps at the retailers are always of size 1 (Proposition 2.1). However, step sizes at the warehouse might be higher. As long as step sizes are 1, we find

$$t_0(N_{0,t}^e) > t_0(N_{0,t}^e - 1) > ... > t_0(1).$$

If, as depicted in Figure 2E.1, step sizes are larger than one and k exists with $t(k-1) >$

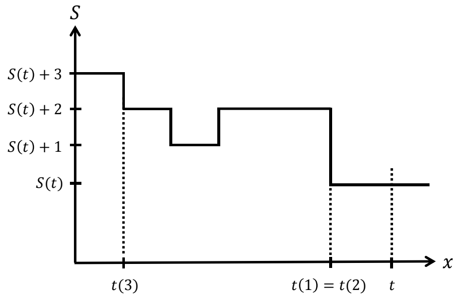


Figure 2E.1: Example for echelon base-stock level at the warehouse.

$t(k)$, we set $t(k-1) := t(k)$. This yields

$$t_0(N^e_{0,t}) \geq t_0(N^e_{0,t} - 1) \geq ... \geq t_0(1).$$

Retailer step times $t_i(n)$ are found analogously. We now define additional random variables $V_0(n)$ and $V_i(n)$ that help us to determine the distribution of overstocks. These random variables give the number of overstocks just before a downward step, i.e.

$$V_0(n) := \lim_{\substack{x \to t_0(n) \\ x < t_0(n)}} IP^e_0(x) - S^e_0(x), \forall\ n \in \{1, ..., N^e_{0,t}\},$$
$$V_i(n) := \lim_{\substack{x \to t_i(n) \\ x < t_i(n)}} IP_i(x) - S_i(x), \forall\ i \in I, \forall\ n \in \{1, ..., N_{i,t}\}.$$

Their distributions are defined recursively by

$$\mathbb{P}(V_0(N^e_{0,t}) = 0) = 1, \tag{2.46}$$

$$\mathbb{P}(V_0(n) = k) = \sum_{l=k-1}^{N^e_{0,t}-(n+1)} \mathbb{P}(V_0(n+1) = l)\mathbb{P}(D^0_{t_0(n+1)}(t_0(n) - t_0(n+1)) = l - (k-1)),$$
$$n \in \{1, ..., N^e_{0,t} - 1\}, k \in \{1, ..., N^e_{0,t} - n\}, \tag{2.47}$$

$$\mathbb{P}(V_0(n) = 0) = \sum_{l=0}^{N^e_{0,t}-(n+1)} \mathbb{P}(V_0(n+1) = l)\mathbb{P}(D^0_{t_0(n+1)}(t_0(n) - t_0(n+1)) \geq l + 1),$$
$$n \in \{1, ..., N^e_{0,t} - 1\}, \tag{2.48}$$

and

$$\mathbb{P}(V_i(N_{i,t}) = 0) = 1, \tag{2.49}$$

$$\mathbb{P}(V_i(n) = k) = \sum_{l=k-1}^{N_{i,t}-(n+1)} \mathbb{P}(V_i(n+1) = l)\mathbb{P}(D^i_{t_i(n+1)}(t_i(n) - t_i(n+1)) = l - (k-1)),$$
$$n \in \{1, ..., N_{i,t} - 1\}, k \in \{1, ..., N_{i,t} - n\}, \tag{2.50}$$

$$\mathbb{P}(V_i(n) = 0) = \sum_{l=0}^{N_{i,t}-(n+1)} \mathbb{P}(V_i(n+1) = l)\mathbb{P}(D^i_{t_i(n+1)}(t_i(n) - t_i(n+1)) \geq l + 1),$$
$$n \in \{1, ..., N_{i,t} - 1\}, \tag{2.51}$$

where we used that $N^e_{0,t_0(n+1)} = N^e_{0,t} - (n+1)$, $N_{i,t_i(n+1)} = N_{i,t} - (n+1)$. Having these probabilities, we finally define the distributions of $OV^e_0(t)$ and $OV_i(t)$ through

$$\mathbb{P}(OV^e_0(t) = m) = \sum_{k=m-1}^{N^e_{0,t}-1} \mathbb{P}(V_0(1) = k)\mathbb{P}(D^0_{t_0(1)}(t - t_0(1)) = k - (m-1)), \tag{2.52}$$

$$\mathbb{P}(OV_i(t) = m) = \sum_{k=m-1}^{N_{i,t}-1} \mathbb{P}(V_i(1) = k)\mathbb{P}(D^i_{t_i(1)}(t - t_i(1)) = k - (m-1)). \tag{2.53}$$

It holds that

$$\mathbb{P}(IP^e_0(t) = S + n) = \mathbb{P}(OV^e_0(t) = n), \qquad \mathbb{P}(IP_i(t) = S + n) = \mathbb{P}(OV_i(t) = n), \tag{2.54}$$

and the expected inventory positions at time t are given through

$$\mathbb{E}[IP^e_0(t)] = S^e_0(t) + \mathbb{E}[OV^e_0(t)], \qquad \mathbb{E}[IP_i(t)] = S_i(t) + \mathbb{E}[OV_i(t)]. \tag{2.55}$$

Chapter 3

Supplier Selection under Failure Risk, Quantity and Business Volume Discounts

We consider a supply chain problem with simultaneous supplier selection and order allocation for multiple products. The suppliers offer quantity and business volume discounts, and they are subject to failure. The buyer aims at minimizing total expected costs. We consider both all-units and incremental quantity discounts and find optimal solutions through a new mixed-integer linear programming formulation. We discuss the trade-off between economies of scale and failure risk and show the cost reduction of our exact approach compared to a previously proposed heuristic.

3.1 Introduction and Related Literature

Natural disasters not only cause horrific human tragedies; they also shatter the economy and influence businesses all around the globe. Japan's 2011 "quadruple disaster" (The Economist, 2011) is one recent example of a supply chain disruption that highlights the importance to hedge supply chains against supplier outage. Multi-sourcing, i.e. having different suppliers for the same item, is one way of approaching this issue. However, this leads to a more complex supply chain (e.g., setting up ordering policies is more difficult), the contract negotiations with different suppliers may be rather time-consuming, or the costs per item may increase due to lower quantity discounts compared to single-sourcing. Balancing the advantages and disadvantages of multi-sourcing is a key issue when it comes to the design of the supply chain and a major challenge for production/inventory systems.

When sourcing an item, a buyer has to consider the risk of a supply chain disruption, which has a certain probability. Dealing with this risk, the buyer needs to decide upon the supplier selection and the allocation of orders between suppliers. Ruiz-Torres and Mahmoodi (2006) studied a problem with multiple suppliers subject to failure risk and used an enumeration to decide about supplier selection and order allocation. Meena et al. (2011) considered disruptions due to catastrophic events and determined the optimal number of suppliers. Sawik (2014) compared single sourcing to dual sourcing strategies in make-to-order systems. Silbermayr and Minner (2016) considered multi-sourcing with different speeds and costs per supplier and examined the trade-off between economies of scale and supply disruptions. Tang (2006), Snyder et al. (2016) and Fahimnia et al. (2015) reviewed supply chains subject to disruptions and their risk management. Thomas and Tyworth (2006) presented a critical literature review on order splitting. Minner (2003) provided an overview of multi-sourcing problems.

Individual contracts with suppliers often allow for quantity discounts, i.e. in addition to considering the disruption risk of a supplier, there are incentives for the buyer to allocate larger orders to suppliers with larger quantity discounts. *Incremental* discounts only apply to items exceeding a certain order quantity while *all-units* discounts apply to all items. Burke et al. (2008) modeled incremental quantity discounts without supplier failure risk. Munson and Rosenblatt (1998) reviewed quantity discounts and Munson and Jackson (2015) provided a more recent, extensive overview on quantity discounts. Another regularly observed price reduction comes from business volume discounts, i.e. the supplier offers a discount on the total amount of sales generated by a buyer. Sadrian and Yoon (1994) and Xia and Wu (2007) considered supplier selection under business volume discounts. An overview on multi-criteria supplier selection was presented in Ho et al. (2010).

Our research is based on Meena and Sarmah (2013) who considered order allocation with supplier failure risk, all-units quantity discounts and deterministic demand among a pre-selected, fixed set of suppliers. The authors proposed a genetic algorithm for solving the optimization problem. However, formulating the problem as a mixed-integer linear program (MILP), we find optimal solutions to all their instances in negligible computation time. Their heuristic solutions deviate from our optimal solutions by up to 4%. We further extend the problem to include integrated supplier selection and multiple products. We model both incremental and all-units quantity discounts and include business volume discounts. In a numerical study, we show that our MILP-approach solves realistic problem sizes.

The chapter is structured as follows: Section 3.2 introduces the problem and presents the optimization model. In Section 3.3, we present numerical results that consist of a) a discussion of the results of Meena and Sarmah (2013) and b) a numerical study of the extended problem. In Section 3.4, we provide conclusions and ideas for further research.

3.2 Model

We consider a risk-neutral decision maker who faces deterministic demand for multiple products and has to decide about the supplier selection and order allocation for a single period. There is a pool of suppliers where every supplier offers both quantity discounts per product and business volume discounts depending on the total sales volume. All suppliers may disrupt and have certain failure probabilities that may be correlated. Further, a super-event that hits all suppliers simultaneously can occur. Each supplier has a maximum capacity for each product and delivers products in fixed lot sizes. From this supplier pool, we want to find a selection of suppliers such that the total expected costs (TEC), consisting of procurement costs, supplier management costs, and shortage penalty costs, are minimized.

If a supplier is selected, fixed supplier management costs that are independent of the number of products ordered from this supplier, arise. For each product ordered at a supplier, a certain minimum share of the total demand has to be allocated. If one selected supplier fails, the remaining selected suppliers try to compensate for this loss by producing up to their respective capacity limits. Note that, for each product, only those suppliers who have been selected to deliver the product in the first place can compensate a loss.

The sequence of events is as follows: the buyer selects suppliers and allocates orders. Selected suppliers might fail; their quantities are reallocated to suppliers who can compensate for the loss. The available share of the orders is delivered, the rest is lost and incurs a penalty cost.

The quantity discounts offered by the suppliers have different levels; the start of one level is denoted as a price break quantity. For all-units discounts, the attained quantity

Table 3.1: Notation.

Sets and parameters	
$p \in P$	products
$i \in N$	suppliers
$j \in J$	quantity discount levels
$b \in B$	business volume discount levels
$\mathcal{P}(N)$	set of supplier selections — power set of N
$m \in \mathcal{B}(N)$	enumeration of selections — $\mathcal{B}(N) = \{1, ..., \lvert\mathcal{P}(N)\rvert\}$
T_{pm}	supplier selection m for product p — $T_{pm} \in \mathcal{P}(N)$
$\mathcal{P}(T_{pm})$	power set of T_{pm}
D_p	deterministic demand for product p
C_p	base price for product p
L_p	per unit penalty cost for not satisfying the demand for product p
F	management costs per supplier (identical for all suppliers $i \in N$)
Q_{pi}	lot size for product p at supplier i
$Q_{pi}^{\max}$	capacity for product p at supplier i
$Q_{pi}^{\min}$	minimum order quantity for product p at supplier i (if supplier i is selected for product p)
Q_{pij}	price break quantity for product p at supplier i at quantity discount level j
d_{pij}	relative quantity discount at price break quantity Q_{pij}
$\tilde{Q}_{ib}$	price break volume at supplier i for business volume discount level b
$\tilde{d}_{ib}$	relative business volume discount at price break volume $\tilde{Q}_{ib}$
d_{pijb}	relative combined quantity and volume discounts, $d_{pijb} = 1 - (1 - d_{pij}) \cdot (1 - \tilde{d}_{ib})$
p^*	probability of super-event that all suppliers fail
X_i	binary random variable — $X_i = 1$ if supplier i fails
a_i, c_m	auxiliary parameters (explained below)
M	large number
Decision variables	
q_{pijb}	quantity of product p at supplier i with quantity discount j and volume discount b
q_{pij}	quantity of pruduct p at supplier i with quantity discount j
q_{pi}	quantity of product p at supplier i
$\tilde{q}_{pib}$	quantity of product p at supplier i with volume discount b
k_{pi}	integer for determining multiple of lot size of product p at supplier i, $q_{pi} = k_{pi} \cdot Q_{pi}$
x_{pij}	indicator whether quantity $q_{pij} > 0$
x_{pi}	indicator whether supplier i is selected for product p, i.e. if $q_{pi} > 0$
x_i	indicator whether supplier i is selected, i.e. $\exists q_{pi} > 0$
$\tilde{x}_{ib}$	indicator whether volume discount b applies at supplier i
b_{pm}	indicator whether supplier selection T_{pm} is chosen
z_p	auxiliary decision variable (explained below)
$u_{pm}, \bar{u}_{pm}$	auxiliary binary decision variables (explained below)

discount is applied to all units ordered, whereas in the case of incremental discounts, the respective discount level only applies to those units exceeding the price break quantity.

A supplier offers a business volume discount if the total sales volume of a buyer exceeds a certain threshold. That is, the buyer's total costs reduce by a certain percentage. For business volume discounts we use the term price break volumes, analogously to price break quantities. Table 3.1 introduces the notation that is used to develop the model.

The TEC have three components: (i) purchasing cost, (ii) supplier management costs, and (iii) expected total penalty costs. First, we have purchasing costs (PC)

$$PC = \sum_{p \in P} C_p \cdot \sum_{i \in N} \sum_{j \in J} \sum_{b \in B} q_{pijb} \cdot (1 - d_{pijb}). \tag{3.1}$$

The total quantity of product p ordered at supplier i, q_{pi}, can be uniquely mapped to one attained quantity discount level j and one attained business volume discount level b. For this combination of quantity and business volume discounts, it holds $q_{pijb} = q_{pi}$, while for all other combinations we have $q_{pijb} = 0$, as will be ensured in the constraints. Therefore, first summing over all quantity and business volume discounts yields the desired summands for finally summing over products and suppliers. Multiplication with $(1 - d_{pijb})$ applies the combined quantity and volume discount per product and supplier.

We further have supplier management costs (SMC), given through

$$SMC = F \cdot \sum_{i \in N} x_i. \tag{3.2}$$

These are linear in the number of suppliers and apply if a supplier is selected for at least one product, regardless of the number of products and the respective quantities.

The last component of the total costs are the expected total penalty costs (ETP). For a fixed set of suppliers $S \subseteq N$, these are obtained by

$$ETP_S = \sum_{p \in P} L_p \cdot \left(p^* \cdot D_p + (1 - p^*) \cdot \sum_{A \in \mathcal{P}(S)} \mathbb{P}(A) \cdot \left[D_p - \min\left(D_p, \sum_{l \in S \setminus A} Q_{pl}^{\max}\right) \right] \right), \tag{3.3}$$

where

$$\mathbb{P}(A) = \mathbb{P}(X_j = 1, \forall j \in A, X_l = 0, \forall l \in S \setminus A). \tag{3.4}$$

For a given supplier selection, A is a subset of failed suppliers. $\mathbb{P}(A)$ represents the joint probability of having a failure of all suppliers in A and no failure by the remaining suppliers $S\backslash A$. If a supplier fails, the remaining suppliers try to compensate the loss through extra production. Following Meena and Sarmah (2013), the purchasing prices for the additional production lots equal the prices of the failed suppliers, i.e. there are no additional costs for the buyer to consider. We explain ETP in more detail but will do so for our broader setting where the supplier selection is part of the optimization, rather than given. Here, the expected total penalty costs depend on the number of selected suppliers, i.e. they are non-linear. We deal with this non-linearity by preprocessing the *expected penalty costs of product p and selection m*

$$V_{pm} = L_p \cdot \sum_{A \in \mathcal{P}(T_{pm})} \mathbb{P}(A) \cdot \Big[D_p - \min\big(D_p, \sum_{l \in T_{pm} \backslash A} Q_{pl}^{\max} \big) \Big] \tag{3.5}$$

for all possible selections $T_{pm} \in \mathcal{P}(N)$, which resembles the last term in the above definition of ETP. For obtaining the expected penalty costs V_{pm} of selection T_{pm}, we sum over all possible sets of jointly failed suppliers $A \in \mathcal{P}(T_{pm})$. For each of these sets A, we calculate the joint failure probability of having failures by suppliers $i \in A$ ($X_i = 1$) and of having no failures by suppliers $i \in T_{pm} \backslash A$ ($X_i = 0$). The resulting probability is multiplied by the total lost demand for product p arising from failure set A, which is the difference between the demand and the sum of the capacities of the non-failed suppliers. With this definition of expected penalty costs V_{pm}, we find the linear expression

$$ETP = \sum_{p \in P} \left(p^* \cdot L_p \cdot D_p + (1 - p^*) \cdot \sum_{m \in \mathcal{B}(N)} b_{pm} \cdot V_{pm} \right), \tag{3.6}$$

with $b_{pm} = 1$ for exactly one selection T_{pm} per product p and zero otherwise. The ETP per product then consist of the probability of the super-event that hits all suppliers, p^*, where all demand is lost, plus the probability that the super-event does not happen, multiplied by the corresponding expected penalty costs of the particular product and supplier selection for that product. Finally, we sum over all products.

