Edition KWV

Marcel Sieke

Supply Chain Contract Management

A Performance Analysis of Efficient Supply Chain Contracts

Springer Gabler

Edition KWV

Die „Edition KWV“ beinhaltet hochwertige Werke aus dem Bereich der Wirtschaftswissenschaften. Alle Werke in der Reihe erschienen ursprünglich im Kölner Wissenschaftsverlag, dessen Programm Springer Gabler 2018 übernommen hat.

Weitere Bände in der Reihe http://www.springer.com/series/16033

Marcel Sieke

Supply Chain Contract Management

A Performance Analysis of Efficient Supply Chain Contracts

Springer Gabler

Marcel Sieke
The Boston Consulting Group
Köln, Germany

Bis 2018 erschien der Titel im Kölner Wissenschaftsverlag, Köln
Dissertation Universität zu Köln, 2008

Edition KWV
ISBN 978-3-658-24381-4 ISBN 978-3-658-24382-1 (eBook)
https://doi.org/10.1007/978-3-658-24382-1

Springer Gabler

This Springer Gabler imprint is published by the registered company Springer Fachmedien Wiesbaden GmbH part of Springer Nature
The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Acknowledgements

And then, that hour the star rose up,
the clearest, brightest star, that always heralds
the newborn light of day, the deep-sea-going ship
made landfall on the island ... Ithaca, at last.
(The Odyssey, Book 13, Homer)

Sometimes the PhD process felt like the Odyssey: You are determined to finish the voyage one day, but sometimes you get blown off your course and you have to find your way through. Fortunately, my Odyssey was significantly shorter and less perilous than the original one and the existence of this book should be evidence enough of the successful completion. However, what this work does not adequately show are the people I met on this voyage. For them I would like to dedicate some words.

First of all, I am very grateful to my supervisor, Prof. Ulrich Thonemann, for guiding me on my way to the PhD and offering me so many opportunities. His interest and effort was of great importance for the successful completion of this dissertation. It can be also attributed to him that I was able to leave for a one year research stay in Switzerland. There, Prof. Ralf Seifert welcomed me at the Ecole Polytechnique Fédérale de Lausanne (EPFL) and IMD. He was a great source of ideas and I had an enjoyable and productive year there. I also would like to thank Prof. Horst Tempelmeier for being the second advisor and Prof. Werner Delfmann for heading the thesis defense.

Beside my academic advisors, the support of friends and colleagues that I have met on my way should not be underestimated. First of all, my fellow PhD students from Münster and Cologne, namely, Alexander Weyers, Andreas Brinkhoff, Felix Papier, Kai Hoberg, Karin Möllering, Karl-Hendrik Magnus, Marc Albers, Michael Becker, Nadine Holtvogt, Peer Kleinau, Raik Özsen, Silke Jütte, and Ulf Merschmann were a constant source of joy and discussions. Also our student aides Jan Wienpahl, Claudia Perl, Eva Wilbuer, Carolin Hüsing, Kathrin Rieger, Katja Henne, Analena Fuchs, and Sebastian Jucken were always there for organizing literature, proof-reading, or going out.

I do not want to forget some friends that were not part of my direct working environment, but that have also played an important role in finishing this thesis. On this note, many thanks go to Florian Gerhardt for sharing his apartment and the incredible lake view with me in Switzerland. Special thanks also go to my good old friends Martin and Karoline Haase for distracting me from work with our regular cooking evenings.

Last but not least, I am very grateful to my family which was always a great support on my way to the PhD. Finally, I have to thank Margarita who was encouraging me especially during the final phase and always got me back on track with her love and some delicious Spanakopita.

Cologne, July 2008 *Marcel Sieke*

Contents

List of Tables

List of Figures

List of Abbreviations

c.c.d.f.	Complementary Cumulative Distribution Function
c.d.f.	Cumulative Distribution Function
cf.	Confer
CR	Critical Ratio
CV	Coefficient of Variation
DCF	Discounted Cash Flow
e.g.	Exempli gratia
et al.	Et alii
HP	Hewlett-Packard
i.e.	Id est
i.i.d.	Independent and identically distributed
IT	Information Technology
MTO	Make-to-Order
p.d.f.	Probability Density Function
POS	Point of Sale
ROI	Return on Investment
RPM	Resale Price Maintenance
SC	Supply Chain
s.t.	Such that

U.S.	United States
VaR	Value-at-Risk
VMI	Vendor-Managed Inventory
VP	Vice President
vs.	Versus
WACC	Weighted Average Cost of Capital
WIP	Work-in-Process

List of Symbols

α	Type I-Service, In-Stock Probability
β	Type II-Service, Fill Rate
B	Expected Backorder Level
BO	Backorder Level
b	Unit Backorder Cost; Buyback Price
C	Total Cost
CR	Critical Ratio
c	Unit Purchasing or Production Cost
c_u	Unit Underage Cost
c_o	Unit Overage Cost
D	Demand Random Variable
$d(\cdot)$	Distribution Cost Function
E	Expectation
$F(\cdot)$	Demand c.d.f.
$f(\cdot)$	Demand p.d.f.
g	Loss-of-Goodwill Cost
γ	Type III-Service
h	Unit Holding Cost

I	Expected Inventory Level
IL	Inventory Level
IP	Inventory Position
K	Fixed Cost
L_x	Lead Time to Echelon x
λ	Profit Share Factor
μ	Mean Demand
N	Number of Retailers
O	Opportunity Cost
OH	On-Hand Inventory
OO	Open Orders (Pipeline Stock)
P	Penalty Cost Function
Π	Total Profits
$\hat{\Pi}$	Reservation Profit
Q	Order Quantity; Batch Size
R	Production Size; Review Period
r	Unit Selling Price
ω	Reservation Factor
S	Order-Up-To Level; Shipment Quantity
SL	Contract Service Level
s	Reorder Point
σ	Standard Deviation
T	Number of Periods
t	Time Index
t_x	Transfer Payment to x

u	Revenue Sharing Fraction
$u(\cdot)$	Variability Cost Function
V	Coalition
v	Salvage Value
$v(\cdot)$	Value Function
w	Wholesale Price
x	Supply Chain Channel Configuration
ξ	Demand Random Variable
y	Base Stock Level
$\emptyset$	Empty Set
$\#$	Number
$\|...\|$	Vector Sum Operator
$\cdot^*$	Optimality (superscript)
$(x)^+$	Positive Part of x, i.e., $\max\{x; 0\}$

Chapter 1

Introduction

In recent years, supply chain management is increasingly receiving attention as an important performance driver for companies. Companies have started to realize that their actions also affect other partners in the supply chain and vice versa. The high focus on shareholder value has played a decisive role in taking a different view on the supply chain. For example, it has been found out empirically that supply chain malfunctions lead to an average stock price loss of 7.5% immediately after the malfunction and 18.5% after the first year (Neale, 2003). Global markets also forced companies to face competition from low cost countries. The resulting cost pressure also led to a higher pressure for low inventory levels. Therefore, efficient inventory management is critical to the success of companies. In particular if we consider that the average cost of holding a dollar of inventory is about 20-40 cents (Neale, 2003).

Yet, low inventory levels cannot be the only objective. In times of powerful customers, the availability of products and customer service is equally important because it can seriously affect the company's bottom line. For example, about 4% of annual sales of a typical U.S. retailer are lost due to abandoned purchases (Corsten and Gruen, 2004). Companies therefore need to provide excellent availability and service to retain cus-

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_1

tomers. At the same time, rising volatility in customer demand has led to more complex planning processes. For example, Cisco had to write-off $2.69 billion in raw materials (half of its quarterly sales) in 2001 because subcontractors had stockpiled semi-finished products. The main cause of the problem was that Cisco rewarded the subcontractors for quickly fulfilling its orders. Accordingly, the subcontractors chose exceptionally high inventory levels because Cisco's rewards could significantly increase their profits and customer demand always had exceeded expectations in the preceding quarters in any case. But an unexpected weakening in customer demand then finally led to the huge write-off (Narayanan and Raman, 2004).

Clearly, the problem in the Cisco example originated from misaligned incentives. Similarly, Lee and Billington (1992) found out that common pitfalls for good supply chain management is the non-existence of common supply chain metrics and an inadequate definition of customer service. Therefore, companies spend a considerable amount of money and time for the specification of contracts that should align incentives. Common metrics are needed to specify contracts that are clear and understandable to both contracting parties. But which metrics should companies use for their contracts? And for which levels of the metrics should a company aim? The definition of good performance metrics and the optimal decisions based on these metrics is the focus of this thesis.

1.1 Motivation

Our research is motivated by a joint study with *McKinsey & Co.* that analyzed the European retail industry (Thonemann et al., 2005). We found out that the most successful companies in this sector have a tight control over the supplier's delivery performance, i.e., they control the supplier service levels (or fill rate) on a regular and short-term basis. Within this group, a number of retailers use financial penalties to enforce the suppliers' performance goals that are fixed in the supply contracts.

Especially in this sector it is a common practice that powerful downstream companies, for example retailers such as Wal-Mart or Carrefour, often play a dominant role in the supply chain and specify contractual terms.

Service level contracts are also used between manufacturers and suppliers. Examples from other industries include Fab-tek, a company that provides titanium products for industrial use, that charges penalty payments for late deliveries (Shapiro et al., 1992). These penalties depend on the number of backorders as well as on the time the orders are outstanding. In the airline industry, penalties are also a common part of contracts when it comes to the on-time delivery of ordered planes. For example, Boeing's late deliveries in 1997 triggered an enormous amount of late fees, and likewise did Airbus when it came to the delivery of its A380 jets in 2008.

Service level contracts are a common way to assure supply continuity of finished products, components, and raw materials. The advantage of service level contracts is the ease of use with respect to performance measurement and enforceability. The service level is monitored closely by the manufacturer and whenever the supplier's delivery performance falls short, appropriate countermeasures can be taken by the manufacturer to ensure future compliance of the supplier. As mentioned above, an easy yet powerful approach is to enforce a penalty payment when the target service level cannot be adhered to by the supplier (Metty et al., 2005). However, in many industries the enforcement of penalties and the measurement of the supplier's performance have largely remained unstructured, i.e., penalty payments are enforced on an irregular basis or in an unforeseeable amount. This leads to enormous overhead costs for dispute resolution and to fuzzy incentives and decision-making on both sides. Taking into account that the total value of shipments in manufacturing is about $4,265 billion in the U.S. in 2007, it is not surprising that this is a major area for improvement (U.S. Census Bureau, 2008). For example, disputed amounts between suppliers and manufacturers are on average 3-5% of total sales in the U.S. high-tech industry (Billington, 2005).

Frequently, building trust and forming partnerships between suppliers and manufacturers is seen as a remedy. However, this approach in itself does not necessarily lead to supply chain coordination, as C. Gopal, VP of supply chain management at Unisys and former VP at Dell Inc., suggests:

> "Trust can only be engendered by considering the risks and having joint metrics, with penalties and incentives." (Beth et al., 2003)

Despite its prevalence in some industries, the application of these service level contracts has not wide spread to other industries or geographical regions. However, controlling and enforcing a superior supplier fill rate is a major contributor to an above average competitive performance according to a benchmark study of the Aberdeen Group (Aberdeen Group, 2004). Similarly, Lee (2004) states that competitive advantage can be achieved by a good performance measurement scheme and that good contracts that are based on these performance measures will align the supply chain partners by making their relationship explicit.

1.2 Research Objectives

In this thesis we focus on efficient contracts, i.e., contracts that coordinate the supply chain. In the first part of the thesis, we analyze service level based contracts. We want to explore the essential characteristics of efficient and enforceable contracts. More precisely, is a higher contract service level always good? How high should the penalty payment be? How responsive to the manufacturer's orders should the supplier be? To gain theoretical insights, we address the following research questions: Firstly, how do different service level contracts influence the supplier's optimal decision with respect to his inventory policy? Secondly, how do these contracts influence the supply chain partner's profits? Thirdly, how can supply chain coordination be ensured?

Next, we analyze if the supplier can show some gaming behavior when service level based contracts are used and how this behavior influences the supply chain performance. The analysis of gaming behavior in supply contracts has been identified as a major area of interest in supply chain management (Thomas, 2007). For example, a supplier that has signed a service level contract with two retailers can try to increase the deliveries to one retailer at the expense of another retailer in order to offset an inferior performance in an earlier period. In addition, we want to identify differences to other inventory allocation policies that have been proposed in literature. We also investigate whether the supplier's gaming behavior might lead to incentive conflicts between supplier and retailers.

Finally, we analyze a supply chain where multiple retailers differ with respect to their demand variability and cost position. We want to identify potential incentives for a manufacturer to exclude certain retailers or distribution channels from the supply chain. If these positive effects exist, we want to suggest an optimal admission policy and find a differential pricing scheme that depends on the demand variability the retailers bring into the supply chain. Further, we want to answer the question of how to allocate supply chain profits fairly among the supply chain partners.

Our analysis will help decision makers to design optimal contracts or to revise existing sourcing and distribution contracts. Additionally, including a perspective on incentives to current literature on service levels in supply chains also enables us to highlight differences between incentive-based and non-incentive-based inventory control policies that use service levels as primary performance measures.

1.3 Outline

The first part of the thesis (Chapters 2 and 3) introduces the basic concepts of our work. Firstly, in Chapter 2 we introduce the foundations of inventory management as far as they are relevant to our work. We present the different motivations for holding inventory, relevant inventory costs, and basic inventory control models. Secondly, Chapter 3 introduces the basic concepts of supply chain contracting. We motivate the need for contracts and introduce the solutions to some commonly used contract types.

In Chapters 4 to 6 we present our main results. The analyses of each chapter are preceded by a short introduction and followed by a conclusion that summarizes our findings. The corresponding proofs can be found at the end of each chapter. In the following, we shortly introduce each chapter in turn:

The analysis in Chapter 4 has been motivated by the observation that supply contracts are used to coordinate the activities of the supply chain partners. In many industries, service level based supply contracts are commonly used. Under such a contract, a company agrees to achieve a certain service level and to pay a financial penalty if it misses it. Although service level based contracts are among the most widely used contracts in practice, they have not been analyzed analytically and we are filling the gap in this chapter. We analyze two types of service level based supply contracts that are designed by a manufacturer and offered to a supplier. The first type of contract is a flat penalty contract under which the supplier pays a fixed penalty to the manufacturer in each period in which the contract service level is not achieved. The second type of contract is a unit penalty contract under which a penalty is due for each unit delivered fewer than specified by the parameters of the contract. We show how the supplier responds to the contracts and how the contract parameters can be chosen such that the supply chain is coordinated. We also derive structural results about optimal values of the contract parameters, provide numerical results, and connect our service level measures to traditional service level measures.

Chapter 5 builds on the results of the previous chapter and extends the service level model to a supply chain with multiple retailers. In practice, optimal inventory allocation policies have a significant impact on profits in the retail industry. A manufacturer ships products to the retailers' stores where the end customer buys the product during the selling season. It has been put forward that it is beneficial for the manufacturer to reserve a certain fraction of the inventory for a second replenishment. Then the manufacturer can replenish the retailers' inventories optimally and can take advantage of the risk pooling effect. In practice, retailers require a certain availability of the product throughout the selling season. Supply contracts are used to coordinate the delivery of products. Under such a contract, the manufacturer agrees to achieve a certain service level and to pay a financial penalty if she misses it. We analyze how a manufacturer responds to a service level contract if she wants to minimize her expected costs. We develop an optimal allocation strategy for the multiple retailer case and show that the traditional inventory balancing approach leads to higher expected costs than our approach. We also show that a service level contract does not necessarily guarantee a minimum availability at the retail stores.

After analyzing the question of *how* to optimally serve multiple retailers, we analyze in Chapter 6 the questions *if* all of the retailers should be served. Managing demand variability is a main challenge that companies face in today's volatile markets. In this chapter we analyze how supply chain performance can be optimized by explicitly taking demand variability into account and making a conscious decision about admitting it to the supply chain. By observing the impact of diversified demand patterns on the efficiency of the supply chain, we derive an optimal distribution channel strategy that trades off benefits of market size against the cost of demand variability. We derive contracts that coordinate the supply chain and lead to the supply chain optimal distribution channel selection.

Chapter 7 concludes the thesis. We summarize the results and point out the major contributions of our research. In addition, we critically review our assumptions and suggest areas for future research.

Chapter 2

Foundations of Inventory Management

In today's business world, inventories play an important role. Inventories are used to hedge against shortages in raw materials, allow a smooth production flow within a production line, or ensure the timely delivery of products to the end customer. But inventories also constitute a liability for a single company and the economy as a whole: Inventories are costly since they represent bounded capital that has to be managed continuously. In the worst case, inventories can become obsolete after a certain time. Considering that the total value of inventories in the U.S. industry is $1.45 trillion in January 2008 (U.S. Census Bureau, 2008), it becomes clear that the efficient management of inventories is of paramount importance for companies. Especially inventory on the way to the end customer, i.e., inventory held at retailers and wholesalers, accounts for about two thirds of the total inventory value (U.S. Census Bureau, 2008). It is therefore not surprising that the optimal management of supply chain inventories has received significant attention in practice and in research likewise.

In this chapter we present the foundations of inventory management.

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_2

In Section 2.1, we first describe the basic concepts of inventory management. We introduce different types of inventory, motivations for holding inventory, and common cost components that are associated with inventories. In Section 2.2 we then present some popular approaches for the management of inventories in a stochastic demand setting. Clearly, it is neither possible nor reasonable to summarize all literature and concepts related to inventory management in this thesis. Therefore, we focus on introducing concepts and models that are relevant to the purpose of this thesis. The interested reader is referred to, for example, Silver et al. (1998), Tempelmeier (2006), or Zipkin (2000) for a more extensive review of this area of research.

2.1 Basic Concepts of Inventory Management

In this section we describe the basic contingencies that influence inventory management decisions. We first introduce the four types of inventory in Subsection 2.1.1. In Subsection 2.1.2 we show some important structures of inventory systems. In Subsection 2.1.3 we present common motives why companies are holding inventory and show in Subsection 2.1.4 which costs have to be taken into account for holding this inventory. Finally, we provide in Subsection 2.1.5 a short discussion of different inventory system characteristics that are relevant to the mathematical modeling of such systems and conclude in Subsection 2.1.6 with the presentation of different approaches to measure the size of inventories.

2.1.1 Inventory Types

To allow for a better management and control of inventories, inventories can be categorized into different types for which customized management approaches exist. For this purpose, inventory is typically categorized into four different types. Nahmias (2005), for example, refers to them as raw materials, components, work-in-process, and finished goods.

1. *Raw materials* are used in the immediate transformation or production activity of the company.

2. *Components* are not ready to be shipped but represent a functional subpart of a product. Nahmias (2005) also refers to them as subassemblies.

3. *Work-in-Process* (WIP) is inventory that is in the system and is ready to be processed by the next production stage. Depending on the definition, raw materials and components are also sometimes referred to as work-in-process.

4. *Finished goods* are goods that do not require any further transformation step and can therefore be shipped the end customer.

Depending on the application, the same inventory can be categorized into different types. For example, a chemical company would consider synthetic granules as finished goods whereas a manufacturer of plastic conduits would consider the granules as raw material.

In our thesis, we mainly focus on finished goods inventory that is on the way to the end customer. However, some of our analyses are also applicable to inventories of the other categories.

2.1.2 Inventory Systems

In this subsection we describe the system where the aforementioned inventory types are held. In this context, the supply chain is the natural environment where inventory management takes places. From a conceptual view, a supply chain consists of all parties that are directly or indirectly connected in order to fulfill a customer request (Chopra and Meindl, 2004). The ultimate goal for inventory management in the supply chain is the optimal management over all stages of the supply chain. Unfortunately, this is not feasible for nearly all practical applications and also the theoretical analysis of such systems becomes too complex for even easy supply chain structures. Therefore, inventory management takes places on a higher abstraction level where only specific parts of the supply chain are considered.

For example, the decision maker can distinguish between a single-product and a multiple-product analysis, or can consider only one party in a single-echelon analysis or a subset of stages of the supply chain in a multi-echelon analysis. In our thesis, we focus on a one product setup. However, it should be noted that the analysis of multiple product systems can lead to significant cost savings as it has been shown by Thonemann et al. (2002).

The structure of the inventory system with respect to the number of parties considered plays an important role because it drives the complexity of the analysis. While single-stage inventory systems have been analyzed extensively in literature, recent research focuses more and more on multi-echelon systems. Figure 2.1 shows different structures for a multi-echelon inventory system. In a serial inventory system, a product flows sequentially through N stages to the end customer. We use this supply chain structure for our analysis in Chapter 4. In a distribution system, one product flows through a tree-like distribution structure. This structure is prevalent in retail systems where one supplier or wholesaler supplies multiple retailers. We use this structure in Chapters 5 and 6. Other structures, as for example a network structure or assembly system, also frequently appear in literature, but are not relevant for the purpose of our thesis.

2.1.3 Reasons for Holding Inventory

Companies have different motivations for holding inventory and numerous reasons have been identified in literature. The most cited reasons in literature are economies of scale, uncertainties, anticipation, transporta-

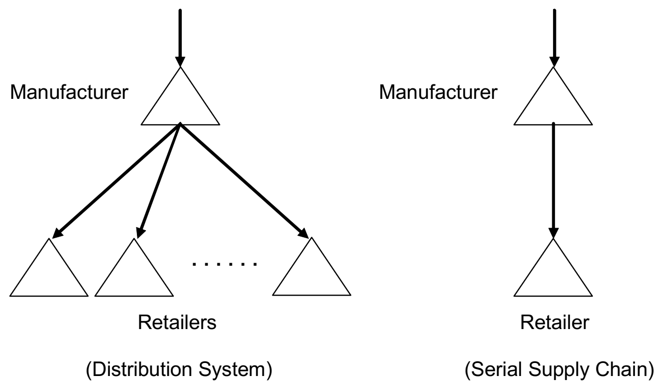

Figure 2.1: Distribution System and Serial Supply Chain

tion, speculation, and control costs (for example, in Nahmias (2005) and Silver et al. (1998)). Next, we shortly introduce each reason in turn.

1. *Economies of scale* A company has an incentive to hold inventory if it faces significant setup costs. For example, consider an inventory system where significant cleaning and readjustment tasks at the machines have to be performed before the production line can be changed from one product to the other. Then it would be beneficial for the company to produce the products in larger lot sizes and to stock inventory in the warehouses. Thereby, the costs of changing from one product to the other can be distributed over a larger number of units.

2. *Uncertainty* In most practical applications, uncertainty plays a significant role when it comes to holding inventory. For example, a company can be uncertain about the next weeks' customer demand, the availability of raw materials, the production yield of their processes, or even the appearance of major problems such as boycotts or wars. In these cases, the company uses inventory to hedge against these risks.

3. *Anticipation* If a company faces a significant change in its environment, it can use inventory to manage these changes. For example, in the toy industry the season with the highest end customer demand is the Christmas season. In this industry, 45% of the annual revenues are earned in November and December (Johnson, 2001). Since production capacities are normally not sufficient to fill all demands during that time, companies start to produce the products in advance and stock them in anticipation of the higher demand at the end of the year.

4. *Transportation* Goods that are transported between companies, between stages of the production process, or to the end customer are not available to fulfill other needs of the company. Therefore, they are considered as inventory. This inventory, which is often referred to as pipeline inventory, often plays a significant role in practical applications. For example, transporting goods with a container vessel from China to Europe takes about 4-6 weeks. During these weeks the inventory is already in possession of the company but cannot be used to fulfill customer demands and is therefore considered as (forced) inventory.

5. *Speculation* The speculation motive can be often found when we consider raw materials that are used in industry as, for example, oil, gold, or platinum. In this case, the company can buy the raw material and keep it on stock in anticipation of a price increase.

6. *Control Costs* If a company holds inventory, the inventory has to be managed. The management, for example counting inventory or determining reorder points, sometimes requires significant capital expenditures. For example, an adequate IT system has to be put in place or sophisticated inventory control models have to be developed. Although this might be justifiable for some products, a sophisticated inventory management is not beneficial for cheap goods, as for example for screws. Therefore, a company would

be better off if it would just store more screws instead of incurring high control costs by using advanced inventory management approaches.

In our research we focus on the uncertainty motive. In our models, the supply chain partners do not exactly know the customer demand and therefore seek to minimize the costs that are connected to demand uncertainty. The types of costs that play an important role in an inventory system are introduced next.

2.1.4 Types of Inventory Costs

In the previous subsection we have discussed the different reasons for holding inventory. Now we give full particulars to the measurable effects of holding inventory. In this section we focus on the most common cost components that exist in inventory systems. The different types of costs that normally are analyzed in inventory management models are holding costs, penalty costs, and order costs (Nahmias, 2005). These costs can be either stable over time or can vary over time depending on the specific application.

Holding costs are costs that are directly related to the holding of inventory. Inventory that is held in a warehouse leads to a variety of costs. Some of the costs are costs for renting the warehouse space, taxes, insurance, breakage, deterioration, obsolescence, and opportunity costs (Nahmias, 2005). Out of all of the aforementioned cost components, opportunity costs often play a major role in practical applications. Opportunity costs are determined by multiplying the inventory value with the company's interest rate. Normally, the weighted average cost of capital (WACC) is used to determine the relevant interest rate. In some industries, as for example the IT industry, obsolescence costs also represent a major part of the holding costs. Obsolescence and rework to meet engineering needs fall into this category. These costs can frequently reach up to 40% of the selling price (Lee and Billington, 1992).

For example, in the personal computer industry the devaluation of component inventory equals 1-4% of the purchasing price each week (Lee and Billington, 1992). If products cannot be sold through the original distribution channels or only at a significantly lower price, the holding costs are the difference between the original purchasing price and the salvage value, i.e., the discounted selling price. In practice, the salvage value is often difficult to measure because it can depend on the demand during the selling season as shown in Cachon and Kök (2007).

Penalty costs can appear in different ways. If customer orders arrive and the inventory is not sufficient to fill all of the orders, the orders are either backordered, i.e., filled in one of the next periods, or lost. If orders are backordered, the company has to spend money for order expediting (leading to higher administrative costs) and also faces a loss-of-goodwill of the customer (Nahmias, 2005). For the case that the demand is lost in a stock-out event, the company loses the profit margin and also incurs an additional loss-of-goodwill cost. For example, Procter&Gamble estimates that it loses the sale with a probability of 29% if the retailer has a stock-out, i.e., the customer substitutes another brand (Neale, 2003). However, the sale might be lost for the manufacturer but not for the retailer. In another study of the American grocery manufacturers, it is estimated in Kraiselburd et al. (2004) that about 60% of customers buy a substitute at the same store. An older study of Emmelhainz et al. (1991) provide even more detailed data: 32% of customers switched brands in case of a stock-out, 41% buy a different size or variety, and 14% go to another store. Clearly, estimating the penalty costs is rather difficult in practice. Therefore it is often substituted by minimum availability requirements as for example by using service levels that are introduced later in this chapter.

Order costs are incurred whenever the company orders or produces new products. They can be differentiated into fixed and variable order costs (Nahmias, 2005). Variable order costs depend on the order size. For example, unit purchasing costs or volume dependent transportation costs fall into this category. On the other side, fixed order costs do not depend on the order size. For example, searching a supplier, getting quotes, placing an order, the controlling and accounting of the incoming order all lead to fixed order costs.

2.1.5 Inventory System Characteristics

In this subsection we describe some characteristics that distinguish inventory systems from each other. The characteristics we discuss are demand, lead time, excess demand management, review pattern, and yield (Nahmias, 2005).

Demand can be either constant or variable. Under constant demand, the demand the company is facing stays on the same level. Variable demand on the other side can represent changes in the demand environment. For example, it can account for trends and seasonality. In addition, we can also distinguish different demand characteristics with respect to uncertainty. Demand can be either known with certainty, i.e., it is deterministic, or demand is random. Then in most cases, demand can be described by a discrete or continuous random variable. In this thesis we focus on random and constant demand. If the demand distribution is not know, for example when lost sales cannot be observed, inventory models tend to become rather complex as for example the model of Tan and Karabati (2004) shows. Demand can also depend on stocking quantities as in Balakrishnan et al. (2004b). Stocking more products at a retailer can increase customer demand because of a higher visibility that might lead to the customers' perception of a higher popularity or better future availability.

Inventory systems can differ with respect to *lead times*. If products are ordered from an outside supplier or are produced in a production line, they are normally not immediately available. If this time delay, also known as lead time, is taken into account, the analysis of the inventory system becomes more accurate. Lead times can be either deterministic or stochastic. Specialized inventory control models exist for both cases

(Zipkin, 2000).

Furthermore, inventory systems differ with respect to how *excess demand* is handled. Demand can be either lost or (partially) backordered. For example, in the lost sales case, if the customer does not find the product on the shelf, he just walks out of the shop or purchases a substitute product as we have shown in the previous subsection. If we assume a backorder setup, excess demand can be delayed to the next period. Then customer demand is not lost, but still a certain penalty cost is incurred as we have seen in the previous subsection.

Review patterns distinguish inventory systems. In literature, there are two types that are typically referred to, continuous review and periodic review (Silver et al., 1998). In a continuous review system, the inventory level, i.e., the number of units on stock, is always known at every time. This can be ensured by an immediate accounting whenever inventory is taken out of the warehouse, for example by using POS-data in a retail environment. In a periodic review system, the inventory levels are only checked after a certain time period, for example, every week or month.

Not all production processes produce perfect products or the deliveries of the supplier can contain broken components. Then the *yield* usually refers to the fraction of usable products with respect to the total order or production size (Silver et al., 1998). The problem of obsolescence also falls into this category. Obsolescence occurs if inventory cannot be held for more than a certain time. For example, most of the grocery products fall into this category, but also spare parts for machines if these machines are not used any more in the field.

2.1.6 Measuring Inventories

In most applications, the quantities on stock drive inventory costs. Therefore, the correct measurement is important for an efficient inventory management. The most intuitive way of measuring inventory is counting the inventory on stock. This quantity is named inventory on-hand OH. Some inventory can be already committed to a specific purpose, for example, to fill some outstanding backorders. The backorder quantity is referred to as BO. Then the net stock or inventory level IL can be computed as

$$IL = OH - BO.$$

Normally, a positive on-hand inventory OH and a positive backorder level BO cannot occur at the same time and consequently the inventory level is positive for a positive on-hand inventory, and negative for a positive backorder level.

The inventory level IL is not a sufficient measure for placing optimal orders since it neglects orders that did not arrive by now, i.e., orders in the pipeline. The inventory position IP takes these open orders OO into account and can be determined by

$$IP = OH - BO + OO = IL + OO.$$

For the zero lead time case, the inventory position always coincides with the inventory level because orders arrive instantaneously.