3.2.1 All-Units Discount

The MILP formulation is given through

$$
\begin{aligned}
\min \quad & TEC = PC + SMC + ETP && && (3.7)\\
\text{s.t.} \quad & Q^{\min}_{pi} \cdot x_{pi} \leq q_{pi} \leq Q^{\max}_{pi} \cdot x_{pi}, && \forall\, p \in P,\ i \in N, && (3.8)\\
& q_{pi} = k_{pi} \cdot Q_{pi}, && \forall\, p \in P,\ i \in N, && (3.9)\\
& \sum_{i \in N} q_{pi} = D_p, && \forall\, p \in P, && (3.10)\\
& x_{pi} \leq x_i, && \forall\, p \in P,\ i \in N, && (3.11)\\
& Q_{pij} \cdot x_{pij} \leq q_{pij} \leq (Q_{pi,j+1} - 1) \cdot x_{pij}, && \forall\, p \in P,\ i \in N,\ j \in J, && (3.12)\\
& \sum_{j \in J} q_{pij} = q_{pi}, && \forall\, p \in P,\ i \in N, && (3.13)\\
& \sum_{j \in J} x_{pij} = x_{pi}, && \forall\, p \in P, i \in N, && (3.14)\\
& \tilde{Q}_{ib} \cdot \tilde{x}_{ib} \leq \sum_{p \in P} C_p \cdot \tilde{q}_{pib} \leq (\tilde{Q}_{i,b+1} - 1) \cdot \tilde{x}_{ib}, && \forall\, i \in N,\ b \in B, && (3.15)\\
& \sum_{b \in B} \tilde{q}_{pib} = q_{pi}, && \forall\, p \in P,\ i \in N, && (3.16)\\
& \sum_{b \in B} \tilde{x}_{ib} = x_i, && \forall\, i \in N, && (3.17)\\
& \sum_{b \in B} q_{pijb} = q_{pij}, && \forall\, p \in P,\ i \in N,\ j \in J, && (3.18)\\
& \sum_{j \in J} q_{pijb} = \tilde{q}_{pib}, && \forall\, p \in P,\ i \in N,\ b \in B, && (3.19)\\
& z_p = \sum_{i \in N} a_i \cdot x_{pi}, && \forall p \in P, && (3.20)\\
& z_p - c_{pm} \leq M \cdot u_{pm}, && \forall\, p \in P,\ m \in \mathcal{B}(N), && (3.21)\\
& z_p - c_{pm} \geq -M \cdot \bar{u}_{pm} && \forall\, p \in P, m \in \mathcal{B}(N), && (3.22)\\
& b_{pm} = 1 - (u_{pm} + \bar{u}_{pm}), && \forall\, p \in P,\ m \in \mathcal{B}(N), && (3.23)\\
& u_{pm} + \bar{u}_{pm} \leq 1, && \forall\, p \in P,\ m \in \mathcal{B}(N), && (3.24)\\
& \sum_{m \in \mathcal{B}(N)} b_{pm} = 1, && \forall\, p \in P, && (3.25)\\
& b_{pm}, u_{pm}, \bar{u}_{pm}, x_{pij}, x_{pi}, \tilde{x}_{ib}, x_i \in \{0,1\}, && \forall\, p \in P,\ i \in N,\ j \in J,\ b \in B, &&
\end{aligned}
$$

$$m \in \mathcal{B}(N), \tag{3.26}$$

$$q_{pijb}, q_{pij}, \tilde{q}_{pib}, q_{pi}, k_{pi} \in \mathbb{N}, \qquad \forall\, p \in P,\ i \in N,\ j \in J,\ b \in B, \tag{3.27}$$

$$z_p \geq 0, \qquad \forall\, p \in P. \tag{3.28}$$

Constraint (3.8) ensures that the quantity q_{pi} of product p ordered from supplier i, if selected, is larger than the minimum quantity and does not exceed the supplier's capacity. (3.9) ensures that q_{pi} is a multiple of the lot size given by supplier i, (3.10) requires the demand for all products to be exactly met. Note that the demand for product p has to be a linear combination of the lot sizes of the selected suppliers, otherwise constraints (3.9) and (3.10) might cause an infeasibility. This is trivially satisfied for all demands if the lot size is one. (3.11) makes sure that, if a supplier is selected for at least one product p, supplier management costs for this supplier arise. (3.12)-(3.14) determine the quantity discount at supplier i for product p. The quantity q_{pij} is only positive if supplier i is selected for product p and discount level j is attained. (3.15)-(3.17) determine the business volume discount at supplier i using the same logic. For $\bar{b} = \sup B$ and $\bar{j} = \sup J$, we define $\tilde{Q}_{i,\bar{b}+1} = \sum_{p \in P} C_p \cdot D_p,\ \forall\, i$, and $Q_{pi,\bar{j}+1} = Q^{\max}_{pi}$. Note that sets B and J are defined such that $\tilde{Q}_{i,1} = 0,\ Q_{pi,1} = 0$, where there might be either no discounts in the first interval, i.e. $d_{pi,1,1} = 0\ \forall\, p, i$, or differences in base prices among suppliers that are modeled through positive discount levels $d_{pi,1,1} > 0$ for some suppliers. (3.18)-(3.19) determine the quantity of product p that is ordered at supplier i at discount levels j and b. $\tilde{q}_{pib}$ and q_{pij} are positive for a maximum of one b and j, respectively. Together, these constraints ensure that supplier i delivers product p at exactly one discount level combination of b and j.

In Table 3.1, we introduced *auxiliary* parameters that are explained in the following. One of the crucial elements of the MILP model is the use of preprocessing to avoid non-linearities. In order to determine the ETP, we computed the expected penalty for every supplier selection of every product, V_{pm}. Now we have to make sure that the respective value of ETP in the objective function for the chosen sets of suppliers is added. That is, we only add one V_{pm} for every product p in equation (3.6), since we only have one selection per product. In other words, $b_{pm} = 1$ for exactly one selection m. Constraints (3.20)-(3.25) make sure that we have exactly one positive decision variable b_{pm} for each product p. In order to achieve this, the parameters a_i are assigned to suppliers $i \in N$

and parameters c_{pm} are assigned to selections $T_{pm} \in \mathcal{P}(N)$ such that

$$c_{pm} = \sum_{i \in T_{pm}} a_i.$$

a_i must be chosen such that $c_{pm_1} \neq c_{pm_2},\ \forall\ m_1, m_2 \in \mathcal{B}(N),\ m_1 \neq m_2$, e.g., $a_i = \frac{1}{2^{i-1}}$. The constraints ensure that z_p from (3.20) equals c_{pm} for exactly one selection $m \in \mathcal{B}(N)$, and for this selection m, the binary decision variable b_{pm} is set to one.

3.2.2 Incremental Discount

When it comes to incremental discounts, different quantity discounts apply to different shares of the purchased quantity (as opposed to all-units discounts, where one discount applies to the entire quantity). Taking this into consideration, we have to determine which discount applies to which share. To do so, we introduce new decision variables, additional constraints and an additional summand in the objective function. Constraints (3.29) - (3.32) ensure that if $q_{pi\tilde{j}b} \geq Q_{pi\tilde{j}}$, the indicator δ_{pijb} is set to one for all $j \in \{2, ..., \tilde{j}\}$. w_{pijb} gives the quantity of product p ordered from supplier i at volume discount level b with lower quantity discount than $d_{pi\tilde{j}b}$. ϵ is a fixed small positive real number.

$$\tilde{q}_{pib} - Q_{pij} \geq -M \cdot (1 - \delta_{pijb}), \qquad \forall\ p \in P,\ i \in N,\ j \in J \backslash \{1\},\ b \in B, \tag{3.29}$$
$$\tilde{q}_{pib} - Q_{pij} + \epsilon \leq M \cdot \delta_{pijb}, \qquad \forall\ p \in P,\ i \in N,\ j \in J \backslash \{1\},\ b \in B, \tag{3.30}$$
$$w_{pijb} = (Q_{pij} - 1) \cdot \delta_{pijb}, \qquad \forall\ p \in P,\ i \in N,\ j \in J \backslash \{1\},\ b \in B, \tag{3.31}$$
$$\delta_{pijb} \in \{0, 1\}, \qquad \forall\ p \in P,\ i \in N,\ j \in J \backslash \{1\}, \tag{3.32}$$
$$w_{pijb} \geq 0, \qquad \forall\ p \in P,\ i \in N,\ j \in J \backslash \{1\},\ b \in B. \tag{3.33}$$

The term IN is added to the objective function and accounts for the additional costs as compared to all-units discounts.

$$IN = \sum_{p \in P} C_p \sum_{i \in N} \sum_{j \in J \backslash \{1\}} \sum_{b \in B} w_{pijb} \cdot (d_{pijb} - d_{pi,j-1,b}). \tag{3.34}$$

3.2.3 Compensation Costs

The assumption made by Meena and Sarmah (2013) that other suppliers compensate for a supplier failure at the same price, including discounts, seems somewhat disputable. In the following, we present an approach for adjusting the model to higher compensation costs. Recall that $C_{pi} = C_p \cdot (1 - d_{pi,1})$ is the base price for product p at supplier i. We solve the following transportation problem (TPP) for all selections $T_{pm} \backslash A$ with $p \in P,\ m \in \mathcal{B}(N), A \in \mathcal{P}(T_{pm})$. The TPP does an optimal allocation of orders for the given (remaining) number of suppliers. It is part of the preprocessing and uses an additional artificial supplier $N+1$, to whom the lost demands are allocated at a cost of $C_{p,N+1} = L_p$.

$$\min \quad C_{T_{pm}\backslash A} = \sum_{i \in T_{pm}\backslash A \cup \{N+1\}} C_{pi} \cdot q_{pi} \tag{3.35}$$

$$s.t. \quad \sum_{i \in T_{pm}\backslash A \cup \{N+1\}} q_{pi} = D_p, \tag{3.36}$$

$$q_{pi} \leq Q_{pi}^{\max}, \qquad \forall\, i \in T_{pm}\backslash A \cup \{N+1\} \tag{3.37}$$

$$q_{pi} \in \mathbb{N}, \qquad \forall\, i \in T_{pm}\backslash A \cup \{N+1\}. \tag{3.38}$$

We set $Q_{p,N+1}^{\max} = D_p$. The objective (3.35) minimizes the total price paid under the selection $T_{pm} \backslash A$. Constraints (3.36)-(3.38) ensure that all demand is met, all suppliers only deliver up to their capacities, and order quantities are non-negative integers. In (3.39), we subtract the optimal costs from the TPP without supplier failure from those with supplier failure. This difference serves as the combined failure and compensation cost of supplier selection T_{pm}.

$$r_{pm}^{A} = C_{T_{pm}\backslash A} - C_{T_{pm}}. \tag{3.39}$$

We replace the term V_{pm} in (3.6) by the *expected penalty and compensation costs of product p and selection m*

$$\tilde{V}_{pm} = \sum_{A \in \mathcal{P}(T_{pm})} \mathbb{P}(A) \cdot r_{pm}^{A}. \tag{3.40}$$

As we do not consider discounts in the TPP, $\tilde{V}_{pm}$ does not include exact compensation costs. Instead, it serves as a close approximation. In particular, it is more realistic than the initial assumption of no compensation costs at all. The computation time for one TPP is less than 0.1 seconds.

3.3 Numerical Results

3.3.1 Discussion of Meena and Sarmah (2013)

We consider the numerical study by Meena and Sarmah (2013) and compare their results with our exact approach. Their problem has a single product, all-units quantity discounts, and supplier failure risk. All of our computations are conducted using Xpress-MP 7.9 on an Intel(R) Core(TM) i7-3770, 3.4 GHz processor with 16 GB RAM.

Order Allocation

Meena and Sarmah (2013) consider only the order allocation part of the problem, i.e. they focus on finding the optimal allocation for a given set of suppliers. They use a genetic algorithm to find solutions that are not always optimal. In Table 3.2, we list the parameters of their base case and three instances from their sensitivity analysis where their order allocation was suboptimal. (In Meena and Sarmah (2013), these are the instances 3 and 6 in their Table 5 and instance 4 in their Table 6.) In these three instances, the problem is to find the best allocation among the 3 predetermined suppliers 8, 9 and 10. We have $C = 10,\ F = 20,\ L = 15,\ p^* = 0.01,\ Q_i^{\min} = 10,\ Q_i = 1\ \forall i$. As there is only one product, we need no product index p. Large deviations of their results from the optimal allocation, as observed in instances 2 and 3 of Meena and Sarmah (2013), cause a large increase in total costs.

Solving the problem introduced by Meena and Sarmah (2013) requires only a small special case of our model with little complexity, no preprocessing effort (as the supplier selection is already given) and negligible computation time. As exact solutions can be obtained that easily, we propose to always apply our exact approach, rather than the heuristic. The relative improvements in Table 2 are up to 4%. Using other examples may increase these improvements even further.

Supplier Selection and Order Allocation

We now apply supplier selection and order allocation simultaneously to the basic problem presented by Meena and Sarmah (2013) with 10 suppliers and 3 quantity discount levels. We find that the optimal solution is to order 10 units at supplier 7 and 90 units at supplier 10, with an objective value of 664.17. For this extended, simultaneous problem, the

Table 3.2: Cases where Meena and Sarmah (2013) find suboptimal results.

$D = 100$			Base case parameters							
Supplier i	$Q_i^{\max}$	p_i	$Q_{i,1}$	$Q_{i,2}$	$Q_{i,3}$	$d_{i,1}$	$d_{i,2}$	$d_{i,3}$		
1	70	0.13	30	45	60	0.11	0.22	0.31		
2	93	0.09	40	55	70	0.07	0.19	0.29		
3	110	0.15	35	45	50	0.09	0.18	0.33		
4	90	0.17	50	60	80	0.14	0.19	0.25		
5	105	0.12	30	40	45	0.10	0.15	0.27		
6	80	0.19	37	52	60	0.17	0.21	0.30		
7	95	0.05	30	45	55	0.13	0.23	0.35		
8	115	0.14	45	50	65	0.10	0.29	0.37		
9	100	0.11	40	55	60	0.15	0.25	0.35		
10	140	0.16	50	60	70	0.20	0.27	0.46		
$D = 150$									**q_i (MS)**	**q_i (OPT)**
8						0.05	0.09	0.12	**15**	**15**
9	base case			base case		0.23	0.26	0.28	**60**	**65**
10						0.22	0.24	0.26	**75**	**70**
								TEC	**1249.9**	**1248.58**
$D = 300$										
8						0.05	0.09	0.12	**90**	**65**
9	base case			base case		0.23	0.26	0.28	**90**	**100**
10						0.22	0.24	0.26	**120**	**135**
								TEC	**2880.9**	**2839.44**
$D = 200$										
8	120	0.14	50	70	90				**100**	**70**
9	90	0.11	70	80	100		base case		**80**	**20**
10	110	0.16	60	90	100				**20**	**110**
								TEC	**1595.9**	**1533.15**

MILP needs 435 Simplex iterations, 97 integer nodes and 17.8 seconds of computation time with default solver settings.

Table 3.3: Selection and allocation for given number of suppliers.

No.	Selection	Allocation	Obj.	MIP nodes	Time (s)
1	10		812.6	1	0.2
2	**(7,10)**	$\boldsymbol{q_{10} = 90}$	**664.17**	**1**	**6.4**
3	(7,9,10)	$q_{10} = 80$	709.55	17	5.3
4	(5,7,9,10)	$q_{10} = 70$	773.6	117	8.5
5	(3,5,7,9,10)	$q_7 = 60$	905.0	0	3.8
6	(2,3,5,7,8,9)	$q_3 = 50$	970.0	1	2.7
7	(3,5,6,7,8,9,10)	$q_6 = 40$	1087.0	7	4.1
8	(1,2,3,4,5,6,7,8)	$q_7 = 30$	1136.0	1	1
9	(1,3,4,5,6,7,8,9,10)	$q_{10} = 20$	1195.0	0	0.1
10	(1,2,3,4,5,6,7,8,9,10)		1215.0	1	0.1

In order to compare the impact of different numbers of suppliers on the total costs, we fix the number of suppliers in the MILP. Table 3.3 shows the resulting selection and allocation decisions. Having few suppliers, e.g., only one, economies of scale arise from quantity discounts. However, this comes at the price of a higher failure risk. If you hedge against failure risk by selecting more suppliers, the economies of scale decrease. Figure 3.1 depicts this trade-off based on the optimal selections from Table 3.3. The decreasing line represents ETP, measured on the right scale, and the increasing line represents the sum of PC and SMC, measured on the left scale.

From this figure, it becomes clear that 2, 3, or 4 suppliers yield lower total costs than having only one supplier. This means that, in this problem, hedging disruption risk has a higher priority than realizing economies of scale.

3.3.2 Full Factorial Design

Data

As we want to evaluate our MILP model for multi-product cases that include business volume discounts, we conduct a numerical study under a full factorial design. We use 10 suppliers, which gives us $1,024$ possible supplier selections. Increasing the number of suppliers to 15 would lead to $32,768$ possible supplier selections, which is the number of

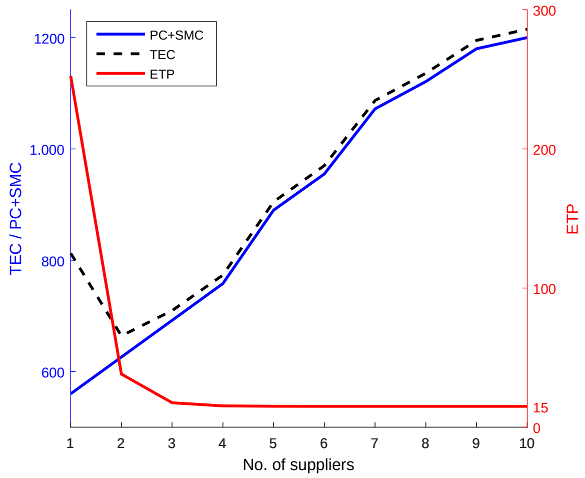


Figure 3.1: Economies of scale vs. failure risk.

binary variables per product required to determine ETP. Nevertheless, having a preselected pool of 10 suppliers before doing the final optimization of selection and allocation seems reasonable.

We have fixed the capacity for each supplier and each product as given in Table 3.4 and vary the parameters as follows.

$$|P| \in \{10, 15\}, |J| \in \{2, 4, 7\}, |B| \in \{0, 3, 5\},$$
$$D \in \{D_1, D_2\}, \mathbb{P} \in \{\mathbb{P}_1, \mathbb{P}_2\}, L \in \{50, 100\}, F \in \{50, 100\},$$

where D_1 and D_2 are deterministic demand vectors for all products. Demands and base prices (in monetary units) for the products are given in Table 3.4. We generated the data similar to Stadtler (2007).

In D_1, demands vary strongly among products ($\mu = 60,\ \sigma = 37$), while in D_2, demands are rather homogeneous ($\mu = 63,\ \sigma = 9$). We assume supplier failures to be independent and the corresponding failure probabilities per supplier to be given through $\mathbb{P}_1$ and $\mathbb{P}_2$ (see Table 3.4). In particular, the failure probabilities under $\mathbb{P}_2$ are twice as high as under $\mathbb{P}_1$. We vary the number of products between 10 and 15. By 10 products we mean the first 10 products. The price breaks and quantity discounts are given in Tables 3A.1-3A.3 in Appendix 3.A. The business volume levels (in monetary units) and the corresponding relative discounts are given in Table 3.5 and kept similar to those presented by Xia and

Table 3.4: Capacity, failure probability, demand and prices.

i	1	2	3	4	5	6	7	8	9	10			
$\mathbb{P}_1(i)$	0.05	0.1	0.15	0.02	0.07	0.12	0.04	0.08	0.14	0.1			
$\mathbb{P}_2(i)$	0.1	0.2	0.3	0.04	0.14	0.24	0.08	0.16	0.28	0.2			
p	$Q_{pi}^{\max}$										D_1	D_2	C_p
1	167	62	111	81	99	101	111	68	135	74	111	74	21
2	111	81	84	80	77	122	96	63	38	78	54	60	15
3	95	111	191	135	93	90	105	51	80	102	127	60	20
4	98	48	83	159	53	89	99	96	60	78	106	77	8
5	93	92	113	80	138	129	60	54	75	71	92	55	13
6	45	242	138	102	137	17	99	173	168	170	11	60	19
7	135	51	138	179	128	152	69	92	119	69	46	69	8
8	146	54	57	69	27	170	120	41	167	29	27	61	14
9	101	95	120	119	66	104	98	134	98	155	65	74	10
10	66	108	51	56	105	174	71	122	95	51	34	70	8
11	74	152	113	53	123	60	42	84	215	117	49	63	9
12	60	146	59	96	48	84	93	99	149	128	97	52	5
13	41	162	44	93	63	111	113	171	92	39	29	66	22
14	120	135	119	21	72	92	138	66	125	59	14	46	12
15	123	102	111	95	62	119	93	141	59	86	41	62	12

Table 3.5: Business volume levels and discounts.

b	1	2	3	4	5
$\tilde{Q}_{ib}$	0	2000	4000	6000	8000
$\tilde{d}_{ib}$	0	0.05	0.15	0.2	0.25

Wu (2007). We vary between 2, 4 and 7 quantity discount intervals and between 0, 3 and 5 business volume discount intervals. For both discount types, there is no discount in the first interval, i.e. the base price of a product is the same for all suppliers. The instances with more quantity or business volume intervals are extensions of the instances with fewer intervals, i.e. $|J| = 4$ means the first four intervals of the seven given in Tables 3A.1-3A.3 are available. According to Munson and Jackson (2015), the maximum quantity discount is usually below 20%. For two and four quantity discount intervals, we stick to this value; for seven intervals, we allow for larger maximum discounts.