In a multi-echelon inventory system, we can distinguish between two inventory measurement approaches: echelon stock and installation stock. In an installation stock measurement, we only take a single installation's stock and pipeline inventory into account as Figure 2.2 shows. This is a simple extension of the one-echelon model to multiple stages. When we measure the echelon stock, we do not only take into account the inventory at the warehouse and in the pipeline to this specific warehouse, but also all inventories that are downstream of that specific warehouse. Thereby, the echelon stock measure takes all *relevant* stock of an echelon into account. This concept was first introduced by Clark and Scarf (1960). In practice the echelon stock measurement is problematic because all downstream parties have to communicate their inventory positions to the upper echelons. Nonetheless, the echelon-stock concept

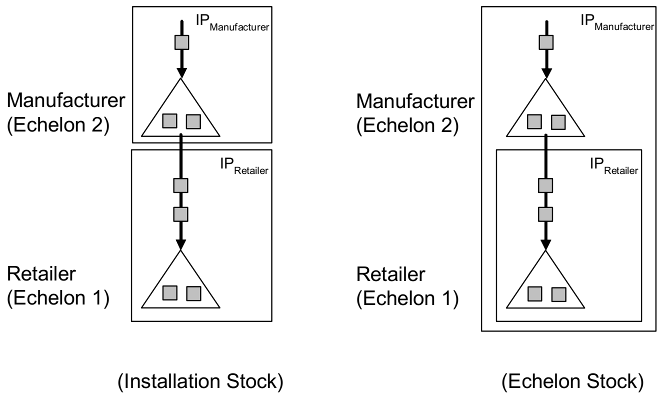

Figure 2.2: Installation and Echelon Stock

has the advantage that it can be used to minimize system-wide costs for a variety of inventory systems (Zipkin, 2000).

2.2 Stochastic Inventory Theory

If the end customer demand is not known with certainty, complex inventory models are used that trade off the effects of stocking too much and too few inventory. In this thesis, we analyze inventory systems that fall into this category, i.e., systems with stationary stochastic demand, positive lead times, and periodic review. To keep our presentation concise, we focus in this section on models that exhibit these characteristics. For models that differ from ours, we refer the reader to Nahmias (2005), Silver et al. (1998), or Zipkin (2000).

First, we discuss some commonly used objective functions for stochastic inventory models in Subsection 2.2.1. Then we introduce a basic inventory model for a single-echelon inventory system in Subsection 2.2.2 and extend the model in Subsection 2.2.3 to a multi-echelon supply chain.

2.2.1 Objective Functions

Management can have different goals with respect to inventories. For example, decision makers want to maximize the availability of a spare part, to minimize the expected costs of their finished product inventories, or to maximize the return-on-investment (ROI). In this section we introduce some popular objective functions used for inventory management. We start with the most prominent one, the expected cost criterion where the decision maker tries to minimize expected inventory costs. Then, we introduce the service level metric where the decision maker tries to achieve a certain availability of the product. Since we extensively use this metric in this thesis, we will discuss it in more detail. Finally, we present some alternative objective functions. In particular, we discuss risk-based objective functions.

2.2.1.1 Expected Cost Criterion

Most inventory models use an expected cost objective function. The objective is to minimize the average long-run costs per period or time unit. Relevant costs in a stochastic demand environment are mainly holding costs, penalty costs, and order costs. The expected cost objective coincides with the well-researched present value or discounted-cash flow (DCF) objective if the time horizon is short and/or the interest rate can be neglected (Zipkin, 2000). From an operational point of view, it is therefore sufficient to concentrate on minimizing the average inventory costs for most practical applications (Silver et al., 1998).

One of the major cost drivers in stochastic inventory models is demand variability. Demand variability makes inventory planning more complex since the exact demand quantities are not known. Within the inventory management literature that deals with the impact of demand variability we can find approaches, such as inventory pooling, constraint ordering, order smoothing, and variability penalties, that try to reduce demand variability as well as to optimize inventory-related costs in the supply chain.

Inventory pooling, introduced in the seminal work of Eppen (1979), aims to reduce the variance of total order quantities received by an upstream supply chain member by aggregating incoming orders from downstream members. Thus, in upstream echelons a reduced variability is expected to be realized. Gerchak and Mossman (1992) found analytically that inventory pooling can reduce expected costs, but does not necessarily reduce optimal inventory levels under all circumstances.

The order forecast smoothing is the determination of the order quantity for the next period taking into consideration the demand of the current period and the past orders. Balakrishnan et al. (2004) state three approaches to control demand variability by applying order smoothing. Thus, they set a limit on retailers' order quantities to prevent variability among consequent orders. Previously, Cachon (1999) proposed a similar approach that constrains the timing of an order. In the variability penalty approach, which is also studied by Porteus (1990), retailers are charged a penalty if their orders show a deviation from a predefined order amount set by the upstream echelon.

We would like to acknowledge the considerable amount of research in multi-channel inventory distribution systems that is related to demand variability. A part of the existing literature states that new channels can reach to customer segments that have not been served by existing channels and expand sales. The other part states that companies must decide among alternative channels to ensure profitability as it is expensive to serve all customer segments in the market. In addition to the benefits and opportunities they promise, distribution channels also impose challenges to the supply chain. Carr and Lovejoy (2000) present a supply chain where a manufacturer has to choose an optimal capacity level and can choose among various demand sources. The authors show how such an optimal demand source selection can be derived by using insights from portfolio theory. Frazier (1999) states that it might be the case that small to moderate-sized customer segments cannot be economically served by a traditional channel strategy, which targets all the customers in the market. The trade-off between market coverage and risk associated with fixed and variable channel investments raises the discussion of operating fewer or more channels (Alptekinoglu and Tang, 2005). In this context, Anderson et al. (1997) claimed in a qualitative analysis that high levels of demand volatility might lead to a large number of channels in operation and on the contrary, low levels of volatility might lead to a smaller number of channels. In the same study, it is also discussed that a multi-channel distribution strategy might reduce involved risk by spreading it among a set of alternative distribution channels. In Chapter 6 we deviate from these studies by allowing a manufacturer to explicitly choose the distribution channels that participate in the supply chain and to derive a pricing scheme such that the system's demand variability can be actively influenced.

2.2.1.2 Service Level-Based Models

For most companies, the availability of products is of paramount importance. In inventory control models, the penalty costs mainly influence the availability of the product. However, the quantification of those penalty costs is not straightforward because it is not easy to estimate the exact loss-of-goodwill costs. Therefore, service levels are used to circumvent the problem of quantifying the penalty costs. Typically, companies try to achieve service levels around 95% whereas even levels below 80% are not uncommon (Neale, 2003). The service level requirements differ among products. Fisher (1997), for example, states that the service level is 98-99% for functional products with predictable demand and can reach 60-90% for innovative products with unpredictable demand. In general, service levels describe the fraction between fulfilled orders and total orders and thereby can also be used as a substitute for measuring customer satisfaction. The definition of fulfilled orders can vary from case to case. For example, in some applications, only orders that can be immediately filled from stock are counted as fulfilled orders whereas in other applications a certain time window exists for successfully fulfilling orders (Silver et al., 1998). In our research we focus on the first case that measures the immediate availability.

In literature, one can mainly find three definitions of service levels, the $\alpha-$, the $\beta-$, and the $\gamma-$service level (Silver et al., 1998). The $\alpha-$service level represents the long-run fraction of periods where no stock-out occurs, i.e.,

$$\alpha = 1 - \frac{\text{\# of periods with stock-out}}{\text{total \# of periods}}.$$

Sometimes the $\alpha-$service level is also referred to as Type I-service level or in-stock probability. The $\alpha-$service level is an appropriate measure, for example, for a grocery store that is focussing on full shelves in the stores.

In contrast to the $\alpha-$service level, the $\beta-$service level takes the number of units short into account. This service level measure is calculated by the fraction of immediately satisfied demands over the total demand, i.e.,

$$\beta = 1 - \frac{\text{\# of units short}}{\text{total demand}}.$$

The $\beta-$service level is popular among practitioners because it indicates the size of the stock-out problem (Nahmias, 2005). It is often also referred to as Type II-service or fill rate.

The $\gamma-$service level is also known as ready rate (Silver et al., 1998). It is computed by taking the time that the net stock is positive. This measure has received little attention in literature because the optimization is more complex than in the other cases. However, the $\gamma-$service level is often used for measuring the service in emergency equipment models (Silver et al., 1998).

2.2.1.3 Alternative Objective Functions

A variety of alternative objective functions exist in literature. For example, risk-aware objective functions are used to optimize inventory policies. Risk-based metrics originate from the area of finance, where the decision maker performs a risk-variance trade-off when choosing an optimal portfolio. Similarly, Chen and Federgruen (2000) analyze different inventory control models with respect to their variance of costs. Using the variance as a risk measure, the variance of costs is computed and plotted against the expected costs for each purchasing quantity. This yields the efficient frontier of possible purchasing quantities. For a given utility function that represents the decision maker's preferences, the utility function is fit to the efficient frontier such that the maximal expected utility is reached. The point where both curves are tangent designates the optimal quantity. The Value-at-Risk measure (VaR) is another popular risk-based approach that measures the risk exposure of an inventory policy. Although the VaR approach originates from the area of corporate finance, it has been successfully applied to inventory problems, for example by Tapiero (2005).

A multitude of other objectives exist, for example, the optimization of the return-on-investment. For more details on these objective functions we refer the reader to the reviews of Beamon (1999) and Gunasekaran et al. (2001) on performance measures in supply chain management.

2.2.2 Single-Echelon Inventory Control Policies

Inventory control policies allow an efficient management of inventories by structuring the ordering decision into two basic decisions: (i) when to order and (ii) how much to order. For deterministic demand, the timing and quantity is always the same and easily predictable. With stochastic demand however, timing and sizes can and do vary in time.

For the ordering decision, only the inventory position IP is relevant because it includes orders in the pipeline, i.e., units that are already ordered but did not arrive yet. In the following we use the notation that is proposed by Silver et al. (1998): s reorder point, Q order quantity, R review period, and S order-up-to level. Based on this notation, we can distinguish four different basic inventory policies.

For *continuous review* inventory systems we have a complete knowledge about the inventory position at all times and we can order accordingly:

- *Fixed order quantity* A control policy that always orders a fixed amount is often referred to as a (s, Q) policy. Whenever the inventory position IP is less or equal s, a fixed amount Q is ordered. s is the reorder point and Q is the batch size. This policy takes holding, backorder, and setup costs into account. Clearly, if the setup costs are small, the batch size can also be small.

- *Variable order quantity* If the order size does not have to be fixed, a (s, S) policy leads to optimal results for the case with holding, backorder, and setup costs. Whenever the inventory position IP reaches the reorder point s, an order is placed such that the inventory position reaches S. If the order size per demand event equals one, as for example in a spare parts setup with Poisson demand, the order size is fixed and equals $S - s$. The optimization of a (s, S) policy is not trivial and requires complex calculations. If the setup costs are comparably small or the availability of the product is of high importance, a so called $(S-1, S)$ policy is used. In this case, a new unit is ordered after every demand event. A $(S - 1, S)$ policy is also called a base stock policy.

For *periodic review* models as in our thesis, we can distinguish the policies with respect to the fixed ordering costs:

- *Without fixed ordering costs* A (R, S) policy can used when fixed ordering costs can be neglected. In this policy, R stands for the review period. The review period is the length of time between two reviews of the inventory position. Therefore, only at these times, the exact quantity is known and orders can be placed. Similar to the base stock policy in the continuous review case, the order will bring the inventory position up to the order-up-to level S. This policy performs well in situations where the order setup costs are small.

- *With fixed ordering costs* We can use a (R, s, S) policy if fixed ordering cost exist. Here we order every R periods and only order if the inventory position IP is less or equal the reorder level s. Then we order up to S units. This policy has been proven to be optimal by Scarf (1959).

In the following we review the basic (R, S) policies for the single-echelon case. First, we present the optimization for the expected cost criterion. Second, we show how service level constraints influence the optimal ordering decision.

2.2.2.1 Cost Objective

The newsvendor model (also known as the newsboy model) is the building block of a (R, S) policy. The newsvendor model is frequently used for the management of products that can be only sold during a fixed selling season as, for example, fashion goods. The decision maker has to decide on the order quantity before the exact demand for the selling season is known. No second replenishment opportunity is permitted.

The goal is to determine the optimal order quantity y^*. The demand D is stochastic and the demand distribution is known and has the p.d.f. $f(\cdot)$ and the c.d.f. $F(\cdot)$. The product is sold to the end customer for r per unit. The unit purchasing costs is c and inventory that is left at the end of the selling season can be sold with a salvage price of $v < c$ per unit. It follows that inventory that is left at the end of the selling season incurs an overage cost of $c_o = c - v$. If the demand was too high, an underage cost $c_u = (r - c) + g$ has to be paid. This cost represents the loss of the profit margin and, if applicable, a loss-of-goodwill cost g. Then the expected cost function for an order quantity y can be written as

$$\begin{aligned} EC^N(y) &= c_o E(y-D)^+ + c_u E(D-y)^+ \\ &= c_o \int_{\xi=0}^{y} (y-\xi) f(\xi)\, d\xi + c_u \int_{\xi=y}^{\infty} (\xi - y) f(\xi)\, d\xi. \end{aligned}$$

It can be easily shown that the expected cost function $EC^N(y)$ is convex in y. The minimization of $EC(y)$ by using the first-order criterion shows that the optimal order size satisfies

$$F(y^*) = \frac{c_u}{c_u + c_o}.$$

The fraction on the right hand side is often called critical ratio (CR). For normally distributed demand it can be shown that the optimal order quantity is

$$y^* = \mu + z\sigma$$

where μ is the mean demand and σ the standard deviation of the demand. The safety factor z can be found with

$$z = F_{01}^{-1}(CR)$$

where $F_{01}^{-1}(CR)$ is the CR-quantile of the standard normal distribution with $\mu = 0$, $\sigma = 1$.

The basic newsvendor model can be easily translated into a periodic review (R, S) policy. For a zero lead time, the inventory position equals the inventory level and the extension to a multiple period inventory model is straightforward. In this case, the demand is assumed to be i.i.d. between periods. Unfilled demand from the previous periods can be filled in subsequent periods when new inventory is available. These backorders cost b per unit backordered. Excess inventory at the end of a period can be carried forward to the next period and incurs a unit inventory holding cost of h. The inventory level y in the newsvendor model becomes the order-up-to level y, i.e., at the beginning of each period, we order such that the inventory position is always y. Then the order quantity Q_t at the beginning of period t is

$$y - OH_{t-1} + BO_{t-1} = y - IL_{t-1} = Q_t.$$

At the end of period t the inventory on hand is $OH_t = (y - D_t)^+$, the number of backorders $BO_t = (D_t - y)^+$, and the inventory level IL_t. Since the demand D_t is i.i.d., we can use the expected cost criterion, i.e.,

$$\begin{aligned} EC^M(y) &= hE(y-D)^+ + bE(D-y)^+ \\ &= h\int_{\xi=0}^{y} (y-\xi) f(\xi)\, d\xi + b\int_{\xi=y}^{\infty} (\xi - y) f(\xi)\, d\xi. \end{aligned}$$

Clearly, the expected cost $EC^M(y)$ resembles the newsvendor model and therefore the optimal order-up-to level is characterized by

$$F(y^*) = \frac{b}{b+h}.$$

Since either there can be a positive on-hand inventory or a positive backorder level, we always order the previous periods demand, i.e., $Q_t = D_{t-1}$.

The model can be extended to positive lead times LT. Our objective is to find the order-up-to level y that minimizes expected costs. But now we have to take the number of outstanding orders OO_t at the end of period t into account. Then the order quantity Q_t in period t is

$$y - OO_{t-1} - OH_{t-1} + BO_{t-1} = y - IP_{t-1} = Q_t.$$

Now it can be shown that the inventory on-hand and backorder level in period t only depends on the demand during the previous $LT+1$ periods

(Thonemann, 2005). Therefore the expected costs can be written as

$$\begin{aligned} EC(y) &= hE(y - D_{LT+1})^{+} + bE(D_{LT+1} - y)^{+} \\ &= h\int_{\xi=0}^{y} (y-\xi) f_{LT+1}(\xi)\, d\xi + b\int_{\xi=y}^{\infty} (\xi - y) f_{LT+1}(\xi)\, d\xi \end{aligned}$$

where D_{LT+1} denotes the demand during the last $LT+1$ periods and f_{LT+1} is the corresponding p.d.f. Again the optimal order-up-to level is characterized by

$$F_{LT+1}(y^{*}) = \frac{b}{b+h}$$

where F_{LT+1} is the demand c.d.f. during $LT+1$ periods.

The solution of the basic (R, S) model stays tractable even if the cost functions are defined in a slightly different way. Chen and Zheng (1993) analyze a backorder cost function that consists of a fixed and a proportional component. Similarly, Rosling (2002) analyzes the performance of inventory systems that are characterized by non-linear shortage costs. In both papers, the objective functions are quasi-convex which allows for an efficient optimization of the inventory system by using the first-order criterion.

2.2.2.2 Service Level Objective

It has been shown before that it is difficult to quantify the backorder cost b in many practical applications. Instead, service level constraints are used to avoid the quantification of the backorder cost. We discuss now how the $\alpha-$ and $\beta-$service levels can be used in the (R, S) periodic review model we have described before.

The α-service level measures the fraction of periods that have no stock-outs. This occurs with a probability of $F_{LT+1}(y)$. The optimization problem then is

$$\begin{aligned} &\min_{y} \; h(y - D_{LT+1})^{+} \\ &\text{s.t.} \;\; F_{LT+1}(y) \geq \alpha. \end{aligned}$$

The $\beta-$service level (or fill rate) measures the fraction of demands that can be filled immediately from stock, i.e.,

$$\beta(y) = 1 - \frac{E\left(\left(D - (y - D_{LT})^{+}\right)^{+}\right)}{\mu}$$

This calculation avoids a double counting of stock-outs if some orders are backordered for more than one period. Johnson et al. (1995) discuss the double counting problem and provide easy expressions for the fill rate formula. However, in practice the probability that orders are back-ordered for more than one period is normally comparably small and can be therefore neglected. Then, the approximate measure

$$\beta(y) = 1 - \frac{E\left(D_{LT+1} - y\right)^{+}}{\mu}$$

can be used. The optimization problem is then

$$\begin{aligned} \min_{y} \quad & h\left(y - D_{LT+1}\right)^{+} \\ \text{s.t.} \quad & 1 - \frac{E\left(D_{LT+1} - y\right)^{+}}{\mu} \geq \beta. \end{aligned}$$

The advantage of the $\alpha-$ and $\beta-$service levels is that they can be easily understood. Also the computation of the optimal order-up-to level y is straightforward and can be determined by taking the smallest value that satisfies the service level constraint. However, these service level measures have a major shortcoming that hinders their direct use in practice: their calculation is based on an infinite-horizon measurement, i.e., we have to measure over an infinite time horizon to come up with a reliable estimation. In practice however, service levels are measured over shorter (finite) time horizons, for example, weeks or months.

Some recent work has analyzed the distribution of service levels if the performance is measured over a finite number of periods. Chen et al. (2003) and Thomas (2005) were one of the first to analyze finite horizon service level models. Chen et al. (2003) study the effect of a finite time horizon on the service level. This approach is of particular interest in practical applications because infinite horizon service levels are difficult to control and companies measure performance over finite time horizons. Chen et al. show that using an infinite horizon to derive optimal base stock levels in a periodic review system leads to a higher than desired short-term service level. Subsequently, Thomas (2005) analyzes the effect of the length of the time horizon on the distribution of service levels. In this thesis, we extend Thomas' work by integrating finite horizon service level measures in the objective function and charging financial penalties if a contract service level is not achieved.

The $\beta-$service level over a finite horizon with T measurement periods and zero lead time can be written as

$$\beta_T(y) = 1 - \frac{(D_1 - y)^+ + ... + (D_T - y)^+}{D_1 + ... + D_T}$$

whereas the standard infinite horizon formulation is

$$\lim_{T\to\infty} \beta_T(y) = 1 - \frac{E(D - y)^+}{\mu}.$$

Clearly, measuring over a finite horizon leads to the observation that the service level becomes a random variable that cannot be known with certainty for a given base stock level y. The time horizon T has a distinct effect on the height of the service level as the following proposition shows:

Proposition 2.1 *(from Thomas (2005)) For all measurement horizon lengths $T > 1$ and base stock levels $y > 0$, we have $E(\beta_1(y)) \geq E(\beta_T(y)) \geq \lim_{T\to\infty} E(\beta_T(y))$.*

Clearly, using an infinite measurement leads to an underestimation of the service level. In our research, we will make extensive use of the finite horizon definition since it facilitates the control of contractual behavior.

2.2.3 Multi-Echelon Inventory Control Policies

There exists a variety of models that consider multi-echelon inventory systems. As before, we will focus on periodic review systems with base stock levels. For a serial supply chain, Clark and Scarf (1960) developed the first model that optimizes a multi-echelon supply chain for a finite horizon. Federgruen and Zipkin (1984b) extended the model to an infinite horizon. In Cachon and Zipkin (1999), we can find a comparably simple presentation of Clark and Scarf's model for a two-echelon serial supply chain that we introduce next. Afterwards we shortly discuss the application of service levels in multi-echelon supply chains and conclude with an overview of inventory allocation policies that are used if multiple retailers are served.

2.2.3.1 Serial Supply Chain Model

To keep our presentation concise we restrict our analysis to a two-echelon serial supply chain where the first echelon faces the end customer demand and the second echelon also holds inventory and replenishes the first echelon. The lead time to the second echelon is L_2 and the lead time between the second echelon and the first echelon is L_1. Unit holding costs are h_1 and $h_2 \leq h_1$ and are charged with respect to the echelon stock position. Unit backorder costs only appear at the first echelon with a cost b. The base stock levels y_1 and y_2 are also based on the echelon stock measure. The cost function for the first echelon is then

$$\begin{aligned} EC_1(y_1) &= h_1 E\left(y_1 - D_{L_1+1}\right)^+ + (b+h_2)\, E\left(D_{L_1+1} - y_1\right)^+ \\ &= h_1 \int_{\xi=0}^{y_1} (y_1 - \xi)\, f_{L_1+1}(\xi)\, d\xi \\ &\quad + (b+h_2) \int_{\xi=y_1}^{\infty} (\xi - y_1)\, f_{L_1+1}(\xi)\, d\xi. \end{aligned}$$

As in the newsvendor model, the cost minimizing base stock level y_1 is characterized by

$$F_{L_1+1}(y_1^*) = \frac{b+h_2}{b+h_1+h_2}.$$

Given the retailer's optimal base stock level y_1^*, we define the induced penalty function as

$$G(x) = EC_1(\min\{y_1^*; x\}) - EC_1(y_1^*).$$

The second echelon expected cost function can then be written as

$$EC_2(y_2) = h_2 E(y_2 - D_{L_2} - \mu) + E(G(y_2 - D_{L_2})).$$

$EC_2(y_2)$ can be numerically minimized with respect to y_2 and gives the optimal second echelon base stock level y_2.

The echelon inventory base stock levels can be translated to installation stock inventory levels under the condition that the inventory system is nested (Axsäter and Rosling, 1993), i.e., $y_1 = y_1^*$ and $y_2 = y_2^* - y_1^*$.

2.2.3.2 Allocation Policies for Distribution Systems

If a manufacturer with limited inventory has to serve multiple retailers, the question of how to allocate the inventory among the retailers becomes important. Initially, Clark and Scarf (1960) had to make an assumption on the inventory allocation when they extended their serial supply chain model that we have introduced above to a distribution system. In order to optimize their distribution system, retailer inventories had to be balanced. Subsequently, inventory balancing received some more attention in inventory allocation studies where a manufacturer has to supply multiple retailers from a central warehouse. One line of research focuses on setups where the central warehouse does not hold any stock. Eppen and Schrage (1981) analyze a model with identical costs at the retailers and develop a cost optimal policy. Federgruen and Zipkin (1984) extend their model to non-identical retailers and develop an approximation that is based on the assumption that the inventory imbalance is negligible. Zipkin (1984) analyzes the case if inventory imbalances are significant and improves the stocking policy for this case. Verrijdt and de Kok (1996) develop an optimal ordering and allocation policy if a certain service level should be guaranteed at the retailers.

In the second line of research, inventory is held back at the central warehouse for a later replenishment of the retailers. Jackson (1988) extends the model of Eppen and Schrage by allowing the central warehouse to hold back some inventory. They use a ship-up-to-S policy and show that the approach leads to high reductions in backorder periods. Jackson and Muckstadt (1989) analyze the case for two replenishment opportunities with identical retailers. Jönsson and Silver (1987) optimize service levels. Since their service measure is based on an infinite-horizon, their analysis results in a backorder minimization. Subsequently, McGavin et al. (1993) present a model similar to ours. For the infinite retailer case they show that a 50/25 policy gives good results, i.e., a second replenishment should take place in the middle of the period and 25% of the inventory should be reserved for this replenishment. van der Heijden (1999) optimizes the allocation if a service level constraint and a fixed shipment frequency exists.

2.2.3.3 Service Levels in Multi-Echelon Systems

In multi-echelon inventory systems, service levels are typically specified at the most downstream echelon and cost minimizing target inventory levels are computed for each echelon. A review of the relevant service level models can be found in Diks et al. (1996). Cohen and Lee (1988), Lee and Billington (1993), and Choi et al. (2004) analyze supply chain models where service levels *between* echelons of a supply chain are relevant. Cohen and Lee (1988) analyze a continuous review inventory system. They use a decomposition approach to model the different sites of a supply chain network and assume that these local target levels are exogenously specified. They provide an approach that optimizes order points and lot sizes using local service level targets as lower bounds on the service level for each site. Lee and Billington (1993) analyze a periodic review multi-echelon supply chain model, which was motivated by an application at Hewlett-Packard where the last echelon must achieve a given customer service level. They develop a heuristic to determine local target service levels for each site. Choi et al. (2004) consider a two-echelon supply chain where the supplier has finite capacity and manages the inventory of his components at the manufacturer's site under a VMI arrangement similar to our model in Chapter 5. The manufacturer uses the components for assembling the final product with a finite production capacity. He faces stochastic end customer demand and requires a certain customer service level. Choi et al. show that specifying target service and backorder levels for the supplier leads to optimal inventory decisions at the supplier echelon.

In this chapter, we have introduced the basic concepts of inventory management. In the next chapter, we build on these approaches to design contracts that lead to a good performance of the supply chain.

Chapter 3

Foundations of Supply Chain Contracts

The inventory models that we have presented in the previous chapter are based on the assumption that all decisions concerning the inventory are made by only one central decision maker that wants to maximize his own profits. In a single-echelon inventory system this clearly does not lead to any conflicts, but in a multi-echelon system the outcome can be a highly inefficient supply chain. Whang (1995) rightly states that a locally optimal solution can be suboptimal from a global point of view when each party pursues its own goals.

Contracts are needed to align the supply chain partners' incentives to a common goal. Normally, this common goal should be the maximization of total supply chain profits. Contracting is a tool that is needed to negotiate the exact terms of this aligned supply chain relationship. Good contracts then do not only reduce costs for solving disputes, but also lead to a more stable relationship between the supply chain partners and reduce transaction costs (Tsay et al., 1999). The following example shows the importance of having clear and easily enforceable contracts: In the U.S. information technology sector, disputed amounts between

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_3

supplier and manufacturer are on average 3-5% of total sales in recent years. This leads to enormous overhead costs for dispute resolution on both sides and to unclear incentives and decisions of the involved parties (Billington, 2005).

Research on contract design in inventory management originated in the seminal work of Spengler (1950) on double marginalization. He shows that a retailer chooses less than optimal order quantities because the retailer does not take the manufacturer's profit margin into account. Subsequently, the topic of contracting and inventory management has received significant attention, especially in recent years, as the literature review of Cachon (2003) shows.

The design of contracts is also strongly influenced by legal issues. However, we do not consider these issues in detail in our work. We will focus on the operational details of contracts as far as they are relevant for inventory management. In Section 3.1 we motivate the use of contracts in an inventory system setup. In Section 3.2 we classify contracts by contractual clauses. In Section 3.3 we introduce some basic contract models in inventory management.

3.1 Motivation for Contracts

A contract makes the exact incentives explicit. Each supply chain partner knows what is expected from him and what is important to the supply chain in total. When we talk about contracts in the area of inventory management, a simplified model with one manufacturer and one retailer is used frequently. Before we describe efficient supply chain contracts in detail, we introduce this basic manufacturer-retailer supply chain that is presented in Figure 3.1. The retailer orders y units for a total price of $w(y)$ from the manufacturer who in turn produces the y units at a cost $c(y)$. After y units have been delivered, the retailer faces a stochastic customer demand D and sells $\min(D; y)$ units to the end customer and receives the revenue $r(\min(D; y))$.

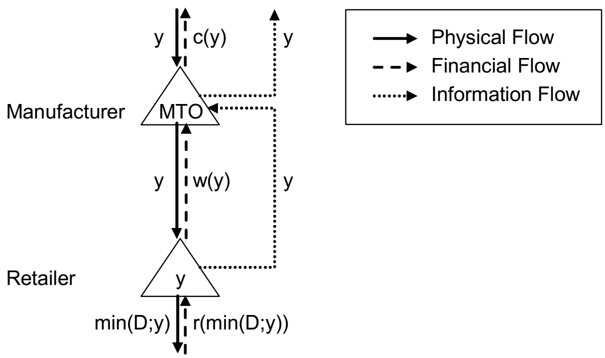

Figure 3.1: Supply Chain with Physical, Financial, and Information Flows

With a centralized decision maker, the expected supply chain profit can be optimized with respect to y and will lead to the supply chain optimal expected profit Π^c. Π^c is often referred to the *first-best* case solution (Tsay et al., 1999). In the decentralized case, the retailer only optimizes his own profits and the supply chain would only realize a profit of Π^d. For this case, Spengler (1950) has shown that for $w > c$, we have $\Pi^d < \Pi^c$, i.e., the supply chain is not coordinated. This is the so-called *double marginalization* effect.

In literature, the goal of designing good contracts for inventory management is that the contracts lead to a coordinated supply chain, i.e., a supply chain where the supply chain profits equal the profits of the centralized solution $\Pi^d = \Pi^c$. Under such contracts, the total supply chain profit can be improved with respect to the uncoordinated case and the improvements can be divided up between the manufacturer and retailer such that a win-win situation results. However, not all contracts divide up the profits fairly among the supply chain partners. For example, the retailer can take all gains from the contract and leave the manufacturer with the same profit level as before. Unfortunately, the definition of fairness is hard to quantify in this context. Still, some approaches exist that take contract fairness into account. For example, Cui et al. (2007) consider a contract fairness measure that takes profit level inequalities into account.