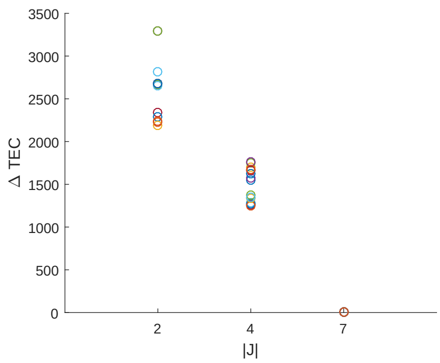
All suppliers allow for any order size within their capacity, i.e. $Q_{pi} = 1$ for all products as long as the order size exceeds the minimum order quantity, which is set to $Q_{pi}^{\min} = 5$ for all suppliers and products. The probability of a super-event hitting all suppliers is set to $p^* = 0.01$. The penalty cost L_p per lost sale is identical for all products and is varied between 50 and 100. So are the supplier management costs. In order to remain in the setting of Meena and Sarmah (2013), we focus on the problem with costs for lost demands but without compensation costs.

Results

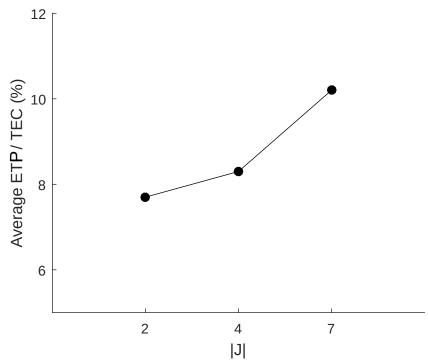
We ran all 288 instances for four hours. 58% of the instances were solved to optimality, 74% of the instances have a gap of less than 2%. The average number of visited MILP nodes is 31,000.

Supplier structure. For most products, there is one main supplier while the other suppliers serve as backup. While the main supplier delivers most of the units of an item, a backup supplier often only delivers the minimum quantity; her main function is to serve as insurance against a failure of the main supplier. However, mostly due to volume discounts, we sometimes see orders that are equally split among suppliers. This helps both suppliers towards attaining their respective discount levels.

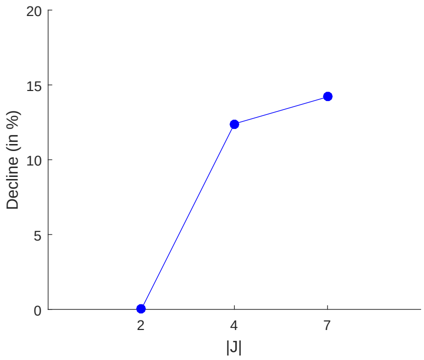
Quantity discounts. Varying the number of quantity discount levels between two, four and seven, we observe a strong sensitivity of the total costs to these changes. If you have more quantity discounts, a decrease in the total costs is natural; however, the decrease from two to four intervals is about 14%, whereas the decrease from two to seven intervals is about 35%, i.e. quite significant. This effect is shown in Figure 3.2a: each circle represents the additional TEC of one product if the number of discount intervals is smaller than seven, i.e. two or four. The explanation is that for a higher



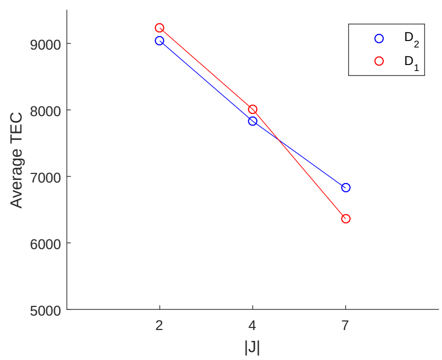
(a) Change of the total costs.



(b) Proportion of ETP in the total costs.



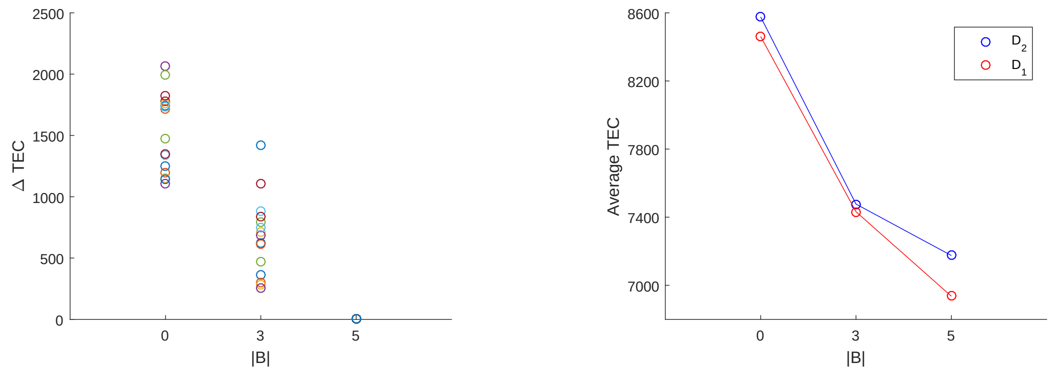
(c) Av. no. of suppliers/product.



(d) Av. total costs of demands.

Figure 3.2: Variation of the number of quantity discounts.

number of discount intervals (and thus, in this study, higher attainable discounts), the buyer's decision tends towards realizing economies of scale and reducing the focus on risk management. This is shown in Figure 3.2b, which depicts the proportion of expected total penalty costs in the total costs for the different numbers of discount intervals. For a higher number of intervals, the proportion of the expected total penalty costs grows. In Figure 3.2c, where we look at the sensitivity of the number of suppliers used per product to changes in the number of intervals, the effects underline the aforementioned findings. If we have more discount intervals, fewer suppliers are used, which increases the overall risk of suffering from supplier failure. In Figure 3.2d, we compare the two demand cases and find that the second demand pattern D_2 can benefit more if two to four intervals are offered, while the first demand pattern D_1 benefits more if seven intervals are offered. The second demand pattern has a lower variation and fewer products have low demands, i.e. more products reach higher discount levels in the first two cases. However, also having fewer high demands, the second demand pattern has fewer chances for fully exploiting discount levels five to seven – if they exist – than the first demand pattern.



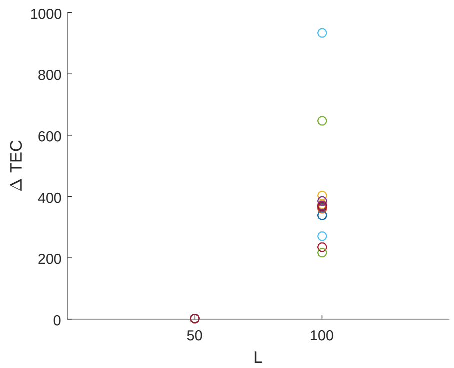
(a) Change of the total costs. (b) Av. total costs per demand.

Figure 3.3: Variation of the number of business volume discounts.

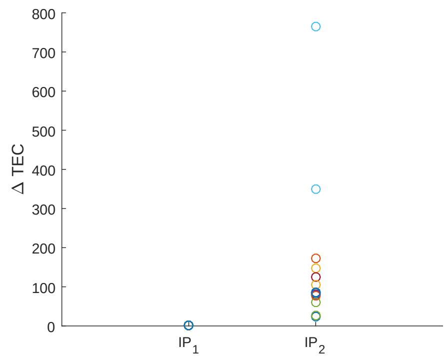
Business volume discounts. Looking at Figure 3.3a, we find that the number of business volume discounts has a huge effect on the total costs, which is in line with the expectations. However, the effect is somewhat smaller than the effect obtained from quantity discounts, which is due to volume discounts being lower than quantity discounts. In Figure 3.3b, we see that D_1 benefits more from business volume discounts than D_2 in all cases. In particular, when having five business volume discount levels, the demand structure of D_1 allows for a better use of business volume discounts than

D_2. Larger discount levels lead to a stronger focus on economies of scale and to more risk, i.e. the above analysis of quantity discounts also applies to volume discounts.

Penalty costs/Failure probabilities. If we double the penalty costs from 50 to 100, we see an increase in the total costs of about 400 monetary units on average (Fig. 3.4a). If we double the failure probabilities from $\mathbb{P}_1$ to $\mathbb{P}_2$, the increase of the total costs is only 100 monetary units (Fig. 3.4b). In particular, if we double failure probabilities instead of penalty costs, the increase in the total costs is much lower.



(a) Change of the total costs (L).



(b) Change of the total costs ($\mathbb{P}$).

Figure 3.4: Variation of penalty costs and failure probabilities.

Risk mitigation strategies. We observe two different strategies when risk (through L or $\mathbb{P}$) is increasing: (i) The buyer adds additional backup suppliers for several products as insurance against a failure by the main supplier. (ii) The buyer does a complete reallocation and uses a less risky but more expensive main supplier. While (i) dilutes economies of scale, (ii) reduces the sheer number of suppliers and focuses more strongly on economies of scale; however, with (ii) one no longer has the cheapest main supplier. Effect (ii) can be observed more often when penalty costs are increased, which explains the higher total costs in those instances compared to an increased failure probability.

Changing the number of products from 10 to 15 increases computation times significantly. Varying supplier management costs between 50 and 100 does not lead to different decisions. This is no surprise since economies of scale from quantity and volume discounts carry larger incentives for focusing on only a few suppliers. For incremental quantity discounts, the total price reduction in the same numerical setup is lower, therefore the sensitivity to the number of quantity discount levels decreases. Computation times increase due to an increased number of binary variables.

3.4 Conclusion

We considered the problem of simultaneous supplier selection and order allocation under quantity and business volume discounts and supplier failure risk. We formulated a mixed-integer linear program that solves realistic problem sizes. Considering the results of a previously published heuristic, we showed potential for improvement and derived the optimal solutions. In a numerical study, we further gained insights into the sensitivity of the optimal decisions with respect to input parameters as, e.g., penalty costs, failure probabilities, or the number of discount levels. We studied changes in total expected costs and expected total penalty costs and analyzed the trade-off between economies of scale and failure risk, finding different supplier selection strategies for an increasing risk.

There are several approaches one might wish to consider for further research. One is to extend the problem at hand and include multiple time periods. Another idea would be to extend the problem to random demand. As we consider a risk-neutral decision maker, another interesting research task would be to incorporate the decision maker's attitude towards risk.

3.A Input: Quantity Discounts

Table 3A.1: Price breaks and discount levels (1).

i	$Q_{pi,1}$	$Q_{pi,2}$	$Q_{pi,3}$	$Q_{pi,4}$	$Q_{pi,5}$	$Q_{pi,6}$	$Q_{pi,7}$	$d_{pi,1}$	$d_{pi,2}$	$d_{pi,3}$	$d_{pi,4}$	$d_{pi,5}$	$d_{pi,6}$	$d_{pi,7}$
$p = 1$														
1	0	22	42	63	78	87	101	0	0.07	0.1	0.17	0.28	0.35	0.38
2	0	27	48	71	78	89	96	0	0.04	0.12	0.19	0.29	0.36	0.36
3	0	21	40	66	83	83	95	0	0.03	0.15	0.19	0.29	0.34	0.37
4	0	23	44	66	78	93	101	0	0.02	0.13	0.21	0.3	0.33	0.38
5	0	23	46	72	77	88	98	0	0.05	0.15	0.21	0.29	0.34	0.38
6	0	25	47	69	83	87	98	0	0.06	0.14	0.22	0.3	0.36	0.37
7	0	21	43	72	75	94	102	0	0.04	0.13	0.19	0.27	0.35	0.41
8	0	26	47	64	80	88	104	0	0.02	0.1	0.18	0.29	0.34	0.39
9	0	27	49	62	83	86	98	0	0.03	0.13	0.17	0.25	0.35	0.41
10	0	13	20	34	38	43	46	0	0.02	0.09	0.22	0.24	0.36	0.38
$p = 2$														
1	0	13	19	31	36	41	51	0	0.03	0.13	0.18	0.32	0.34	0.41
2	0	9	21	32	35	45	50	0	0.04	0.13	0.19	0.26	0.34	0.36
3	0	12	22	34	36	43	48	0	0.04	0.14	0.2	0.29	0.33	0.39
4	0	11	20	31	37	41	49	0	0.04	0.15	0.18	0.32	0.32	0.4
5	0	13	22	35	39	45	50	0	0.06	0.16	0.21	0.25	0.33	0.37
6	0	11	20	33	40	44	46	0	0.04	0.14	0.2	0.26	0.32	0.41
7	0	12	21	31	36	41	50	0	0.07	0.13	0.17	0.27	0.34	0.4
8	0	12	22	34	39	45	50	0	0.05	0.15	0.22	0.31	0.32	0.4
9	0	10	24	33	40	41	48	0	0.07	0.13	0.17	0.27	0.36	0.39
10	0	26	55	71	91	105	119	0	0.02	0.08	0.18	0.26	0.35	0.4
$p = 3$														
1	0	23	56	73	88	102	114	0	0.05	0.09	0.18	0.25	0.32	0.38
2	0	21	54	81	91	100	115	0	0.03	0.15	0.2	0.27	0.32	0.37
3	0	24	49	72	85	105	117	0	0.05	0.12	0.17	0.27	0.33	0.38
4	0	31	48	71	90	98	120	0	0.05	0.15	0.19	0.25	0.34	0.39
5	0	22	53	77	91	106	115	0	0.01	0.1	0.17	0.27	0.33	0.4
6	0	24	56	74	90	106	116	0	0.03	0.12	0.17	0.25	0.32	0.37
7	0	23	56	75	87	106	115	0	0.02	0.13	0.22	0.29	0.35	0.38
8	0	20	45	80	87	100	120	0	0.05	0.08	0.18	0.24	0.33	0.4
9	0	26	51	81	85	100	119	0	0.06	0.13	0.21	0.29	0.35	0.39
10	0	18	47	65	76	90	97	0	0.03	0.11	0.24	0.3	0.34	0.41
$p = 4$														
1	0	18	43	59	76	88	101	0	0.06	0.08	0.19	0.26	0.34	0.37
2	0	18	43	62	77	82	92	0	0.05	0.12	0.22	0.28	0.32	0.4
3	0	24	39	64	78	90	94	0	0.01	0.1	0.22	0.29	0.32	0.39
4	0	18	43	62	74	80	93	0	0.05	0.09	0.19	0.24	0.32	0.37
5	0	26	43	65	71	84	93	0	0.03	0.1	0.21	0.28	0.34	0.38
6	0	23	39	62	73	84	92	0	0.06	0.09	0.17	0.29	0.33	0.38
7	0	18	42	60	78	84	90	0	0.01	0.1	0.23	0.26	0.35	0.39
8	0	25	43	68	71	81	95	0	0.04	0.08	0.17	0.31	0.35	0.4
9	0	26	39	65	72	84	93	0	0.04	0.13	0.18	0.32	0.36	0.37
10	0	22	38	53	65	77	82	0	0.04	0.1	0.22	0.31	0.34	0.4
$p = 5$														
1	0	19	33	57	63	74	83	0	0.06	0.12	0.2	0.28	0.33	0.4
2	0	22	41	58	69	73	81	0	0.03	0.14	0.22	0.26	0.33	0.38
3	0	23	38	56	64	73	85	0	0.06	0.12	0.19	0.3	0.34	0.38
4	0	19	33	56	65	76	82	0	0.04	0.12	0.18	0.26	0.35	0.39
5	0	15	39	56	66	75	85	0	0.01	0.12	0.16	0.32	0.33	0.4
6	0	22	36	60	66	75	83	0	0.02	0.09	0.21	0.29	0.34	0.4
7	0	19	33	56	62	70	85	0	0.05	0.12	0.19	0.29	0.35	0.4
8	0	22	34	56	61	73	87	0	0.04	0.15	0.2	0.25	0.36	0.38
9	0	18	34	59	62	74	83	0	0.02	0.15	0.21	0.25	0.33	0.38
10	0	3	5	7	8	8	10	0	0.03	0.1	0.16	0.26	0.35	0.41

Table 3A.2: Price breaks and discount levels (2).