3.2 Classification of Contracts

In literature we can find a variety of different contracts that can be used to manage inventory optimally. These contracts can be classified into different classes. Tsay et al. (1999) provide a classification that assigns contracts to different classes based on their contract clauses. They suggest contract clauses that are concerned with decision rights, pricing, minimum purchasing commitments, quantity flexibility, buyback and return, allocation, lead time, and quality.

1. *Decision rights* If the contract specifies decision rights, a manufacturer for example can exert direct control over a retailer. For example, the manufacturer can dictate the retailer's order size or the retail price under a Resale Price Maintenance (RPM) agreement. Clearly, this contract clause requires high negotiation power of one supply chain partner.

2. *Pricing* The pricing determines how physical transactions between the two contract parties translate into financial streams. For example, the wholesale price can be written as $w(x) = K + xk'$ for an order quantity x. Then K is the fixed cost and k' the variable cost factor. This wholesale price contract is then called a two-part tariff because it includes both a fixed payment and a variable payment. In other situations, quantity discounts are possible for high order quantities or the manufacturer can subsidize the retailers if their inventory loses value. This is frequently used in price protection agreements (Lee et al., 2000) that are prevalent in the personal computer industry, as Callioni et al. (2005) show for an

example at HP. Lariviere and Porteus (2001) also study price-only contracts in a newsvendor setting. They show that a lower coefficient of variation (CV) leads to a higher wholesale price.

3. *Minimum Purchasing* The minimum purchasing clause determines the minimum quantity a company will buy from a supplier. Prominent examples can be found in the computer industry where a manufacturer has to commit to a certain quantity because the supplier has to build up capacity well in advance (Cachon, 2004).

4. *Quantity Flexibility* This contract clause allows one of the parties to deviate from a previously committed quantity. For example, a manufacturer wants to change his previously committed purchasing quantity of a component because additional knowledge of the potential customer demand has become available. How much and at which cost a supply chain partner can deviate from this quantity is determined by a quantity flexibility contract (Tsay, 1999).

5. *Buyback and Return* With this contract clause, the retailer can return unsold products to the manufacturer. For example, Pasternack (1985) shows that the supply chain can be coordinated with a buyback contract if the supplier has infinite capacity.

6. *Allocation* The allocation clause specifies how the manufacturer's inventory is distributed among the retailers if the manufacturer's inventory is limited.

7. *Lead Time* With a lead time clause, the promised lead time can be specified. For example, Dell requires an immediate availability of components from their suppliers (Kapuscinski et al., 2004) or a company can specify a time-window in which deliveries should arrive. This contract clause is often connected with a pricing clause as for example in Cachon and Zhang (2006) where a manufacturer has to pay late-fees if the product is not immediately available.

8. *Quality* The quality clause determines the characteristics of the product. Also the functionality or process quality can be specified with this clause.

The contracts of this thesis fall into the categories pricing, lead time, and allocation. In the following section we therefore introduce the basic supply chain contract models that are closely related to our research.

3.3 Contracting Models

In this section we first introduce basic one-period contracting models and show how the supply chain can be coordinated (Subsection 3.3.1). Then, we extend the analysis to multiple-period base stock models (Subsection 3.3.2).

3.3.1 One-Period Contracting Models

In the most basic case, a manufacturer delivers her product to a retailer. The manufacturer uses a make-to-order (MTO) strategy and the retailer determines his optimal inventory level with the newsvendor model. The manufacturer's unit production cost is c, the wholesale price is w per unit, the salvage value is v per unit, and the unit selling price is r per unit. Demand is stochastic and follows the known p.d.f. $f(\cdot)$. Then for a given wholesale price w, the retailer's optimal order quantity y^r satisfies

$$F(y^r) = \frac{r-w}{r-v}.$$

Clearly, we can see from Section 2.2.2 that the *first-best* solution y^* that maximizes the total expected supply chain profit is characterized by

$$F(y^*) = \frac{r-c}{r-v}$$

It follows that the supply chain with a wholesale price contract only earns an inferior expected profit compared to a centralized supply chain.

Therefore, a simple wholesale price contract only coordinates the supply chain if the retailer takes all profits, i.e., $w = c$. In practice, a manufacturer would not accept such a contract since it violates his participation constraint to realize a certain minimal profit. Clearly, the manufacturer does not have an incentive to participate in such a setup.

In literature, different solutions have been proposed to solve the double marginalization problem. We will introduce the most important ones next, the buyback and revenue sharing contract.

A contract type that is widespread in the high-tech and publishing industry is the buyback contract. In a buyback contract, a manufacturer buys back excessive inventories from the retailers or gives them a subsidy that sets off potential losses by discounted selling. With this contract, the manufacturer pays the retailer b per unit of excess inventory at the end of the selling season. The supply chain is coordinated if the order size of the decentralized system equals the order size of the centralized system, i.e., when the critical ratios are equal. Therefore,

$$\frac{r-c}{r-v} = \frac{r-w}{r-b}$$

has to hold. In other words, if the buyback price b equals

$$b = -\frac{c-v}{r-c}r + \frac{r-v}{r-c}w$$

the supply chain is coordinated (Thonemann, 2005). There exist an infinite number of potential buyback contracts that coordinate the supply chain. The wholesale price can be used to allocate profits between the manufacturer and the retailer. If $w = c$ the retailer takes all profits as in the wholesale price contract, whereas the manufacturer takes all profits if $w = r$.

Another contract type is the revenue-sharing contract (Cachon and Lariviere, 2005). This contract type has been successfully applied in the American video rental business (Cachon and Lariviere, 2001). In this contract, the retailer pays a fraction $(1 - u)$ of the revenue to the manufacturer. It follows that the retailer keeps ur per unit sold. The supply chain is coordinated if

$$\frac{r-c}{r-v} = \frac{ur-w}{ur-v}.$$

Therefore, coordinating contracts can be found by using

$$u = -\frac{v\,(r-c)}{r\,(c-v)} + \frac{r-v}{r\,(c-v)}w$$

with

$$\frac{r-c}{r-v}v \leq w \leq c.$$

Again, there exist an infinite number of coordinating contracts. As before, the wholesale price can be used to allocate profits between the supply chain partners.

Other contract types that are based on the previously presented base model include quantity flexibility, quantity discounts, or sales rebates clauses. Details about these contracts can be found in Cachon (2003).

3.3.2 Base Stock Contracts

The models of the last subsection were based on a one-period analysis. In this subsection, we introduce contracting models that deal with a multiple period setup where the upstream party only has limited capacity. In this case the upstream party does not follow a make-to-order strategy as before, but also follows a base stock policy, i.e., make-to-stock.

This setup is analyzed by Chen (1999), Lee and Whang (1999), Cachon and Zipkin (1999), and Porteus (2000). Chen (1999) analyzes a decentralized multi-echelon supply chain. He shows that the supply chain can be coordinated by an inventory accounting scheme where the *accounting inventory level* is used instead of the traditional echelon stock level. The accounting inventory level of the downstream echelon includes all orders placed by the downstream echelon with the upstream echelon, regardless of whether the upstream echelon can deliver the order or backorders it. Lee and Whang (1999) analyze a decentralized serial supply chain where the upper echelon faces inventory holding costs, but no backorder penalty costs. Based on the main ideas of Clark and Scarf (1960), they develop an optimal non-linear incentive scheme that incentivizes all echelons to choose the supply chain optimal base stock levels by penalizing certain echelons for shortages and subsidizing others for holding inventory. Porteus (2000) builds on this approach by introducing responsibility tokens. These responsibility tokens are given to the downstream echelon as a replacement for real units whenever the order cannot be fully filled. The issuer of the responsibility token bears all the financial consequences of delayed deliveries at lower echelons. Porteus shows that the supply chain can be coordinated with this approach. Cachon and Zipkin (1999) show how supply chain coordination can be achieved in a serial two-echelon supply chain with a payment scheme that depends linearly on the backorders of the two echelons and the retailer's on-hand inventory. They show that there exists a linear contract that leads to supply chain coordination. This contract is based on linear transfer payments t_I, t_B^r, and t_B^s where t_I is paid by the manufacturer to the retailer for every unit on the retailer's stock. Therefore it serves as an inventory subsidy. t_B^r has to be paid by the retailer for every backorder at the retail store. Thereby, the supplier can ensure a certain availability of his product. Finally, t_B^s is paid by the manufacturer to the supplier for every backorder at the manufacturer. This ensures that the manufacturer has sufficient stock to meet the retailer's orders. Cachon and Zipkin (1999) demonstrate that the solution is a Nash equilibrium, which implies that it is optimal for both echelons to choose the supply chain optimal solution.

Chapter 4

Coordinating Service Level Contracts for a Serial Supply Chain

4.1 Introduction

The coordination of the partners of a supply chain is governed by supply contracts. There are several types of supply contracts that are commonly used in practice. A particularly popular type of supply contract is the service level contract (Behrenbeck et al., 2007). Under a service level contract, companies specify a service level that a supplier must achieve and the consequences of missing it. Since most companies measure their own service levels and many companies also measure the service levels of their suppliers, the information required for implementing a service level contract is typically available at both partners of the supply chain. In the consumer goods industry, for instance, essentially all manufacturers measure their own service levels and 70% of the retailers measure the service levels of the manufacturers (Thonemann et

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_4

al., 2005). In this industry, manufacturers and retailers typically agree on a service level that a manufacturer is expected to achieve. However, many of these agreements are informal and the retailers frequently do not specify the sanctions for missing the contract service level in their contracts. Those who specify sanctions typically use financial penalty payments that the manufacturer must pay to the retailer if the contract service level is not met.

One of the main challenges when negotiating a service level contract is agreeing on the contract parameters, i.e., the values of the contract service level and the penalty payment. This issue has not been addressed in the literature and we are filling the gap in this chapter. We analyze a supply chain operating under a service level contract and show how the optimal parameter values of the contract can be determined. Our interest is in parameter values of the contract that coordinate the supply chain, i.e., that ensure that under decentralized decision making the supply chain optimal solution is achieved. We also derive structural insights about optimal parameter value combinations, such as the convexity of the optimal penalty cost in the contract service level. The results of our analyses can be used by decision makers to design optimal service level contracts and provide a solid foundation for contract negotiations.

The remainder of the chapter is organized as follows: In Section 4.2, we develop a mathematical model of a two-echelon supply chain that is governed by a service level contract. We analyze two types of service level contracts, a flat penalty contract and a unit penalty contract. Under a flat penalty contract, a fixed penalty payment is due in each period in which the contract service level is not achieved. Under a unit penalty contract, a penalty payment is due for each unit delivered fewer than specified by the parameters of the contract. In Section 4.3, we analyze the optimal response of the supplier to both types of contracts. In Section 4.4, we build on the optimal response functions and derive contracts that coordinate the supply chain. We show that the supply chain can be coordinated for any value of the contract service level if the value of the penalty payment and the wholesale price are chosen properly. However, values of the contract service levels that match traditional α and β service levels are particularly appealing to decision makers. In Section 4.5, we therefore analyze contracts with this property in some detail and prove that they always exist. In Section 4.6, we conclude. All proofs of this chapter are contained in Section 4.7.

4.2 Model Description

We consider a two-echelon supply chain with one supplier (indexed by s) and one manufacturer (indexed by m). Both companies operate under periodic review base stock policies with base stock levels y_s and y_m, respectively. Excess demand is backordered. The sequence of events during a period is the same at both companies: At the beginning of a period, shipments arrive and orders are placed. Then, backorders are filled. Finally, demands arrive and are filled.

The unit inventory holding costs of the supplier and manufacturer are h_s and h_m and are charged against the inventory left over at the end of a period. To restrict our model to non-trivial solutions, we require $h_s < h_m$ (Zipkin, 2000). The manufacturer receives r for every unit sold and encounters a backorder penalty cost b_m for each unit that is backordered at the end of a period. This cost is interpreted as usual and includes losses in customer good-will (Porteus, 1990). In Section 4.6, we discuss how our approach can be used in situations where the manufacturer uses a service level constraint as opposed to a backorder penalty cost. The supplier encounters penalty costs if the fraction of demand that the supplier fills is below the contract service level SL ($0 < SL \leq 1$). The lead time of the supplier is $L_s > 0$ and the lead time between the supplier and the manufacturer is $L_m > 0$.

Demand is stochastic, stationary, continuous, and independent between periods. Demand can be arbitrarily distributed as long as the p.d.f. is strongly unimodal or logconcave. This property holds for most theoretical distributions that are relevant for modeling demand, such as the Normal, truncated Normal, Gamma with shape parameter $\alpha > 1$, Beta(α, β) with parameters $\alpha \geq 1$, $\beta \geq 1$, and the Uniform distributions. For an in-depth treatment of logconcave distributions and their application to inventory control we refer the reader to Rosling (2002). To keep our analyses concise, we will focus on distributions with infinite non-negative support. Distributions with finite support can be treated analogously, but they require the definition of feasible regions for the parameter values of the supply contracts, which makes the analysis much more complex and adds little value.

We denote the demand over t periods by D_t and the corresponding p.d.f. and c.d.f. by $f_t(\cdot)$ and $F_t(\cdot)$, respectively. Note that the logconcave property is inherited to demand convolutions (Karlin and Proschan, 1960), i.e., if the demand over a single period is logconcave, then the demand over the lead time is also logconcave. As usual, the demand distribution is known to the supplier and the manufacturer (Lee and Whang, 1999).

We analyze two types of supply contracts, which we refer to as flat penalty and unit penalty contracts. Under a *flat penalty contract*, the supplier pays the manufacturer a fixed amount p for each period in which the contract service level SL is not met, i.e., for each period in which the supplier does not fill at least a fraction SL of the manufacturer's orders. Let D denote the demand of the current period and let D_{L_s} denote the demand over the previous L_s periods. Then, the inventory available for filling demand of the current period is $y_s - D_{L_s}$ and the flat penalty function can be written as

$$P_f(y_s, p, SL, D, D_{L_s}) = \begin{cases} p & \text{if } SL\ D > y_s - D_{L_s} \\ 0 & \text{if } SL\ D \leq y_s - D_{L_s}. \end{cases}$$

Note that we implicitly require that backorders are filled before the demand of the current period and that the supplier is charged a penalty in each period in which all or some backorders are not filled. This approach ensures that the supplier does not build up backorders.

Under the *unit penalty contract* we analyze, the supplier is charged a penalty of p for each unit she delivers less than $SL\ D$ if she fills at least some of the demand of the current period. If she fills no demands of the current period, she is charged a penalty of p for each unit of demand of the current period. With this contract structure, we penalize unfilled demands of periods in which no demands are filled higher than unfilled demands of periods in which at least some demands are filled. The unit penalty function can be written as

$$P_u(y_s, p, SL, D, D_{L_s}) = \begin{cases} pD & \text{if } D_{L_s} \geq y_s \\ p(y_s - D_{L_S} - SL\ D)^- & \text{if } D_{L_s} < y_s. \end{cases}$$

The first case holds if no inventory is available at the beginning of the period. The second case holds if some inventory is available at the beginning of the period. Alternatively to this penalty function, we could use a penalty function where we penalize unfilled demands of periods in which some demands are filled equally to unfilled demands of periods in which no demands are filled. The disadvantage of this penalty function is that the resulting expected profit function is neither concave nor quasi-concave which makes the optimization of the contract parameters analytically intractable. We note that in most practical situations the service levels of a supplier are reasonable high. In the consumer goods industry, for instance, essentially all companies achieve service levels of greater than 90% (Behrenbeck et al., 2007). Then, some demands can almost always be filled and the penalty cost of the contract we analyze is very close to the penalty cost of the contract with the alternative penalty function, i.e., both contracts are very similar. However, in situations with very low service levels at the supplier (< 50%), the penalties of the contracts can differ substantially.

We consider a Stackelberg game in which the manufacturer is the Stackelberg leader and the supplier is the Stackelberg follower (Tomlin, 2003; Wang et al., 2004). The objective of both parties is maximizing their individual expected profits. The manufacturer has information on the lead time L_s, the inventory holding cost h_s, and the unit cost c of the supplier. He designs a service level contract that he offers to the supplier. The service level contract specifies the wholesale price w, the type of penalty function (flat or unit penalties), the penalty cost factor p, and the contract service level SL. Given this contract, the supplier decides on the base stock level y_s that maximizes her expected profit, i.e.,

$$\begin{aligned} E\Pi_s^*(w,p,SL) &= \max_{y_s} E[(w-c)D - h_s(y_s - D_{L_s+1})^+ \\ &\quad -P(y_s,p,SL,D,D_{L_s})], \end{aligned}$$

where $P(y_s,p,SL,D,D_{L_s}) = P_f(y_s,p,SL,D,D_{L_s})$ under a flat penalty contract and where $P(y_s,p,SL,D,D_{L_s}) = P_u(y_s,p,SL,D,D_{L_s})$ under a unit penalty contract. The decision variable of the supplier is the supplier's base stock level y_s.

The manufacturer's objective is to maximize his expected profit. The decision variables are the contract type and parameters, and the manufacturer's base stock level y_m, i.e.,

$$\begin{aligned} \Pi_m^* &= \max_{y_m,w,p,SL} E\left[(r-w)D + P(y_s,p,SL,D,D_{L_s})\right] \\ &\qquad - h_m I_m(y_m,y_s) - b_m B_m(y_m,y_s) \\ &\quad s.t. \quad E\Pi_s^*(w,p,SL) \geq \hat{\Pi}_s. \end{aligned}$$

The first two terms of the manufacturer's objective function model the two streams of income the manufacturer generates. The first income stream is the contribution generated by selling products to end customers. The second income stream is the penalty payment the manufacturer receives from the supplier.

The third term of the objective function is the expected inventory holding cost, where $I_m(y_m,y_s)$ denotes the average inventory level at the end of a period. It can be computed as

$$I_m(y_m, y_s) = F_{L_s+1}(y_s)i_m(y_m) + \int_{x=y_s}^{\infty} i_m(y_m + y_s - x)f_{L_s+1}(x)dx,$$

where

$$i_m(z) = \int_{\delta=0}^{z} (z - \delta)f_{L_m+1}(\delta)d\delta$$

is the manufacturer's average on-hand inventory and z denotes the manufacturer's on-hand inventory at the beginning of the period. Cachon (2003) uses a similar approach.

The fourth term of the objective function is the expected backorder penalty cost, where $B_m(y_m, y_s)$ denotes the average backorder level at the end of the period. It can be computed as

$$B_m(y_m, y_s) = F_{L_s+1}(y_s)b_m(y_m) + \int_{x=y_s}^{\infty} b_m(y_m + y_s - x)f_{L_s+1}(x)dx,$$

where

$$b_m(z) = \int_{\delta=z}^{\infty} (\delta - z)f_{L_m+1}(\delta)d\delta.$$

is the manufacturer's average backorders and z denotes the manufacturer's on-hand inventory at the beginning of the period.

The constraint $E\Pi_s^*(w, p, SL) \geq \hat{\Pi}_s$ ensures that the supplier achieves an expected profit that is greater than or equal to her reservation profit $\hat{\Pi}_s$. If the expected profit of the supplier is below the reservation profit, she will not accept the contract. Therefore, the manufacturer designs only contracts that guarantee the supplier a minimal expected profit $\hat{\Pi}_s$. Corbett et al. (2004) discuss the use of reservation profits in detail.

4.3 Supplier Response

In this section, we analyze how the expected profit functions of the supplier are affected by the type and parameters of the supply contract and show that the expected profit functions are quasi-concave in the

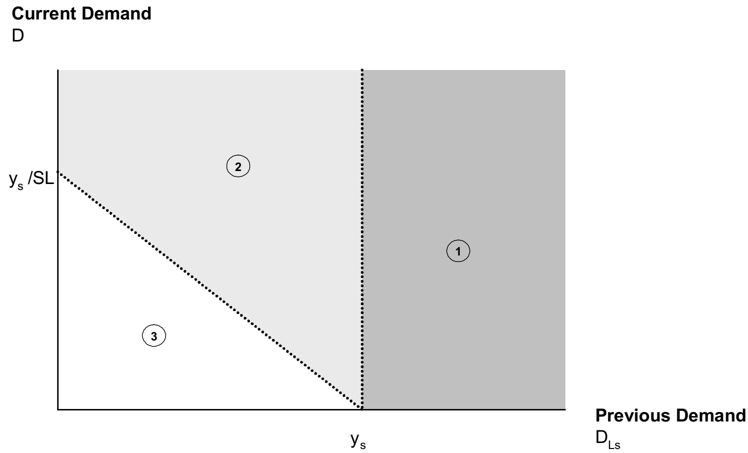

Figure 4.1: Demand Space Partitioning

supplier's base stock level. We also derive the optimality conditions for both contract types.

4.3.1 Flat Penalty Contract

Under a flat penalty contract, the supplier incurs a penalty charge of p in each period in which the contract service level is not met, i.e., in each period in which $SL\ D > y_s - D_{L_s}$. To determine the probability of this event, we partition the demand space into the three areas shown in Figure 4.1. In area 1, the demand over the previous L_s periods was greater than the base stock level y_s. The inventory is insufficient to fill the backorders of previous periods, the service level is zero, and the supplier incurs a penalty payment of p. In area 2, all previous backorders are filled, but the inventory is less than $SL\ D$ and the supplier incurs a penalty of p. In area 3, all previous backorders are filled, the inventory is sufficient to fill at least $SL\ D$ of the demand, and no penalty is incurred.

The probability that the supplier incurs *no* penalty charge is

$$\Pr(SL\ D \leq y_s - D_{L_s}) = \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{SL}\right) dx$$

and the probability that the supplier incurs a penalty charge is

$$\Pr(SL\ D > y_s - D_{L_s}) = 1 - \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{SL}\right) dx.$$

The expected penalty cost per period is $p \Pr(SL\ D > y_s - D_{L_s})$ and the expected profit of the supplier can be computed as

$$\begin{aligned} E\Pi_s^f(y_s) &= E[(w-c)D - h_s(y_s - D_{L_s+1})^+] - \\ &\quad p \Pr(SL\ D > y_s - D_{L_s}) \\ &= (w-c)\mu - h_s \int_{x=0}^{y_s} (y_s - x) f_{L_s+1}(x) dx - \\ &\quad p\left(1 - \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{SL}\right) dx\right). \end{aligned} \tag{4.1}$$

Figure 4.2 illustrates the cost terms for truncated normally distributed demand with $\mu = 20$, $\sigma = 5$, $L_s = 2$, and $h_s = 1$, $p = 10$, and $SL = 0.9$. From Figure 4.2 it can be seen that the expected cost function is not convex in the base stock level y_s. Similarly, it can be demonstrated that the expected profit function is not concave in y_s. However, Proposition 4.1 states that the expected profit function $E\Pi_s^f(y_s)$ is quasi-concave in the base stock level y_s and states the optimality condition.

Proposition 4.1 *Under a flat penalty contract, the supplier's expected profit function $E\Pi_s^f(y_s)$ is quasi-concave in y_s. The optimal base stock level satisfies*

$$-h_s F_{L_s+1}(y_s) + \frac{p}{SL} \int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s - x}{SL}\right) = 0.$$

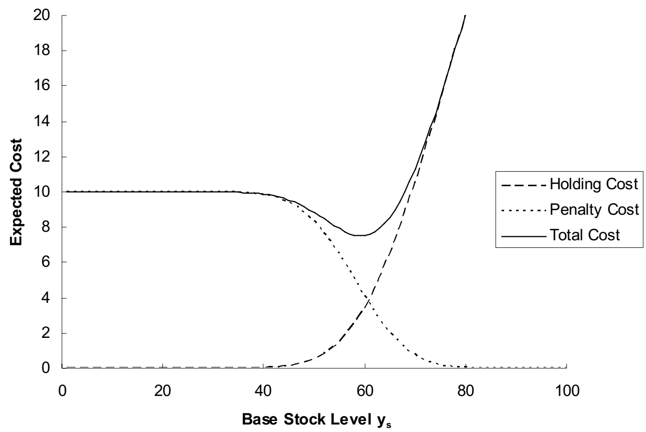

Figure 4.2: Holding Cost, Penalty Cost, and Total Cost for a Flat Penalty Contract

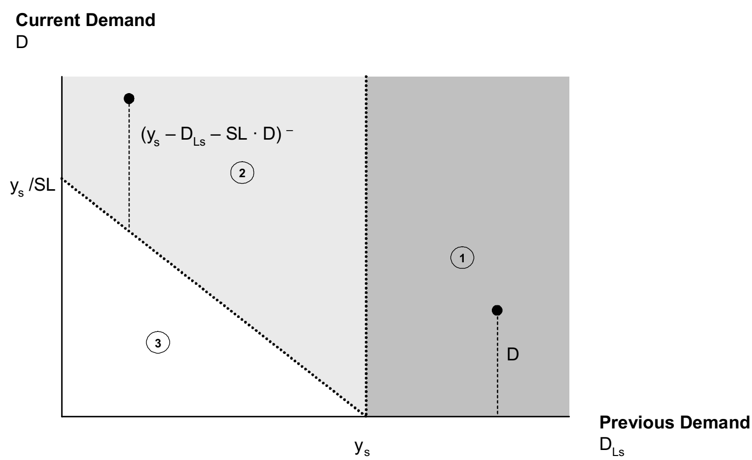

Figure 4.3: Weighted Demand Partition

4.3.2 Unit Penalty Contract

Under a unit penalty contract, the supplier incurs a penalty charge of p for each unit of demand that is backordered and that exceeds the number of units required by the service level contract. To determine the expected penalty per period, we use the demand space partitioning of Figure 4.3. In area 3, all backorders and demands are filled and it suffices to analyze areas 1 and 2. In area 1, the demand over the previous L_s periods was greater than the base stock level y_s. The inventory is insufficient to fill the backorders of the previous periods and the supplier incurs a penalty payment of pD. In area 2, all previous backorders are filled, but the inventory is less than $SL\ D$ and the supplier incurs a penalty of $p(y_s - D_{L_S} - SL\ D)^-$. So, the expected penalty charge per period is

$$E\left[P_u(y_s, SL, D, D_{L_s})\right] =$$

$$p\left(\int_{x=0}^{y_s}\int_{\varphi=\frac{y_s-x}{SL}}^{\infty}\left(\varphi - \frac{y_s-x}{SL}\right) f(\varphi) f_{L_s}(x) d\varphi dx + \int_{x=y_s}^{\infty} \mu f_{L_s}(x)dx\right)$$

$$= p\left(\int_{x=0}^{y_s} b_s\left(\frac{y_s-x}{SL}\right) f_{L_s}(x)dx + (1 - F_{L_s}(y_s))\,\mu\right),$$

where

$$b_s(z) = \int_{\delta=z}^{\infty} (\delta - z) f(\delta) d\delta$$

is the expected backorder level. The expected profit function is

$$E\Pi_s^u(y_s^u) = (w-c)\mu - h_s \int_{x=0}^{y_s} (y_s - x) f_{L_s+1}(x)dx$$

$$- p\left(\int_{x=0}^{y_s} b_s\left(\frac{y_s-x}{SL}\right) f_{L_s}(x)dx + (1 - F_{L_s}(y_s))\,\mu\right). \quad (4.2)$$

Figure 4.4 illustrates the cost terms for truncated normally distributed demand with $\mu = 20$, $\sigma = 5$, $L_s = 2$, and $h_s = 1$, $p = 1$, and $SL = 0.9$. As before, the expected total cost function is not convex in y_s and the

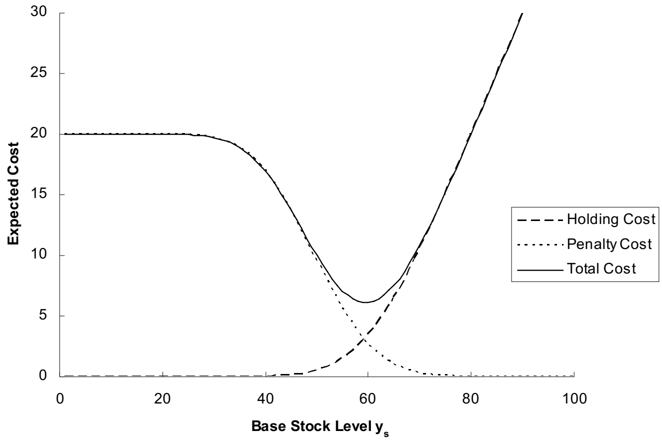

Figure 4.4: Holding Cost, Penalty Cost, and Total Cost for a Unit Penalty Contract

expected profit function is not concave in y_s. However, Proposition 4.2 states that the expected profit function $E\Pi_s^u(y_s)$ is quasi-concave and states the optimality condition.

Proposition 4.2 *Under a unit penalty contract, the supplier's expected profit function $E\Pi_s^u(y_s)$ is quasi-concave in y_s. The optimal base stock level satisfies*

$$-h_s F_{L_s+1}(y_s) - p\int_{x=0}^{y_s} \frac{\left(F\left(\frac{y_s-x}{SL}\right)-1\right)}{SL} f_{L_s}(x)dx = 0.$$

We have seen how the supplier responds to flat and unit penalty contracts. Next, we analyze how contract parameters can be determined that coordinate the supply chain.

4.4 Coordinating Contracts

Our interest is in solutions that coordinate the supply chain, i.e., in solutions that ensure that the maximum expected supply chain profit of a centralized solution is achieved. In the following, we will first characterize the optimal centralized solution and then present coordinating flat and unit penalty contracts.

4.4.1 Optimal Centralized Solution

Clark and Scarf (1960) and Federgruen and Zipkin (1984) have shown how the base stock levels of the optimal centralized solution can be computed. We refer to the optimal base stock levels as $y_s^{SC^*}$ and $y_m^{SC^*}$ and denote the corresponding expected supply chain profit by $E\Pi^*_{SC}(y_m^{SC^*}, y_s^{SC^*})$. In our setting, the manufacturer's objective is to maximize his expected profit subject to a constraint that the supplier's expected profit is at least $\hat{\Pi}_s$. This implies that the maximum expected profit of the manufacturer is

$$E\Pi^*_m(y_m^{SC^*}, y_s^{SC^*}) = E\Pi^*_{SC}(y_m^{SC^*}, y_s^{SC^*}) - \hat{\Pi}_s.$$

To achieve this profit, the manufacturer must use a base stock level $y_m = y_m^{SC^*}$ and must design a contract that (i) incentivizes the supplier to choose a base stock level of $y_s = y_s^{SC^*}$ and (ii) results in an expected profit of $\hat{\Pi}_s$ at the supplier, i.e., the manufacturer is only interested in contracts with

$$\begin{aligned} E\Pi_s(y_s^{SC^*}) &= (w-c)\mu - h_s \int_{x=0}^{y_s^{SC^*}} (y_s^{SC^*} - x) f_{L_s+1}(x)dx \\ &\quad - E[P(y_s^{SC^*}, SL, p, D, D_{L_s})] \\ &= \hat{\Pi}_S. \end{aligned}$$

Solving the equation for the wholesale price w, we obtain the optimal wholesale price of the coordinated supply chain:

$$\begin{aligned} w^*(y_s^{SC^*}, SL, p) &= c + \frac{h_s}{\mu} \int_{x=0}^{y_s^{SC^*}} (y_s^{SC^*} - x) f_{L_s+1}(x) dx \\ &+ \frac{E[P(y_s^{SC^*}, SL, p, D, D_{L_s})]}{\mu} + \frac{\hat{\Pi}_S}{\mu}. \end{aligned}$$

The last equation shows us that the supply chain optimal wholesale price $w^*(y_s^{SC^*}, SL, p)$ is equal to the sum of unit cost, expected unit inventory holding cost at the supplier, expected unit penalty cost, plus the unit reservation profit $\hat{\Pi}_S/\mu$. To compute the optimal wholesale price, the manufacturer must compute the optimal $y_s^{SC^*}$ using the Clark and Scarf model and must specify SL and p such that the supplier uses a base stock level that is equal to the base stock level of the coordinating solution, i.e., such that $y_s = y_s^{SC^*}$. We will show next how the manufacturer can specify SL and p for our two contract types such that the supply chain is coordinated.