i	$Q_{pi,1}$	$Q_{pi,2}$	$Q_{pi,3}$	$Q_{pi,4}$	$Q_{pi,5}$	$Q_{pi,6}$	$Q_{pi,7}$	$d_{pi,1}$	$d_{pi,2}$	$d_{pi,3}$	$d_{pi,4}$	$d_{pi,5}$	$d_{pi,6}$	$d_{pi,7}$
$p = 6$														
1	0	2	4	6	8	9	10	0	0.05	0.1	0.19	0.31	0.34	0.39
2	0	3	4	6	8	8	10	0	0.02	0.13	0.22	0.31	0.34	0.4
3	0	2	5	6	8	9	10	0	0.05	0.13	0.22	0.3	0.36	0.37
4	0	2	4	7	8	9	10	0	0.02	0.11	0.17	0.3	0.34	0.39
5	0	2	5	7	8	9	10	0	0.07	0.1	0.17	0.26	0.33	0.39
6	0	3	4	7	8	9	10	0	0.03	0.16	0.17	0.29	0.35	0.41
7	0	2	4	6	8	9	10	0	0.06	0.09	0.16	0.3	0.34	0.36
8	0	2	4	6	7	9	10	0	0.02	0.09	0.19	0.27	0.33	0.39
9	0	2	4	6	8	9	9	0	0.03	0.09	0.21	0.32	0.35	0.39
10	0	11	21	28	34	39	43	0	0.02	0.09	0.22	0.25	0.36	0.36
$p = 7$														
1	0	7	17	29	31	35	40	0	0.04	0.13	0.2	0.27	0.36	0.41
2	0	11	17	29	34	36	43	0	0.05	0.13	0.17	0.28	0.33	0.4
3	0	8	20	29	32	37	42	0	0.04	0.08	0.21	0.24	0.33	0.38
4	0	9	16	29	32	38	39	0	0.04	0.15	0.17	0.3	0.32	0.4
5	0	11	16	29	31	38	43	0	0.05	0.14	0.17	0.28	0.34	0.37
6	0	11	18	26	34	36	43	0	0.05	0.14	0.17	0.28	0.32	0.39
7	0	8	20	28	33	35	41	0	0.05	0.09	0.17	0.31	0.34	0.37
8	0	10	21	26	32	35	43	0	0.05	0.15	0.17	0.31	0.36	0.38
9	0	11	16	29	33	36	40	0	0.07	0.15	0.18	0.3	0.34	0.4
10	0	6	12	15	19	21	25	0	0.02	0.16	0.19	0.27	0.33	0.37
$p = 8$														
1	0	7	11	16	18	20	23	0	0.05	0.15	0.19	0.28	0.33	0.38
2	0	6	11	17	19	22	25	0	0.02	0.14	0.18	0.3	0.34	0.38
3	0	5	12	15	18	21	25	0	0.02	0.12	0.18	0.25	0.36	0.38
4	0	5	11	15	18	23	23	0	0.05	0.09	0.23	0.25	0.32	0.4
5	0	5	11	17	20	21	25	0	0.04	0.11	0.22	0.26	0.34	0.41
6	0	5	10	15	20	23	25	0	0.04	0.09	0.2	0.28	0.35	0.37
7	0	5	10	17	19	22	26	0	0.05	0.08	0.17	0.32	0.32	0.37
8	0	7	12	17	19	22	24	0	0.06	0.16	0.18	0.3	0.35	0.4
9	0	6	11	15	18	22	25	0	0.03	0.1	0.17	0.26	0.36	0.38
10	0	14	28	41	42	49	58	0	0.05	0.1	0.23	0.26	0.33	0.41
$p = 9$														
1	0	10	24	37	45	51	58	0	0.03	0.11	0.22	0.31	0.33	0.41
2	0	11	28	39	49	52	60	0	0.06	0.12	0.2	0.31	0.35	0.39
3	0	15	24	36	43	53	61	0	0.06	0.13	0.19	0.29	0.36	0.4
4	0	15	28	37	47	55	56	0	0.03	0.08	0.17	0.26	0.34	0.38
5	0	10	23	39	46	55	60	0	0.05	0.15	0.21	0.25	0.36	0.39
6	0	13	27	41	46	49	59	0	0.04	0.12	0.24	0.31	0.33	0.41
7	0	11	26	37	45	53	61	0	0.04	0.15	0.17	0.29	0.35	0.39
8	0	16	25	41	46	53	56	0	0.06	0.11	0.18	0.32	0.33	0.41
9	0	12	24	40	45	52	57	0	0.03	0.12	0.19	0.24	0.34	0.4
10	0	5	15	19	24	27	30	0	0.03	0.08	0.17	0.29	0.33	0.38
$p = 10$														
1	0	7	15	21	25	26	30	0	0.02	0.09	0.21	0.26	0.35	0.39
2	0	8	14	19	23	26	29	0	0.07	0.13	0.19	0.31	0.34	0.4
3	0	6	12	19	25	26	30	0	0.05	0.11	0.24	0.26	0.35	0.37
4	0	6	14	19	23	29	31	0	0.04	0.15	0.19	0.28	0.35	0.38
5	0	6	13	22	23	28	30	0	0.05	0.09	0.21	0.27	0.32	0.41
6	0	7	15	22	24	27	31	0	0.04	0.16	0.17	0.31	0.36	0.38
7	0	8	13	19	25	28	31	0	0.05	0.12	0.19	0.29	0.34	0.39
8	0	7	14	20	24	28	32	0	0.04	0.14	0.17	0.26	0.35	0.41
9	0	6	14	19	22	28	29	0	0.05	0.16	0.22	0.27	0.35	0.41
10	0	10	17	32	35	39	45	0	0.04	0.1	0.23	0.25	0.35	0.39

Table 3A.3: Price breaks and discount levels (3).

i	$Q_{pi,1}$	$Q_{pi,2}$	$Q_{pi,3}$	$Q_{pi,4}$	$Q_{pi,5}$	$Q_{pi,6}$	$Q_{pi,7}$	$d_{pi,1}$	$d_{pi,2}$	$d_{pi,3}$	$d_{pi,4}$	$d_{pi,5}$	$d_{pi,6}$	$d_{pi,7}$
$p = 11$														
1	0	10	21	28	33	39	44	0	0.07	0.11	0.19	0.29	0.33	0.37
2	0	8	21	30	35	39	42	0	0.02	0.12	0.21	0.29	0.34	0.36
3	0	10	19	29	34	41	45	0	0.02	0.14	0.18	0.25	0.34	0.39
4	0	8	18	27	36	38	45	0	0.02	0.15	0.2	0.25	0.36	0.39
5	0	10	19	31	35	41	45	0	0.01	0.09	0.23	0.28	0.35	0.41
6	0	11	21	29	34	37	43	0	0.03	0.09	0.21	0.31	0.36	0.4
7	0	10	21	31	33	38	44	0	0.04	0.11	0.19	0.28	0.34	0.39
8	0	10	20	27	34	41	45	0	0.03	0.08	0.18	0.24	0.34	0.39
9	0	8	22	28	36	38	46	0	0.06	0.12	0.2	0.24	0.35	0.37
10	0	15	37	62	63	76	89	0	0.05	0.11	0.19	0.3	0.32	0.41
$p = 12$														
1	0	16	35	59	69	74	91	0	0.06	0.09	0.19	0.28	0.36	0.39
2	0	21	40	56	64	74	89	0	0.07	0.1	0.2	0.27	0.34	0.36
3	0	22	36	56	67	76	92	0	0.07	0.15	0.22	0.3	0.32	0.39
4	0	24	42	58	68	76	92	0	0.02	0.13	0.19	0.27	0.35	0.39
5	0	16	41	62	63	80	85	0	0.02	0.12	0.19	0.28	0.33	0.36
6	0	18	38	62	64	77	89	0	0.05	0.15	0.17	0.3	0.33	0.4
7	0	21	34	58	72	77	90	0	0.02	0.09	0.16	0.31	0.34	0.38
8	0	18	35	59	68	74	87	0	0.04	0.14	0.18	0.27	0.33	0.37
9	0	23	37	60	67	81	91	0	0.04	0.14	0.19	0.29	0.33	0.37
10	0	5	12	17	19	25	25	0	0.06	0.12	0.21	0.32	0.34	0.38
$p = 13$														
1	0	5	11	18	20	25	27	0	0.04	0.09	0.24	0.25	0.34	0.37
2	0	4	11	17	20	25	27	0	0.03	0.13	0.23	0.29	0.34	0.39
3	0	5	12	17	20	22	25	0	0.05	0.1	0.2	0.27	0.34	0.36
4	0	6	13	19	19	25	25	0	0.05	0.09	0.18	0.26	0.33	0.37
5	0	6	13	19	19	23	25	0	0.04	0.1	0.22	0.26	0.33	0.37
6	0	5	12	18	19	22	25	0	0.03	0.15	0.22	0.3	0.32	0.4
7	0	5	13	18	22	23	26	0	0.02	0.09	0.22	0.32	0.33	0.38
8	0	4	13	19	21	24	26	0	0.05	0.1	0.22	0.25	0.35	0.4
9	0	7	10	17	21	24	26	0	0.03	0.08	0.17	0.3	0.33	0.39
10	0	3	5	9	10	12	13	0	0.01	0.12	0.21	0.26	0.35	0.4
$p = 14$														
1	0	3	5	9	10	11	12	0	0.06	0.08	0.2	0.32	0.35	0.37
2	0	3	6	9	10	11	13	0	0.02	0.15	0.18	0.3	0.35	0.41
3	0	2	5	9	10	12	13	0	0.04	0.1	0.17	0.27	0.35	0.4
4	0	2	5	9	10	11	13	0	0.05	0.09	0.23	0.3	0.33	0.36
5	0	2	6	8	10	12	13	0	0.03	0.1	0.17	0.28	0.33	0.38
6	0	2	5	8	10	11	12	0	0.05	0.12	0.17	0.3	0.34	0.38
7	0	3	6	9	10	11	13	0	0.03	0.09	0.21	0.27	0.33	0.39
8	0	2	6	8	9	11	12	0	0.05	0.16	0.23	0.25	0.35	0.37
9	0	3	6	8	10	11	13	0	0.05	0.11	0.2	0.29	0.35	0.39
10	0	9	14	24	28	31	37	0	0.04	0.1	0.22	0.31	0.34	0.39
$p = 15$														
1	0	7	15	23	30	34	37	0	0.01	0.08	0.17	0.26	0.33	0.39
2	0	9	14	25	31	31	38	0	0.03	0.1	0.24	0.27	0.35	0.38
3	0	8	15	25	29	31	37	0	0.04	0.08	0.2	0.3	0.33	0.36
4	0	10	14	24	27	32	39	0	0.03	0.12	0.21	0.24	0.34	0.39
5	0	7	17	26	27	31	39	0	0.02	0.14	0.16	0.32	0.34	0.38
6	0	6	16	26	29	34	39	0	0.06	0.13	0.22	0.3	0.34	0.37
7	0	7	15	24	28	34	36	0	0.04	0.09	0.22	0.28	0.36	0.38
8	0	6	15	25	28	32	36	0	0.06	0.09	0.17	0.25	0.35	0.38
9	0	10	15	24	29	31	38	0	0.03	0.14	0.2	0.28	0.36	0.39
10	0	9	22	35	38	45	50	0	0.06	0.15	0.19	0.28	0.33	0.4

Chapter 4

Forecast Evolution in the Final Order Problem with Product Returns

We consider the final order problem of a spare parts provider who faces customer demands and receives product returns. The returns can be remanufactured or disposed/salvaged. There are forecasts for both demands and returns. We consider an evolution of forecasts and examine its influence on remanufacturing policies and costs. We prove the structure of the optimal policy. Using stochastic dynamic programming, we find the following results: (i) In many instances, there exists a flexibility effect of forecast evolution. Under this effect, we place larger final orders and keep more returns than without forecast evolution. This gives us the ability to respond to new information. (ii) There is a pull-away-from-center effect if new information is ignored. If we update the forecasts, we obtain more balanced buying decisions, i.e. we buy less at lower costs and more at higher costs than without forecast evolution. In a numerical study, we find that a consideration of forecast evolution yields significant cost savings, up to 14.4%; especially if we have time-dependent demands. The largest part of the savings is obtained from updating the demand forecasts; updating the return forecasts has only a small influence.

4.1 Introduction

Service parts and after-sales services have a significant influence on the economy. In the U.S., they represent 8% of the gross domestic product (Cohen et al., 2006). In the automotive sector, they account for 36% of all revenues, and for more than 50% in the technology sector (Guajardo et al., 2015). It is estimated that revenues from service parts will have grown from EUR 20 billion in 2012 to almost EUR 100 billion in 2020 in the Chinese automotive industry (McKinsey, 2013).

Product lifecycles, in particular those of electronic products, have become very short and lie in the range of months (Graves and Willems, 2008). The main reason for this development is the speed of the technological progress: companies frequently release new generations of their products in order to remain competitive. Although there is no demand for *new* products after the end of their lifecycle, the sold products are in use for the coming years. That is, due to product failures, a company faces ongoing demand for spare parts throughout a *service period* that can range between 4 and 30 years (Teunter and Klein Haneveld, 2002). As the production of the components of a product usually terminates with the end of the product lifecycle, the company sources a *final order* or *last-time buy*, or produces a *final lot* to satisfy future spare part demand. Customers return failed products for repair and the company replaces the defective part by a spare part. If the replaced part is repairable, it can satisfy a future spare part demand. Customers might also return intact products, e.g., if they switch to a new device. That is, product returns constitute an additional source of supply after the time of the final order. Thus, the company's inventory management comprises the decision on the size of the final order, but also the decision on how to handle product returns: Which share of product returns will be remanufacturable? Is it worth to remanufacture the returned products? Should remanufacturing take place immediately upon the arrival of the product returns? How many product returns should be kept, how many returns should be disposed?

There are three different approaches that deal with the final order problem: (i) the service-driven approach, (ii) the cost-driven approach and (iii) forecasting-based approaches. The aim of the service-driven approach is to determine the inventory decisions such that a certain service level is satisfied. The cost-driven approach does not consider a service level but rather assigns backorder or penalty costs to situations where demand cannot be met. The aim is to minimize costs, i.e. to balance the actual costs from purchasing, holding, remanufacturing or disposal, and those from backorders or product shortages. The forecasting-based approaches do not focus on the inventory problem but on finding the best forecast for the demand over the service period. This forecast is used as the final order. We use a cost-driven approach to model our problem.

We consider a company that manufactures electronic products. The company holds an inventory of spare parts so as to assure a high level of service for its own operations or those of its customers, as might be prescribed by a service contract. We assume that there is a time when the supplier of a spare part stops the production, the end-of-production (EOP). At EOP, the company can place one final order with the supplier to satisfy future demands over a given service period. Returned units can serve as an additional source of supply for the company after EOP, provided that they can be remanufactured. The company has to decide on a) the final order (once), b) the remanufacturing of returned items, and c) the disposal of returned items.

The inventory control policies for the remaining service period depend on the demand and return distributions. We assume that we have unbiased *forecasts* (or *expected values*) for future demands and returns, e.g., from monitoring the installed base. These forecasts are dynamic in that they are updated as the company gains more information on future demands and returns. Previous research on the final order problem predominantly assumes static forecasts. That is, the forecasts available at the time of the final order remain unchanged throughout the remaining service period. Instead of using static forecasts, we assume that the company can update and improve the forecasts whenever there is new information.

For modeling the forecast changes, we use the Martingale Model of Forecast Evolution (MMFE) introduced by Graves et al. (1986) and Heath and Jackson (1994). We consider forecast updates for both returns and demands. The initial return and demand forecasts are time-varying over the remaining service period. We further assume that more distant forecasts are less accurate, and that any forecast update improves the forecast accuracy.

This paper makes several contributions.

(i) We prove that a remanufacture-up-to and dispose-down-to policy is optimal. The remanufacturing is bounded by the available product returns. The remanufacture-up-to and the dispose-down-to levels depend on the current updated forecasts.

(ii) As forecast evolution reduces uncertainty, one would intuitively expect less required inventory, i.e. the final order quantity to decrease compared to the same problem without forecast evolution. We show, however, that this intuition is not always right. In fact, we find a *flexibility effect* of forecast evolution. This effect is twofold:

(a) In many cases, the final order in the *forecast evolution problem* is larger than in the *basic problem* without forecast evolution. This can be observed especially if final order costs are high and remanufacturing costs are low. The basic problem tries to save costs by placing a small final order, betting on a sufficiently high number of product returns in the future and at the same time accepting a high risk of shortages. The forecast evolution problem has a more flexible remanufacturing policy due to more advance information from the forecast updates. This flexibility allows for a reduction of shortage risk; however, if the forecasts become higher through the updates, realizing the policy's flexibility requires having a higher inventory. In one out of four test problems of our numerical study, a larger final order is placed to ensure the availability of the higher inventory.

(b) For the same reason as in (a), the forecast evolution problem keeps more returned items than the basic problem: if increased forecasts require more inventory to exploit the policy's full flexibility, a higher stock of returns allows for raising the inventory sufficiently and attaining the desired inventory levels.

(iii) We generally tend to place larger final orders if purchasing costs are low, and smaller final orders if purchasing costs get higher. But are these decisions adequate in their magnitude? In our numerical analyses, we find that the basic problem exhibits a *pull-away-from-center effect*. In fact, *with* forecast evolution we tend to buy less at low purchasing costs than *without* forecast evolution, and more at high purchasing costs than *without* forecast evolution. That is, if purchasing costs are low, an inventory manager who does not include new future forecast information tends to overrate the savings potential of a large final buy. Similarly, she tends to underrate the threat of shortages if costs are high.

(iv) The evolution of forecasts yields cost savings of up to 14.4%, with an average of 4.5%. The largest cost savings are found if demands are time-dependent. The main part of the cost savings is due to updating the demand forecasts. Updating the return forecasts, however, only has a small influence.

The remainder of the paper is organized as follows. In Section 2, we discuss literature that is related to the final order problem, to forecast evolution, and to remanufacturing systems. In Section 3, we introduce both the forecast evolution model and the basic model, and characterize the optimal policy. In Section 4, we describe the full factorial design of our numerical study and discuss the results. Section 5 concludes the paper.

4.2 Related Literature

Our work touches three bodies of literature: (1) the final order problem, (2) the evolution of forecasts, and (3) inventory systems with remanufacturing and disposal.

4.2.1 Final Order Problem

We review the literature on the cost-driven approach and refer the reader to Pourakbar et al. (2012) for a review on the service-driven and forecasting-based approaches. Note that we assume a dynamic forecast process, but forecasting itself is not part of our work.

Works that set the foundation of the problem include Teunter and Fortuin (1999), Cattani and Souza (2003) and Teunter and Klein Haneveld (2002). Teunter and Fortuin (1999) consider a final order problem with product returns. They derive a closed formula that yields a final order that is close to optimal. They further investigate the reduction of stock to a remove-down-to level. Cattani and Souza (2003) consider the option of delaying the time of the final order. They show that, while the buyer benefits from a delay, the supplier requires additional incentives for agreeing to a delay. Teunter and Klein Haneveld (2002) consider a problem where it is possible to order products throughout the final service period but with an increasing purchasing price. They determine an order-up-to policy with time-dependent, decreasing order-up-to levels.

Pourakbar et al. (2012) consider spare parts for electronic consumer products. Every instance of demand is tied to a defective part. Up to a certain time, defective parts are repaired. After that time, parts are no longer repaired; instead, substitutes are provided. The authors determine both the final order quantity and the optimal time for switching from repair to substitution.

If a customer replaces an old machine that is still intact by a new one, this is called a phase-out. Returns from such phase-outs can be used as spare parts for other old machines that are still in place. Inderfurth and Kleber (2013) consider both the remanufacturing of returned items and the resuming of regular production. Returns are independent of demands and include both repairable defective items and items from phase-out returns. They propose a heuristic that finds the near-optimal final order, remanufacture-up-to and manufacture-up-to levels. In their numerical examples, forecasts have a constant coefficient of variation. That is, the uncertainty of the forecasts does not increase for more distant events.

Pourakbar et al. (2014) consider spare parts for capital goods with both returns of defective items and phase-out returns. They assume deterministic timing and quantities of phase-out returns (and discuss how one could extend their approach to uncertainty). Demand depends on the installed base and is therefore non-stationary. The demand between two consecutive phase-out returns is stationary, though. Every demand is tied to the return of a defective item, which is repairable with a fixed probability. The authors determine the time-varying optimal repair threshold for defective items.

Behfard et al. (2015) also examine the final order problem for capital goods. They consider imperfect repair of returned items; that is, the repair is not successful for all items. In their model, returns are a function of the demand for spare parts. Under imperfect repair, a base-stock policy is no longer optimal. They show that base-stock policies still yield near-optimal approximations. For a broad and structured overview of the literature on the final order problem, the reader is referred to their paper, as well.

4.2.2 Forecast Evolution

The MMFE was first introduced by Graves et al. (1986) and Heath and Jackson (1994). A comprehensive introduction can also be found in Toktay and Wein (2001).

Güllü (1996) applied additive forecast evolution to a capacitated production/inventory system. Initial demand forecasts are identical for all periods. The forecast changes are stationary and forecasts are updated once, i.e. in the period before demand is realized. He showed that a modified forecast-corrected base-stock policy is optimal. The term *modified base-stock policy* is used for a base-stock policy that is subject to capacity constraints. A *forecast-corrected base-stock level* is the value up to which we order after subtracting the future forecasts from our inventory. He further compared a model with forecast evolution to the basic model without forecast evolution and assessed the value of information. He proved that the expected costs in the model with forecast evolution are always lower than without forecast evolution. He also showed that, under mild conditions, the forecast-corrected base-stock level in the model with forecast evolution is smaller than or equal to the one in the basic model.

Toktay and Wein (2001) applied additive forecast evolution to a single-item production stage with stochastic capacity. They considered stationary demand and a general forecast horizon. They found that a modified forecast-corrected base-stock policy is optimal. The forecast-corrected base-stock level depends on all (updated) future forecasts and is quite complex. Therefore, they considered a less complex policy: a modified state-independent forecast-corrected base-stock policy that is independent of the future forecasts. They used heavy traffic approximations to find the optimal base-stock level for this policy.

Iida and Zipkin (2006) considered an *uncapacitated* inventory model and applied both additive and multiplicative forecast evolution. They allowed initial forecasts, forecast changes and parameters to be *non-stationary*. The authors stated that, here too, a base-stock policy that depends on all future forecasts is optimal. Further, they developed bounds for the optimal base-stock levels. They also considered the special case of stationary forecast changes and parameters. They found that a myopic forecast-corrected base-stock policy is optimal for this special case.

4.2.3 Remanufacturing and Disposal

This section briefly reviews general inventory systems with random demands and random returns. In every period, the decision-maker takes three decisions: (i) on manufacturing, (ii) on remanufacturing, and (iii) on disposal.

Simpson (1978) considered a single-stage inventory system with zero lead time. He showed that the optimal policy is fully characterized by manufacture-up-to, remanufacture-up-to and dispose-down-to levels. Inderfurth (1997) extended the problem to positive lead times. He found manufacture-up-to, remanufacture-up-to and dispose-down-to levels to be optimal if manufacturing and remanufacturing lead times are identical, and if the manufacturing lead time exceeds the remanufacturing lead time by one period.

The remanufacturing quantity is limited by the number of available returns. That is, a pure remanufacturing system can be viewed as a capacitated production-inventory system. Ciarallo et al. (1994) and Iida (2002) considered a capacitated production-inventory system with non-stationary demand. They showed that a modified base-stock policy is optimal. DeCroix (2006) included remanufacturing into the multi-echelon model of Clark and Scarf (1960) and derived optimal policies; for all stages but the most upstream one, a base-stock policy is optimal. Atasu et al. (2008) give an overview on product reuse.

4.3 Forecast Evolution Model

We consider a single-product single-stage inventory system at the end-of-production. We have a finite time horizon T, the service period, which is divided into discrete periods $t \in \{1, ..., T\}$. That is, our initial period, when we can place the final order, is period 1. We face random demands D_t and receive random returns R_t in every period. As part of the final order, items can be purchased at a unit cost c_F. After the final order, product returns are the only source of supply. These can be stored at holding cost h_r per period, remanufactured at cost c_R or scrapped at disposal cost c_U. The returned units can have a salvage value instead of a disposal cost, i.e. c_U can be negative. W_t denotes the stock of returned products at the beginning of period t, i.e., before any decision is made and before any demands or returns are realized. I_t denotes the stock of serviceable inventory at the beginning of period t. The holding cost for a unit of serviceable inventory is h_s per period. If a demand in period $t < T$ cannot be served, it is backordered at cost b per period. In period T, unsatisfied demand results in a shortage penalty p per unit. Though not all product returns might be well enough preserved for remanufacturing, the return forecasts and return distributions used in the following refer to *remanufacturable* product returns.

We denote μ_t as the initial demand forecast for time t, and λ_t as the initial return forecast for time t. We denote the demand forecast for D_{t+i}, $i \in \{0, ..., T-t\}$, at time t by $D_{t,t+i}$. Similary, the return forecast for R_{t+i} at time t is denoted by $R_{t,t+i}$. We have $D_{t,t} = D_t$ and $R_{t,t} = R_t$. At the beginning of the service period, we have $D_{1,t} = \mu_t$, $R_{1,t} = \lambda_t$ for all t.

In each period, we update all forecasts within a forecast horizon of H periods. The demand/return forecast change for the current period is the difference between the latest demand/return forecast and the actual demand/return realization. We briefly explain the notation of the forecast changes. The last forecast change for demand D_t is denoted by $\epsilon_{t,1}$. This includes the period when it is realized (t) and indicates that it is the last forecast change (1). The *second to last* forecast change for D_t, taking place in period $t-1$, is denoted by $\epsilon_{t-1,2}$, and so on. That is, we define the forecast change $\epsilon_{t,i}$ to be the update made in time period t to the forecast for period $t+i-1$. It holds

$$D_{t,t} = D_t = D_{t-1,t} + \epsilon_{t,1}. \tag{4.1}$$

$$D_{t-H+i,t} = \mu_t, \qquad \text{if} \quad i \leq 0, \tag{4.2}$$

$$D_{t-H+i,t} = \mu_t + \sum_{j=1}^{i} \epsilon_{t-H+j,H+1-j}, \qquad \text{if} \quad i > 0. \tag{4.3}$$

Similarly, we denote the forecast changes for returns by $\rho_{t,i}$. We have

$$R_{t,t} = R_t = R_{t-1,t} + \rho_{t,1}. \tag{4.4}$$

$$R_{t-H+i,t} = \lambda_t, \qquad \text{if} \quad i \leq 0, \tag{4.5}$$

$$R_{t-H+i,t} = \lambda_t + \sum_{j=1}^{i} \rho_{t-H+j,H+1-j}, \qquad \text{if} \quad i > 0. \tag{4.6}$$

We denote the demand forecast vector at the beginning of period t, prior to the forecast update, by

$$\vec{D_t} = (D_{t-1,t}, ..., D_{t-1,t-2+H}). \tag{4.7}$$

The vector is of length $H-1$, it consists of all forecasts for future demands from period $t-1$. However, if t is close to the end of the planning horizon, i.e. $t-2+H > T$, the vector only contains forecasts up until time T and is of length $T-t+1 < H-1$. The return forecast vector is analogously defined as

$$\vec{R_t} = (R_{t-1,t}, ..., R_{t-1,t-2+H}). \tag{4.8}$$

The demand and return forecast *change* vectors in period t are

$$\vec{\epsilon_t} = (\epsilon_{t,1}, \epsilon_{t,2}, ..., \epsilon_{t,H}), \quad \vec{\rho_t} = (\rho_{t,1}, \rho_{t,2}, ..., \rho_{t,H}), \tag{4.9}$$

respectively. Since they include both the forecast change for the current period t and the forecast changes for the next $H-1$ periods, these vectors are of length H. We assume that the updates within a time period are independent and that the forecast process is a martingale. That is,

$$(i) \qquad \mathbb{E}[\epsilon_{t,i}] = \mathbb{E}[\rho_{t,i}] = 0, \ \forall t, i, \tag{4.10}$$

$$(ii) \qquad Var(\epsilon_{t,i}) = \sigma_{t,i}^2, \quad Var(\rho_{t,i}) = \tau_{t,i}^2, \ \forall t, i, \tag{4.11}$$

$$(iii) \qquad Cov(\epsilon_{s,i}, \epsilon_{t,j}) = 0, \quad Cov(\rho_{s,i}, \rho_{t,j}) = 0, \ \forall i, j, s \neq t. \tag{4.12}$$

Note that the variances are time-dependent. That is, we allow the forecast changes to be non-stationary. This enables us to account for a higher uncertainty for events occuring in a more distant future.

Table 4.1: Notation.

T	Number of periods
t	Time index
I_t	Inventory level at the beginning of period t
W_t	Capacity of remanufacturable returns at the beginning of period t
c_F	Purchasing cost per item in final order
c_R	Remanufacturing cost per item
c_U	Disposal cost/salvage value per item
h_s	Inventory holding cost per item per period
h_r	Return holding cost per item per period
b	Backorder cost per item per period
p	Penalty cost per item
y	Final order quantity at the beginning of period 1
r_t	Remanufacturing quantity at the beginning of period t
u_t	Disposal quantity at the beginning of period t
D_t	Demand in period t
R_t	Return in period t
μ_t	Initial demand forecast for period t
λ_t	Initial return forecast for period t
$D_{t,t+i}$	Demand forecast for period $t+i$ in period t
$R_{t,t+i}$	Return forecast for period $t+i$ in period t
$\epsilon_{t,i}$	Demand forecast change in period t for demand D_{t+i-1}
$\rho_{t,i}$	Return forecast change in period t for return R_{t+i-1}
$\sigma^2_{t,i}$	Variance of demand forecast change $\epsilon_{t,i}$
$\tau^2_{t,i}$	Variance of return forecast change $\rho_{t,i}$

We assume zero lead times for both the delivery of the final order and remanufacturing. At the beginning of period 1, we decide on the final order y. In period $t \in \{2, ..., T\}$, we decide on both the remanufacturing quantity r_t and the disposal quantity u_t. The sequence of events in each period t is as follows. (i) The final order decision is taken ($t = 1$)/remanufacturing and disposal decisions are taken simultaneously ($t > 1$); (ii) forecast updates, demands and returns are realized simultaneously. That is, in each period we make the sourcing and disposal decisions prior to seeing the realization of demand and of the product returns, as well as the forecast updates. Thus, a return from period t can only be remanufactured starting from period $t + 1$. The state transitions for inventory and returns are

$$I_{t+1} = I_t + r_t - D_t = I_t + r_t - D_{t-1,t} - \epsilon_{t,1}, \qquad t \in \{1, .., T-1\}, \tag{4.13}$$

$$W_{t+1} = W_t - r_t - u_t + R_t = W_t - r_t - u_t + R_{t-1,t} + \rho_{t,1}, \qquad t \in \{1, .., T-1\}. \tag{4.14}$$

The notation is summarized in Table 4.1.

Next, we introduce two dynamic programming models: 1) one with forecast evolution, 2) one without forecast evolution (basic model). We will also use a third model that only updates the demand forecasts; however, we will not introduce it separately.

4.3.1 Forecast Evolution Model: Functional Equations

We define the direct expected holding and backorder/penalty costs for period $t \in \{1, ..., T\}$ as

$$L_t(x) = h_s \int_{-\infty}^{x} (x-z) \mathrm{d}\mathbb{P}_{\epsilon_{t,1}}(z) + v \int_{x}^{\infty} (z-x) \mathrm{d}\mathbb{P}_{\epsilon_{t,1}}(z), \tag{4.15}$$

where x is the inventory level after purchasing the final order ($t = 1$)/remanufacturing ($t > 1$) minus the forecast for the demand of the current period. We have $x = I_t + r_t - D_{t-1,t}$ for times $t \in \{2, ...T\}$ and $x = y - \mu_1$ for $t = 1$. We have backorder costs of $v = b$ for times $t \in \{1, ..., T-1\}$ and penalty costs of $v = p$ for $t = T$. $\mathbb{P}_{\epsilon_{t,1}}$ is the distribution function of $\epsilon_{t,1}$ with mean zero and variance $\sigma^2_{t,1}$. We consider the functional equations (4.16)-(4.19). The value function V_t denotes the minimum expected cost from period t until the end of the horizon if optimal decisions are taken in all periods and all states.

In period T, we only need to consider the direct costs, namely the expected holding and penalty costs, the costs for remanufacturing, and the costs for the disposal of returns. We assume that all remaining returns are scrapped or salvaged at the end of the service period. Their disposal takes place as soon as we have decided on the remanufacturing quantity r_T, and thus no holding costs are incurred in period T as part of equation (4.16).

In period $t \in \{1, ..., T-1\}$, we consider both direct costs and cost-to-go. The transitions to future states depend on the forecast changes. The expected costs of the possible future states yield the cost-to-go of the current period. At the beginning of period 1, all forecasts are equal to the initial forecast; there is no need to consider them in the state space. Without loss of generality, we assume that we have no initial serviceable inventory and no initial returns.

Returns in the last period cannot be used within the T period horizon. As a consequence, return forecasts are not part of the state space in period T. The only difference between equations (4.17) and (4.18) is that there are no return forecasts in the cost-to-go function of (4.17).

Period T

$$V_T(I_T, W_T, \vec{D}_{T-1}) = \min_{r_T} \Big\{ c_R \cdot r_T + c_U \cdot (W_T - r_T) + L_T(I_T + r_T - D_{T-1,T}) \Big\} \tag{4.16}$$

Period $t = T - 1$

$$\begin{aligned} V_t(I_t, W_t, \vec{D}_t, \vec{R}_t) = \min_{r_t, u_t} \Big\{ & c_R \cdot r_t + c_U \cdot u_t + L_t(I_t + r_t - D_{t-1,t}) \\ & + h_r \cdot (W_t - r_t - u_t) + \mathbb{E}\big[V_{t+1}(I_t + r_t - D_{t,t}, W_t - r_t - u_t + R_{t,t}, \vec{D}_{t+1})\big] \Big\} \end{aligned} \tag{4.17}$$

Period $t \in \{2, ..., T - 2\}$

$$\begin{aligned} V_t(I_t, W_t, \vec{D}_t, \vec{R}_t) = \min_{r_t, u_t} \Big\{ & c_R \cdot r_t + c_U \cdot u_t + L_t(I_t + r_t - D_{t-1,t}) \\ & + h_r \cdot (W_t - r_t - u_t) + \mathbb{E}\big[V_{t+1}(I_t + r_t - D_{t,t}, W_t - r_t - u_t + R_{t,t}, \vec{D}_{t+1}, \vec{R}_{t+1})\big] \Big\} \end{aligned} \tag{4.18}$$

Period 1

$$V_1 = \min_{y} \Big\{ c_F \cdot y + L_1(y - \mu_1) + \mathbb{E}\big[V_2(y - D_{1,1}, R_{1,1}, \vec{D}_2, \vec{R}_2)\big] \Big\}. \tag{4.19}$$

The expectation of the cost-to-go function is taken with respect to the demand and return forecast change vectors $\vec{\epsilon_t}$ and $\vec{\rho_t}$ of period t. The optimization in each period is subject to

$$r_t + u_t \leq W_t \ \forall t > 1, \tag{4.20}$$

$$y, r_t, u_t \geq 0 \ \forall t. \tag{4.21}$$

We are particularly interested in analyzing the optimal remanufacturing and disposal policies. The following theorem presents insights into their structure.

Theorem 4.1. *The optimal policy for remanufacturing and disposal at time t with forecasts $\vec{D}_t$ and $\vec{R}_t$ is a modified remanufacture-up-to and dispose-down-to policy with remanufacture-up-to level $S_t^{\vec{D},\vec{R}} := S_t(\vec{D}_t, \vec{R}_t)$ and dispose-down-to level $Q_t^{\vec{D},\vec{R}} := Q_t(\vec{D}_t, \vec{R}_t)$. That is, the optimal decisions are given through*

$$r_t(I_t, W_t, \vec{D}_t, \vec{R}_t) = \begin{cases} 0, & \text{if } I_t \geq S_t^{\vec{D},\vec{R}}, \\ S_t^{\vec{D},\vec{R}} - I_t, & \text{if } I_t < S_t^{\vec{D},\vec{R}},\ I_t + W_t \geq S_t^{\vec{D},\vec{R}}, \\ W_t, & \text{if } I_t + W_t < S_t^{\vec{D},\vec{R}}. \end{cases} \tag{4.22}$$

$$u_t(I_t, W_t, \vec{D}_t, \vec{R}_t) = \begin{cases} 0, & \text{if } I_t + W_t \leq Q_t^{\vec{D},\vec{R}}, \\ \min\{W_t, I_t + W_t - Q_t^{\vec{D},\vec{R}}\}, & \text{if } I_t + W_t > Q_t^{\vec{D},\vec{R}}. \end{cases} \tag{4.23}$$

Proof. See Appendix 4.A. □

Theorem 4.1 states that the optimal inventory control policy after EOP consists of two parameters: the remanufacture-up-to level and the dispose-down-to level. These parameters depend on i) the time, ii) the demand forecasts, and iii) the return forecasts. That is, they depend on the *state* of the system, since demand and return forecasts are part of the state space of the stochastic dynamic program. However, the two parameters are invariant to changes in the remaining states, i.e. the states of serviceable inventory and remanufacturable returns, provided that the forecast states remain unchanged.

4.3.2 Basic Model: Functional Equations

In the following, we present the functional equations for the basic model without forecast updates. We use the basic model as a comparison. In theory, we could formulate it as a forecast evolution model with one forecast change both for demands and returns. However, we present the general form: the random variables are demands and returns, rather than forecast changes. The direct cost in the basic model is given by

$$\tilde{L}_t(x) = h_s \int_{-\infty}^{x} (x-d)\mathrm{d}\mathbb{P}_{D_t}(d) + v \int_{x}^{\infty} (d-x)\mathrm{d}\mathbb{P}_{D_t}(d). \tag{4.24}$$

We have $x = I_t + r_t$ for $t \in \{2, ...T\}$ and $x = y$ for $t = 1$. We have backorder costs of $v = b$ for $t \in \{1, ..., T-1\}$ and penalty costs of $v = p$ for $t = T$. $\mathbb{P}_{D_t}$ is the distribution function of D_t.

Period T

$$\tilde{V}_T(I_T, W_T) = \min_{r_T} \Big\{ c_R \cdot r_T + c_U \cdot (W_T - r_T) + \tilde{L}_T(I_T + r_T) \Big\} \tag{4.25}$$

Period $t \in \{2, ..., T-1\}$

$$\tilde{V}_t(I_t, W_t) = \min_{r_t, u_t} \Big\{ c_R \cdot r_t + c_U \cdot u_t + \tilde{L}_t(I_t + r_t) + h_r \cdot (W_t - r_t - u_t) \tag{4.26}$$
$$+ \mathbb{E}\big[\tilde{V}_{t+1}(I_t + r_t - D_t, W_t - r_t - u_t + R_t)\big] \Big\}$$

Period 1

$$\tilde{V}_1 = \min_{y} \Big\{ c_F \cdot y + \tilde{L}_1(y) + \mathbb{E}\big[\tilde{V}_2(y - D_1, R_1)\big] \Big\}. \tag{4.27}$$

The optimization is again subject to (4.20) and (4.21).

4.4 Numerical Study

We conduct a numerical study to assess the effect of forecast evolution. Our aim is to find insights into the different behavior of the problem when forecast evolution is applied. What are the differences in expected costs between the models? What is the difference in the final order decisions and why? What is the difference in remanufacturing and disposal?

4.4.1 Complexity Reduction

Similar to Güllü (1996), we update demand and return forecasts one period before events are realized. Iida and Zipkin (2006) stated that this update, one period before the events, yields the greatest value of information. Thus, our forecast horizon is $H = 2$. This allows us to keep both the state space and the outcome space comparatively small. Güllü (1996) showed that the state of serviceable inventory and the demand forecast for the current period can be collapsed and the result remains optimal. He defined the *forecast-corrected* inventory level $\hat{I}_t = I_t - D_{t-1,t}$. (Note that the expression in Güllü (1996) is *modified* inventory level instead of *forecast-corrected* inventory level. However, we use the term *modified* as it is used in Toktay and Wein (2001), i.e. to highlight the modification by the capacity constraint.) By forecast-corrected, we mean corrected by the *demand forecast*. This allows for the reduction of the state space by one dimension. This state space reduction can also be applied for $H > 2$, i.e. in the general model that we presented in Section 4.3.1. However, for simplicity and in the interest of the similarity of return and demand forecast vectors, we chose to introduce the general version instead of a collapsed one, as the state space reduction cannot be applied to returns. In fact, the capacity of remanufacturable returns, W_t, and the updated return forecast for period t, $R_{t-1,t}$, remain two separate states. Our resulting state space at time t has three dimensions, the outcome space of random events has four dimensions. They are given by

$$(\hat{I}_t, W_t, R_{t-1,t}) \text{ and } (\epsilon_{t,2}, \epsilon_{t,1}, \rho_{t,2}, \rho_{t,1}), \tag{4.28}$$

respectively. In the following, we apply Theorem 4.1 and present the detailed policies for $H = 2$.