4.4.2 Coordinating Flat Penalty Contract

For a flat penalty contract, Proposition 4.3 states for which combinations of SL and p the supplier chooses the supply chain optimal base stock level.

Proposition 4.3 *Under a flat penalty contract the supplier's optimal base stock level is equal to* $y_s^{SC^*}$ *if*

$$p(SL) = \frac{h_s F_{L_s+1}(y_s^{SC*})}{\frac{1}{SL} \int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\left(\frac{y_s^{SC*}-x}{SL}\right) dx}, 0 < SL \leq 1. \qquad (4.3)$$

For a given contract service level SL, we can compute the coordinating penalty cost $p(SL)$ using Equation (4.3). Figure 4.5 shows numerical results for three examples with truncated normally distributed demand with $\mu = 20$, $\sigma = 5$, $h_s = 1$, $L_s = 2$, and $L_m = 4$. To analyze a variety of situations, we used inventory holding cost and backorder penalty cost combinations of $(h_m, b_m) = \{(1.7, 1.5), (55, 55), (1500, 1500)\}$, resulting in optimal centralized solutions of $(y_s^{SC*}, y_m^{SC*}) = \{(30, 109), (50, 100), (60, 100)\}$.

Figure 4.5 illustrates that the coordinating penalty cost $p(SL)$ is increasing in the contract service level SL if the base stock level at the supplier is low ($y_s^{SC*} = 30$). Similarly, $p(SL)$ is decreasing in the contract service level SL if the base stock level y_s^{SC*} is high ($y_s^{SC*} = 60$). For an intermediate base stock level ($y_s^{SC*} = 50$), the coordinating penalty cost $p(SL)$ is decreasing-increasing in the contract service level SL. Proposition 4.4 states general results on the effect of the contract service level SL on the coordinating penalty cost $p(SL)$.

Proposition 4.4 *Under a flat penalty contract the penalty cost factor p is quasi-convex in the contract service level SL.*

To understand the rationale behind the impact of the contract service level SL on the coordinating penalty cost $p(SL)$, recall that the supplier trades off marginal savings in expected inventory holding cost against marginal increases in expected penalty payments when deciding on the base stock level y_s. From Equation (4.1), it can be seen that in a coordinated supply chain

$$h_s \frac{d}{dy_s} E(y_s - D_{L_s+1})^+ \Big|_{y_s = y_s^{SC*}} = -p \frac{d}{dy_s} \Pr(SL\ D > y_s - D_{L_s}) \Big|_{y_s = y_s^{SC*}} .$$

Note that the marginal expected inventory holding cost (left hand side) does not depend on the contract service level SL, but that the marginal expected penalty cost (right hand side) does. Now consider a situation where the *marginal penalty payment probability,* i.e., d/dy_s $\Pr(SL\ D > y_s - D_{L_s}) \leq 0$ is increasing in SL, such as for $y_s^{SC*} = 30$ in our numerical example. To keep the marginal expected penalty payment constant in such a setting, we must increase the penalty cost p if we increase the contract service level SL. In other words, in situations in

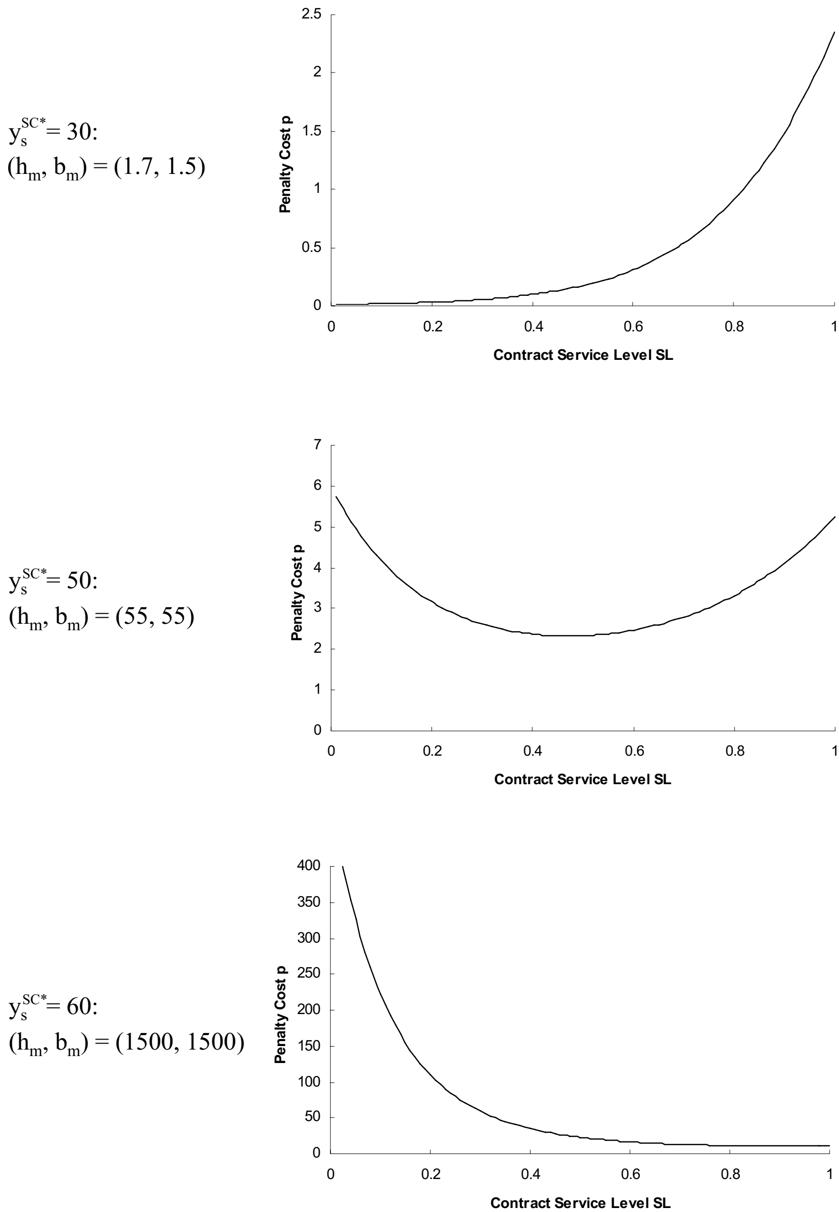

Figure 4.5: Coordinating Flat Penalty Contracts

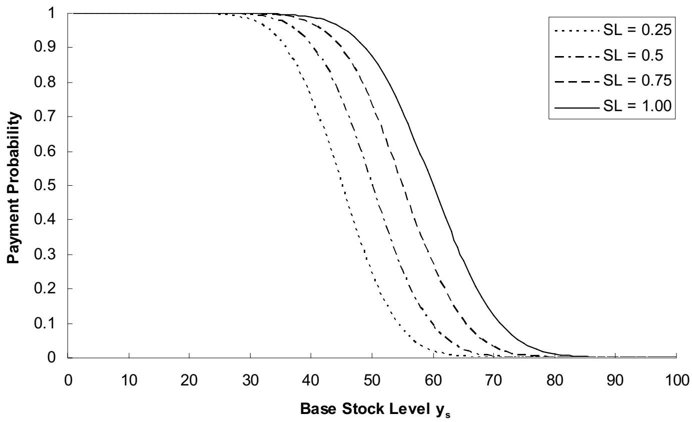

Figure 4.6: Penalty Payment Probabilities for Different Service Levels

which the marginal penalty payment probability is increasing in SL, the coordinating penalty cost $p(SL)$ is increasing in the contract service level SL. Similarly, in situations in which the marginal penalty payment probability is decreasing in SL, the coordinating penalty cost $p(SL)$ is decreasing in contract service level SL.

The question remaining is when the marginal penalty payment probability decreases in SL and when it increases in SL? To answer this question, we re-write the penalty payment probability as

$$\Pr(SL\ D > y_s - D_{L_s}) = \Pr(D_{L_s} + SL\ D > y_s),$$

which shows that the payment probability is a complementary cumulative distribution function (ccdf.). For our numerical example, this ccdf. is shown in Figure 4.6 for various contract service levels SL. Since demand is logconcave distributed, the marginal penalty payment probability, i.e., the derivative of the ccdf., is increasing for small base stock levels y_s and is decreasing for large y_s, which explains why the coordinating penalty cost $p(SL)$ is increasing for small base stock levels y_s

and decreasing for large base stock levels y_s.

4.4.3 Coordinating Unit Penalty Contract

Analogously to Proposition 4.3, Proposition 4.5 states for which combinations of SL and p the supplier chooses the supply chain optimal base stock level under a unit penalty contract.

Proposition 4.5 *Under a unit penalty contract the supplier's optimal base stock level is equal to* $y_s^{SC^*}$ *if*

$$p(SL) = \frac{h_s F_{L_s+1}(y_s^{SC*})}{\int_{x=0}^{y_s^{SC*}} \frac{\left(1-F\left(\frac{y_s^{SC*}-x}{SL}\right)\right)}{SL} f_{L_s}(x)dx}, 0 < SL \leq 1. \tag{4.4}$$

For a given contract service level SL, we can compute the coordinating penalty cost $p(SL)$ using Equation (4.4). Figure 4.7 shows numerical results for our three numerical examples that are similar to the results of the flat penalty contract. For a low base stock level ($y_s^{SC*} = 30$) the coordinating penalty cost $p(SL)$ is increasing in the contract service level SL, for a medium base stock level ($y_s^{SC*} = 50$) it is decreasing-increasing, and for a large base stock level ($y_s^{SC*} = 60$) it is decreasing. As for the flat penalty contracts, we can state general results on the effect of the contract service level SL on the coordinating penalty cost $p(SL)$ in Proposition 4.6.

Proposition 4.6 *Under a unit penalty contract the penalty cost factor* p *is convex in the contract service level* SL.

As before, we can explain the effect of the contract service level SL on the coordinating penalty cost $p(SL)$ by recalling that the supplier trades off marginal savings in expected inventory holding cost against marginal increases in expected penalty payments when deciding on the base stock level y_s. From Equation (4.2), it can be seen that in a

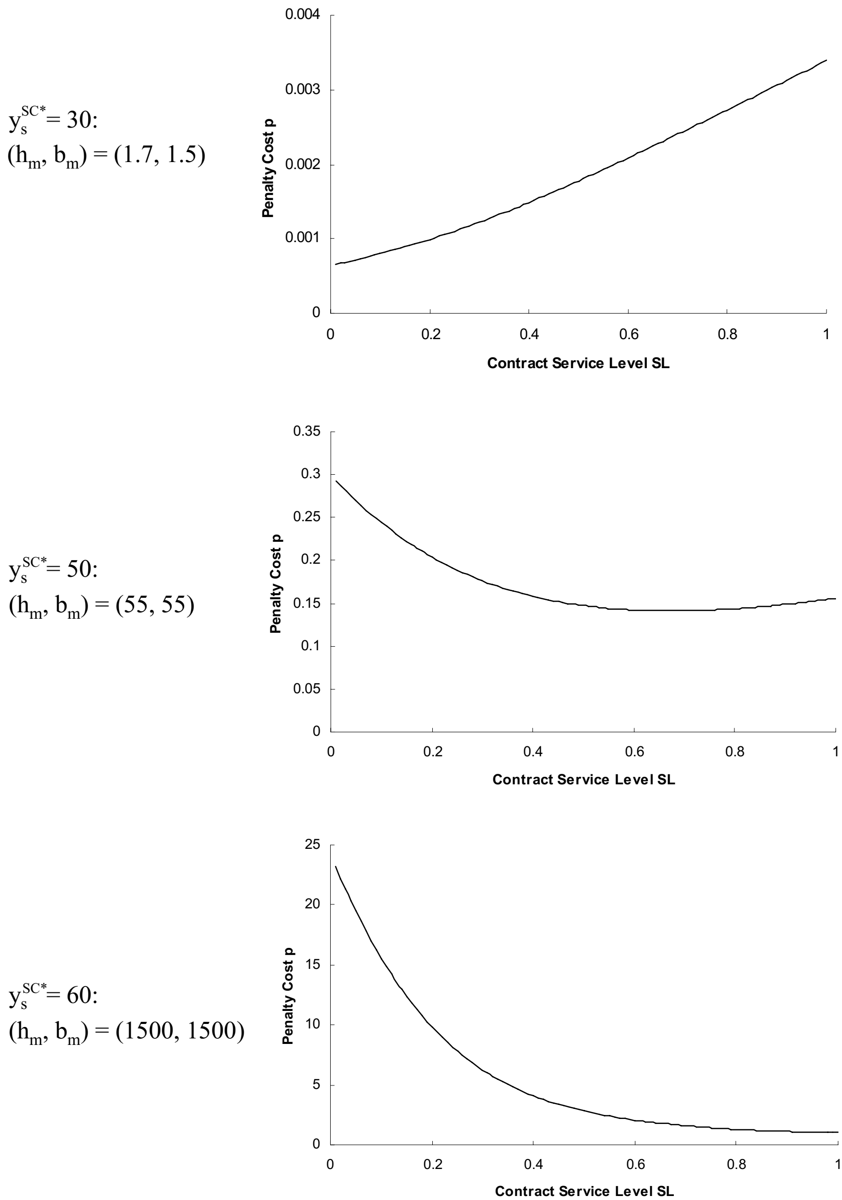

Figure 4.7: Coordinating Unit Penalty Contracts

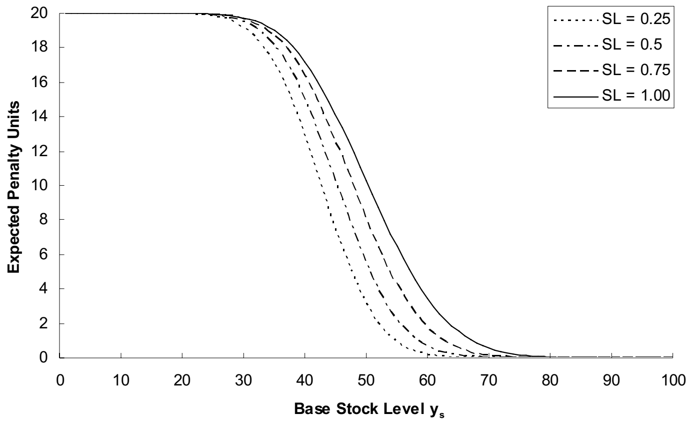

Figure 4.8: Expected Number of Units Short for Different Service Levels

coordinated supply chain

$$h_s \frac{d}{dy_s} E(y_s - D_{L_s+1})^+ \Big|_{y_s = y_s^{SC*}} =$$
$$- p \, \frac{d}{dy_s} \left(\int_{x=0}^{y_s} b_s \left(\frac{y_s - x}{SL} \right) f_{L_s}(x) dx + (1 - F_{L_s}(y_s)) \, \mu \right) \Big|_{y_s = y_s^{SC*}} .$$

Using the same arguments as in Subsection 4.4.2, we see that in situations in which the *marginal expected number of units short* , i.e., $d/dy_s \left(\int_{x=0}^{y_s} b_s \left(\frac{y_s - x}{SL} \right) f_{L_s}(x) dx + (1 - F_{L_s}(y_s)) \, \mu \right)$, is increasing in the contract service level SL, the coordinating penalty cost $p(SL)$ is increasing in SL. In situations in which the marginal expected number of units short is decreasing in SL, the coordinating penalty cost $p(SL)$ is decreasing in the contract service level SL. Figure 4.8 shows that the marginal expected number of units short is increasing in the contract service level SL for small base stock levels y_s and is decreasing for large y_s. As before, this observation explains the effect that for a unit penalty contract the coordinating penalty cost $p(SL)$ is increasing for small base stock levels y_s and decreasing for large base stock levels y_s.

The manufacturer must decide which combination of SL and p to specify in the contract. Since any point on the curves of Figures 4.5 and 4.7 coordinates the supply chain, the manufacturer has infinitely many combinations to choose from. However, there exists one point on each curve that is particularly attractive. We refer to these points as *contract consistent points*, because at these points the contract service level is equal to the traditional α or β service levels. We discuss this issue next.

4.5 Contract Consistent Service Levels

In inventory theory and in practice, two commonly used service levels are the α and β service levels. The α service level specifies the fraction of periods in which demand is completely filled. The β service level specifies the expected fraction of demand that is filled in a period. These measures are based on an infinite horizon analysis of the inventory systems, i.e., an infinite number of periods is used for measuring the service level, whereas we measure the service level in each period. The lengths of the time horizons over which the service levels are measured matter. Thomas (2005) shows that service level measures with long and short time horizons might differ significantly and that the expected finite horizon service level is always greater than the infinite horizon service level.

We use a time horizon of one period in this chapter, i.e., we evaluate the service level each period. The corresponding infinite horizon service level measures are the α and β service levels: The traditional α service level measure is closely related to the contract service level SL of a flat penalty contract. The traditional β service level is closely related to the contract service level SL of a unit penalty contract.

In a coordinated supply chain, the supplier chooses a base stock level $y_s^{SC^*}$ and we can compute the corresponding infinite horizon α and β

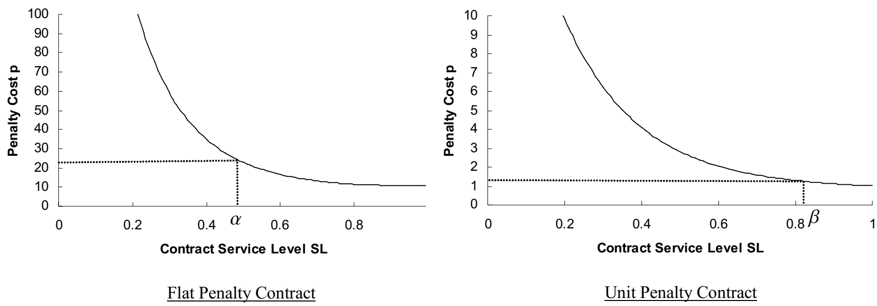

Figure 4.9: Consistent Contracts with $\alpha = 50\%$ and $\beta = 82.75\%$

service levels as (Sobel, 2004)

$$\alpha = F_{L_s+1}\left(y_s^{SC^*}\right) \tag{4.5}$$

and

$$\beta = 1 - \frac{E\left(\left[D - \left(y_s^{SC^*} - D_{L_s}\right)^+\right]^+\right)}{\mu} = \frac{\int_{x=0}^{y_s^{SC^*}} \left[F_{L_s}(x) - F_{L_s+1}(x)\right] dx}{\mu}. \tag{4.6}$$

From a managerial perspective, it would be attractive to use supply contracts where the contract service level SL is equal to the traditional α or β service level, because the supplier can then focus on achieving the service level by using well-known methods from inventory management, i.e., by using Equations (4.5) and (4.6).

Figure 4.9 shows how we can design such contracts for our example with $y_s^{SC*} = 60$ and $\alpha = 50\%$ and $\beta = 82.75\%$. If we choose $SL = \alpha = 50\%$ and $p = 22.86$ for a flat penalty contract or $SL = \beta = 82.75\%$ and $p = 1.24$ for a unit penalty contract, the supply chain is coordinated. Proposition 4.7 states that such service level consistent contracts always exist.

Proposition 4.7 *For all contract service levels* $0 < SL \leq 1$*, there exists a flat penalty contract and a unit penalty contract that coordinate the supply chain.*

4.6 Conclusion

Service levels are commonly used in theory and practice for evaluating supplier performance. In many supply contracts, service levels are specified as well as the consequences of not achieving them. In this chapter, we have analyzed two types of such service level based supply contracts, flat penalty and unit penalty contracts. We have shown that for any service level $0 < SL \leq 1$, there exists a coordinating penalty cost $p(SL)$ and a wholesale price w such that the supply chain is coordinated, i.e., the supply chain optimal solutions are chosen by the supplier and the manufacturer. The supplier achieves an expected profit that is equal to her reservation profit and the manufacturer maximizes his expected profit. We have also derived some structural properties about coordinating contracts, such as the (quasi-)convexity of the coordinating penalty cost $p(SL)$ in the contract service level SL, and have provided numerical results. Finally, we have compared our service level measures with the traditional service level measures. The results of our analyses can support decision makers in specifying the parameters of service level based supply contracts.

Our model can be easily modified to a setting where the manufacturer uses a minimal customer service level instead of a backorder penalty cost. Then, the expected profit function at the manufacturer changes and therefore our centralized solution has to be modified accordingly. Van Houtum et al. (1996) show how the supply chain optimal base stock levels for this setup can be determined. We can use their approach for determining the supply chain optimal base stock levels, from which everything follows.

In the literature, service level based supply contracts have not been analyzed analytically, although these contracts are very popular in practice. With this chapter, we take a first step towards filling this gap. The results of our analyses allow for a better understanding of the effect of service level based supply contracts on supplier, manufacturer, and supply chain performance. In our models, the objective is maximizing expected profits and we have shown that the supply chain is coordinated for any combination $(SL, p(SL))$, assuming that the optimal wholesale price is chosen. We have not analyzed the effect of these combinations on the variability of profits for each company. Taking the variability of profits into account, we can select a combination $(SL, p(SL))$ that coordinates the supply chain at a minimal variability of profits of the supplier or the manufacturer. We have analyzed this issue numerically, but have so far not been able to derive structural results and therefore leave the analysis for future research.

4.7 Proofs

Proof of Proposition 4.1. We want to show that $E\Pi_s^f(y_s)$ is quasi-concave in y_s and a unique maximum y_s^f exists. Consider the derivative

$$\frac{dE\Pi_s^f(y_s)}{dy_s} = -h_s F_{L_s+1}(y_s) + \frac{p}{SL}\int_{x=0}^{y_s} f_{L_s}(x)\, f\left(\frac{y_s - x}{SL}\right) dx.$$

Assume that the derivative $\frac{dE\Pi_s^f(y_s)}{dy_s}$ is non-negative for $0 \leq y_s \leq y_s^f$. Then

$$\begin{aligned} -h_s F_{L_s+1}(y_s) + \frac{p}{SL}\int_{x=0}^{y_s} f_{L_s}(x)\, f\left(\frac{y_s - x}{SL}\right) dx \quad &\geq \quad 0 \\ \Leftrightarrow \frac{F_{L_s+1}(y_s)}{\frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x)\, f\left(\frac{y_s-x}{SL}\right) dx} \quad &\leq \quad \frac{p}{h_s}. \end{aligned}$$

We can rewrite the last line as

$$\begin{aligned}
\frac{F_{L_s+1}(y_s)}{\frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s-x}{SL}\right)dx} &= \frac{F_{L_s+1}(y_s)}{\frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s-x}{SL}\right)dx} \cdot \\
&\quad \frac{f_{L_s+1}(y_s)}{f_{L_s+1}(y_s)} \\
&= \frac{F_{L_s+1}(y_s)}{f_{L_s+1}(y_s)} \cdot \\
&\quad \frac{f_{L_s+1}(y_s)}{\frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s-x}{SL}\right)dx} \\
&\leq \frac{p}{h_s}.
\end{aligned} \tag{4.7}$$

For logconcave frequency functions $f(x)$, the fraction $\frac{F_{L_s+1}(y_s)}{f_{L_s+1}(y_s)}$ is non-decreasing in y_s (Rosling, 2002).

For the second term $\frac{f_{L_s+1}(y_s)}{\frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x)f\left(\frac{y_s-x}{SL}\right)dx}$, we find that it is non-decreasing in y_s since the logconcavity of $f(x)$ implies monotone convolution ratios (Rosling, 2002), i.e., $\frac{f_n(y_s)}{f_m(y_s)}$ is non-decreasing in y_s for $n \geq m$ with $f_n(y_s) = f_{L_s+1}(y_s)$ and with $f_m(y_s) = \frac{1}{SL}\int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s-x}{SL}\right)dx$ being the frequency function of the partial convolution $D_{Ls} + SL \cdot D$.

From $(g(x)h(x))' = g(x)h'(x) + g'(x)h(x)$ the LHS of Equation (4.7) is non-decreasing in y_s. It follows that there exists only one positive area of $\frac{dE\Pi_s^f(y_s)}{dy_s}$ in at most one subset of y_s and the sign of $\frac{dE\Pi_s^f(y_s)}{dy_s}$ changes at most once from + to − and thus the objective function $E\Pi_s^f(y_s)$ is quasi-concave in y_s. Then, an optimal base stock level y_s^f is unique. Hence it is sufficient to set the first derivative to zero.

We also see that a higher penalty cost p leads to a higher base stock level y_s because the positive area is increasing due to a higher p/h_s. ■

Proof of Proposition 4.2. We want to show that $E\Pi_s^u(y_s)$ is quasi-concave in y_s and a unique maximum y_s^u exists. Consider the derivative

$$\frac{dE\Pi_s^u(y_s)}{dy_s} = -h_s F_{L_s+1}(y_s) - p\int_{x=0}^{y_s} \frac{\left(F\left(\frac{y_s-x}{SL}\right)-1\right)}{SL} f_{L_s}(x)dx.$$

Assume that the derivative $\frac{dE\Pi_s^f(y_s)}{dy_s}$ is non-negative for $0 \leq y_s \leq y_s^u$. Then

$$\begin{aligned} -h_s F_{L_s+1}(y_s) - p\int_{x=0}^{y_s} \frac{\left(F\left(\frac{y_s-x}{SL}\right)-1\right)}{SL} f_{L_s}(x)dx &\geq 0 \\ \Leftrightarrow \frac{\int_{x=0}^{y_s} \frac{\left(1-F\left(\frac{y_s-x}{SL}\right)\right)}{SL} f_{L_s}(x)dx}{F_{L_s+1}(y_s)} &\geq \frac{h_s}{p}. \end{aligned}$$

In the following we will analyze the term $\frac{\int_{x=0}^{y_s}\left(1-F\left(\frac{y_s-x}{SL}\right)\right)f_{L_s}(x)dx}{F_{L_s+1}(y_s)}$. Reorganizing the term results in

$$\begin{aligned} \frac{\int_{x=0}^{y_s}\left(1-F\left(\frac{y_s-x}{SL}\right)\right)f_{L_s}(x)dx}{F_{L_s+1}(y_s)} &= \frac{F_{L_s}(y_s) - \int_{x=0}^{y_s} F\left(\frac{y_s-x}{SL}\right)f_{L_s}(x)dx}{F_{L_s+1}(y_s)} \\ &= \frac{F_{L_s}(y_s)}{F_{L_s+1}(y_s)} \cdot \\ &\quad \left(1 - \frac{\int_{x=0}^{y_s} F\left(\frac{y_s-x}{SL}\right)f_{L_s}(x)dx}{F_{L_s}(y_s)}\right). \end{aligned}$$

We know that $\frac{F_{L_s}(y_s)}{F_{L_s+1}(y_s)}$ is non-increasing in y_s and ≥ 0. Also the term $\left(1 - \frac{\int_{x=0}^{y_s} F\left(\frac{y_s-x}{SL}\right)f_{L_s}(x)dx}{F_{L_s}(y_s)}\right)$ is non-increasing in y_s and we see that $\frac{\int_{x=0}^{y_s} F\left(\frac{y_s-x}{SL}\right)f_{L_s}(x)dx}{F_{L_s}(y_s)} \leq 1$ (Rosling, 2002).

From $(g(x)h(x))' = g(x)h'(x) + g'(x)h(x)$ we can see that the term $\frac{F_{L_s}(y_s)}{F_{L_s+1}(y_s)}\left(1 - \frac{\int_{x=0}^{y_s} F\left(\frac{y_s-x}{SL}\right)f_{L_s}(x)dx}{F_{L_s}(y_s)}\right)$ is also non-increasing in y_s and $\frac{d}{dy_s}\frac{\int_{x=0}^{y_s}\left(1-F\left(\frac{y_s-x}{SL}\right)\right)f_{L_s}(x)dx}{F_{L_s+1}(y_s)} \leq 0$. It follows that $\frac{\int_{x=0}^{y_s}\frac{\left(1-F\left(\frac{y_s-x}{SL}\right)\right)}{SL}f_{L_s}(x)dx}{F_{L_s+1}(y_s)}$ is non-increasing in y_s. Then the derivative is positive in at most one subset of y_s and the sign of $\frac{dE\Pi_s^u(y_s)}{dy_s}$ changes at most once from + to − and hence the objective function is quasi-concave in y_s. Thus, an optimal base stock level y_s^u is unique and it is sufficient to set the first derivative to zero.