Forecast Evolution Model

As we consider the forecast-corrected inventory level $\hat{I}_t$, we adjust the dynamic program from Section 4.3 accordingly. Thus, the state space is three-dimensional and instead of the remanufacture-up-to level $S_t(D_{t-1,t}, R_{t-1,t})$ and the dispose-down-to level $Q_t(D_{t-1,t}, R_{t-1,t})$, we determine the *forecast-corrected* remanufacture-up-to level $\hat{S}_t(R_{t-1,t})$ and the *forecast-corrected* dispose-down-to level $\hat{Q}_t(R_{t-1,t})$. $\hat{S}_t(R_{t-1,t})$ is actually independent of the return forecast change $R_{t-1,t}$. This is quite intuitive. Returns from period t cannot be used until period $t + 1$. Further, we can carry over as many returns from W_t to period $t + 1$ as we want, and there is no setup cost for remanufacturing. That is, we have $\hat{S}_t := \hat{S}_t(\lambda_t) = \hat{S}_t(R_{t-1,t})$. From Theorem 4.1, we find the remanufacturing decision

$$r_t = \min\big\{ \max\{\hat{S}_t - \hat{I}_t, 0\}, W_t\big\}. \tag{4.29}$$

We remanufacture the difference between the forecast-corrected remanufacture-up-to level $\hat{S}_t$ and the forecast-corrected inventory level $\hat{I}_t$. If this difference is larger than the remanufacturing capacity W_t, we remanufacture only W_t items.

The forecast-corrected dispose-down-to level $\hat{Q}_t(R_{t-1,t})$ actually depends on $R_{t-1,t}$. In fact, $\hat{Q}_t(R_{t-1,t})$ is linear in the return forecast change $\rho_{t-1,2}$. That is, $\hat{Q}_t(R_{t-1,t}) = \hat{Q}_t(\lambda_t + \rho_{t-1,2}) = \hat{Q}_t(\lambda_t) - \rho_{t-1,2}$. This is also intuitive. If we expect to have one more return available in period $t + 1$, we dispose of an additional unit in period t. We set

$\hat{Q}_t := \hat{Q}_t(\lambda_t)$. From Theorem 4.1, we find the disposal decision

$$u_t = \min\big\{\max\{\hat{I}_t + W_t + \rho_{t-1,2} - \hat{Q}_t, 0\}, W_t\big\}. \tag{4.30}$$

That is, we want to dispose of remanufacturable units such that the sum of the forecast-corrected inventory level $\hat{I}_t$ and the return level W_t is no more than $\hat{Q}_t - \rho_{t-1,2}$. As we do not dispose of serviceable inventory, the disposal quantity is bounded by the return level W_t.

Basic Model

Similar to the forecast evolution model, we consider a *forecast-corrected* inventory level $\tilde{I}_t := I_t - \mu_t$ and adjust the dynamic program from Section 4.3.2 accordingly. We consider the *forecast-corrected* remanufacture-up-to level $\tilde{S}_t$ and the *forecast-corrected* dispose-down-to level $\tilde{Q}_t$. As forecasts do not change in the basic model, the levels depend only on the initial forecasts, i.e. $\tilde{S}_t := \tilde{S}_t(\lambda_t)$ and $\tilde{Q}_t := \tilde{Q}_t(\lambda_t)$. We use $\tilde{S}_t$ at time t to find the remanufacturing decision:

$$r_t = \min\big\{\max\{\tilde{S}_t - \tilde{I}_t, 0\}, W_t\big\}. \tag{4.31}$$

That is, we remanufacture the difference between the forecast-corrected remanufacture-up-to level $\tilde{S}_t$ and the forecast-corrected inventory level $\tilde{I}_t$. If this difference is larger than the remanufacturing capacity W_t, we remanufacture only W_t items.

We use the dispose-down-to level $\tilde{Q}_t$ at time t to find the disposal decision:

$$u_t = \min\big\{\max\{\tilde{I}_t + W_t - \tilde{Q}_t, 0\}, W_t\big\}. \tag{4.32}$$

We dispose of remanufacturable units such that the sum of the forecast-corrected inventory level $\tilde{I}_t$ and the return level W_t is no higher than $\tilde{Q}_t$. The disposal quantity is bounded by the return level W_t.

Forecast Model Specification

For the forecast model, we expect that the accuracy of the initial demand and return forecasts is less if the respective events lie farther in the future. Thus, we allow for both non-stationary initial forecasts and non-stationary forecast changes. We assume that the forecast changes are discrete and follow a discretized normal distribution with zero mean. We assume that the variances of the forecast changes grow linearly with time. Thus, we set

$$Var(D_t) = Var(\epsilon_{t-1,2}) + Var(\epsilon_{t,1}) = \sigma^2_{t-1,2} + \sigma^2_{t,1} = t \cdot \bar{\sigma}^2, \tag{4.33}$$

$$Var(R_t) = Var(\rho_{t-1,2}) + Var(\rho_{t,1}) = \tau^2_{t-1,2} + \tau^2_{t,1} = t \cdot \bar{\tau}^2, \tag{4.34}$$

with constants $\bar{\sigma}$ and $\bar{\tau}$. From this linearity assumption, it follows that:

$$Var(\epsilon_{t,1}) = 1 \cdot \bar{\sigma}^2 \ \forall t, \quad Var(\rho_{t,1}) = 1 \cdot \bar{\tau}^2 \ \forall t, \tag{4.35}$$

and as a consequence

$$Var(\epsilon_{t-1,2}) = (t-1) \cdot \bar{\sigma}^2 \ \forall t, \quad Var(\rho_{t-1,2}) = (t-1) \cdot \bar{\tau}^2 \ \forall t. \tag{4.36}$$

4.4.2 Experimental Design

For our numerical study, we consider a remaining service period of six years, which we divide into $T = 6$ periods of one year each. In order to evaluate the models, we consider different forecast patterns. That is, we consider both constant and decreasing demand patterns. We do not consider increasing demand patterns, as such a characteristic can be rarely found. For returns, however, we consider constant, decreasing and increasing patterns. With decreasing demand patterns, the return patterns, too, might decrease. However, the opposite can also be observed: the return patterns might increase, e.g., upon the release of a new product. We assume the magnitude of the overall returns to be smaller than the one of demands, or of equal size. We fix demand and return patterns for all cases, i.e. two patterns for demands, and six patterns for returns. We combine every demand pattern with every return pattern, which yields twelve (six by two) different patterns. The patterns are represented by expected mean demands/returns. For demand patterns, we use Γ, for returns patterns, we use Ξ.

$$\Gamma_1 = [10, 10, 10, 10, 10, 10],$$

$$\Gamma_2 = [16, 15, 13, 9, 5, 2],$$

$$\Xi_1 = [5, 5, 5, 5, 5, 5],$$

$$\Xi_2 = [8, 7, 6, 5, 3, 1],$$
$$\Xi_3 = [1, 3, 5, 6, 7, 8],$$
$$\Xi_4 = [10, 10, 10, 10, 10, 10],$$
$$\Xi_5 = [16, 15, 13, 9, 5, 2],$$
$$\Xi_6 = [2, 5, 9, 13, 15, 16].$$

Both demand patterns add up to 60. The *low* return patterns Ξ_1 to Ξ_3 add up to 30. That is, we have less returns than demands. The *high* return patterns Ξ_4 to Ξ_6 add up to 60. That is, the amount of returns resembles the amount of demands. We assume that each product has one of the twelve *characteristic* (Γ, Ξ) pattern combinations. However, the individual forecasts for different products with the same characteristic pattern may be different. Therefore, we generate the initial forecasts, which are the basis for our numerical study, by drawing random numbers from the above patterns. In order to do so, we assume that the above patterns are expected values of Poisson distributions. That is, we have T-dimensional vectors of Poisson rates $\Gamma = (\gamma_1, ..., \gamma_T)$ and $\Xi = (\xi_1, ..., \xi_T)$. An initial demand forecast μ_t is generated by drawing one realization from a Poisson distribution with mean γ_t. An initial return forecast λ_t is generated by drawing one realization from a Poisson distribution with mean ξ_t. From each combination, we draw ten samples. That means we have 120 sets of initial forecasts. For each of these forecasts, we conduct a numerical experiment with the following full factorial design.

Time horizon	$T = 6$
Costs per unit in final order	$c_F \in \{7, ..., 17\}$
Remanufacturing costs	$c_R \in \{8, 12, 16\}$
Salvage value/disposal costs	$c_U \in \{-2, 0, 2\}$
Backorder costs	$b \in \{5, 10\}$
Holding costs	$h_s = 1$
Holding costs of returns	$h_r \in \{0.2, 0.5, 0.8\}$
Penalty costs	$p = 40$
Standard deviation of one-period demand forecast change	$\bar{\sigma} \in \{0.5, 1\}$
Standard deviation of one-period return forecast change	$\bar{\tau} \in \{0.5, 1\}$

All demands and returns are assumed to be independent, and so are the forecast changes. In sum, this yields 285,120 test problems. Each test problem is solved by 1) the forecast evolution model, 2) the basic model, and 3) a model where demand forecasts are updated, but return forecasts are not.

4.4.3 Policy Results

The application of forecast evolution reduces the inventory system's uncertainty. That is, we might expect reduced safety stocks. We expect this reduction to take place in a) the remanufacture-up-to levels and b) the final order quantity. Thus, intuitively, one would assume that final orders are always smaller in the forecast evolution model than in the basic model. However, this intuition does not come true. We briefly explain this in the following example.

Example: Flexibility Effect

We consider a service period with length $T = 2$ and the following parameters:

$$c_F = 10, \qquad c_R = 10, \qquad c_U = 0, \qquad h_s = 1, \qquad h_r = 0.8, \qquad b = 5, \qquad p = 40.$$

Returns are deterministic, demands are stochastic, with

$$\mu_1 = 15, \qquad \mu_2 = 14, \qquad \lambda_1 = 9, \qquad \lambda_2 = 0.$$

That is, $\tau^2_{t,i} \equiv 0 \; \forall t, i \in \{1, 2\}$. We assume a simple structure of demand forecast changes, i.e., we assume $\epsilon_{1,1}, \epsilon_{1,2}, \epsilon_{2,1} \in \{-1, 0, 1\}$ with probabilities

$$\begin{aligned}
&\mathbb{P}(\epsilon_{t,i} = -1) = 0.3, \\
&\mathbb{P}(\epsilon_{t,i} = 0) = 0.4, \\
&\mathbb{P}(\epsilon_{t,i} = 1) = 0.3, \qquad\qquad \forall\, (t, i) \in \{(1, 1), (1, 2), (2, 1)\}.
\end{aligned}$$

For the basic model, this means

$$\begin{aligned}
&\mathbb{P}(D_1 = 14) = 0.3, \qquad\qquad \mathbb{P}(D_2 = 12) = \mathbb{P}(D_2 = 16) = 0.09, \\
&\mathbb{P}(D_1 = 15) = 0.4, \qquad\qquad \mathbb{P}(D_2 = 13) = \mathbb{P}(D_2 = 15) = 0.24,
\end{aligned}$$

$$\mathbb{P}(D_1 = 16) = 0.3, \qquad \mathbb{P}(D_2 = 14) = 0.34.$$

In the forecast evolution model, the optimal final order quantity is $y = 22$ with total expected costs of 308.12. In the basic model, it is $y = 21$ with total expected costs of 311.45. In the forecast evolution model, the optimal policy for the second period is to remanufacture up to the (updated) forecast plus one extra unit. That is, the remanufacture-up-to level depends on $\epsilon_{1,2}$ and is 14 for $\epsilon_{1,2} = -1$, 15 for $\epsilon_{1,2} = 0$, or 16 for $\epsilon_{1,2} = 1$. In the basic model, the optimal policy is to remanufacture up to the (initial) forecast plus one extra unit, i.e. the remanufacture-up-to level is 15. The remaining returns are disposed of in both models.

In the forecast evolution model, we can adjust the second period's remanufacture-up-to level after the first part of the uncertainty has been revealed. This policy is more flexible and yields savings in total expected costs. However, we need to be able to exploit this flexibility in the first place. That is, we need the capacity to raise the inventory level sufficiently if our forecast increased, i.e. $\epsilon_{1,2} = 1$.

The expected costs for ordering (a suboptimal) $y = 21$ in the forecast evolution model are 309.02. In particular, if $\epsilon_{1,1} = 1$ (i.e., $D_1 = 16$), and $\epsilon_{1,2} = 1$ (i.e., $D_{1,2} = 15$), we have a remaining inventory of 5 at the beginning of the second period. We can only remanufacture up to 14, since the capacity from returns is 9. The optimal policy, however, would remanufacture up to 16 in order to avoid high expected shortage costs. If we order (the optimal) $y = 22$, we have an inventory of 6 at the beginning of the second period and can remanufacture up to 15 instead of 14. We benefit from this flexibility: the expected shortage costs decrease. In particular, the savings of expected shortage costs are higher than the costs for purchasing an additional unit in the final order.

This explains that the final order in the forecast evolution model can be greater than in the basic model. We will call this the *flexibility effect* of forecast evolution.

The Size of the Final Order

The purpose of this subsection is to show two effects: the flexibility effect and the pull-away-from-center effect. In order to do so, we analyze the size of the final order in all test problems with respect to the main drivers of the effects, namely the final order cost and the cost of remanufacturing.

In 29.1% of the 285,120 test problems, the final order is smaller in the forecast evolution model than in the basic model. In 44.1%, the final order has the same size in the forecast evolution as in the basic model. And in 26.8%, the final order is greater in the forecast evolution model than in the basic model.

Table 4.2: Differences in final order sizes with respect to final order costs.

c_F	$y_{FE} < y_B$	$y_{FE} = y_B$	$y_{FE} > y_B$
7	67.84%	21.20%	10.96%
8	58.37%	27.76%	13.87%
9	50.00%	29.82%	20.18%
10	43.26%	36.48%	20.26%
11	34.10%	42.01%	23.89%
12	25.42%	47.69%	26.89%
13	18.38%	49.20%	32.42%
14	12.06%	54.55%	33.39%
15	6.88%	56.39%	36.73%
16	2.82%	60.38%	36.80%
17	1.00%	59.14%	39.86%
avg	29.10%	44.06%	26.84%

Table 4.2 presents the difference in the final order size with respect to the final order costs c_F. We denote the final order in the forecast evolution model by y_{FE} and the final order in the basic model by y_B. In general, both models take advantage of smaller final order costs by ordering more. Similarly, they try to order less if final order costs increase. However, the forecast evolution model tends to have smaller final orders than the basic model if final order costs are low (see Table 4.2). The opposite also holds: if final order costs are high, the forecast evolution model tends to order more than the basic model. That means the basic model (i) *overrates* the savings potential of low final order costs by buying *too much*; (ii) *underrates* the threat of buying *too little* when final order costs are high. The forecast evolution model, in contrast, (i) buys less safety stock if final order costs are low; (ii) invests into *flexibility* when final order costs are high. That means it can react to forecast changes because it possesses sufficient stock to do so. Overall, its buying decisions are more balanced. We conclude that, if compared to the forecast evolution model, the basic model exhibits a *pull-away-from-center* effect. Figure 4.1 shows a sketch of this effect. That is, the order range of the forecast evolution model varies less in the final order costs than the one of the basic model.

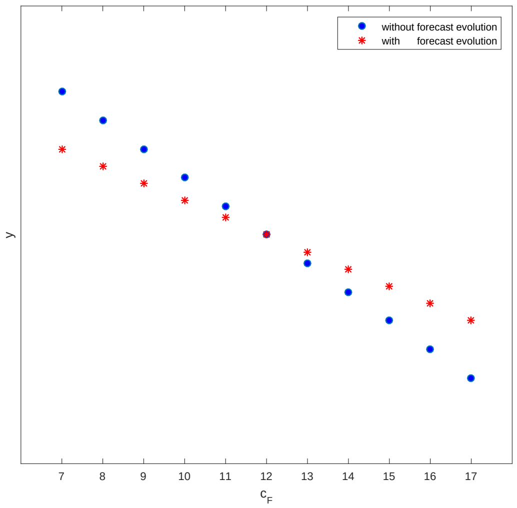


Figure 4.1: Sketch of pull-away-from-center effect.

Figure 4.2 presents the differences in the final order sizes with respect to the remanufacturing costs. We make similar observations. If remanufacturing is cheap, the basic model tends to place smaller final orders than the forecast evolution model. While the basic model tries to save costs on the final order, the larger final order in the forecast evolution model yields a higher flexibility.

Remanufacturing and Disposal Levels

Next, we examine the difference in the serviceable inventories of both models and the difference in returns of products to be kept. We show the second characteristic of the flexibility effect, which is based on the differences in the product returns. The forecast-corrected remanufacture-up-to level can be interpreted as *safety stock* of the serviceable inventory; we compare the safety stocks of both models.