We also see that a higher penalty cost p leads to a higher base stock level y_s because the positive area increases due to a lower h_s/p. ■

Proof of Proposition 4.3. The supplier solves

$$\begin{aligned} E\Pi_s^*(y_s) &= \max_{y_s} \; (w^*(y_s^{SC^*}, SL, p) - c)\mu \\ & \quad -h_s \int_{x=0}^{y_s} (y_s - x) f_{L_s+1}(x) dx \\ & \quad -E[P_f(y_s, SL, p, D, D_{L_s})]. \end{aligned}$$

Differentiation of the expected profit function with respect to y_s yields

$$\frac{d}{dy_s} E\Pi_s(y_s) = -h_s F_{L_s+1}(y_s) + \frac{p}{SL} \int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s - x}{SL}\right) dx. \quad (4.8)$$

With Equation (4.3) in Equation (4.8) we see that $d/dy_s E\Pi_s(y_s) = 0$ for $y_s = y_s^{SC^*}$. Since the expected profit function is quasi-concave, this corresponds to the optimal solution. ∎

Proof of Proposition 4.4. The proof follows by the unimodal property of the demand distribution. Consider the derivative

$$\frac{d}{dSL} p(SL) = -\frac{h_s F_{L_s+1}(y_s) \frac{d}{dSL}\left(\frac{1}{SL}\int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\left(\frac{y_s^{SC*}-x}{SL}\right) dx\right)}{\left(\frac{1}{SL}\int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\left(\frac{y_s^{SC*}-x}{SL}\right) dx\right)^2}.$$

The term $\frac{d}{dSL}\left(\frac{1}{SL}\int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\left(\frac{y_s^{SC*}-x}{SL}\right) dx\right)$ is the derivation of the frequency function of the convolution $D_{L_s}+SL\cdot D$. From $F_n(x) \leq F_m(x)$ with $n \geq m$ and $n = m + \varepsilon$ we can see that $f_n(x) \leq f_m(x)$ for $x \leq \bar{y}$ with $\bar{y}$ being the modal value of the convolution $D_{L_s} + SL \cdot D$ with $m = L_s + SL$, and $f_n(x) \geq f_m(x)$ for $x \geq \bar{y}$. Thus, for a given $y_s^{SC^*}$ the sign of $\frac{d}{dSL}\left(\frac{1}{SL}\int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\left(\frac{y_s^{SC*}-x}{SL}\right) dx\right)$ can change at most one time from $+$ to $-$ and it follows immediately that the sign of $\frac{dp(SL)}{dSL}$ can only change at most once from $-$ to $+$. This concludes our proof. ∎

Proof of Proposition 4.5. The supplier solves

$$\begin{aligned} E\Pi_s^*(y_s) &= \max_{y_s} \ (w^*(y_s^{SC^*}, SL, p) - c)\mu \\ &\quad -h_s \int_{x=0}^{y_s} (y_s - x) f_{L_s+1}(x)dx \\ &\quad -E[P_u(y_s, SL, p, D, D_{L_s})]. \end{aligned}$$

Differentiation of the expected profit function with respect to y_s yields

$$\frac{d}{dy_s} E\Pi_s(y_s) = -h_s F_{L_s+1}(y_s) - p \int_{x=0}^{y_s} \frac{\left(F\left(\frac{y_s - x}{SL}\right) - 1\right)}{SL} f_{L_s}(x)dx. \quad (4.9)$$

With Equation (4.4) in Equation (4.9) we see that $d/dy_s E\Pi_s(y_s) = 0$ for $y_s = y_s^{SC^*}$. Since the expected profit function is quasi-concave, this corresponds to the optimal solution. ■

Proof of Proposition 4.6. In the coordinated solution the penalty cost factor equals

$$p(SL) = -\frac{h_s F_{L_s+1}(y_s^{SC*})}{\int_{x=0}^{y_s^{SC*}} \frac{\left(F\left(\frac{y_s^{SC*} - x}{SL}\right) - 1\right)}{SL} f_{L_s}(x)dx}, 0 < SL \leq 1.$$

First, we will analyze the denominator

$$-\int_{x=0}^{y_s^{SC*}} \frac{\left(F\left(\frac{y_s^{SC*} - x}{SL}\right) - 1\right)}{SL} f_{L_s}(x)dx.$$

By substitution with $z = \frac{y_s^{SC*} - x}{SL}$ we get

$$\begin{aligned} &\int_{z=\frac{y_s^{SC*}}{SL}}^{0} (F(z) - 1) f_{L_s}(y_s^{SC*} - zSL)dz \\ = &\int_{z=0}^{\frac{y_s^{SC*}}{SL}} (1 - F(z)) f_{L_s}(y_s^{SC*} - zSL)dz = \hat{F}\left(\frac{y_s^{SC*}}{SL}\right). \end{aligned}$$

From Rosling (2002) we know that $1 - F(z)$ is logconcave if $f(z)$ is logconcave, and that logconcavity of $1 - F(z)$ is closed under convolution. Thus, $\hat{F}\left(\frac{y_s^{SC*}}{SL}\right)$ is logconcave in $\frac{y_s^{SC*}}{SL}$ and consequently

$$\bar{F}\left(\frac{y_s^{SC*}}{SL}\right) = \frac{1}{\hat{F}\left(\frac{y_s^{SC*}}{SL}\right)}$$

is logconvex in $\frac{y_s^{SC*}}{SL}$. Now, let $t(SL)$ be the transformation function $t(SL) = \frac{y_s^{SC*}}{SL}$. Clearly, $t(SL)$ is convex in the contract service level SL $(0 < SL \leq 1)$. Then

$$\bar{F}\left(\frac{y_s^{SC*}}{SL}\right) = \bar{F}\left(t(SL)\right).$$

From Bagnoli and Bergstrom (2005), Theorem 7, we know that if $\bar{F}$ is logconvex and t is a convex function, the composition $\bar{F}\left(t(x)\right)$ is logconvex.

From Boyd and Vandenberghe (2004), we see that logconvexity implies convexity. Scaling with a constant factor preserves convexity. Thus, $p(SL)$ is convex in the contract service level SL. ■

Proof of Proposition 4.7. For all y_s and $0 < SL \leq 1$, there always exists a penalty factor $p(SL)$ that leads to the optimal base stock level if the support of the underlying distribution function has infinite non-negative support. Then, there always exists a flat penalty and a unit penalty contract that achieves the centralized solution. Note that for demand distributions that have a finite support, for instance, the Beta or Uniform distribution, we may not always find such a contract. ■

Chapter 5

Optimal Inventory Allocation for Multiple Retailers under Service Level Contracts

5.1 Introduction

The decision on how to allocate inventory among multiple retailers during a selling season can have a significant impact on the costs of a manufacturer. The manufacturer faces two decisions: First, she has to decide if she wants to distribute all inventory at the beginning of the selling season, i.e., to follow a ship-all policy, or if she rather reserves a certain fraction of the inventory for a second delivery later in the season. Second, she has to decide on an optimal allocation of inventory among the retailers. It has been shown in literature that it is beneficial for the manufacturer to reserve a certain fraction for a second delivery. The second delivery then should try to balance the retailers' inventories as

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_5

much as possible. This approach is widely accepted in practice. For example, an international sports equipment manufacturer ships only 70% of the initial inventory to the stores at the beginning of the selling season. The other 30% is distributed later on during the selling season depending on the real demands that have been observed at the stores.

In practice, the retailers have a natural interest in a certain availability of the manufacturer's product throughout the selling season. Therefore, the manufacturer's performance is controlled by supply contracts with specific financial consequences (Fry et al., 2001). There are several types of supply contracts that are commonly used. A particularly popular type of supply contract in literature and practice is the service level contract. Under a service level contract, retailers specify a service level that the manufacturer must achieve and the financial consequences of missing it. Since most companies measure their own service levels and many companies also measure the service levels of their suppliers, the information required for implementing a service level contract is typically available at both partners of the supply chain. In the consumer goods industry, for instance, essentially all manufacturers measure their own service levels and 70% of the retailers measure the service levels of the manufacturers. In this industry, manufacturers and retailers typically agree on a service level that a manufacturer is expected to achieve (Behrenbeck et al., 2007).

One of the main challenges for the manufacturer is the optimal fulfillment of multiple service level contracts. This issue has not been analyzed in literature and we are filling the gap in this chapter. We analyze a supply chain with one manufacturer and multiple retailers. The manufacturer has signed a service level contract with each retailer. Our interest is in optimal allocation policies for a given set of service level contracts. We analyze how a manufacturer responds to multiple service level contracts optimally. We derive decision rules that allow the manufacturer to trade off between different contracts and show that the traditional inventory balancing approach leads to suboptimal cost levels. The results of our analyses provide a foundation for decision makers when they have to allocate scarce inventory among multiple retailers.

The remainder of the chapter is structured as follows: In Section 5.2, we present a mathematical model of a two-echelon supply chain that is governed by service level contracts. In Section 5.3, we analyze the one-retailer case. We derive the probability that a penalty has to be paid and show that a ship-all policy is optimal. We extend the analysis to multiple retailers in Section 5.4 where we trade off among multiple contracts. We show how the manufacturer can make an optimal allocation at the second replenishment and compare our approach to an inventory balancing policy. In Section 5.5 we analyze how the manufacturer's optimal policy affects the retailers' performance. We conclude in Section 5.6. All proofs of this chapter are contained in Section 5.7.

5.2 Supply Chain Model

We consider a two-echelon supply chain as in Figure 5.1 with one manufacturer that produces one product. The manufacturer's objective is to minimize expected costs. The product can only be sold during one selling season and it is delivered to N identical retailers (indexed by $i = 1, \ldots, N$). Our setup is similar to a VMI system where the manufacturer is responsible for the management of her products at the retailers' sites. Under such an agreement, the manufacturer decides on the shipment quantities and bears all inventory-related costs of her replenishment decisions.

The retailers face stochastic, continuous, and stationary end customer demand that is assumed to be independent between periods and retailers. Demand can be arbitrarily distributed as long as the p.d.f. is strongly unimodal with a non-negative skewness. We have already seen before that this property holds for most theoretical distributions that are relevant for modeling demand, such as the Normal, truncated Normal, Gamma with shape parameter $\alpha > 1$, Beta(α, β) with parameters

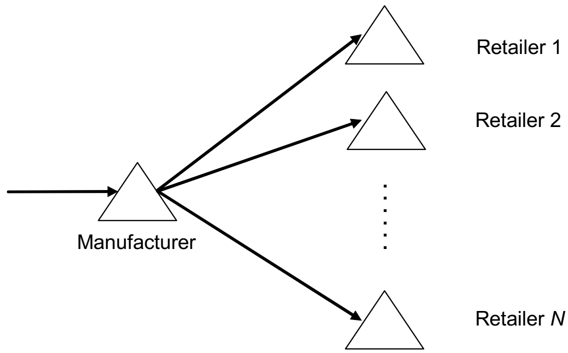

Figure 5.1: Supply Chain Setup

$\alpha \geq \beta \geq 1$, and the Uniform distributions. For an in-depth treatment of logconcave distributions and their application to inventory control we refer the reader to Rosling (2002). Again, we focus on distributions with infinite non-negative support in order to keep our analysis concise. Distributions with finite support can be treated analogously, but they require the definition of feasible regions for the parameter values of the supply contracts, which makes the analysis much more complex and adds little value.

The manufacturer produces R units of the product in $t = 0$, i.e., at the beginning of the selling season, without knowing the exact customer demand. She has two replenishment opportunities, $t = 1$ and $t = 2$, to ship the product to the retailers' stores. In literature, the assumption of only two replenishment opportunities is widely used because it offers a good trade-off between the additional fixed delivery costs and the potential savings due to the risk pooling effect (Güllü and Erkip, 1996). We also assume that the replenishments of the manufacturer arrive at the retail stores instantaneously, i.e., with zero lead time.

The sequence of events during the selling season is as follows: First, the manufacturer produces the product and allocates the initial inventory to the retailers at the beginning of the first period $t = 1$. The retailers then receive the customer orders and start shipping them. If the inventory is not sufficient to fill all customer orders, the retailer backorders and fills the orders after new inventory has arrived at $t = 2$. The backorder assumption is reasonable in cases where the customers do not have an immediate substitute for the product or where the retailer can quote different lead times to the customer. The latter case is widely used at online retailers. For example, Amazon.com can backorder customer orders into the next replenishment period by quoting longer lead times on their web page.

In $t = 2$, the supplier can replenish the inventories at the retail stores. The retailers fill open backorders first and start fulfilling new orders if inventory is still available. Demand in $t = 2$ that cannot be filled is partially backordered. At the end of $t = 2$, excess inventory at the retailer is either transshipped to a retailer in order to fill open backorders or otherwise is salvaged or carried forward to another selling season. In the latter case the manufacturer has to pay an overage cost h per unit of excess inventory (after transshipments) in the supply chain.

To ensure an adequate performance, the retailers require that the manufacturer guarantees a certain service level for the product. This is a common approach in VMI systems as it has been suggested by Choi et al. (2004) or Fry et al. (2001). Choi et al. (2004), for example, use as β-service level (also Type-II service) the infinite-horizon measure

$$1 - \beta = 1 - \frac{\text{expected unsatisfied demand per unit time}}{\text{expected demand per unit time}}.$$

Clearly, a retailer would not wait for an infinite (or large) number of periods to measure the service level and to enforce a penalty payment. In contrast, we focus on contracts that have a finite measurement horizon. The retailers measure the service level only during the two replenishment periods $t = 1, 2$. In this case, the service level is stochastic. That means that it is not known with certainty which service level can be achieved during a selling season for given replenishments in $t = 1, 2$.

Under such a penalty contract, the manufacturer pays the retailer a fixed amount p for a selling season in which the contract service level $0 < SL \leq 1$ is not met, i.e., for a season in which the manufacturer's replenishments do not allow the retailer to fill at least a fraction SL of the customer demand immediately from stock. In $t = 1$, the number of units short is $(\xi_{i1} - S_{i1})^+$ where S_{it} denotes the shipment quantity from the manufacturer to retailer i in period t and ξ_{it} is the stochastic end customer demand in period t at retailer i. In $t = 2$, only the additional units short are counted in order to avoid a double-counting of backorders that were already accounted for in $t = 1$. The additional units short in the second period are $\left(\xi_{i2} - (S_{i1} + S_{i2} - \xi_{i1})^+\right)^+$. Then, the penalty cost function can be written as

$$P_i\left(S_{i1}, S_{i2}, \xi_{i1}, \xi_{i2}\right) = \begin{cases} p & \text{if } \frac{(\xi_{i1} - S_{i1})^+ + \left(\xi_{i2} - (S_{i1} + S_{i2} - \xi_{i1})^+\right)^+}{\xi_{i1} + \xi_{i2}} > 1 - SL \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

In the following we denote the corresponding demand p.d.f. and c.d.f. by $f_t(\cdot)$ and $F_t(\cdot)$, respectively. To keep our analysis concise, we assume that the replenishment periods t have equal lengths and that the demand distributions have a non-negative support, i.e., demand will always be non-negative. Replenishment periods with different lengths can be treated analogously, but would only complicate the notation without adding value to our analysis.

The cost function of the manufacturer equals

$$\begin{aligned} C\left(\bar{S}_1, \bar{S}_2, \bar{\xi}_1, \bar{\xi}_2\right) &= \sum_{i=1}^{N} P_i\left(S_{i1}, S_{i2}, \xi_{i1}, \xi_{i2}\right) \\ &\quad + h\left(\sum_{i=1}^{N}\left(S_{i1} + S_{i2} - \xi_{i1} - \xi_{i2}\right)\right)^+ \\ \text{with } |\bar{S}_1| + |\bar{S}_2| &= R \end{aligned} \tag{5.2}$$

where $\bar{S}_t = (S_{1t}, \ldots, S_{Nt})$ and $\bar{\xi}_t = (\xi_{1t}, \ldots, \xi_{Nt})$ denote the shipment and demand vectors in periods $t = 1, 2$. The operator $|\ldots|$ sums the components of a vector, for instance, $|\bar{S}_1| = \sum_{i=1}^{N} S_{i1}$.

The manufacturer's main concern is to avoid penalty payments by fulfilling the service level contracts and to keep inventories at a low level at the same time. At the beginning of the selling season, the manufacturer is the most unsure about the service level performance that she will finally achieve during that season. At the end of the first replenishment period $t = 1$, the manufacturer already has better knowledge of her performance with respect to the service level contracts. Therefore, she can adjust her allocation decision in $t = 2$. For example, she could deliver more products to a retailer where the performance was suboptimal at the expense of another retailer where the performance in $t = 1$ was superior. Thereby, the manufacturer can actively influence the penalty probabilities for each contract.

In the following section, we gain some important insights on the behavior of a service level contract by analyzing the single-retailer case.

5.3 Single-Contract Allocation

In the single contract case, the manufacturer only serves one retailer i. The manufacturer wants to minimize expected costs, i.e.,

$$\begin{aligned} \min_{R,S_i} EC\left(S_{i1}, S_{i2}, \xi_{i1}, \xi_{i2}\right) &= P_i\left(S_{i1}, S_{i2}, \xi_{i1}, \xi_{i2}\right) \\ &\quad + h\left(S_{i1} + S_{i2} - \xi_{i1} - \xi_{i2}\right)^+ \\ \text{with } S_{i1} + S_{i2} &= R. \end{aligned}$$

Her decision consists of the production quantity R and the initial allocation quantity S_{i1}. The overage costs, i.e., costs driven by the net inventory left at the end of the selling season, only depend on R because $(S_{i1} + S_{i2} - \xi_{i1} - \xi_{i2})^+ = (R - \xi_{i1} - \xi_{i2})^+$. The penalty costs however have a more complex behavior that depends on the production quantity

R and the allocation quantities S_{i1} and $S_{i2} = R - S_{i1}$. Therefore, we will analyze them next.

Equation (5.1) shows that a penalty p has to be paid by the manufacturer if the service level cannot be met, i.e., if

$$\frac{(\xi_{i1} - S_{i1})^+ + \left(\xi_{i2} - (S_{i1} + S_{i2} - \xi_{i1})^+\right)^+}{\xi_{i1} + \xi_{i2}} > 1 - SL$$

or

$$(\xi_{i1} - S_{11})^+ + \left(\xi_{i2} - (S_{i1} + S_{i2} - \xi_{i1})^+\right)^+ - (1 - SL)\,(\xi_{i1} + \xi_{i2}) > 0.$$

Since we are interested in the expected penalty costs, we need to derive the probability of a penalty event.

We can derive necessary conditions for a penalty event by partitioning the demand space (ξ_{i1}, ξ_{i2}). The demand partitions correspond to certain stock-out events during the selling season. The following proposition summarizes the conditions for a penalty payment:

Proposition 5.1 *A penalty payment p has to be paid to the retailer if*

- $\xi_{i2} > \frac{R+S_{i1}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1}$ *for* $R > \xi_{i1} > S_{i1}$ *and* $\xi_{i1} + \xi_{i2} > R$, *or*
- $\xi_{i2} > \frac{S_{i1}}{SL} - \xi_{i1}$ *for* $\xi_{i1} > R \geq S_{i1}$ *and* $\xi_{i1} + \xi_{i2} > R$, *or*
- $\xi_{i2} > \frac{R}{SL} - \xi_{i1}$ *for* $\xi_{i1} \leq S_{i1}$ *and* $\xi_{i1} + \xi_{i2} > R$, *or*
- $\xi_{i2} < -\frac{S_{i1}}{(1-SL)} + \frac{SL}{1-SL}\xi_{i1}$ *for* $\xi_{i1} > S_{i1}$ *and* $\xi_{i1} + \xi_{i2} \leq R$.

Clearly, the manufacturer will never have to pay a penalty if $\xi_{i1} \leq S_{i1}$ and $\xi_{i1} + \xi_{i2} \leq R$ because customer demand could always be met.

Figure 5.2 shows the demand space for a service level $0 < SL < 1$ and a low initial allocation $S_{i1} \ll R$. The grey Areas 1–3 indicate the demand space where a penalty payment has to be paid. In Area 1, the initial allocation S_{i1} was sufficient to fill all demands in the first period,

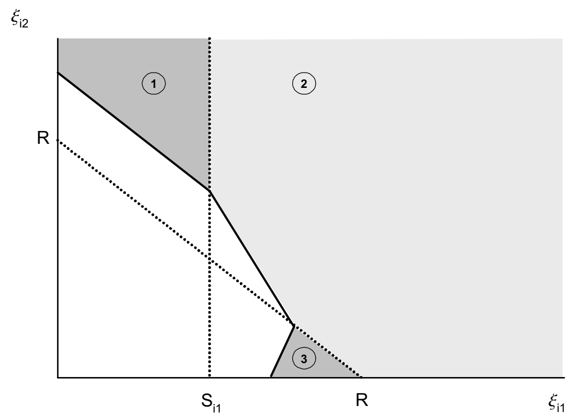

Figure 5.2: Demand Partitions for $SL < 1$ and $S_{i1} \ll R$

but the second period demand was too high to achieve the contract service level. In Area 2, not all customer demands could be filled in the first period and also at the end of the second period some customer orders are still outstanding. In Area 3, the initial inventory S_{i1} was not sufficient to fill all customer orders in $t = 1$. However, the inventory level at the beginning of the second period was high enough to fill all backorders from the first period and a sufficient number of demands in the second period. Interestingly, for a given first period demand ξ_{i1}, a high second period demand ξ_{i2} would lead to no penalty payment whereas a low demand ξ_{i2} would lead to a penalty. This is due to the fact that a perfect fulfillment cannot offset the inferior performance in the first period if the second period demand is too low because the service level is a weighted average of both periods' stock-outs. An exceptional case can appear when the first period demand cannot be fulfilled although the retailer received a high initial allocation $S_{i1} \leq R$. Then Figure 5.3 shows that the slope at the bottom left of Area 4 equals the slope of Area 1. In this case the demand in $t = 1$ was higher than the total

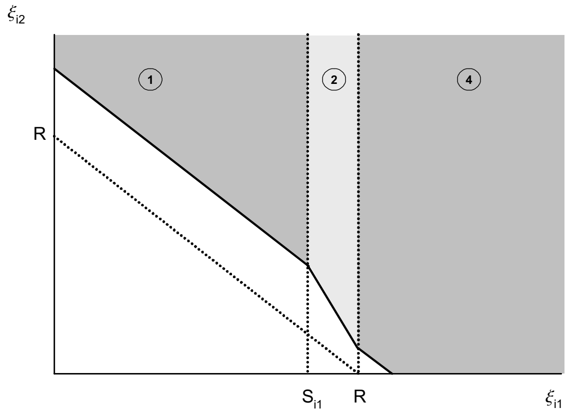

Figure 5.3: Demand Partition for $SL < 1$ and $S_{i1} \leq R$

available inventory R and not a single second period demand could be filled.

By integrating the demand probabilities over the demand partitions where a penalty is due, we can derive the probability of a penalty event.

Proposition 5.2 *The manufacturer has to pay a penalty p with probability*

$$\Pr(P_i > 0) = \int_{\xi_{i1}=0}^{S_{i1}} \left(1 - F\left(\frac{R}{SL} - \xi_{i1}\right)\right) f\left(\xi_{i1}\right) d\xi_{i1}$$
$$+ \int_{\xi_{i1}=S_{i1}}^{\min(S_{i1}+(1-SL)R,R)} \left(1 - F\left(\frac{R+S_{i1}}{SL} - \left(\frac{1}{SL}+1\right)\xi_{i1}\right)\right) f\left(\xi_{i1}\right) d\xi_{i1}$$
$$+ \int_{\xi_{i1}=S_{i1}/SL}^{\min(S_{i1}+(1-SL)R,R)} F\left(-\frac{S_{i1}}{(1-SL)} + \frac{SL}{(1-SL)}\xi_{i1}\right) f\left(\xi_{i1}\right) d\xi_{i1}$$
$$+1 - F\left(\min\left(S_{i1} + (1-SL)\,R, R\right)\right)$$
$$- \int_{\xi_{i1}=R}^{S_{i1}/SL} F\left(\frac{S_{i1}}{SL} - \xi_{i1}\right) f\left(\xi_{i1}\right) d\xi_{i1}.$$

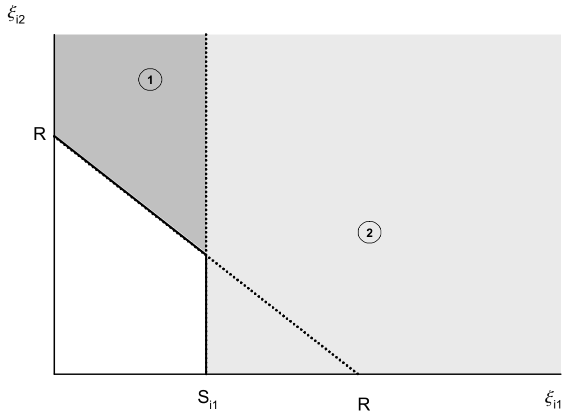

Figure 5.4: Demand Partition for $SL = 1$

For a contract service level $SL = 1$ we get a simple stock-out contract where a penalty is due if in any of the two periods at least one stock-out occurred. Figure 5.4 shows the case for a service level of $SL = 1$. In this special case the probability of a penalty simplifies to

$$\Pr(P_i > 0) = \int_{\xi_{i1}=0}^{S_{i1}} (1 - F(R - \xi_{i1})) f(\xi_{i1}) d\xi_{i1} + 1 - F(S_{i1}).$$

Based on the penalty probability from Proposition 5.2, we can identify the optimal inventory reservation policy that is summarized in the following proposition.

Proposition 5.3 *For the single-contract case, $S_{i1} = R$, i.e., a ship-all policy is optimal.*

A ship-all policy is optimal for one retailer because the manufacturer does not have an incentive to hold back inventory because she cannot realize any risk pooling effects. Neither are there any additional overage costs because the overage costs do not differ between manufacturer and retailers in our model. However, it would be beneficial for the manufacturer to hold back some inventory for a second replenishment if holding costs were significantly higher at the retailers' warehouses than at the manufacturer's site.

Now we can derive the manufacturer's optimal production quantity R^s under a ship-all policy that minimizes her expected costs

$$\min_R EC(R) = h \int_{x=0}^{R} (R - x)(f_1 * f_2)(x)\, dx + p \Pr(P_i > 0)$$

where $f_1 * f_2$ denotes the demand convolution over the periods $t = 1, 2$.

Proposition 5.4 *The unique cost-minimizing production quantity R^s satisfies*

$$-\frac{p}{SL} \frac{d(F_1 * F_2)}{dx} \left(\frac{R^s}{SL} \right) + h(F_1 * F_2)(R^s) \stackrel{!}{=} 0.$$

In this section, we have shown how the penalty probability can be derived for the one-retailer case and how the optimal production quantity can be determined. One result was that a ship-all policy is optimal because there are no positive effects of holding back inventory at the manufacturer. In a multiple retailer case, these effects do exist and they can be significant. Therefore, we show in the next section how the manufacturer can take advantage of the risk pooling effect by holding back inventory and allocating it optimally among the retailers.

5.4 Multiple-Contract Allocation

In the multiple-contract case, the supplier is managing multiple retailer accounts. That means that the manufacturer has signed a service level contract with each of the retailers and wants to minimize his expected costs

$$\begin{aligned}\min_{\bar{S}_1,\bar{S}_2} EC\left(\bar{S}_1,\bar{S}_2,\bar{\xi}_1,\bar{\xi}_2\right) &= \sum_{i=1}^{N} EP_i\left(S_{i1},S_{i2},\xi_{i1},\xi_{i2}\right) \\ &\quad +hE\left(\sum_{i=1}^{N}\left(S_{i1}+S_{i2}-\xi_{i1}-\xi_{i2}\right)\right)^{+} \\ \text{with } |\bar{S}_1|+|\bar{S}_2| &= R.\end{aligned} \tag{5.3}$$

As in the single-retailer model, the allocation policy does not influence the overage costs. The supplier faces two decisions: (i) How much inventory to reserve for the second replenishment and (ii) how to allocate the inventory among the retailers in each replenishment period.

The inventory reservation decision allows the supplier to hold back a fraction $\omega = \frac{|\bar{S}_2|}{R}$ for a second replenishment. This inventory can be used to replenish the retailers' inventory optimally. Holding back inventory leads to positive risk pooling effects because it gives the manufacturer a higher degree of freedom to equalize different demand realizations at the retailers.

The allocation decision, i.e., the quantity each retailer receives at each replenishment, takes the retailers characteristics into account. In the first replenishment, retailers are identical and receive the same quantities. In the second period, the retailers normally differ with respect to their first period performance. Therefore, a manufacturer would prefer to stock more inventory at a retailer where the improvement in the expected penalty costs is the highest.

In literature, several approaches for an optimal solution of the two decisions exist. A common approach that is frequently used is *inventory balancing* that we will shortly introduce next.

5.4.1 Inventory Balancing

Inventory balancing is used to minimize expected backorder costs. The manufacturer has to balance the inventories after observing the first period's demand. For identical retailers this is done by simply equalizing the inventory levels at the retailers (McGavin et al., 1993). For non-identical manufacturers, Zipkin (1984) proposed an equal-fractile rule. Note that it is not always possible to rebalance retailer inventories if excess stock at the retailer cannot be transshipped or returned to the manufacturer.

The inventory balancing decision is based on two state variables: the available inventory ωR for the second replenishment, and the vector of inventory levels $IL_{t=2}$ at the retailers at the beginning of $t = 2$. Based on this information, the inventory balancing is minimizing expected backorders of the last period.

The inventory balancing decision rule only looks into the future because the backorder cost function for the last period does not depend on the performance in $t = 1$. In contrast, the penalty function of a service level contract includes past performance in period 1 and the uncertainty about the future performance in period 2. Therefore, a simple inventory balancing approach does not necessarily lead to optimal results. In the next section we will therefore show, how an optimal inventory allocation can be found.

5.4.2 Contract Balancing

With a *contract balancing* approach, we minimize the manufacturer's expected costs from Equation (5.3). Our focus will be the comparison of the inventory balancing and the contract balancing approach. In order to focus on the allocation decision and to keep our analysis tractable we will make two assumptions: First, the manufacturer's production quantity R is determined by a single-retailer ship-all policy, i.e., $R = N \cdot R^s$ with N being the number of retailers and R^s being determined from Proposition 5.4. As other authors already have stated, the optimization of R is not straightforward, but we will show later how it can be optimized numerically where $R = N \cdot R^s$ serves as an upper bound. Second, we assume the reservation factor ω is determined in a second step or is given by a heuristic. For example, McGavin et al. (1993) have derived an optimization approach for the infinite-retailer case and found out that the 50/25-heuristic leads to good results in the finite-retailer model for a variety of parameter sets.

The discussion in Section 5.3 has shown that the probability for a penalty event depends on the demands in period $t = 1$ and $t = 2$. Therefore, we cannot base our allocation decision in $t = 2$ solely on the current inventory levels at the retailers, but we also have to take the first period performance into account, i.e., the demand ξ_{i1} and the number of customer backorders $(\xi_{i1} - S_{i1})^+$.

Since the retailers are identical, the first replenishment is $S_{i1} = \frac{(1-\omega)R}{N}$ for retailer i. Now the first period demand ξ_{i1} is observed and the retailers use their initial inventory S_{i1}. The manufacturer then has to decide on the optimal replenishment at the second replenishment opportunity. Some retailers can already be excluded from a replenishment decision if there is no possibility of avoiding a penalty payment any more:

Proposition 5.5 *A penalty is always due at retailer i if the first period demand*

$$\xi_{i1} > \frac{S_{i1} + \omega R}{SL}.$$

For a given first period demand ξ_{i1}, the manufacturer can determine the probability of a penalty event for each service level contract. These penalty probabilities can be influenced by the second replenishment S_{i2}. To simplify the solution of our model, we exclude the probabilities $F\left(-\frac{S_{i1}}{(1-SL)} + \frac{SL}{(1-SL)}\xi_{i1}\right)$ and $F\left(\frac{S_{i1}}{SL} - \xi_{i1}\right)$ from our analysis because they do not depend on the second replenishment S_{i2}. These probabilities are small and only relevant if some retailers have negative inventory levels after the first period (see Proposition 5.1). This leads to a small overestimation of the improvement of penalty probabilities in some cases. However, a stock-out after the first period is not highly probable in practical applications (for example, if the 50/25-heuristic is used). Therefore, we suggest that the error from our approximation is small. Then the total available inventory after the replenishment for retailer i is $R_i = S_{i1} + S_{i2}$ and the corresponding penalty probabilities are

$$\begin{aligned}
&\Pr\left(P_i > 0 \mid \xi_{i1}, S_{i2}\right) \\
&\qquad = \begin{cases} 1 - F\left(\frac{R_i}{SL} - \xi_{i1}\right) & \text{for } \xi_{i1} \leq S_{i1} \\ 1 - F\left(\frac{R_i + S_{i1}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1}\right) & \text{for } \xi_{i1} > S_{i1} \end{cases} \\
&\qquad\quad = \begin{cases} 1 - F\left(\frac{S_{i1} + S_{i2}}{SL} - \xi_{i1}\right) & \text{for } \xi_{i1} \leq S_{i1} \\ 1 - F\left(\frac{2S_{i1} + S_{i2}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1}\right) & \text{for } \xi_{i1} > S_{i1}. \end{cases}
\end{aligned}$$

Next, we define the *contract inventory level* as

$$I_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right) = \begin{cases} \frac{S_{i1} + S_{i2}}{SL} - \xi_{i1} & \text{for } \xi_{i1} \leq S_{i1} \\ \frac{2S_{i1} + S_{i2}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1} & \text{for } \xi_{i1} > S_{i1}. \end{cases} \tag{5.4}$$

The contract inventory level is used as a substitute for the inventory level $S_{i1} - \xi_{i1}$ because only the contract inventory level drives the probability of penalty events. The marginal penalty probability then reduces to

$$\begin{aligned}
&\frac{d\Pr\left(P_i > 0 \mid \xi_{i1}, S_{i2}\right)}{dI_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)} \\
&\qquad = \begin{cases} -\frac{1}{SL} f\left(\frac{S_{i1} + S_{i2}}{SL} - \xi_{i1}\right) & \text{for } \xi_{i1} \leq S_{i1} \\ -\frac{1}{SL} f\left(\frac{2S_{i1} + S_{i2}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1}\right) & \text{for } \xi_{i1} > S_{i1} \end{cases} \\
&\qquad = -f\left(I_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)\right).
\end{aligned}$$

It follows that all contracts lie on the same marginal trade-off curve $-f(\cdot)$. Figure 5.5 shows an example *before* the second replenishment (i.e., with $S_{i2} = 0$) for eight retailers with truncated-normally distributed demand.

Even before the optimization of the second period replenishment, we can identify the retailers where a replenishment will have the greatest

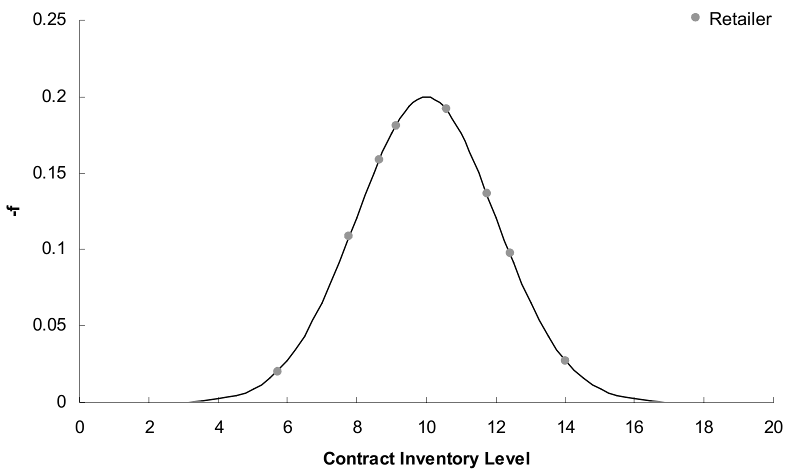

Figure 5.5: Contract Inventory Level

impact on the reduction of expected penalty costs:

Proposition 5.6 *Replenishing the retailers with the smallest absolute differences* $|I_i^c(S_{i1}, 0, \xi_{i1}) - \nu|$ *leads to the highest marginal penalty probability where* ν *denotes the modal value of* f.

The following proposition allows us to characterize the optimal solution:

Proposition 5.7 *In the optimal second period allocation solution, the marginal penalty probability and contract inventory level is equal for all* i *and* j *with* $S_{i2}, S_{j2} > 0$, *i.e.,*

$$\begin{aligned}
\frac{d\Pr\left(P_i > 0 \mid \xi_{i1}, S_{i2}\right)}{dI_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)} &= \frac{d\Pr\left(P_j > 0 \mid \xi_{j1}, S_{j2}\right)}{dI_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)}, \\
I_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right) &= I_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right).
\end{aligned}$$

For the optimization of the second period allocation decision, we have to adapt the steepest-descent approach. The steepest-descent approach would lead to the optimal solution if all contracts would lie on the convex part of the demand c.d.f. Unfortunately, retailers can also lie in the concave part of the demand c.d.f. as Figure 5.5 shows. The following algorithm extends the steepest-descent approach and finds the optimal solution:

Solution-Algorithm:

1. Sort the retailers in ascending order $o : i \to k$ indexed by k with respect to $I_i^c\left(S_{i1}, 0, \xi_{i1}\right)$

2. $M = L = \arg\min_k I_k^c\left(S_{k1}, 0, \xi_{k1}\right) - \nu$ with $I_k^c\left(S_{k1}, 0, \xi_{k1}\right) - \nu \geq 0$ where ν is the modal value of the distribution function f

3. $\tau = 0, \Delta = 0, A = \omega R, S = \emptyset, I^{c*} = I_L^c\left(S_{L1}, 0, \xi_{L1}\right)$

4. While $A \geq 0$ Then

   (a) $M = M + 1; \tau = \tau + 1$

   (b) $U = \min(\tau SL \cdot \left(I_M^c\left(S_{M1}, 0, \xi_{M1}\right) - I_{M-1}^c\left(S_{(M-1)1}, 0, \xi_{(M-1)1}\right)\right),\ A)$

   (c) $A = A - U; I^{c*} = I^{c*} + \frac{U}{\tau SL}$

   (d) $\Delta = \Delta + \tau\left(F\left(I_M^c\left(S_{M1}, \frac{U}{\tau SL}, \xi_{M1}\right)\right) - F\left(I_M^c\left(S_{M1}, 0, \xi_{M1}\right)\right)\right)$

5. $S = S \cup (L, I^{c*}, \Delta)$

6. If $(L > 1)$ Then $M = L = L - 1$, $\tau = 0, \Delta = 0, A = \omega R, I^{c*} = I_L^c\left(S_{L1}, 0, \xi_{L1}\right)$ and go back to Step 4.

7. $I^{c*} = (0,1,0)^T \left\{(L, I^{c*}, \Delta) | (L, I^{c*}, \Delta) \in S, \Delta = \Delta^*\right\}, L^* = I_L^c\left(S_{L1}, 0, \xi_{L1}\right)$ with $L = (1,0,0)^T\{(L, I^{c*}, \Delta)|(L, I^{c*}, \Delta) \in S, \Delta = \Delta^*\}$ where $\Delta^* = \arg\min_\Delta\{\Delta|(L, I^{c*}, \Delta) \in S\}$.

The optimal replenishments S_{i2} can be computed from the contract

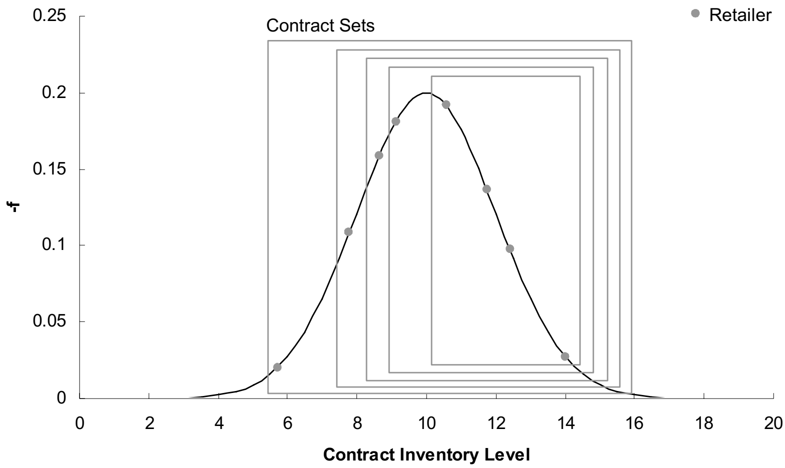

Figure 5.6: Contract Sets for Optimization

inventory levels I^{c*} and L^*. From Equation (5.4), we get

$$S_{i2}^* = \begin{cases} SL\left(I^{c*} + \xi_{i1}\right) - S_{i1} & \text{for} \quad \begin{array}{l} \xi_{i1} \le S_{i1}, \\ I^{c*} \ge I_i^c\left(S_{i1}, 0, \xi_{i1}\right) \ge L^*, \end{array} \\ SL \cdot I^{c*} - 2S_{i1} + \left(1 + SL\right)\xi_{i1} & \text{for} \quad \begin{array}{l} \xi_{i1} > S_{i1}, \\ I^{c*} \ge I_i^c\left(S_{i1}, 0, \xi_{i1}\right) \ge L^*, \end{array} \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 5.8 *The Solution-Algorithm minimizes the manufacturer's expected penalty costs for a given first period demand* ξ_{i1}*.*

The intuition behind the algorithm is to increase the convex contract set by sequentially adding contracts from the concave part. The contract sets have been highlighted in Figure 5.6. Since the solutions for the contract sets are also locally optimal, we get several local optima under which we find the optimal solution.

Figure 5.7 shows the difference between contract balancing and inventory balancing for retailers that are sorted with respect to their in-

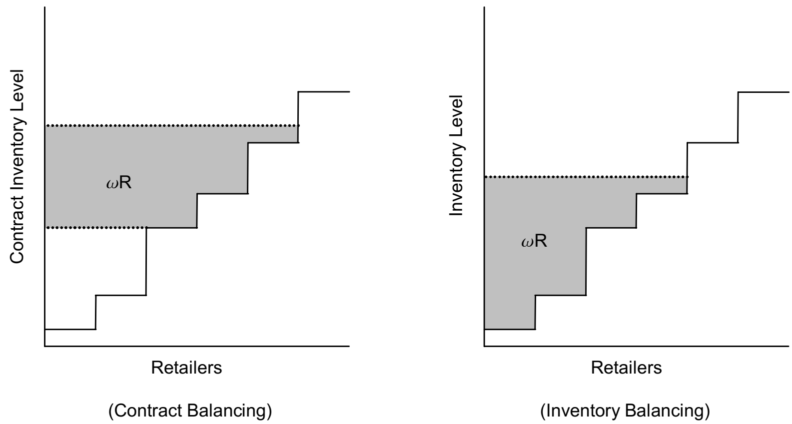

Figure 5.7: Contract Balancing vs. Inventory Balancing

ventory levels. In our contract balancing approach, we replenish retailers that have a contract inventory level above a certain threshold until the available inventory ωR is depleted. A retailer with a lower contract inventory level does not receive a replenishment because a penalty cannot be avoided with a high probability. Therefore, only retailers with a high reduction potential receive a replenishment in the second period. As we could see from Proposition 5.6, the contract inventory levels of these contracts are close the modal value of f. In contrast, the traditional inventory balancing approach starts with the retailer that has the lowest inventory level and fills up the retailers' inventories until the centrally held stock is depleted.

Figure 5.8 shows an example for different R and ω with four retailers and truncated-normal demand. Demand is characterized by $\mu = 10$ and $\sigma = 4$. The contract service level SL is 0.9 and $p = 20, h = 1$. The optimal production quantity $R^* = 94$ for this example can be found numerically. This quantity then leads to an optimal reservation level of $\omega = 0.25$. Setting $R < 94$ leads to higher costs and tends to have

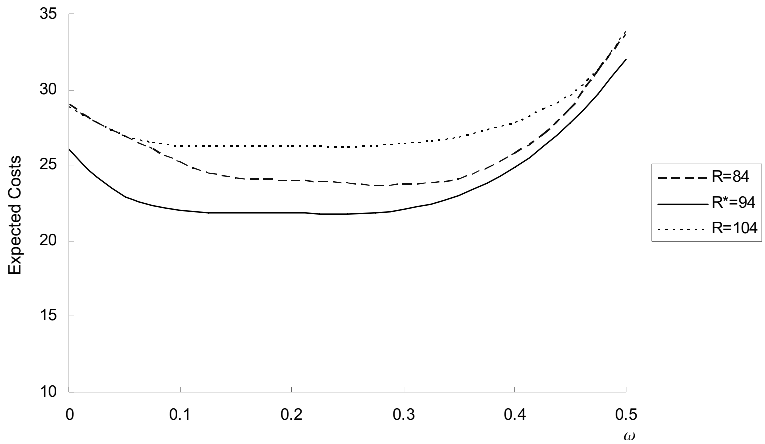

Figure 5.8: Expected Costs for Different ω and R

a slightly higher reservation level ω. Since R is low, the manufacturer ships less inventory in $t = 1$ and instead ships more in the second period to the retailers in order to prevent penalty payments. A $R > 94$ also leads to higher expected costs, but the optimal ω is slightly lower in our example. Since the production size can fulfill nearly all demands in the second period, we only need to care about the first period and ship enough inventory to the retailers. In the example, we can also see that the optimal R^* can be found by optimizing the R for $\omega = 0$, i.e., by assuming a ship-all policy. Therefore, our assumption of using the sum of the individual solution would lead to good results. The example also shows that the cost savings can be significant. Reserving 25% of R^* for later delivery lowers the manufacturer's expected costs by 18%.

We have seen that contract balancing minimizes the manufacturer's expected costs. How the manufacturer's cost-minimizing gaming behavior affects the retailers' performance will be analyzed in the next section.

5.5 Retailer Performance Analysis

In the contract balancing approach, the manufacturer minimizes her expected costs by allocating the inventory optimally among the retailers. Until now we have not analyzed how the optimal policy influences the retailers.

For the retailers, backorder levels or number of stock-outs are an important performance measure because they immediately drive the retailer's costs: A backorder in the first replenishment period leads to backorder costs, for instance, due to order expediting or loss-of-goodwill. A second period backorder can also lead to lost sales if the backorders cannot be filled by transshipped units from other retailers. Therefore, it is important to analyze the expected backorders at the retailers.

The first period expected backorder level at retailer i is

$$B_1\left(S_{i1}\right) = \int_{\xi_{i1}=S_{i1}}^{\infty} \left(\xi_{i1} - S_{i1}\right) f\left(\xi_{i1}\right) d\xi_{i1}$$

and the second period expected backorder level given the first period demand is

$$B_2\left(S_{i2}|\xi_{i1}\right) = \int_{\xi_{i2}=\left(S_{i1}+S_{i2}-\xi_{i1}\right)^+}^{\infty} \left(\xi_{i2} - S_{i1} - S_{i2} + \xi_{i1}\right) f\left(\xi_{i2}\right) d\xi_{i2}.$$

In the first period, all retailers have the same inventory level and the performance does not differ among the retailers. In the second period however, some retailers receive a replenishment and some not. Retailers with a low first period demand do not receive a replenishment because the availability is still guaranteed by the excess stock from the first period. Retailers with a very high first period demand do not receive a replenishment because additional stock would not reduce the probability of a penalty significantly as shown in Proposition 5.8.

In the system optimal solution

$$\begin{aligned}\min_{\bar{S}_1,\bar{S}_2} EC\left(\bar{S}_1,\bar{S}_2,\bar{\xi}_1,\bar{\xi}_2\right) &= b\sum_{i=1}^{N}\left(B_1\left(S_{i1}\right)+B_2\left(S_{i2}\right)\right) \\ &\quad +hE\left(\sum_{i=1}^{N}\left(S_{i1}+S_{i2}-\xi_{i1}-\xi_{i2}\right)\right)^{+} \\ \text{with } |\bar{S}_1|+|\bar{S}_2| &= R\end{aligned} \tag{5.5}$$

a service level contract is only a transfer pricing scheme and can therefore be neglected. The manufacturer would try to minimize the expected backorders in the retailer echelon. Since it is well-known that $B_2\left(S_{i2}|\xi_{i1}\right)$ is decreasing and convex in S_{i2}, the manufacturer would follow an inventory balancing approach where she replenishes the retailers with the lowest inventory levels first. The retailer with the lowest inventory level would benefit the most of a replenishment. We can see that there exists a conflict of interest between the manufacturer's and the retailers' goals because inventory balancing and contract balancing lead to different inventory allocations in $t = 2$.

Figure 5.9 shows the expected backorders for the retailer echelon and the manufacturer's expected cost levels for different contract service levels SL for an example with four retailers and truncated-normal demand with $\mu = 10$, $\sigma = 4$. The manufacturer produces $R^* = 94$, and the supply chain cost factors are $p = 20, h = 1$. As expected, the inventory balancing approach leads to less backorders than contract balancing. The difference is the highest, when the contract service level SL is high because then the manufacturer is most focused on the penalty payments whereas a low contract service level SL leads to more balanced inventories. On the other hand, the contract balancing approach leads to lower expected costs for the manufacturer than the inventory balancing because the manufacturer minimizes expected penalty payments by the contract balancing approach. In our example, the contract balancing approach can reduce expected costs by about 14% compared to the in-

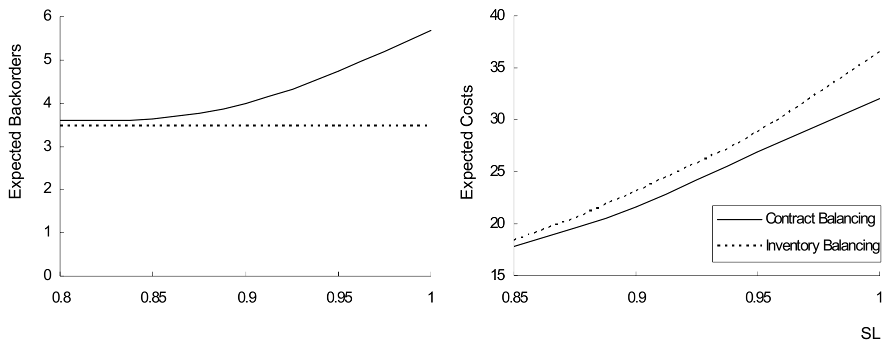

Figure 5.9: Expected Backorders and Expected Costs

ventory balancing approach for a contract service level $SL = 1$. But this contract also leads to the highest differences in the expected backorders.

An important insight for practitioners is that a service level contract alone does not guarantee a good performance for the retailers if the manufacturer serves *multiple* retailers. Our results extend the observation of Choi et al. (2004) who analyzed the one retailer case. They found out that a minimum service level constraint on the manufacturer does not always lead to a minimum retailer service level and that an additional backorder contract parameter is needed to ensure a minimum retailer performance.

5.6 Conclusion

Service level contracts are commonly used in theory and practice for evaluating supplier performance. In many supply contracts, service levels are specified as well as the consequences of not achieving them. In this chapter, we have analyzed a two-echelon supply chain with multiple retailers. We have shown that if the manufacturer manages each retailer individually then a ship-all policy is optimal and the optimal production size can be determined easily. We have also derived results for the case that the retailers are managed jointly. Then the manufacturer can realize risk pooling effects and can reduce her expected costs by reserving inventory for a second replenishment. We have shown that the traditional inventory balancing approach leads to suboptimal results for the manufacturer. The contract balancing approach minimizes the manufacturer's expected costs by taking the contract specification into account. We have seen that the solution is not trivial and we had to rely on an optimal solution algorithm. Finally, we have analyzed the impact of the manufacturer's fulfillment policy on the retailers and have shown that it might not guarantee a good performance for the retailers. The results of our analyses can support decision makers in fulfilling service level based supply contracts. Our model can be modified to a setting with non-identical retailers. Unfortunately, the solution of this problem is more complex than in our model and we had to rely on a pure numerical solution. Therefore, we leave this setting for future research.

In literature, service level based supply contracts in inventory allocation decisions have not been analyzed analytically, although these contracts are very popular in practice. With this chapter, we take a first step towards filling this gap. The results of our analyses allow for a better understanding of the effect of service level based supply contracts on manufacturers and retailers.

5.7 Proofs

Proof of Proposition 5.1. We partition the demand space (ξ_{i1}, ξ_{i2}) with respect to two different cases: (i) a stock-out in period $t = 1$ and (ii) a stock-out at the end of period $t = 2$. That corresponds to the cases $\xi_{i1} > (\leq) S_{i1}$ for (i) and $\xi_{i1} + \xi_{i2} > (\leq)\, R$ for (ii). For all four combinations of (i) and (ii) the reformulation of the penalty condition $(\xi_{i1} - S_{11})^+ + \left(\xi_{i2} - (S_{i1} + S_{i2} - \xi_{i1})^+\right)^+ - (1 - SL)(\xi_{i1} + \xi_{i2}) > 0$ then leads to the following conditions:

For $R > \xi_{i1} > S_{i1}$ and $\xi_{i1} + \xi_{i2} > R$,

$$\begin{aligned} \xi_{i1} - S_{i1} + \xi_{i2} - S_{i1} - S_{i2} - \xi_{i1} - (1 - SL)(\xi_{i1} + \xi_{i2}) &> 0 \\ (1 + SL)\,\xi_{i1} + SL\xi_{i2} - S_{i1} - R &> 0 \\ \frac{R + S_{i1}}{SL} - \left(1 + \frac{1}{SL}\right)\xi_{i1} &< \xi_{i2}. \end{aligned}$$

For $\xi_{i1} > R > S_{i1}$ and $\xi_{i1} + \xi_{i2} > R$,

$$\begin{aligned} \xi_{i1} - S_{i1} + \xi_{i2} - (1 - SL)(\xi_{i1} + \xi_{i2}) &> 0 \\ SL\xi_{i1} + SL\xi_{i2} - S_{i1} &> 0 \\ \xi_{i2} &> \frac{S_{i1}}{SL} - \xi_{i1}. \end{aligned}$$

For $\xi_{i1} \leq S_{i1}$ and $\xi_{i1} + \xi_{i2} > R$,

$$\begin{aligned} \xi_{i2} - S_{i1} + S_{i2} - \xi_{i1} - (1 - SL)(\xi_{i1} + \xi_{i2}) &> 0 \\ \xi_{i2} &> \frac{R}{SL} - \xi_{i1}. \end{aligned}$$

For $\xi_{i1} > S_{i1}$ and $\xi_{i1} + \xi_{i2} \leq R$,

$$\begin{aligned} \xi_{i1} - S_{11} - (1 - SL)(\xi_{i1} + \xi_{i2}) &> 0 \\ \xi_{i2} &< -\frac{S_{i1}}{1 - SL} + \frac{SL}{1 - SL}\xi_{i1}. \end{aligned}$$

For $\xi_{i1} \leq S_{i1}$ and $\xi_{i1} + \xi_{i2} \leq R$ no penalty has to be paid because all demands can be satisfied. ■

Proof of Proposition 5.2. The penalty probability can be computed by adding the demand partition probabilities for which a penalty occurs.

For $\xi_{i1} \leq S_{i1}$ and $\xi_{i1} + \xi_{i2} > R$ (Area 1), the penalty probability equals

$$\int_{\xi_{i1}=0}^{S_{i1}} \left(1 - F\left(\frac{R}{SL} - \xi_{i1}\right)\right) f\left(\xi_{i1}\right) d\xi_{i1}.$$

For $\min\left(S_{i1} + \left(1 - SL\right) R, R\right) \geq \xi_{i1} > S_{i1}$ and $\xi_{i1} + \xi_{i2} > R$ (Area 2), the penalty probability equals

$$\int_{\xi_{i1}=S_{i1}}^{S_{i1}+(1-SL)R} \left(1 - F\left(\frac{R + S_{i1}}{SL} - \left(\frac{1}{SL} + 1\right)\xi_{i1}\right)\right) f\left(\xi_{i1}\right) d\xi_{i1}.$$

For $\xi_{i1} > \min\left(S_{i1} + \left(1 - SL\right) R, R\right)$ a penalty is always due and occurs with probability $1 - F\left(\min\left(S_{i1} + \left(1 - SL\right) R, R\right)\right)$.

For $\min\left(S_{i1} + \left(1 - SL\right) R, R\right) \geq \xi_{i1} > S_{i1}/SL$ and $\xi_{i1} + \xi_{i2} \leq R$ (Area 3), the penalty probability is

$$\int_{\xi_{i1}=S_{i1}/SL}^{S_{i1}+(1-SL)R} F\left(-\frac{S_{i1}}{(1 - SL)} + \frac{SL}{(1 - SL)}\xi_{i1}\right) f\left(\xi_{i1}\right) d\xi_{i1}.$$

For $R \leq \xi_{i1} \leq S_{i1}/SL$ (Area 4) we have to subtract the probability

$$\int_{\xi_{i1}=R}^{S_{i1}/SL} F\left(\frac{S_{i1}}{SL} - \xi_{i1}\right) f\left(\xi_{i1}\right) d\xi_{i1}.$$

Note that ξ_{i1} has to be greater than S_{i1}/SL because demand is non-negative. We also make use of the fact that $F\left(0\right) = 0$ for demand distributions with non-negative support.

Adding the probabilities yields the penalty probability

$$\Pr(P_i > 0) = \int_{\xi_{i1}=0}^{S_{i1}} \left(1 - F\left(\frac{R}{SL} - \xi_{i1}\right)\right) f(\xi_{i1})\, d\xi_{i1}$$
$$+ \int_{\xi_{i1}=S_{i1}}^{\min(S_{i1}+(1-SL)R,R)} \left(1 - F\left(\frac{R+S_{i1}}{SL} - \left(\frac{1}{SL}+1\right)\xi_{i1}\right)\right) f(\xi_{i1})\, d\xi_{i1}$$
$$+ \int_{\xi_{i1}=S_{i1}/SL}^{\min(S_{i1}+(1-SL)R,R)} F\left(-\frac{S_{i1}}{(1-SL)} + \frac{SL}{(1-SL)}\xi_{i1}\right) f(\xi_{i1})\, d\xi_{i1}$$
$$+1 - F(\min(S_{i1} + (1-SL)R, R))$$
$$- \int_{\xi_{i1}=R}^{S_{i1}/SL} F\left(\frac{S_{i1}}{SL} - \xi_{i1}\right) f(\xi_{i1})\, d\xi_{i1}.$$

■

Proof of Proposition 5.3. A ship-all policy with $R = S_{i1}$ is optimal if it leads to $\frac{d\Pr(P_i)}{dS_{i1}} < 0$ for all $S_{i1} \leq R$. Since the holding costs only depend on R, the lowest penalty costs also minimize the expected costs of the manufacturer. To see that $\frac{d\Pr(P_i)}{dS_{i1}} < 0$ consider the demand partitioning from Proposition 5.1. Increasing S_{i1} then leads to a larger non-penalty area. It follows that $\frac{d\Pr(P_i)}{dS_{i1}} < 0$. ■

Proof of Proposition 5.4. We assume a ship-all policy, i.e., $R = S_{i1}$. The no-penalty probability then equals

$$\int_{\xi_{i1}=0}^{R/SL} F\left(\frac{R}{SL} - \xi_{i1}\right) f(\xi_{i1})\, d\xi_{i1}.$$

Then

$$\frac{d\Pr(P_i > 0)}{dR} = -\int_{\xi_{i1}=0}^{R/SL} \frac{f\left(\frac{R}{SL} - \xi_{i1}\right)}{SL} f(\xi_{i1})\, d\xi \leq 0.$$

The term corresponds to the probability that $R < SL(\xi_{i1} + \xi_{i2})$ or $\frac{R}{SL_i} < \xi_{i1} + \xi_{i2}$. Then the probability of a penalty equals

$$\Pr(P_i > 0) = 1 - (F_1 * F_2)\left(\frac{R}{SL}\right).$$

The expected holding costs are convex and increasing in R. The optimal solution then satisfies

$$-\frac{p}{SL}\left(F_1 * F_2\right)'\left(\frac{R}{SL}\right) + h\left(F_1 * F_2\right)(R) \stackrel{!}{=} 0$$

with $(..)'$ being the derivative. By using log-concave properties, it easy to show that the cost function is quasi-convex in R and a unique minimum must exist:

$$\begin{aligned}
-\frac{p}{SL}\left(F_1 * F_2\right)'\left(\frac{R}{SL}\right) + h\left(F_1 * F_2\right)(R) &< 0 \\
h\left(F_1 * F_2\right)(R) &< -\frac{p}{SL}\left(F_1 * F_2\right)'\left(\frac{R}{SL}\right) \\
\frac{\left(F_1 * F_2\right)(R)}{\frac{1}{SL}\left(F_1 * F_2\right)'\left(\frac{R}{SL}\right)} &< -\frac{p}{h} \\
\frac{\left(F_1 * F_2\right)(R)}{\left(F_1 * F_2\right)'(R)}\frac{\left(F_1 * F_2\right)'(R)}{\frac{1}{SL}\left(F_1 * F_2\right)'\left(\frac{R}{SL}\right)} &< -\frac{p}{h} \\
\frac{\left(F_1 * F_2\right)(R)}{\left(F_1 * F_2\right)'(R)}\frac{\left(F_1 * F_2\right)'(R)}{\left(F_1 *_{SL} F_2\right)'(R)} &< -\frac{p}{h} \qquad (5.6)
\end{aligned}$$

with $*_{SL}$ denoting the partial convolution $SL\xi_{i1}+SL\xi_{i2}$. For logconcave frequency functions $f(x)$, the fraction $\frac{(F_1 * F_2)(R)}{(F_1 * F_2)'(R)}$ is non-decreasing in R (Rosling, 2002). For the second term $\frac{(F_1 * F_2)'(R)}{(F_1 *_{SL} F_2)'(R)}$, we find that it is non-decreasing in R since the logconcavity of $f(x)$ implies monotone convolution ratios (Rosling, 2002), i.e., $\frac{f_n(R)}{f_m(R)}$ is non-decreasing in R for $n \geq m$ with $f_n(R) = \left(F_1 * F_2\right)'(R)$ and with $f_m(R) = \left(F_1 *_{SL} F_2\right)'(R)$ being the frequency function of the partial convolution $SL\xi_{i1} + SL\xi_{i2}$.