Our analyses are based on the parameters of the forecast evolution model and the basic model, namely the forecast-corrected remanufacture-up-to levels $\tilde{S}_t$ and $\hat{S}_t$ and dispose-down-to levels $\tilde{Q}_t$ and $\hat{Q}_t$. The new return level after remanufacturing and disposal is $\tilde{Q}_t - \tilde{S}_t$ in the basic model and $\hat{Q}_t - \hat{S}_t$ in the forecast evolution model. We define the absolute differences

$$\Delta_t = \tilde{S}_t - \hat{S}_t, \ \forall t, \tag{4.37}$$

$$\delta_t = \tilde{Q}_t - \hat{Q}_t, \ \forall t, \tag{4.38}$$

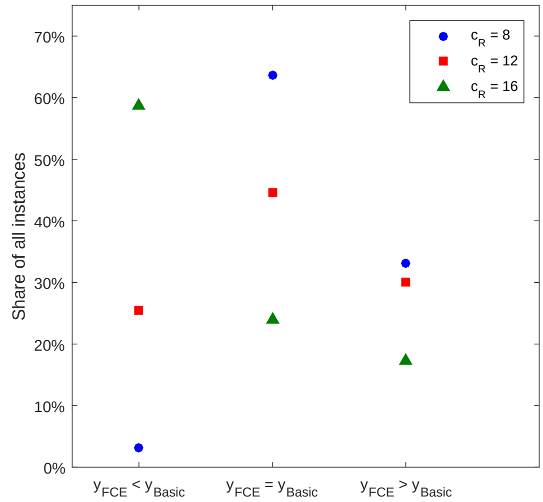


Figure 4.2: Differences in final order sizes with respect to remanufacturing costs.

$$\omega_t = (\tilde{Q}_t - \tilde{S}_t) - (\hat{Q}_t - \hat{S}_t) = \delta_t - \Delta_t, \ \forall t. \tag{4.39}$$

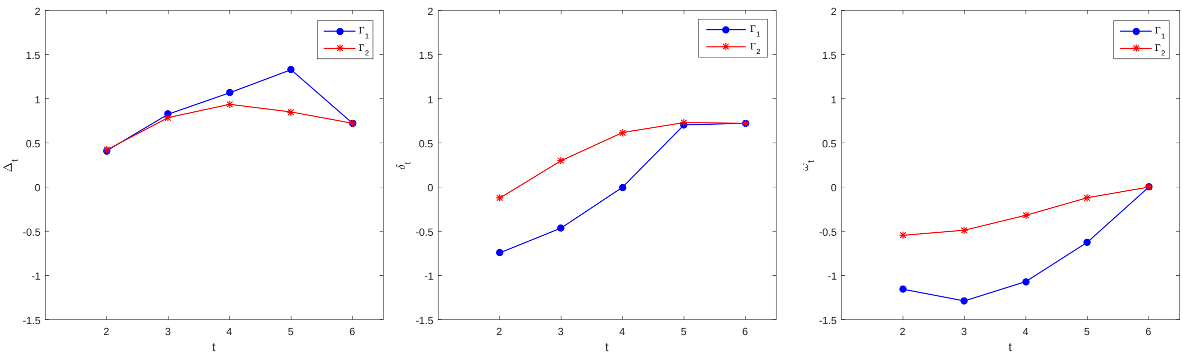
Remanufacturing and disposal take place in periods $t \in \{2, ..., 6\}$. Within each of the two demand patterns Γ_1 and Γ_2, there is almost no difference between the return patterns $\Xi_i, \ i \in \{1, .., 6\}$. Therefore, our analysis focuses on Γ_1 and Γ_2.

In Figure 4.3a, the difference Δ_t between the (higher) remanufacture-up-to level from the basic model and the (lower) level from the forecast evolution model is presented. The results are in line with Güllü (1996) who proved that the forecast-corrected base-stock levels in the forecast evolution model are lower than or equal to the ones from the basic model. That is, $\Delta_t \geq 0 \ \forall t$.

In period $T = 6$, each unit in W_T is either remanufactured or disposed of. That is, we dispose down to the forecast-corrected dispose-down-to level. We have $\hat{Q}_T = \hat{S}_T$ and $\tilde{Q}_T = \tilde{S}_T$, and it holds that $\delta_T = \Delta_T$. This can be observed when you compare Figures 4.3a and 4.3b.

We further observe that the absolute differences between $\tilde{S}_T$ and $\hat{S}_T$ are equal for both demand patterns Γ_1 and Γ_2. Consequently, the optimization only considers direct costs in the last period. Due to the definition of the forecast changes, the direct costs have identical uncertainty structures for $\mu_T = 10$ and $\mu_T = 2$.

For the other periods $t \in \{2, ..., 6\}$, if demand is stationary (Γ_1), Δ_t increases from



(a) Average Δ_t. (b) Average δ_t. (c) Average ω_t.

Figure 4.3: Comparison of remanufacturing and disposal parameters.

time 2 to time 5. While the forecast evolution model can use new information, the basic model suffers from an increasing uncertainty. The increasing uncertainty leads to increasing safety stocks; an increasing Δ_t is the direct consequence. If demand is non-stationary (Γ_2), this effect is curtailed, since most demand takes place in the first periods. This means that the increasing uncertainty of later periods influences only a smaller share of the demand, causing a smaller difference in safety stocks.

As we mentioned earlier, the *forecast-corrected* remanufacture-up-to levels in the forecast evolution model are always smaller than the ones in the basic model. Consequently, the average serviceable inventory in the forecast evolution model is smaller than in the basic model. That is, there is less safety stock.

If demand forecast changes are positive, the remanufacture-up-to levels in the forecast evolution model increase. (Note that this does not imply an increase of the *forecast-corrected* remanufacture-up-to levels.) In order to attain these levels, the model requires more capacity for remanufacturing. That is, it keeps more returns than the basic model (if sufficient returns are already available). In Figure 4.3c, we can see that $\omega_t < 0$, $t \in \{2, 3, 4, 5\}$, i.e. the forecast evolution model has more remaining returns after remanufacturing and disposal than the basic model in all relevant periods. (The remaining returns for $t = 6$ are zero in both models.) This is the second main characteristic of the flexibility effect of forecast evolution we introduced in Section 4.4.3.

Lastly, we consider the joint quantity of forecast-corrected serviceable inventory and product returns. In periods 2, 3, and 4, δ_t is negative if demand is stationary (Figure 4.3b), i.e. the forecast evolution model keeps a higher joint quantity than the basic model. That is, the difference between the product returns that are kept in the two models (flexibility) is larger than the difference between the forecast-corrected serviceable inventories of the two models (safety stock). At time 5, however, δ_t turns to a positive value. The (time-increasing) safety stock of the basic model now surpasses the required flexibility of the forecast evolution model. For the non-stationary demand structure, with earlier demands, this happens even earlier.

4.4.4 Cost Results

This subsection makes cost comparisons. The decision relevant costs of the forecast evolution model and the basic model are compared. We investigate how the individual demand and return forecast updates influence the outcome and present parameter combinations that yield a high cost difference. Lastly, we determine those cost components that are the main drivers of the cost differences between the models.

If we consider the overall cost structure, we find that final order costs constitute more than 50% throughout all test problems. The predominant part of the final order remains fixed in both the forecast evolution model and the basic model. In particular, the minimum of y_{FE} and y_B remains fixed in both models, and so are the costs $c_F \min\{y_{FE}, y_B\}$. In order to compare both models, we set our focus on those costs that are actually influenced by our decision. We define the *decision relevant costs*

$$V_1^D = V_1 - c_F \min\{y_{FE}, y_B\}, \quad \tilde{V}_1^D = \tilde{V}_1 - c_F \min\{y_{FE}, y_B\}, \tag{4.40}$$

where V_1 and $\tilde{V}_1$ are the minimum expected costs of the dynamic programs in Sections 4.3.1 and 4.3.2, respectively. Note that a positive salvage value, i.e. $c_U = -2$, yields a profit that causes a cost reduction. That is, V_1^D and $\tilde{V}_1^D$ can be very small and a comparison in terms of relative cost savings is not meaningful. (In fact, for low final order costs, high remanufacturing costs, and large returns, V_1^D and $\tilde{V}_1^D$ become negative, i.e. they represent a profit.) Thus, we focus our analysis of decision relevant costs on test problems with disposal costs $c_U \geq 0$. Figure 4.4 shows the average and maximum savings. The average savings range between 3.1% and 6.5% among the patterns, with a mean of 4.5%. The maximum savings range between 9.5% and 14.4%. Further, we can observe that the savings potential of non-stationary demand (Γ_2) is larger.

In order to include the test problems with a positive salvage value into our analyses,

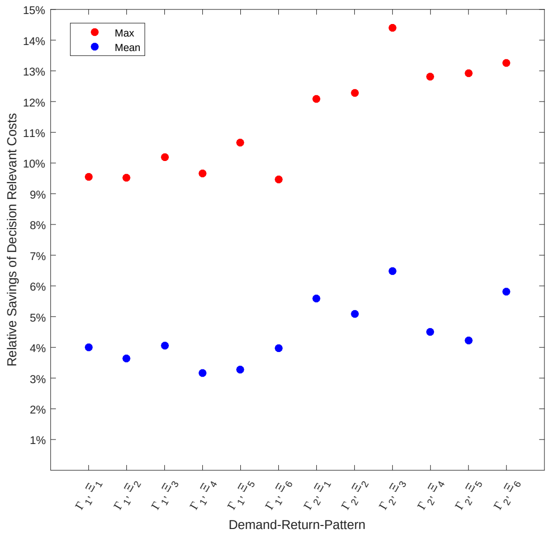


Figure 4.4: Relative savings of decision relevant costs.

we also consider total expected costs V_1 and $\tilde{V}_1$ instead of decision relevant costs V_1^D and $\tilde{V}_1^D$. We find that the average cost savings of forecast evolution are even higher for test problems with a positive salvage value ($c_U = -2$) than for problems with $c_U \geq 0$. In fact, test problems with a positive salvage value yield on average 0.5% higher cost savings than problems with positive disposal costs ($c_U = 2$). That is, our previous conclusion from Figure 4.4 can be underlined if the analysis is extended to a positive salvage value.

We further look at the parameter combinations where forecast evolution has the highest savings potentials. We find that low remanufacturing costs ($c_R = 8$) and a positive salvage value ($c_U = -2$) yield the highest proportion of high savings. Low remanufacturing costs and zero salvage value are also promising, at least if final order costs are low. So is the combination of medium-sized remanufacturing costs ($c_R = 12$) and positive disposal costs ($c_U = 2$).

We also solved each test problem under the assumption that we only update the demand forecasts. By comparing this setting to the case where we update both the demand and return forecasts, we can assess the incremental value from updating the returns forecasts. We find that updating the return forecasts has only a small influence. In fact, return forecast updates on average make up for only one tenth of the cost savings; the remaining nine tenth are caused by updating the demand forecasts. The explanation for this observation is that the return forecast has only an implicit effect on

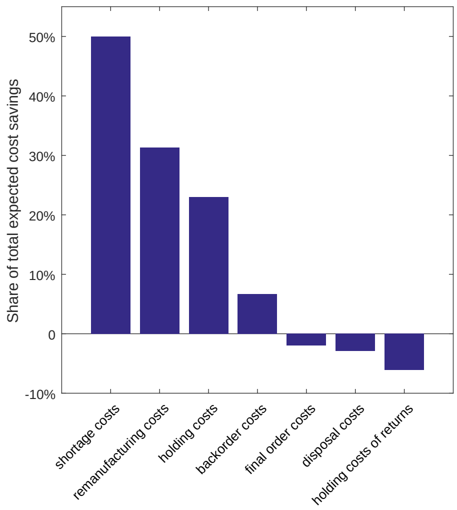


Figure 4.5: Share of total expected cost savings (I).

the decisions. It indicates whether the capacity for remanufacturing in the next period will be smaller or larger. This influences the number of remanufacturable items that are kept or disposed. However, if the capacity is not used up to its limit, the change of the return forecast is not very likely to influence the remanufacturing and disposal decisions to a high degree.

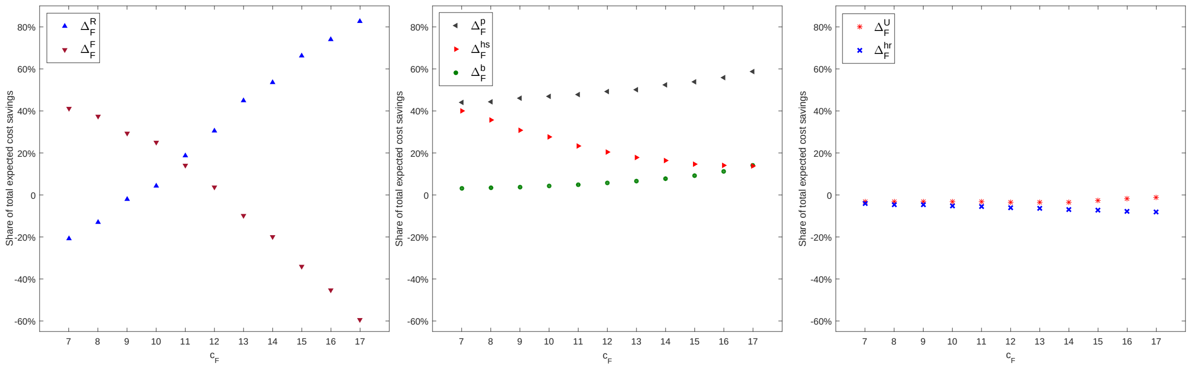
We examine the savings from forecast evolution in more detail. For each test problem, we determine the relative amount of the cost savings from forecast evolution that is attributed to each of the types of costs. In Figure 4.5, we report their averages over all test problems.

(i) Reduced penalty costs yield the highest per centage of cost savings. The explanation is the same as in the example in Section 4.4.3: the higher flexibility of the forecast evolution model allows for a reduction of shortages in period T. The same reasoning can be applied to the savings in backorder costs.

(ii) There are large percent savings in both holding and remanufacturing costs. Both effects are due to the lower remanufacture-up-to levels in the forecast evolution model.

(iii) In the previous section, we observed that the forecast evolution model holds more returns in order to preserve flexibility. Consequently, holding costs of returns are larger in the forecast evolution model, and the difference in Figure 4.5 is negative.

We observed earlier that, for low final order costs, the forecast evolution model tends to place smaller final orders than the basic model; for high final order costs, the forecast



(a) Expected c_F, c_R. (b) Expected p, b, h_s. (c) Expected c_U, h_r.

Figure 4.6: Share of total expected cost savings (II).

evolution model tends to place larger final orders than the basic model. This explains why, if the average over all instances is taken, the final order costs of the forecast evolution model are higher than in the basic model (the negative difference for final order costs in Figure 4.5). Similarly, a positive salvage value has a larger impact than positive disposal costs (the negative difference for disposal costs).

Figure 4.6 presents the share of the cost savings per category with respect to different final order costs. We consider the expected savings of backorder costs (Δ_F^b), holding costs (Δ_F^{hs}), holding costs of returns (Δ_F^{hr}), penalty costs (Δ_F^p), remanufacturing costs (Δ_F^R), disposal costs (Δ_F^U), and final order costs (Δ_F^F) for the forecast evolution model. We use subscript F as opposed to subscript t in Section 4.4.3.

For $c_F \leq 12$, the forecast evolution model orders less than the basic model and saves final order costs. For $c_F \geq 13$, the forecast evolution model orders more. For remanufacturing, the effect is vice versa: with small c_F (and thus a smaller final order), there is more need for remanufacturing in the forecast evolution model than in the basic model (Figure 4.6a). The basic model already possesses more inventory due to its pull-away-from-center effect.

Savings for holding costs of returns and disposal costs are (slightly) negative for all c_F (Figure 4.6c). Savings for penalty, backorder, and holdings costs are positive for all c_F (Figure 4.6b). The expected savings in shortage costs are higher than 40% for all c_F. As we explained above, the basic model underrates the threat of shortages when costs for the final order increase; consequently, the expected shortage costs increase in the final order costs.

4.5 Conclusion

We considered the final order problem of an electronics service provider with random demands, random returns, remanufacturing and disposal options. We examined the impact of applying forecast evolution by comparing a model with forecast evolution to a basic model without. We considered forecast updates for both returns and demands. Both initial return and demand forecasts were non-stationary, as were the forecast changes: we considered a higher uncertainty for more distant events.

Our main results were: (i) We proved that a modified remanufacture-up-to dispose-down-to policy is optimal in both the forecast evolution model and the basic model. (ii) We found a flexibility effect of forecast evolution: In 26.8% of 285,120 test problems, the optimal final order is larger in the forecast evolution model than in the basic model. Further, the forecast evolution model keeps more remanufacturable returns than the basic model. Both facets of the effect enable the forecast evolution model to exploit its additional information by applying a more flexible policy. (iii) We found that neglecting new information exhibits a pull-away-from-center effect. If final order costs are low, both models tend to place larger final orders; if final order costs are high, both models tend to place smaller final orders. However, without forecast evolution, the final order size at low costs is too large, and at high costs, it is too small. Forecast evolution yields more balanced buying decisions. (iv) Forecast evolution yielded decision relevant cost savings of up to 14.4%. The average savings of decision relevant costs were 4.5%. The largest impact of forecast evolution can be found if demands are non-stationary. While updating the demand forecasts influences the total expected costs significantly, the effect of updating the return forecasts is only small.

Further research might consider a correlation between demands and returns. Another interesting approach might be to extend the problem by adding a production option, which would serve as an additional source of inventory. As the stochastic dynamic program suffers from the curse of dimensionality for a larger forecast horizon, the investigation of approximate methods would also be of interest.

4.A Proof of Theorem 4.1

The proof is similar to the one in Simpson (1978). Without loss of generality, we only consider one particular demand forecast $\vec{D}_t$ and one particular return forecast $\vec{R}_t$. Similar to V_t, we define

$$\begin{aligned}\mathcal{H}(r_t, u_t) = c_R \cdot r_t + c_U \cdot u_t + L_t(I_t + r_t - D_{t-1,t}) + h_r \cdot (W_t - r_t - u_t) \\ + \mathbb{E}\big[V_{t+1}(I_t + r_t - D_{t,t}, W_t - r_t - u_t + R_{t,t}, \vec{D}_{t+1}, \vec{R}_{t+1})\big].\end{aligned} \tag{4.41}$$

We define the *Lagrangian function*

$$\mathcal{L}(r_t, u_t, l_t) = \mathcal{H}(r_t, u_t) + l_t(r_t + u_t - W_t), \tag{4.42}$$

where l_t is the *Lagrangian operator*. In the following, we apply the Kuhn-Tucker saddle point conditions (Kuhn and Tucker, 1951). The Langrangian function is convex in the decision variables r_t and u_t, as both direct costs and cost-to-go in $\mathcal{H}$ are convex. Further, the Lagrangian function is linear in the Lagrangian operator l_t. Therefore, the saddle point conditions are sufficient for a global minimum. Let r_t^*, u_t^* and l_t^* denote the optimal solutions to the minimization problem. For ease of notation, we set

$$I_t^* = I_t + r_t^*, \tag{4.43}$$

$$W_t^* = W_t - r_t^* - u_t^*, \tag{4.44}$$

$$\mathbb{A}_{t+1}(I_t^*, W_t^*, \vec{D}_{t+1}, \vec{R}_{t+1}) = \mathbb{E}\big[V_{t+1}(I_t^* - D_{t,t}, W_t^* + R_{t,t}, \vec{D}_{t+1}, \vec{R}_{t+1})\big], \tag{4.45}$$

$$\mathbb{L}_t(I_t^*) = L_t(I_t^* - D_{t-1,t}). \tag{4.46}$$

For any function $f(\alpha, \beta)$, we define the partial derivatives $f^\alpha(\alpha, \beta) := \frac{\partial f(\alpha,\beta)}{\partial \alpha}$ and $f^\beta(\alpha, \beta) := \frac{\partial f(\alpha,\beta)}{\partial \beta}$. That is, superscript α denotes the partial derivative with respect to the first position and superscript β the partial derivative with respect to the second. We further define $f^{\alpha-\beta}(\alpha, \beta) := f^\alpha(\alpha, \beta) - f^\beta(\alpha, \beta)$. In our setting, the first position is $\alpha = r_t$ and the second position is $\beta = u_t$. The saddle point conditions are

$$\text{if } r_t^* = 0, \text{ then } \mathbb{L}_t^\alpha(I_t^*) + \mathbb{A}_{t+1}^{\alpha-\beta}(I_t^*, W_t^*, \vec{D}_{t+1}, \vec{R}_{t+1}) \geq -c_R + h_r - l_t^*, \tag{4.47}$$

$$\text{if } r_t^* > 0, \text{ then } \mathbb{L}_t^\alpha(I_t^*) + \mathbb{A}_{t+1}^{\alpha-\beta}(I_t^*, W_t^*, \vec{D}_{t+1}, \vec{R}_{t+1}) = -c_R + h_r - l_t^*, \tag{4.48}$$

$$\text{if } u_t^* = 0, \text{ then } l_t^* - \mathbb{A}_{t+1}^{\beta}(I_t^*, W_t^*, \vec{D}_{t+1}, \vec{R}_{t+1}) + c_U - h_r \geq 0, \tag{4.49}$$

$$\text{if } u_t^* > 0, \text{ then } l_t^* - \mathbb{A}_{t+1}^{\beta}(I_t^*, W_t^*, \vec{D}_{t+1}, \vec{R}_{t+1}) + c_U - h_r = 0, \tag{4.50}$$

$$\text{if } l_t^* = 0, \text{ then } W_t^* \geq 0, \tag{4.51}$$

$$\text{if } l_t^* > 0, \text{ then } W_t^* = 0. \tag{4.52}$$

The optimal remanufacture-up-to levels $S_t(\vec{D}_t, \vec{R}_t)$ and dispose-down-to levels $Q_t(\vec{D}_t, \vec{R}_t) = S_t(\vec{D}_t, \vec{R}_t) + q_t(\vec{D}_t, \vec{R}_t)$ are such that

$$\mathbb{L}_t^{\alpha}\big(S_t(\vec{D}_t, \vec{R}_t)\big) + \mathbb{A}_{t+1}^{\alpha}\big(S_t(\vec{D}_t, \vec{R}_t), q_t(\vec{D}_t, \vec{R}_t), \vec{D}_t, \vec{R}_t\big) = -c_R + c_U, \tag{4.53}$$

$$\mathbb{A}_{t+1}^{\beta}\big(S_t(\vec{D}_t, \vec{R}_t), q_t(\vec{D}_t, \vec{R}_t), \vec{D}_t, \vec{R}_t\big) = -h_r + c_U. \tag{4.54}$$

If there is no $q_t(\vec{D}_t, \vec{R}_t)$ satisfying equation (4.54), then it holds that $q_t(\vec{D}_t, \vec{R}_t) = 0$ and $S_t(\vec{D}_t, \vec{R}_t)$ is defined by equation (4.53) only. (In period T, there is no u_T and the decision on disposal is made implicitly; there is no cost-to-go function either. The above equations are adjusted accordingly.) There are six different regions:

(1) Remanufacture all returns, if
$I_t + W_t < S_t(\vec{D}_t, \vec{R}_t)$,

(2) Remanufacture some returns, if
$S_t(\vec{D}_t, \vec{R}_t) < I_t + W_t \leq Q_t(\vec{D}_t, \vec{R}_t) \ \wedge \ I_t < S_t(\vec{D}_t, \vec{R}_t)$,

(3) Remanufacture some returns and dispose of some returns, if
$I_t + W_t > Q_t(\vec{D}_t, \vec{R}_t) \ \wedge \ I_t < S_t(\vec{D}_t, \vec{R}_t)$,

(4) Do nothing, if
$I_t \geq S_t(\vec{D}_t, \vec{R}_t) \ \wedge \ I_t + W_t \leq Q_t(\vec{D}_t, \vec{R}_t)$,

(5) Dispose of some returns, if
$S_t(\vec{D}_t, \vec{R}_t) \leq I_t < Q_t(\vec{D}_t, \vec{R}_t) \ \wedge \ I_t + W_t > Q_t(\vec{D}_t, \vec{R}_t)$,

(6) Dispose of all returns, if
$I_t \geq S_t(\vec{D}_t, \vec{R}_t) \ \wedge \ I_t + W_t > Q_t(\vec{D}_t, \vec{R}_t)$.

From applying induction, starting at period T, we find that the proposed policy satisfies the saddle point conditions for all regions and all time periods. As this is analogous to Simpson (1978), it is not shown. As there were no limitations on $\vec{D}_t$ and $\vec{R}_t$, the result holds for all forecasts. $\square$

Chapter 5

Conclusion

5.1 Summary

In many industries, inventory management is a crucial factor towards promoting a company's success. This thesis aimed at improving the inventory management by addressing three important topics.

First, we considered inventory systems that face time-dependent Poisson demand. We applied unit-tracking, that is, we decomposed the problem into single-unit subproblems, in order to find time-dependent ordering policies. For a single-stage inventory system, we found a decision rule that yields the optimal policy, i.e. a base-stock policy with varying base-stock levels. For the one-warehouse multi-retailer system, we found a myopic decision rule that yields time-dependent base-stock levels and performs very well in the majority of cases. We compared its costs to the costs of repeatedly applying a time-independent policy and found large savings. In order to derive these results, observations on the arrival process were necessary: as retailers, due to their stock adjustments, do not reorder every customer demand, the demand at the warehouse is not Poisson distributed. Therefore, we characterized the distribution of arrival times at the warehouse. Lastly, we found a simple alternative method that is based on the stationary model with adjustments by a simple time-shift and yields very good results. This can be of special interest for practitioners as it allows them to bypass the rather technical derivations of non-stationary arrival distributions at all sites.

The second problem addressed in this thesis focused on the selection of suppliers and the allocation of orders. A buyer, who we might think of as a manufacturer, needs to procure different products. When taking the sourcing decisions, the buyer aims at negotiating large discounts. There were two types of discounts: quantity discounts, i.e. the discount applies to the particular product if the order for that product is sufficiently large; and business volume discounts, that is, the discount takes the entire sales volume of the supplier, i.e. all products, into account. In a deterministic world, the buyer's result would be an optimal strategy with very few suppliers and large economies of scale. However, if there is a chance that some suppliers will fail to deliver, which is a real threat and cannot be ignored, the strategy needs to be adjusted for the loss that accompanies a supplier failure. We developed a MILP model that finds the optimal supplier selection and order allocation strategy that accounts for both economies of scale and failure risk, and at the same time balances the trade-off between them.

The third problem addressed in this thesis dealt with the placement of a final order for a product at the end of its production. This final order is used for satisfying future demand; however, it can be complemented by the remanufacturing of future product returns. The inventory manager has to decide on the size of the final order, the remanufacturing and the disposal policies for product returns. These decisions are subject to uncertainty originating from two sources: future demands and returns. Our approach to the problem included an evolution of forecasts, i.e. the additional information arising in every period is used for adjusting the demand and return forecasts for this new information. We proved the structure of the optimal policy for this problem. We compared our approach to an approach that does not update forecasts, i.e. retains a higher future uncertainty, and found both significant cost savings from using forecast evolution and interesting effects on the size of the final order, the remanufacturing and disposal policies. In particular, we found a flexibility effect of forecast evolution. That is, in many problem instances, it is optimal to keep more inventory, either through a larger final order or through fewer disposals, because this provides the flexibility to make use of the additional information gained by updating the forecasts. We further found a pull-away-from-center effect: if forecasts are not updated, the final order size is too large if costs are low, and too small if costs are high. This means that the decision maker overestimates the savings potential of final order costs if she does not include forecast updates.

5.2 Limitations and Further Research

For a one-warehouse multi-retailer problem, the optimal inventory policy is still unknown. Even if the problem is stationary, existing solutions require assumptions such as first-come first-served or a balance assumption, all of which make them suboptimal. Finding the optimal policy would be of great interest for both stationary and non-stationary demand.

For further research, one might consider the final order problem with supplier selection and failure risk. Without failure risk, the buyer places the final order as late as possible (Cattani and Souza, 2003) with a single supplier. However, this might change if the supplier has a dynamic risk of failure. Under a single-supplier strategy, the buyer might start building up inventory for the final phase earlier, which reduces the size of the final order. In the case of a multi-supplier strategy, the final order might be split among several suppliers. In particular, if we include forecast evolution, updates of the expected failure risk might cause a switch from a single-supplier to a multi-supplier strategy, along with a change in the timing and size of the final order, the remanufacturing and the disposal policies.

Bibliography

Abhyankar, HS and SC Graves (2001). Creating an inventory hedge for Markov-modulated Poisson demand: An application and model. *Manufacturing & Service Operations Management* 3(4), pp. 306–320.

Arthur D. Little (2008). Automotive after sales 2015. *Arthur D. Little.*

Atasu, A, VDR Guide, and LN Van Wassenhove (2008). Product reuse economics in closed-loop supply chain research. *Production and Operations Management* 17(5), pp. 483–496.

Axsäter, S (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research* 38(1), pp. 64–69.

— (2006). *Inventory Control.* Springer, Berlin.

— (2007). On the first come–first served rule in multi-echelon inventory control. *Naval Research Logistics* 54(5), pp. 485–491.

Axsäter, S and J Marklund (2008). Optimal position-based warehouse ordering in divergent two-echelon inventory systems. *Operations Research* 56(4), pp. 976–991.

Baller, S, A Di Battista, S Dutta, and B Lanvin (2016). The networked readiness index 2016. *The global information technology report 2016.* Ed. by S Baller, S Dutta, and B Lanvin. World Economic Forum, Geneva, pp. 3–37.

Behfard, S, MC van der Heijden, A Al Hanbali, and WHM Zijm (2015). Last time buy and repair decisions for spare parts. *European Journal of Operational Research* 244(2), pp. 498–510.

Beichelt, F (2006). *Stochastic processes in science, engineering and finance.* Chapman & Hall, Boca Raton.

Bohner, C and S Minner (2017a). Inventory systems with non-homogeneous Poisson demand. *Working paper.* TUM School of Management. Technical University of Munich.

— (2017b). Supplier selection under failure risk, quantity and business volume discounts. *Computers & Industrial Engineering* 104, pp. 145–155.

Bohner, C, SC Graves, and S Minner (2017). Forecast evolution in the final order problem with product returns. *Working paper*. TUM School of Management. Technical University of Munich.

Burke, GJ, SS Erenguc, and AJ Vakharia (2008). Optimal requirement allocation among quantity-discount quoting suppliers. *Operations Management Research* 1(1), pp. 53–60.

Cattani, KD and GC Souza (2003). Good buy? Delaying end-of-life purchases. *European Journal of Operational Research* 146(1), pp. 216–228.

Chen, F and J-S Song (2001). Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research* 49(2), pp. 226–234.

Ciarallo, FW, R Akella, and TE Morton (1994). A periodic review, production planning model with uncertain capacity and uncertain demand – optimality of extended myopic policies. *Management Science* 40(3), pp. 320–332.

Clark, AJ and H Scarf (1960). Optimal policies for a multi-echelon inventory problem. *Management Science* 6(4), pp. 475–490.

Cohen, MA, N Agrawal, and V Agrawal (2006). Winning in the aftermarket. *Harvard Business Review* 84(5), pp. 129–138.

DeCroix, GA (2006). Optimal policy for a multiechelon inventory system with remanufacturing. *Operations Research* 54(3), pp. 532–543.

Fahimnia, B, CS Tang, H Davarzani, and J Sarkis (2015). Quantitative models for managing supply chain risks: A review. *European Journal of Operational Research* 247(1), pp. 1–15.

Graves, SC (1985). A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science* 31(10), pp. 1247–1256.

Graves, SC and SP Willems (2008). Strategic inventory placement in supply chains: Nonstationary demand. *Manufacturing & Service Operations Management* 10(2), pp. 278–287.

Graves, SC, HC Meal, S Dasu, and Y Qui (1986). Two-stage production planning in a dynamic environment. *Multi-Stage Production Planning and Inventory Control*. Ed. by S Axsäter, C Schneeweiss, and E Silver. Springer, Berlin, pp. 9–43.

Guajardo, JA, MA Cohen, and S Netessine (2015). Service competition and product quality in the U.S. automobile industry. *Management Science* 62(7), pp. 1860–1877.

Güllü, R (1996). On the value of information in dynamic production/inventory problems under forecast evolution. *Naval Research Logistics* 43(2), pp. 289–303.

Heath, DC and PL Jackson (1994). Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. *IIE Transactions* 26(3), pp. 17–30.

Hillestad, RJ (1982). *Dyna-METRIC: Dynamic multi-echelon technique for recoverable item control*. Rand Corporation.

Ho, W, X Xu, and PK Dey (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research* 202(1), pp. 16–24.

Howard, C and J Marklund (2011). Evaluation of stock allocation policies in a divergent inventory system with shipment consolidation. *European Journal of Operational Research* 211(2), pp. 298–309.

Iida, T (2001). The infinite horizon non-stationary stochastic multi-echelon inventory problem and near-myopic policies. *European Journal of Operational Research* 134(3), pp. 525–539.

— (2002). A non-stationary periodic review production-inventory model with uncertain production capacity and uncertain demand. *European Journal of Operational Research* 140(3), pp. 670–683.

Iida, T and PH Zipkin (2006). Approximate solutions of a dynamic forecast-inventory model. *Manufacturing & Service Operations Management* 8(4), pp. 407–425.

Inderfurth, K (1997). Simple optimal replenishment and disposal policies for a product recovery system with leadtimes. *OR Spektrum* 19(2), pp. 111–122.

Inderfurth, K and R Kleber (2013). An advanced heuristic for multiple-option spare parts procurement after end-of-production. *Production and Operations Management* 22(1), pp. 54–70.

Karlin, S (1960). Dynamic inventory policy with varying stochastic demands. *Management Science* 6(3), pp. 231–258.

Kuhn, HW and AW Tucker (1951). Nonlinear programming. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Ed. by J Neyman. University of California Press, Berkeley, pp. 481–492.

Levi, R, R Roundy, VA Truong, and X Wang (2017). Provably near-optimal balancing policies for multi-echelon stochastic inventory control models. *Mathematics of Operations Research* 42(1), pp. 256–276.

McKinsey (2013). The road to 2020 and beyond: What's driving the global automotive industry? *McKinsey & Company*.

Meena, PL and SP Sarmah (2013). Multiple sourcing under supplier failure risk and quantity discount: A genetic algorithm approach. *Transportation Research Part E: Logistics and Transportation Review* 50, pp. 84–97.

Meena, PL, SP Sarmah, and A Sarkar (2011). Sourcing decisions under risks of catastrophic event disruptions. *Transportation Research Part E: Logistics and Transportation Review* 47(6), pp. 1058–1074.

Minner, S (2003). Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics* 81, pp. 265–279.

Morton, TE and DW Pentico (1995). The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Science* 41(2), pp. 334–343.

Muckstadt, JA (1973). A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Science* 20(4-part-i), pp. 472–481.

Muharremoglu, A and JN Tsitsiklis (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research* 56(5), pp. 1089–1103.

Munson, CL and J Jackson (2015). Quantity Discounts: An Overview and Practical Guide for Buyers and Sellers. *Foundations and Trends (R) in Technology, Information and Operations Management* 8(1-2), pp. 1–130.

Munson, CL and MJ Rosenblatt (1998). Theories and realities of quantity discounts: An exploratory study. *Production and Operations Management* 7(4), pp. 352–369.

Oliver Wyman (2015). *System Profit 2035: Automakers Need to Reinvent Sales System.* URL: `http://www.oliverwyman.com/media-center/2015/oliver-wyman-study-on--automotive-sales-of-the-future--system-pr.html` (visited on 03/03/2017).

Pinçe, Ç, JBG Frenk, and R Dekker (2015). The role of contract expirations in service parts management. *Production and Operations Management* 24(10), pp. 1580–1597.

Pourakbar, M, JBG Frenk, and R Dekker (2012). End-of-life inventory decisions for consumer electronics service parts. *Production and Operations Management* 21(5), pp. 889–906.

Pourakbar, M, E Laan, and R Dekker (2014). End-of-life inventory problem with phaseout returns. *Production and Operations Management* 23(9), pp. 1561–1576.

Ruiz-Torres, AJ and F Mahmoodi (2006). A supplier allocation model considering delivery failure, maintenance and supplier cycle costs. *International Journal of Production Economics* 103(2), pp. 755–766.

Sadrian, AA and YS Yoon (1994). A procurement decision support system in business volume discount environments. *Operations Research* 42(1), pp. 14–23.

Sawik, T (2014). Joint supplier selection and scheduling of customer orders under disruption risks: Single vs. dual sourcing. *Omega* 43, pp. 83–95.

Sethi, SP and F Cheng (1997). Optimality of (s, S) policies in inventory models with Markovian demand. *Operations Research* 45(6), pp. 931–939.

Shang, KH (2012). Single-stage approximations for optimal policies in serial inventory systems with nonstationary demand. *Manufacturing & Service Operations Management* 14(3), pp. 414–422.

Sherbrooke, CC (1968). METRIC: A multi-echelon technique for recoverable item control. *Operations Research* 16(1), pp. 122–141.

Silbermayr, L and S Minner (2016). Dual sourcing under disruption risk and cost improvement through learning. *European Journal of Operational Research* 250(1), pp. 226–238.

Simchi-Levi, D and Y Zhao (2012). Performance evaluation of stochastic multi-echelon inventory systems: A survey. *Advances in Operations Research.*

Simpson, VP (1978). Optimum solution structure for a repairable inventory problem. *Operations Research* 26(2), pp. 270–281.

Snyder, LV, Z Atan, P Peng, Y Rong, AJ Schmitt, and B Sinsoysal (2016). OR/MS models for supply chain disruptions: A review. *IIE Transactions* 48(2), pp. 89–109.

Song, J-S and PH Zipkin (1993). Inventory control in a fluctuating demand environment. *Operations Research* 41(2), pp. 351–370.

— (1996). Evaluation of base-stock policies in multiechelon inventory systems with state-dependent demands. Part II: State-dependent depot policies. *Naval Research Logistics* 43(3), pp. 381–396.

Stadtler, H (2007). A general quantity discount and supplier selection mixed integer programming model. *OR Spectrum* 29(4), pp. 723–744.

Tang, CS (2006). Perspectives in supply chain risk management. *International Journal of Production Economics* 103(2), pp. 451–488.

Teunter, RH and L Fortuin (1999). End-of-life service. *International Journal of Production Economics* 59(1), pp. 487–497.

Teunter, RH and WK Klein Haneveld (2002). Inventory control of service parts in the final phase. *European Journal of Operational Research* 137(3), pp. 497–511.

The Economist (2011). *Broken Links*. URL: http://www.economist.com/node/18486015 (visited on 03/29/2017).

Thomas, DJ and JE Tyworth (2006). Pooling lead-time risk by order splitting: A critical review. *Transportation Research Part E: Logistics and Transportation Review* 42(4), pp. 245–257.

Toktay, LB and LM Wein (2001). Analysis of a forecasting-production-inventory system with stationary demand. *Management Science* 47(9), pp. 1268–1281.

Xia, W and Z Wu (2007). Supplier selection with multiple criteria in volume discount environments. *Omega* 35(5), pp. 494 –504.

Zipkin, PH (1989). Critical number policies for inventory models with periodic data. *Management Science* 35(1), pp. 71–80.

— (2000). *Foundations of inventory management*. McGraw-Hill, New York.

van Houtum, G-J and WHM Zijm (1991). Computational procedures for stochastic multi-echelon production systems. *International Journal of Production Economics* 23(1–3), pp. 223–237.