From $(g(x)h(x))' = g(x)h'(x) + g'(x)h(x)$ the LHS of Equation (5.6) is non-decreasing in R. It follows that there exists only one negative area of $\frac{dEC(R)}{dR}$ in at most one subset of R and the sign of $\frac{dEC(R)}{dR}$ changes at most once from $-$ to $+$ and thus the objective function $EC(R)$ is quasi-convex in R. Then, an optimal R^s is unique. Hence it is sufficient to set the first derivative to zero. ■

Proof of Proposition 5.5. The proof follows by the non-negativity of demands. It follows that a penalty is incurred if the first period demand is outside the maximum no-penalty area if the retailer is replenished with the maximum available stock $S_{i1} + \omega R$, i.e.,

$$\xi_{i1} > \frac{S_{i1} + \omega R}{SL}.$$

■

Proof of Proposition 5.6. By considering the derivative

$$\frac{d \Pr\left(P_i > 0 \mid \xi_{i1}, S_{i2}\right)}{dI_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)}$$

it is easy to see that the marginal value $\frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dS_{i2}}$ of increasing S_{i2} is the greatest around the modal value of the distribution function f. Therefore, the potential for improvement $\sum_i \frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dS_{i2}}$ for all retailers is the greatest around the modal value ν of f. ■

Proof of Proposition 5.7. In an optimal solution, the marginal values of all contracts have to be equal, i.e.,

$$\frac{d \Pr\left(P_i > 0 \mid \xi_{i1}, S_{i2}\right)}{dI_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)} = \frac{d \Pr\left(P_j > 0 \mid \xi_{j1}, S_{j2}\right)}{dI_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)}.$$

Otherwise for $\frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dI_i^c(S_{i1}, S_{i2}, \xi_{i1})} > \frac{d \Pr\left(P_j > 0 \mid \xi_{j1}, S_{j2}\right)}{dI_i^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)}$ a reallocation of inventory with $S_{i2}{+}\varepsilon$ and $S_{j2}{-}\varepsilon$ or with $\frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dI_i^c(S_{i1}, S_{i2}, \xi_{i1})} > \frac{d \Pr\left(P_j > 0 \mid \xi_{j1}, S_{j2}\right)}{dI_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)}$ $S_{i2}{-}\varepsilon$ and $S_{j2}{+}\varepsilon$ would lead to a better solution. Since the demand distribution is unimodal, these marginal improvement values $\frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dI_i^c(S_{i1}, S_{i2}, \xi_{i1})}$ can be reached left of the distribution's modal value ν and right of it likewise. To show that in an optimal solution the $I_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right)$-values have to be equal, we show that a solution with equal marginal values $\frac{d \Pr(P_i > 0 \mid \xi_{i1}, S_{i2})}{dI_i^c(S_{i1}, S_{i2}, \xi_{i1})} = \frac{d \Pr\left(P_j > 0 \mid \xi_{j1}, S_{j2}\right)}{dI_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)}$ but $I_i^c\left(S_{i1}, S_{i2}, \xi_{i1}\right) \neq I_j^c\left(S_{j1}, S_{j2}, \xi_{j1}\right)$ and $S_{i2}, S_{j2} > 0$ can never be optimal. Consider the two contracts i and j with $\frac{d\Pr(P_i>0|\xi_{i1},S_{i2})}{dI_i^c(S_{i1},S_{i2},\xi_{i1})} = \frac{d\Pr\left(P_j>0|\xi_{j1},S_{j2}\right)}{dI_j^c\left(S_{j1},S_{j2},\xi_{j1}\right)}$ where i is left and j is right of the modal value ν. For a non-negative skewed distribution functions as in our model, $S_{i2}+\varepsilon$ and $S_{j2}-\varepsilon$ would lead to a at least as good solution as before because the first order differences $\frac{d\Pr(P_i>0|\xi_{i1},S_{i2})}{dI_i^c(S_{i1},S_{i2},\xi_{i1})}$ in S_{i2} are greater for retailer i. Therefore, in the optimal solution $\frac{d\Pr(P_i>0|\xi_{i1},S_{i2})}{dI_i^c(S_{i1},S_{i2},\xi_{i1})} = \frac{d\Pr\left(P_j>0|\xi_{j1},S_{j2}\right)}{dI_j^c\left(S_{j1},S_{j2},\xi_{j1}\right)}$ and $I_i^c\left(S_{i1},S_{i2},\xi_{i1}\right) = I_j^c\left(S_{j1},S_{j2},\xi_{j1}\right)$. ■

Proof of Proposition 5.8. From Proposition 5.7 we have seen that

$$\begin{aligned}
\frac{d\Pr\left(P_i>0\mid\xi_{i1},S_{i2}\right)}{dI_i^c\left(S_{i1},S_{i2},\xi_{i1}\right)} &= \frac{d\Pr\left(P_j>0\mid\xi_{j1},S_{j2}\right)}{dI_j^c\left(S_{j1},S_{j2},\xi_{j1}\right)}\\
I_i^c\left(S_{i1},S_{i2},\xi_{i1}\right) &= I_j^c\left(S_{j1},S_{j2},\xi_{j1}\right)
\end{aligned}$$

has to hold in the optimal solution and the optimal solution is characterized by I^{c*}. If only contracts from the convex part of F with $I_i^c\left(S_{i1},S_{i2},\xi_{i1}\right)\geq\nu$ with ν being the modal value of f are taken, the solution immediately follows by the steepest-descent argument. However, contracts from the concave area with $I_i^c\left(S_{i1},S_{i2},\xi_{i1}\right)<\nu$ can also lead to an optimal solution. Therefore, increasing the contract solution set S sequentially into the concave contract set leads to a number of locally optimal solution I^{c*} that satisfy the optimality criteria from Proposition 5.7. From Proposition 5.7 also follows that other optimal solution cannot exist. From these local solutions, the optimal solution can be found by taking the maximal value of the expected penalty reduction Δ. The first contract to be replenished has a contract inventory level $I_i^c\left(S_{i1},S_{i2},\xi_{i1}\right)$ of at least L^*. All contracts with $I_i^c\left(S_{i1},0,\xi_{i1}\right)>I^{c*}$ do not receive a replenishment, i.e., $S_{i2}=0$. ■

Chapter 6

Optimal Channel Selection and Efficient Contracts

6.1 Introduction

Across industry, customer heterogeneity and product variety are proliferating rapidly (Fisher, 1997). Volkswagen AG, the leading German automotive company, reported that the number of product segments in the automotive industry quadrupled from 9 in 1985 to 40 in 2005 in parallel with increased diversity in the market (Volkswagen, 2005). Market diversity and its inherent demand variability make it difficult and expensive to match demand with supply. Back in the early 90s, as an example, the personal computer market experienced rapid growth. Companies like IBM Corporation and Apple Computer went through deep shortages and extensive overstocks in their different product lines. As a result, they experienced costly write-offs and their market shares deteriorated due to the mismatch between inventory levels and actual

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_6

customer demand patterns (cf. Fisher et al. (1997) or Schweitzer and Cachon (2000)).

Nevertheless, many companies have started to seek for approaches in order to better match supply with diversified and volatile demand arising in the market. Prominent examples of the approaches that have been implemented by the companies are Hewlett-Packard's mass-customization approach, Campbell Soup's continuous replenishment approach, Sports Obermeyer's quick response approach, and Zara's rapid order fulfillment approach. Through these, companies achieved reduced inventory levels, increased sales, and improved customer satisfaction levels.

In theory as well as in practice, the approach of addressing diversified customers and fulfilling their needs efficiently by a multi-channel distribution strategy has been put forward (Kotler, 1980). In practice, companies appear to pursue a strategy to launch as many distribution channels as possible in order to distribute market risk as far as possible (Anderson et al., 1997). However, this might not be favorable particularly for distribution channels that induce high levels of volatility. High demand volatility deteriorates the accuracy of forecasts which, in turn, gives rise to high supply chain costs. Moreover, setting up a new distribution channel requires considerable capital investment and a long-term adaptability to the market. Likewise, multi-channel distribution systems experience conflicts among alternative channels and difficulties for the determination of pricing and channel specific strategies. Frazier (1999) states that firms need an optimal multi-channel distribution strategy which will enable them to coordinate their distribution efforts in an improved way.

Yet quantitative research on this topic has been limited, providing insufficient guidance to firms to support decision making in a multi-channel context in response to diversified and volatile demand patterns. To fill this gap, this chapter will show that the supply chain earns inferior profits if a manufacturer optimizes her own profits by selecting retail channels. To remedy this situation, we propose different contracts that lead to a supply chain optimal retail channel selection. These contracts differ with respect to the profit allocation among the retailers and the manufacturer. With an equal share contract, which is based on the expected profit margin, a retailer receives a certain share of the supply chain profit whereas a fair share contract also takes the retailer's demand variability into account. Subsequently, we show for our model that an optimal retail channel selection leads to higher supply chain profits than an unselective strategy where all possible channels are served by the manufacturer.

The remainder of this chapter is organized as follows. In Section 6.2, we develop a mathematical model of a two-echelon supply chain that we use in the subsequent analyses. In Section 6.3, we show how a manufacturer optimizes her own profits by choosing an optimal retail channel configuration. In Section 6.4, we derive the supply chain optimal solution and show that the manufacturer's retail channel configuration does not necessarily lead to supply chain coordination. In Section 6.5, we build on the centralized solution to derive different contracts that lead to supply chain coordination. In Section 6.6, we analyze the impact of loss-of-goodwill costs on supply chain performance. In Section 6.7, we combine our approach with traditional contracts that lead to supply chain coordination with respect to stocking levels. In Section 6.8, we conclude. All proofs of this chapter are contained in Section 6.9.

6.2 Supply Chain Model

We consider a two-echelon supply chain as in Figure 6.1 with one manufacturer (indexed by M) and $N \geq 1$ retailers that are indexed by $i = 1, ..., N$. The manufacturer produces a single product and sells it to the retailers. The retailers then sell the product to the end customers. The manufacturer's and the retailers' objective is to maximize expected profits.

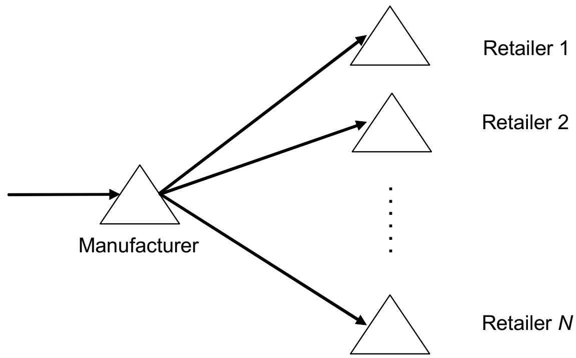

Figure 6.1: Supply Chain Setup

End customer demand is stochastic, continuous, and independent between periods. The demand D_i at retail store i follows the p.d.f. $f_i(\cdot)$ and the c.d.f. $F_i(\cdot)$. We assume that the demand is identically distributed among the retailers and is parameterized by the mean demand μ_i and the standard deviation σ_i for retailer i. Demand is not correlated between retailers. Transshipments between retailers are not considered in our model.

The manufacturer decides which retailers she serves and which she excludes from the supply chain. To describe the specific supply chain configuration, we use the binary vector $x = (x_1, ..., x_N)$ with $x_i \in \{0, 1\}$ where $x_i = 1$ means that the retailer i is served by the manufacturer and $x_i = 0$ that retailer i is excluded from the supply chain. For a supply chain configuration x, the manufacturer faces stochastic retailer orders with mean demand $\mu_{SC}(x) = \sum_{i=1}^{N} x_i \mu_i$ and variance $\sigma_{SC}^2(x) = \sum_{i=1}^{N} x_i \sigma_i^2$. We assume that there is no loss-of-goodwill for not serving a specific retailer. We will show later in this chapter how such a loss-of-goodwill cost could be included into our model.

Inventory at a retail store is managed according to a infinite horizon periodic review base stock policy with a base stock level y_i. Retailer

i receives r_i for every unit sold to the end customer. Unmet demand is backordered and incurs a backorder cost b_i and excess stock results in a physical holding cost h_i. The retailer will order exactly the previous period's demand from the manufacturer in order to replenish his inventory and to fill outstanding backorders.

The manufacturer applies a make-to-order policy, i.e., she exactly produces the retailers' orders. There is no immediate limitation in the quantity the manufacturer can produce every period. However, we assume that the production cost depends on the total demand and demand variability generated by the retailers. In our model the manufacturer's total cost function is

$$C\left(\mu_{SC}, \sigma^2_{SC}\right) = K + c\mu_{SC} + u(\sigma^2_{SC})$$

where the manufacturer faces a fixed cost K per period, a variable production cost c per unit, and a variability cost u that depends on the retailers' aggregate demand variability. The fixed cost K does not have an immediate impact on the optimization problem because it is independent of the supply chain configuration, but it is important for deriving optimal contracts as it will be shown later. The variability cost is increasing in the total variability of the retailers' orders, i.e., we use $u\left(\sigma^2_{SC}\right)$ with $\frac{du(\sigma^2_{SC})}{d\sigma^2_{SC}} \geq 0$. We also assume that $u(\sigma^2_{SC})$ is concave in σ^2_{SC}. Using this assumption, our model can account for two important aspects: Firstly, a higher variability makes production and capacity planning more difficult for the manufacturer and therefore leads to higher costs. For example, planning the manufacturer's production capacity with a newsvendor-type model leads to planning costs that are increasing in the variability of the retailers' orders. We will use this specific setup in our numerical examples. Secondly, the procurement process of the manufacturer becomes more variable if the retailers' orders are more volatile. This then leads to the well-known bullwhip effect at more upstream stages of the supply chain. Thereby, purchasing costs tend to be higher for the manufacturer.

After producing the units, the manufacturer can satisfy the retailers' orders and receives the wholesale price w_i for every unit sold to retailer i. In our supply chain, the manufacturer has sufficient bargaining power to set the wholesale price. Therefore, she is the Stackelberg leader and offers the retailers a take-or-leave contract. A retailer will only accept a contract if he can realize a non-negative expected profit.

Then the expected profit for retailer i is equal to

$$E\Pi_i(y_i) = (r_i - w)\,\mu_i - h_i \int_{\xi=0}^{y_i} (y_i - \xi) f_i(\xi) d\xi - b_i \int_{\xi=y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi.$$

Optimization of $E\Pi_i(y_i)$ with respect to y_i leads to the well-known newsvendor solution $y_i^* = F_i^{-1}\left(\frac{b_i}{b_i+h_i}\right)$. Total expected profits for the entire retailer echelon are

$$E\Pi_R(x, y^*) = \sum_{i=1}^{N} x_i E\Pi_i(y_i^*)$$

where y^* denotes the optimal base stock vector $y^* = \left(y_{1,}^*..., y_N^*\right)$ and x the retail channel configuration $x = (x_1, ..., x_N)$.

The expected profit of the manufacturer is

$$E\Pi_M(x) = \left(\sum_{i=1}^{N} x_i \mu_i w_i\right) - C\left(\mu_{SC}\left(x\right), \sigma_{SC}^2\left(x\right)\right).$$

for a supply chain configuration x, and total expected supply chain profits equals $E\Pi_{SC}(x, y) = E\Pi_M(x) + E\Pi_R(x, y)$. In the following, we characterize the manufacturer's optimal solution if she maximizes her individual expected profit.

6.3 Manufacturer Solution

The manufacturer chooses the optimal wholesale price w^* and retail channel configuration x^* to maximize her expected profits. In this setup, the wholesale price is identical for all retailers. By choosing retail channels, the manufacturer has to trade off between the additional expected revenue from serving a retailer, i.e., $w\mu_i$, and the cost effect of a different manufacturer production cost $C\left(\mu_{SC}\left(x\right), \sigma_{SC}^2\left(x\right)\right)$. To set an optimal wholesale price w^*, the manufacturer also has to take the retailers' participation constraint into account, i.e., a non-negative expected profit.

The optimization of x^* and w^* cannot be done independently. Increasing w^* for a given x^* could eliminate some retail channels because they might have negative expected profits. It follows that x^* cannot be optimal any more. On the other hand, a higher wholesale price would make some profitable retail channels more attractive for the manufacturer because the higher wholesale price could outweigh higher production costs. Then, the optimal configuration x^* would also include these channels.

To find the optimal solution (x^*, w^*) we will first show how a manufacturer determines the optimal configuration x^* for a given wholesale price w. Then we will show how the manufacturer will choose the optimal contract (x^*, w^*) that maximizes her expected profits.

For a given wholesale price w, the manufacturer's objective is

$$\max_x E\Pi_M(x) = \left(\sum_{i=1}^{N} x_i \mu_i w\right) - C\left(\mu_{SC}\left(x\right), \sigma_{SC}^2\left(x\right)\right).$$

To compute the optimal x^*, the retailers have to be ordered with respect to their revenue contribution in relation to their order variability. Since the manufacturer charges the retailers an identical wholesale price w, it is sufficient to order the retailers with respect to their mean-to-variance ratio, i.e.,

$$\frac{\mu_1}{\sigma_1^2} \geq \frac{\mu_2}{\sigma_2^2} \geq \ldots \geq \frac{\mu_{p_m}}{\sigma_{p_m}^2} \geq \ldots \geq \frac{\mu_N}{\sigma_N^2}$$

where $p_m \in \{1, \ldots, N\}$ indexes the retailers in the new order. For this ordering, we can make an important observation for the optimal solution x^*: If retailer l is included in the supply chain, all retailers that have a mean-to-variance ration that is higher than l are also included. The following proposition summarizes this observation:

Proposition 6.1 *In the optimal solution x^*, if $x_k^* = 1$ also $x_l^* = 1$ for $l \in \{1, 2, \ldots, k-1\}$.*

We can compute the optimal solution by increasing the retail channel set sequentially and choosing the retailer p that maximizes expected manufacturer profits. Then all other retailers that have a higher mean-to-variance ratio than retailer p are also served by the manufacturer. The following proposition states the optimality condition.

Proposition 6.2 *For the ordered retailer set, the optimal supply chain configuration x^* for a given wholesale price w satisfies*

$$\left(\sum_{l \le p^*} x_l w \mu_l\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \ge$$
$$\left(\sum_{l \le p} x_l w\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \;\; \forall p \backslash p^*$$

where $x_i(l) = 1$ if $l \le p$ and $x_i(l) = 0$ otherwise.

Now we show how to optimize (x^*, w^*) simultaneously. First note that for each retailer there exists a maximal wholesale price $\hat{w}_i$ that gives him a non-negative expected profit, i.e.,

$$\begin{aligned} E\Pi_i(y_i^*) &= (r_i - \hat{w}_i)\,\mu_i - h_i \int_{\xi=0}^{y_i^*} (y_i^* - \xi) f_i(\xi) d\xi \\ &\quad - b_i \int_{\xi=y_i^*}^{\infty} (\xi - y_i^*) f_i(\xi) d\xi \\ &= 0 \end{aligned}$$

Increasing the wholesale price beyond $\hat{w}_i$ would exclude retailer i because he would not accept the contract. In consequence, the set of retailers the manufacturer can choose from would be reduced. Ordering the retailers with respect to their maximum $\hat{w}_i$ gives us retailer sets that are decreasing in the wholesale price w. The following proposition states the optimal solution.

Proposition 6.3 *The optimal wholesale price w^* has to be from the wholesale price set $(\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_N)$. The optimal contract (x^*, w^*) satisfies*

$$E\Pi_M^M = E\Pi_M(x^*\left(\hat{w}_i^*\right), \hat{w}_i^*) \geq E\Pi_M(x\left(\hat{w}_i\right), \hat{w}_i)\ \forall \hat{w}_i.$$

The expected supply chain profit equals

$$E\Pi_{SC}^M(x^*, y^*, w^*) = E\Pi_M^M(x^*, w^*) + E\Pi_R(x^*, y^*, w^*)$$

In the next section we show how the supply chain optimal solution can be determined.

6.4 Centralized Solution

In the centralized solution, we consider an integrated supply chain. That means that we optimize the total supply chain profits. As before, our decision is based on two aspects: Adding a retailer (i) generates more revenue by enlarging the served market, and (ii) increases costs by higher production costs and higher demand variability. We do not have to specify a wholesale price in the centralized case and can add retailers as long as they contribute positively to the expected supply chain profits.

The centralized expected profit function is

$$
\begin{aligned}
&E\Pi_{SC}(x,y) = \\
&\left(\sum_i x_i \left(r_i\mu_i - h_i \int_{\xi=0}^{y_i} (y_i-\xi) f_i(\xi) d\xi - b_i \int_{\xi=y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi\right)\right) \\
&-C\left(\mu_{SC}(x), \sigma_{SC}^2(x)\right) \\
&= \left(\sum_i x_i \left(r_i\mu_i - d_i(y_i)\right)\right) - C\left(\mu_{SC}(x), \sigma_{SC}^2(x)\right)
\end{aligned}
$$

where

$$
d_i(y_i) = h_i \int_{\xi=0}^{y_i} (y_i-\xi) f_i(\xi) d\xi + b_i \int_{\xi=y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi
$$

denotes the expected distribution costs, i.e., the sum of inventory holding and backorder costs.

Including a retailer has three distinct effects: Firstly, it increases revenue. Secondly, distribution costs are incurred at the retailer. Thirdly, the production costs at the manufacturer are increasing. To solve this optimization problem we order the retailers with respect to their revenue contribution compared to their variance, i.e.,

$$
\begin{aligned}
\frac{\mu_1(r_1-c) - d_1(y_1)}{\sigma_1^2} &\geq \frac{\mu_2(r_2-c) - d_2(y_2)}{\sigma_2^2} \geq \cdots \\
&\geq \frac{\mu_N(r_N-c) - d_N(y_N)}{\sigma_N^2}
\end{aligned}
$$

where $p_m \in \{1,\ldots,N\}$ indexes the retailers in the new order. Again, we can make an important observation for the optimal solution x^*: If retailer l is included in the supply chain, all retailers that have a revenue-contribution-to-variance ratio that is higher than l are also included. The following proposition summarizes this observation.

Proposition 6.4 *In the optimal solution x^*, if $x_k^* = 1$ also $x_l^* = 1$ for $l \in \{1, 2, \ldots, k-1\}$.*

We can compute the optimal solution by increasing the retail channel set sequentially and choosing the retailer p that maximizes expected supply chain profits. Then all other retailers that have a higher revenue-contribution-per-variance ratio than retailer p are served in the coordinated supply chain. The following proposition states the optimality condition.

Proposition 6.5 *For the ordered retailer set, the optimal coordinating supply chain configuration x^* satisfies*

$$\begin{aligned} E\Pi_{SC}^{CE} &= \left(\sum_{l \le p^*} x_l \left(r_l \mu_l - d_l\left(y_l\right)\right)\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \ge \\ & \left(\sum_{l \le p} x_l \left(r_l \mu_l - d_l\left(y_l\right)\right)\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \;\; \forall p \backslash p^* \end{aligned}$$

where $x_i(l) = 1$ if $l \le p$ and $x_i(l) = 0$ otherwise.

The optimal solution in the centralized case is different from the manufacturer solution because the manufacturer takes only her profit from the wholesale price contract into account whereas we consider the total supply chain profit in the centralized case. This effect resembles the well-known double marginalization effect that is frequently discussed in the contracting literature.

In the following we analyze the magnitude of the double marginalization effect in the context of choosing retail channels. We compute the differences between the manufacturer solution and the centralized solution as the benchmark. The efficiency of the supply chain is defined as the ratio of the manufacturer solution and the coordinated solution, i.e., $E\Pi_{SC}^{M}/E\Pi_{SC}^{CE}$.

In our numerical example, the manufacturer can serve up to four retailers. Demand at the retailers is truncated normal distributed with different mean and standard deviations. The unit revenue r_i differs between retailers and holding and backorder costs at the retailers are identical with $h_i = 2$ and $b_i = 3.25$. Under the assumption of truncated normally distributed demand this approximately leads to distribution costs of $d_i(y_i) = \sigma_i$, i.e., they solely depend on the standard deviation the retailers faces. Since we want to focus on the effect of demand variability on supply chain performance, we assume that the manufacturer only has to plan the production capacity in advance and uses a newsvendor trade-off. Thus, the production cost function depends on the standard deviation of retailer orders the manufacturer sees. In our case we use the unit production cost function $C\left(\mu_{SC}(x), \sigma_{SC}^2(x)\right) = u\left(\sigma_{SC}^2(x)\right) = 150\sqrt{(\sigma_{SC}^2(x))}$.

Table 6.1 shows that in many cases the manufacturer would choose the supply chain optimal solution. Then the total supply chain profits equal the centralized solution and the total profits are divided among manufacturer and retailers. However, the table also shows that in some examples, the manufacturer would choose another configuration than in the centralized case. Then the supply chain efficiency, i.e., the ratio $E\Pi_{SC}^{M}/E\Pi_{SC}^{CE}$, can be significantly lower, even below 25%.

In conclusion, a wholesale price contract that the manufacturer derives by optimizing her own expected profits has two main drawbacks: Firstly, the wholesale price treats every retailer similarly. That leads to situations where a highly profitable retail channel is hurt because another channel with high variability significantly hurts supply chain performance but does not bring in additional revenue. Secondly, retail channels are excluded arbitrarily because they have negative profits due to the wholesale price setting. However, in the centralized solution they would have been profitable.

To maximize supply chain profits, the manufacturer should take the end customer prices and the retailers' demand characteristics into account when she makes a decision for or against a retailer. Next we discuss how we can incentivize the manufacturer to choose the supply chain optimal retailer configuration and how to increase her own expected profits at the same time.

r_1	r_2	r_3	r_4	μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4	$E\Pi_{SC}^M$	$E\Pi_{SC}^{CE}$	$E\Pi_{SC}^M/E\Pi_{SC}^{CE}$
22	21	19	18	21	23	27	29	2	2.5	3.5	4	460.27	1037.27	44.37%
22	21	19	18	23	24	26	27	2	2.5	3.5	4	1047.27	1047.27	100.00%
22	21	19	18	25	25	25	25	2	2.5	3.5	4	1057.27	1057.27	100.00%
22	21	19	18	27	26	24	23	2	2.5	3.5	4	1067.27	1067.27	100.00%
22	21	19	18	29	27	23	21	2	2.5	3.5	4	922.49	1077.27	85.63%
22	21	19	18	21	23	27	29	3	3	3	3	1068.00	1068.00	100.00%
22	21	19	18	23	24	26	27	3	3	3	3	1078.00	1078.00	100.00%
22	21	19	18	25	25	25	25	3	3	3	3	1088.00	1088.00	100.00%
22	21	19	18	27	26	24	23	3	3	3	3	675.58	1098.00	61.53%
22	21	19	18	29	27	23	21	3	3	3	3	562.60	1108.00	50.78%
22	21	19	18	21	23	27	29	4	3.5	2.5	2	1037.27	1037.27	100.00%
22	21	19	18	23	24	26	27	4	3.5	2.5	2	1047.27	1047.27	100.00%
22	21	19	18	25	25	25	25	4	3.5	2.5	2	1057.27	1057.27	100.00%
22	21	19	18	27	26	24	23	4	3.5	2.5	2	1067.27	1067.27	100.00%
22	21	19	18	29	27	23	21	4	3.5	2.5	2	1077.27	1077.27	100.00%
24	22	18	16	21	23	27	29	2	2.5	3.5	4	1017.27	1017.27	100.00%
24	22	18	16	23	24	26	27	2	2.5	3.5	4	828.49	1037.27	79.87%
24	22	18	16	25	25	25	25	2	2.5	3.5	4	1057.27	1057.27	100.00%
24	22	18	16	27	26	24	23	2	2.5	3.5	4	735.27	1077.27	68.25%
24	22	18	16	29	27	23	21	2	2.5	3.5	4	805.27	1097.27	73.39%
24	22	18	16	21	23	27	29	3	3	3	3	1048.00	1048.00	100.00%
24	22	18	16	23	24	26	27	3	3	3	3	759.58	1068.00	71.12%
24	22	18	16	25	25	25	25	3	3	3	3	1088.00	1088.00	100.00%
24	22	18	16	27	26	24	23	3	3	3	3	577.60	1108.00	52.13%
24	22	18	16	29	27	23	21	3	3	3	3	647.60	1128.00	57.41%
24	22	18	16	21	23	27	29	4	3.5	2.5	2	1017.27	1017.27	100.00%
24	22	18	16	23	24	26	27	4	3.5	2.5	2	1037.27	1037.27	100.00%
24	22	18	16	25	25	25	25	4	3.5	2.5	2	1057.27	1057.27	100.00%
24	22	18	16	27	26	24	23	4	3.5	2.5	2	1077.27	1077.27	100.00%
24	22	18	16	29	27	23	21	4	3.5	2.5	2	812.95	1097.27	74.09%
27	24	17	13	21	23	27	29	2	2.5	3.5	4	622.77	987.27	63.08%
27	24	17	13	23	24	26	27	2	2.5	3.5	4	700.27	1022.27	68.50%
27	24	17	13	25	25	25	25	2	2.5	3.5	4	777.77	1057.27	73.56%
27	24	17	13	27	26	24	23	2	2.5	3.5	4	855.27	1092.27	78.30%
27	24	17	13	29	27	23	21	2	2.5	3.5	4	932.77	1127.27	82.75%
27	24	17	13	21	23	27	29	3	3	3	3	465.10	1018.00	45.69%
27	24	17	13	23	24	26	27	3	3	3	3	542.60	1053.00	51.53%
27	24	17	13	25	25	25	25	3	3	3	3	620.10	1088.00	56.99%
27	24	17	13	27	26	24	23	3	3	3	3	697.60	1123.00	62.12%
27	24	17	13	29	27	23	21	3	3	3	3	775.10	1158.00	66.93%
27	24	17	13	21	23	27	29	4	3.5	2.5	2	987.27	987.27	100.00%
27	24	17	13	23	24	26	27	4	3.5	2.5	2	335.95	1022.27	32.86%
27	24	17	13	25	25	25	25	4	3.5	2.5	2	252.77	1057.27	23.91%
27	24	17	13	27	26	24	23	4	3.5	2.5	2	535.24	1092.27	49.00%
27	24	17	13	29	27	23	21	4	3.5	2.5	2	612.74	1127.27	54.36%

Table 6.1: Supply Chain Efficiency

6.5 Supply Chain Optimal Contracts

In this section we discuss how optimal contracts can be designed such that the supply chain is optimized. We have seen that the manufacturer's retail channel optimization leads to another retail channel selection than in the centralized case. Next, we derive a contracting scheme that leads to the supply chain optimal solution. Firstly, we show how the supply chain can be coordinated with a wholesale price contract. Secondly, we derive a contract that distributes the expected supply chain profits fairly among the retailers.

6.5.1 Wholesale Price Contract

The manufacturer will only choose the supply chain optimal solution if she can optimize her own expected profits. Therefore, a wholesale price contract has to ensure that this optimal configuration is chosen.

Proposition 6.6 *The supply chain is coordinated for a wholesale price*

$$w_i^{simple} = \begin{cases} r_i - \frac{d_i(y_i^*)}{\mu_i} & \text{if } x_i^* = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The manufacturer will choose the centralized solution x^*. Then the retailers have a zero profit and the supplier takes the whole supply chain profit.

We can see that the wholesale price contracts behave similarly to the wholesale price contract in the traditional contracting literature. There the retailer also realizes the entire supply chain profits and the manufacturer only realizes a zero profit. Similarly, in our contract, no profit is left to the retailer and all gains from supply chain coordination go the manufacturer. We can derive an incentive compatible scheme that can arbitrarily distribute the gains from coordination to the retailers such that the optimal distribution channel selection is preserved. There are two different types of contract that allocate a certain fraction to the retailer echelon that we will introduce in the following. They differ in the way the individual retailers are treated, either equally or fairly with respect to their expected revenue and demand variability.

6.5.2 Equal Share Contract

To model different profit allocations in the supply chain, we use δ to denote the share of expected supply chain profits that go to the manufacturer.

Proposition 6.7 *The wholesale price contract*

$$w_i^{equal}(\delta) = \begin{cases} \delta\left(r_i - \frac{d_i(y_i)}{\mu_i}\right) + (1-\delta)\,\frac{C\left(\mu_{SC}(x^*),\sigma_{SC}^2(x^*)\right)}{\mu_{SC}(x^*)} & \text{if } x_i^* = 1, \\ 0 & \text{otherwise} \end{cases}$$

coordinates the supply chain and gives the manufacturer the expected profit $E\Pi_M^{equal} = \delta E\Pi_{SC}^{CE}$ *and the retailer echelon the expected profit* $E\Pi_R^{equal} = (1-\delta)\,E\Pi_{SC}^{CE}$.

We see that the manufacturer can give the retailers a discount but charges them for some part of the variability costs. For $\delta = 1$ we get the traditional wholesale price contract that allocates all profits to the supplier. For $\delta = 0$, all profits are allocated to the retailer echelon. The retailers then only bear the manufacturer's unit production costs. In some cases, the participation constraint of the retailers can restrict the range of feasible values of δ because some retailers would earn a negative profit if a high wholesale price would be chosen by the manufacturer. Clearly, this should not happen because also these retailers are important for the total supply chain.

Channels with high variability might therefore profit from sharing the risk with all other channels. It might be interesting to modify the wholesale price scheme such that it takes the retailer's variability and their revenue contribution to the expected supply chain profit into account. We present this approach next.

6.5.3 Fair Share Contract

In this section we develop an incentive scheme where every retailer receives a fair share of the total supply chain profits. In the preceding wholesale price model we have seen that all retail channels influence the manufacturer's production costs. Then retailers with high revenues that only contribute little variability to the supply chain may *suffer* from other retailers that induce a high variability into the supply chain. Therefore we will fairly allocate the expected retail channel profits $(1-\delta)\, E\Pi_{SC}^{CE}$ among the retailers.

The basic idea of the fair share contract is as follows: We try to find a fair profit reallocation among the retailers by choosing an individual wholesale price for each retailer. This can be also interpreted as a discount for *good* demand sources. The manufacturer will finally effectuate the reallocation by choosing the optimal wholesale prices. It is important that we achieve the fair reallocation and do not change the manufacturer's optimal retailer choice, i.e., with the new wholesale prices the manufacturer still has to choose the optimal supply chain solution. We achieve this solution by reallocating the retailer share efficiently such that the expected manufacturer profit stays constant.

In cooperative game theory we find similar problems in coalition games. There, the game setup is as follows: a coalition of players cooperates, and realizes a certain overall value from that coalition (Meyerson, 2000). This overall value has to be fairly distributed among the players with respect to their individual contribution to the overall value. Thereby, the distribution of value depends on the individual importance of each player to the overall performance of the coalition. Clearly, this setting resembles our profit allocation problem.

For our analysis, we use the Shapley value that assigns individual utilities π to the participants in a coalition with n players. It is based on a value function $v(C)$ that depends on the coalition C (Meyerson, 2000). To determine the contribution of each player in the optimal supply chain, the Shapley value equals

$$\pi_i = \sum_{i \notin V \subseteq N} \frac{|V|!(n-|V|-1)!}{n!} \left(v\left(V \cup \{i\}\right) - v\left(V\right) \right)$$

where the summation includes all subsets V of all possible coalition N not containing player i. This allocation can be explained by noting that including the player i into the coalition contributes the value $v(V \cup \{i\}) - v(V)$ to the coalition, i.e., a fair compensation. Averaging over all possible permutations in which the coalition can be formed yields the players individual contribution to the coalition.

The Shapley value has the property of being a *fair* distribution in the sense that it is the only distribution with the following properties that are highly relevant for our supply chain problem: (i) $\pi_i \geq v(\{i\})$ for every player i in N, i.e., every player receives at least as much as he would have received without collaboration, and (ii) $\sum_{i \in N} \pi_i = v(N)$, i.e., all profit is distributed among the retailers (Meyerson, 2000).

For our model the following proposition states the supply chain optimal wholesale price contract that allocates expected profits fairly among the retailers.

Proposition 6.8 *The wholesale price contract with*

$$w_i^{fair}\left(\delta\right) = \begin{cases} r_i - \frac{\sum_{i \notin V \subseteq N} \frac{|V|!(n-|V|-1)!}{n!} \left(v(V \cup \{i\}) - v(V)\right) + d_i(y_i^*)}{\mu_i} & \textit{if } x_i^* = 1, \\ 0 & \textit{otherwise} \end{cases}$$

where

$$\begin{aligned} v\left(V\right) &= \left(1-\delta\right) E\Pi_{SC}(x = V, y^*) \\ &= \left(1-\delta\right) \left(\left(\sum_i x_i^* \left(r_i \mu_i - d_i\left(y_i^*\right)\right) \right) - C\left(\mu_{SC}\left(x^*\right), \sigma_{SC}^2\left(x^*\right)\right) \right) \end{aligned}$$

δ	w_1^{equal}	w_2^{equal}	w_3^{equal}	w_4^{equal}
0.00	9.31	9.31	9.31	9.31
0.10	10.77	10.57	10.16	9.96
0.20	12.23	11.82	11.02	10.62
0.30	13.69	13.08	11.87	11.27
0.40	15.15	14.34	12.73	11.93
0.50	16.61	15.60	13.59	12.58
0.60	18.07	16.86	14.44	13.23
0.70	19.53	18.12	15.30	13.89
0.80	20.99	19.38	16.15	14.54
0.90	22.45	20.64	17.01	15.20
1.00	23.91	21.90	17.87	15.85

Table 6.2: Coordinating Equal Share Contracts

coordinates the supply chain and leads to a fair allocation of expected profits. The manufacturer receives $E\Pi_M^{fair} = \delta E\Pi_{SC}^{CE}$ *and the retailer echelon* $E\Pi_R^{fair} = (1-\delta)\, E\Pi_{SC}^{CE}$.

The following example illustrates how this fair allocation can be achieved. We use the same data as before. The unit revenue is $r = \{24, 22, 18, 16\}$, the mean retailer demands are $\mu = \{23, 24, 26, 27\}$, and the standard deviations are $\sigma = \{2, 2.5, 3.5, 4\}$. In Table 6.1 we have seen that the supply chain profits in the manufacturer solution are 828.49 with $x = \{1, 1, 1, 0\}$. In the centralized case the optimal solution is $x = \{1, 1, 1, 1\}$ with an expected supply chain profit of 1037.27.

For this example, Table 6.2 shows the wholesale prices of a coordinating equal share contract. For $\delta = 0$ all profits go to the retailers. Then the wholesale price equals the production costs of the manufacturer. For $\delta = 1$ the manufacturer receives all supply chain profits. Then the wholesale price extracts all retailer profits as can be seen in Table 6.3.

From Table 6.3 we can also see that the manufacturer is able to increase her expected profits compared to the manufacturer solution of 828.49. However, we can also identify the shortcomings of an equal share contract. For example, Retailer 1 only receives 32.38% of retailer profits

δ	$E\Pi_{SC}^{CE}$	$E\Pi_M$	$E\Pi_1^{equal}$	$E\Pi_2^{equal}$	$E\Pi_3^{equal}$	$E\Pi_4^{equal}$	$E\Pi_R^{equal}$
0.00	1037.27	0.00	335.93	302.13	222.51	176.70	1037.27
0.10	1037.27	103.73	302.34	271.91	200.26	159.03	933.55
0.20	1037.27	207.45	268.75	241.70	178.01	141.36	829.82
0.30	1037.27	311.18	235.15	211.49	155.76	123.69	726.09
0.40	1037.27	414.91	201.56	181.28	133.51	106.02	622.36
0.50	1037.27	518.64	167.97	151.06	111.26	88.35	518.64
0.60	1037.27	622.36	134.37	120.85	89.00	70.68	414.91
0.70	1037.27	726.09	100.78	90.64	66.75	53.01	311.18
0.80	1037.27	829.82	67.19	60.43	44.50	35.34	207.45
0.90	1037.27	933.55	33.59	30.21	22.25	17.67	103.73
1.00	1037.27	1037.27	0.00	0.00	0.00	0.00	0.00

Table 6.3: Expected Profits of Coordinating Equal Share Contracts

although he contributes 27.88% of customer revenues with a very low variability, whereas Retailer 4 still receives 17.04% of the total retailers' profits although he only contributes 21.82% of revenues with a high variability. Coordinating fair share contracts address this issue and allocate the retailer profits fairly by also taking the demand variability into account. In Table 6.4 the wholesale prices of fair share contracts are shown.

Then, for example, Retailer 1 pays a lower wholesale price and consequently receives a higher share of expected supply chain profits. Under this fair share contract, Retailer 1 receives 41.08% of the total expected retailer profit. On the other hand, Table 6.5 shows that Retailer 4 now pays a higher wholesale price and thus has a lower expected profit due to his higher demand variability.

To summarize, both contract types can coordinate the supply chain. The equal share contract however, only assigns the wholesale prices with respect to the revenue potential to the supply chain, but does not take the variability-induced costs for the total supply chain into account. The fair share contract remedies this shortcoming by assigning wholesale prices by basing them on the profit potential of a specific retailer.

In many practical applications the manufacturer cannot select the

δ	w_1^{fair}	w_2^{fair}	w_3^{fair}	w_4^{fair}
0.00	5.39	7.15	10.96	12.98
0.10	7.24	8.62	11.65	13.26
0.20	9.09	10.10	12.34	13.55
0.30	10.94	11.57	13.03	13.84
0.40	12.80	13.05	13.72	14.13
0.50	14.65	14.52	14.41	14.41
0.60	16.50	16.00	15.10	14.70
0.70	18.36	17.47	15.79	14.99
0.80	20.21	18.95	16.48	15.28
0.90	22.06	20.42	17.17	15.56
1.00	23.91	21.90	17.87	15.85

Table 6.4: Coordinating Fair Share Contracts

δ	$E\Pi_{SC}^{CE}$	$E\Pi_M$	$E\Pi_1^{fair}$	$E\Pi_2^{fair}$	$E\Pi_3^{fair}$	$E\Pi_4^{fair}$	$E\Pi_R^{fair}$
0.00	1037.27	0.00	426.11	354.00	179.54	77.62	1037.27
0.10	1037.27	103.73	383.50	318.60	161.59	69.86	933.55
0.20	1037.27	207.45	340.89	283.20	143.63	62.09	829.82
0.30	1037.27	311.18	298.28	247.80	125.68	54.33	726.09
0.40	1037.27	414.91	255.67	212.40	107.72	46.57	622.36
0.50	1037.27	518.64	213.06	177.00	89.77	38.81	518.64
0.60	1037.27	622.36	170.45	141.60	71.82	31.05	414.91
0.70	1037.27	726.09	127.83	106.20	53.86	23.29	311.18
0.80	1037.27	829.82	85.22	70.80	35.91	15.52	207.45
0.90	1037.27	933.55	42.61	35.40	17.95	7.76	103.73
1.00	1037.27	1037.27	0.00	0.00	0.00	0.00	0.00

Table 6.5: Expected Profits of Coordinating Fair Share Contracts

supply chain optimal solution because eliminating some channels might incur significant losses-of-goodwill. We discuss this case next.

6.6 Channel Opportunity Costs

In practice, the decision for or against serving some retailers is constrained by other important factors. Firstly, serving a large number of retail channels might increase profits due to risk pooling effects. Secondly, manufacturers fear that not serving certain channels might incur a loss-of-goodwill from the customer or the retailer side if they have sufficient negotiating power. To address this issue, we next compare our selective model to an unselective supply chain, and refine our model in order to include a loss-of-goodwill cost for certain retailers into our model.

First, we compare our centralized solution x^* with an unselective solution $x^u = \{1, \ldots, 1\}$, i.e., when the manufacturer serves all retailers. In Table 6.6 we have used the same data set as before. Table 6.6 shows that the difference between our selective and an unselective supply chain strategy has a significant impact on expected supply chain profits. Especially in supply chains where retailers have different unit revenues r_i, we can observe significant differences between selective and unselective strategies. On the other hand, in supply chains with identical retailers both strategies lead to the same results.

Various authors argue that increasing the number of retailers increases supply chain profits due to risk pooling effects (Anderson et al., 1997). However, traditional risk pooling compares a situation before and after integration, but where all retailers are present in both situation. In contrast, in our selection model, we can select certain retailers in order to maximize supply chain profits.

The difference $E\Pi_{SC}(x^*) - E\Pi_{SC}(x^u)$ can be understood as the opportunity cost of including all retailers, i.e., if the loss of goodwill is greater than the difference, we would include all retailers. But what

r_1	r_2	r_3	r_4	μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4	$E\Pi_{SC}(x^u)$	$E\Pi_{SC}(x^*)$	Difference
20	20	20	20	23	24	26	27	1	2	4	5	292.42	377.98	29.26%
20	20	20	20	25	25	25	25	1	2	4	5	292.42	437.98	49.78%
20	20	20	20	27	26	24	23	1	2	4	5	292.42	497.98	70.30%
20	20	20	20	23	24	26	27	3	3	3	3	488.00	488.00	0.00%
20	20	20	20	25	25	25	25	3	3	3	3	488.00	488.00	0.00%
20	20	20	20	27	26	24	23	3	3	3	3	488.00	488.00	0.00%
20	20	20	20	23	24	26	27	5	4	2	1	292.42	497.98	70.30%
20	20	20	20	25	25	25	25	5	4	2	1	292.42	437.98	49.78%
20	20	20	20	27	26	24	23	5	4	2	1	292.42	377.98	29.26%
22	21	19	18	23	24	26	27	1	2	4	5	282.42	447.98	58.62%
22	21	19	18	25	25	25	25	1	2	4	5	292.42	512.98	75.43%
22	21	19	18	27	26	24	23	1	2	4	5	302.42	577.98	91.12%
22	21	19	18	23	24	26	27	3	3	3	3	478.00	478.00	0.00%
22	21	19	18	25	25	25	25	3	3	3	3	488.00	488.00	0.00%
22	21	19	18	27	26	24	23	3	3	3	3	498.00	498.00	0.00%
22	21	19	18	23	24	26	27	5	4	2	1	282.42	417.98	48.00%
22	21	19	18	25	25	25	25	5	4	2	1	292.42	362.98	24.13%
22	21	19	18	27	26	24	23	5	4	2	1	302.42	307.98	1.84%
24	22	18	16	23	24	26	27	1	2	4	5	272.42	517.98	90.14%
24	22	18	16	25	25	25	25	1	2	4	5	292.42	587.98	101.08%
24	22	18	16	27	26	24	23	1	2	4	5	312.42	657.98	110.61%
24	22	18	16	23	24	26	27	3	3	3	3	468.00	468.00	0.00%
24	22	18	16	25	25	25	25	3	3	3	3	488.00	488.00	0.00%
24	22	18	16	27	26	24	23	3	3	3	3	508.00	508.00	0.00%
24	22	18	16	23	24	26	27	5	4	2	1	272.42	337.98	24.07%
24	22	18	16	25	25	25	25	5	4	2	1	292.42	292.42	0.00%
24	22	18	16	27	26	24	23	5	4	2	1	312.42	312.42	0.00%
27	24	17	13	23	24	26	27	1	2	4	5	257.42	622.98	142.01%
27	24	17	13	25	25	25	25	1	2	4	5	292.42	700.48	139.55%
27	24	17	13	27	26	24	23	1	2	4	5	327.42	777.98	137.61%
27	24	17	13	23	24	26	27	3	3	3	3	453.00	453.00	0.00%
27	24	17	13	25	25	25	25	3	3	3	3	488.00	488.00	0.00%
27	24	17	13	27	26	24	23	3	3	3	3	523.00	523.00	0.00%
27	24	17	13	23	24	26	27	5	4	2	1	257.42	257.42	0.00%
27	24	17	13	25	25	25	25	5	4	2	1	292.42	292.42	0.00%
27	24	17	13	27	26	24	23	5	4	2	1	327.42	327.42	0.00%

Table 6.6: Unselective vs. Selective Supply Chain

happens if only some retailers have an opportunity cost? We discuss this issue next.

To account for loss-of-goodwill in certain retail channels, we introduce a loss-of-goodwill cost into our model. The term O_i is the loss-of-goodwill cost of not serving a channel and is constant for a retailer i. The manufacturer's expected profit function changes to

$$\begin{aligned} E\Pi_M^O(x) &= \left(\sum_{i=1}^{N} \left(x_i \left(r_i \mu_i - d_i \left(y_i\right)\right) + \left(x_i - 1\right) O_i\right)\right) \\ &\quad -C\left(\mu_{SC}\left(x\right), \sigma_{SC}^2\left(x\right)\right) \\ &= \sum_{i=1}^{N} \left(x_i \left(r_i + \frac{O_i}{\mu_i}\right) \mu_i - d_i\left(y_i\right)\right) - \sum_{i=1}^{N} O_i \\ &\quad -C\left(\mu_{SC}\left(x\right), \sigma_{SC}^2\left(x\right)\right). \end{aligned}$$

Only the first term depends on the supply chain configuration x. The loss-of-goodwill cost artificially increases the retailer's revenue potential to $\tilde{r}_i = r_i + \frac{O_i}{\mu_i}$ and thus makes it more attractive for inclusion into the supply chain. Clearly, if the loss-of-goodwill cost O_i becomes large enough, the manufacturer clearly has an incentive to include the retailer into the supply chain. The maximization of $E\Pi_M^O(x)$ works as before. First we derive the new ordering

$$\begin{aligned} \frac{\mu_1\left(\tilde{r}_1 - c\right) - d_1\left(y_1\right)}{\sigma_1^2} &\geq \frac{\mu_2\left(\tilde{r}_2 - c\right) - d_2\left(y_2\right)}{\sigma_2^2} \geq \cdots \\ &\geq \frac{\mu_N\left(\tilde{r}_N - c\right) - d_N\left(y_N\right)}{\sigma_N^2}. \end{aligned}$$

The optimal solution is then given by the following proposition.

Proposition 6.9 *In the optimal solution x^*, if $x_k^* = 1$ also $x_l^* = 1$ for $l \in \{1, 2, \ldots, k-1\}$.*

We can compute the optimal solution by increasing the retail channel set sequentially and choosing the retailer p that maximizes expected supply chain profits. Then all other retailers that have a higher revenue contribution than retailer p are served in the coordinated supply chain. The following proposition states the optimality condition.

Proposition 6.10 *For the ordered retailer set, the optimal coordinating supply chain configuration x^* satisfies*

$$
\begin{aligned}
E\Pi_{SC}^{O} &= \left(\sum_{l \leq p^*} x_l \left(\tilde{r}_l \mu_l - d_l\left(y_l\right)\right)\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \geq \\
& \left(\sum_{l \leq p} x_l \left(r_l \tilde{\mu}_l - d_l\left(y_l\right)\right)\right) - C\left(\mu_{SC}\left(x(l)\right), \sigma_{SC}^2\left(x(l)\right)\right) \;\; \forall p \backslash p^*
\end{aligned}
$$

where $x_i(l) = 1$ if $l \leq p$ and $x_i(l) = 0$ otherwise.

In the optimal solution, we can see that $E\Pi_M^O(x_O^*) \leq E\Pi_M(x^*)$. Loss-of-goodwill costs that depend on the mean demand in the retail channel where the dependency is identical between retail channels lead to the same solution. However, if some channels are more important than others the optimal configuration might change. We will next analyze the effect on total supply chain profits if we increase the loss-of-goodwill cost of retailer j. Assume that retailer j was not in the optimal configuration x^* before (otherwise the solution clearly does not change). Then there are three effects if we include retailer j: (i) the expected profit of the other retailers is decreasing since the total demand variability is increasing. (ii) other retailers might be excluded from the supply chain.

To ensure coordination in the decentralized setting, we have to adjust the wholesale prices accordingly. This is a problem in the equal share contracts because all other retailers have to bear the higher supply chain costs due to this inclusion although the loss-of-goodwill cost only affects the manufacturer. Therefore an adjusted fair share contract can ensure participation of all other retailers. In this contract, the existing retailers receive the same expected profit as without the new retailer. On the other hand, the new retailer will bear all additional costs.

Proposition 6.11 *The wholesale price contract with*

$$w_i(\delta) = \begin{cases} r_i - \frac{\sum_{i \notin V \subseteq N} \frac{|V|!(n-|V|-1)!}{n!}(v(V \cup \{i\}) - v(V)) + d_i(y_i)}{\mu_i} & \text{if } x_i^* = 1, \\ 0 & \text{otherwise} \end{cases}$$

where

$$\begin{aligned} v(V) &= E\Pi_{SC}(x \in V, y) = (1-\delta)\left(\sum_{i=1}^{N}\left(x_i\left(r_i + \frac{O_i}{\mu_i}\right)\mu_i - d_i(y_i)\right)\right. \\ &\quad \left. - \sum_{i=1}^{N} O_i - C\left(\mu_{SC}(x), \sigma_{SC}^2(x)\right)\right) \end{aligned}$$

coordinates the supply chain with loss-of-goodwill costs and leads to a fair allocation of expected profits. The manufacturer receives $E\Pi_M^{opp} = \delta E\Pi_{SC}$ *and the retailer echelon* $E\Pi_R^{opp} = (1-\delta)\, E\Pi_{SC}$.

Until now, we have assumed that the retailers' holding costs do not depend on the wholesale price. In the next section we analyze the case where they depend directly on the manufacturer's pricing decision.

6.7 Alternative Holding Costs

In this section we assume that the wholesale price affects the holding costs of the retailer. In literature the holding costs are based on the physical inventory holding cost h_i and the discount rate $0 \leq \alpha \leq 1$. Then we get the modified holding cost factor

$$h_i = h_i' + (1-\alpha)\, w_i.$$

Substitution into the expected profit function yields

$$\begin{aligned} E\Pi_i(y_i) &= (r_i - w)\,\mu_i - \left(h_i' + (1-\alpha)\,w_i\right)\int_{\xi=0}^{y_i}(y_i - \xi)f_i(\xi)d\xi \\ &\quad - \left(b_i - (1-\alpha)\,w_i\right)\int_{\xi=y_i}^{\infty}(\xi - y_i)f_i(\xi)d\xi. \end{aligned}$$

The critical ratio changes to

$$y_i^* = F_i^{-1}\left(\frac{b_i - (1-\alpha)\,w_i}{b_i + h_i}\right).$$

Clearly, the optimal base stock level does not necessarily equal the critical ratio of the centralized solution

$$y_i^*(x) = F_i^{-1}\left(\frac{b_i - (1-\alpha)\,\frac{C\left(\mu_{SC}(x),\sigma_{SC}^2(x)\right)}{\mu_{SC}(x)}}{b_i + h_i}\right).$$

To solve this model, we cannot use our previous approach any more since the problem structure has changed. Therefore, we have to rely on numerical optimization to solve the optimization problem

$$\max_x \sum_i x_i \left(\left(r_i - c\left(\sigma_{SC}^2(x)\right)\right)\mu_i - d_i(y_i)\right).$$

In the next step, the manufacturer has to offer the retailers a contract that will lead to the profit allocation that we have developed in the previous section *and* incentivizes the retailer to choose the supply chain optimal base stock level at the same time. In literature, several contracts have been proposed that lead to supply chain coordination. We follow the idea of Cachon and Zipkin (1999) and use transfer payments for inventory on stock at the retailer, that is charged t_h per unit on stock, and for backorders that lead to a payment of t_b. Then the following proposition shows coordinating contracts that lead to the same supply chain solution and the same base stock levels as in the centralized solution.

Proposition 6.12 *For our supply chain a three parameter contract with*

$$\begin{aligned}
w_i &= (1-\lambda_i)\, r_i - \lambda_i \frac{C\left(\mu_{SC}(x^*), \sigma^2_{SC}(x^*)\right)}{\mu_{SC}(x^*)}, \\
t_{hi} &= (1-\alpha)(\lambda_i - 1)\frac{C\left(\mu_{SC}(x^*), \sigma^2_{SC}(x^*)\right)}{\mu_{SC}(x^*)} \\
t_{bi} &= (\lambda_i - 1)(b_i + h_i) + (1-\alpha)(1-\lambda_i)\frac{C\left(\mu_{SC}(x^*), \sigma^2_{SC}(x^*)\right)}{\mu_{SC}(x^*)} \\
&\quad \textit{and} \\
\lambda_i &= \frac{E\Pi_i^{fair}(x^*, y^*)}{E\Pi_{SC}^{CE}(x^*, y^*)}
\end{aligned}$$

coordinates the supply chain and leads to the supply chain optimal retailer selection and base stock levels.

We can see that t_h is negative and therefore the manufacturer subsidizes inventory on stock. On the other hand the backorder penalty t_b can be positive or negative depending on the specific situation.

6.8 Conclusion

We have analyzed how supply chain performance can be improved by managing demand variability optimally. We have shown that the well-known double marginalization effect also exists in supply chains where a manufacturer can choose among multiple retailers. Our analyses have shown that the effect of this double marginalization can be quite substantial as it is in the traditional contracting literature. To remedy this effect, we have developed contracts that coordinate the supply chain.

We have also shown that the approach of launching as many distribution channels as possible is often sub-optimal. We have presented cases where it might be even favorable to exclude a profitable channel if the total supply chain suffers from the high variability that this channel would import into the total supply chain.

In this chapter, we have not analyzed correlated retailer demands. Since these demand behavior can be encountered frequently in the context of fashion goods and high-tech articles, an analysis of this problem would improve the applicability of our model. Unfortunately, the analytical structure of our model does not hold any more in this context. Therefore, we leave this area for future research.

6.9 Proofs

Proof of Proposition 6.1. The proof follows the lines of Shen et al. (2003). Assume for contradiction that $x_l^* = 0$ for some $l \in \{1, 2, \ldots, k-1\}$. Define

$$x_i' = \begin{cases} 1, & \text{if } i = l \\ x_i^*, & \text{otherwise,} \end{cases}$$

and

$$x_i'' = \begin{cases} 0, & \text{if } i = k \\ x_i^*, & \text{otherwise,} \end{cases}$$

for each i. Let z^*, z', and z'' be the objective values of x^*, x', and x''. From optimality we know that $z^* - z'' \geq 0$. Then

$$z' - z^* = (w - c)\,\mu_l - \left(u\left(\bar{\sigma}^2 + \sigma_k^2 + \sigma_l^2\right) - u\left(\bar{\sigma}^2 + \sigma_k^2\right)\right)$$

and

$$z^* - z'' = (w - c)\,\mu_k - \left(u\left(\bar{\sigma}^2 + \sigma_k^2\right) - u\left(\bar{\sigma}^2\right)\right).$$

By concavity of

$$\frac{u\left(\bar{\sigma}^2 + \sigma_k^2 + \sigma_l^2\right) - u\left(\bar{\sigma}^2 + \sigma_k^2\right)}{\sigma_l^2} \leq \frac{u\left(\bar{\sigma}^2 + \sigma_k^2\right) - u\left(\bar{\sigma}^2\right)}{\sigma_k^2}$$

we see that

$$\begin{aligned}\frac{z' - z^*}{\sigma_l^2} &= (w - c)\frac{\mu_l}{\sigma_l^2} - \frac{u\left(\bar{\sigma}^2 + \sigma_k^2 + \sigma_l^2\right) - u\left(\bar{\sigma}^2 + \sigma_k^2\right)}{\sigma_l^2} \\ &\geq (w - c)\frac{\mu_k}{\sigma_k^2} - \frac{u\left(\bar{\sigma}^2 + \sigma_k^2\right) - u\left(\bar{\sigma}^2\right)}{\sigma_k^2} = \frac{z^* - z''}{\sigma_k^2}.\end{aligned}$$

It follows by $z' - z^* \geq 0$ that y' is also an optimal solution. ■

Proof of Proposition 6.2. From Proposition 6.1 we can see that we only have to search sequentially over p and no other combinations can be optimal. Therefore, it suffices to search over p and choose the cut-off level p^* that maximizes the manufacturer's expected profits. ■

Proof of Proposition 6.3. For every region of $\hat{w}_i$, the maximal $\hat{w}_i$ is relevant for the computation of x because the expected manufacturer profits are increasing in w. This follows from the optimality condition. Clearly, only the maximal $\hat{w}_i$ are relevant. ■

Proof of Proposition 6.4. The proof follows the lines of Shen et al. (2003). Assume for contradiction that $x_l^* = 0$ for some $l \in \{1, 2, \ldots, k-1\}$. Define

$$x_i' = \begin{cases} 1, & \text{if } i = l \\ x_i^*, & \text{otherwise,} \end{cases}$$

and

$$x_i'' = \begin{cases} 0, & \text{if } i = k \\ x_i^*, & \text{otherwise,} \end{cases}$$

for each i. Let z^*, z', and z'' be the objective values of x^*, x', and x''. From optimality we know that $z^* - z'' \geq 0$ Then

$$z' - z^* = \mu_l\left(r_l - c\right) - d_l\left(y_l\right) - \left(u\left(\bar{\sigma}^2 + \sigma_k^2 + \sigma_l^2\right) - u\left(\bar{\sigma}^2 + \sigma_k^2\right)\right)$$

and

$$z^* - z'' = \mu_k\left(r_k - c\right) - d_k\left(y_k\right) - \left(u\left(\bar{\sigma}^2 + \sigma_k^2\right) - u\left(\bar{\sigma}^2\right)\right).$$

By concavity of

$$\frac{u\left(\bar{\sigma}^2+\sigma_k^2+\sigma_l^2\right)-u\left(\bar{\sigma}^2+\sigma_k^2\right)}{\sigma_l^2} \leq \frac{u\left(\bar{\sigma}^2+\sigma_k^2\right)-u\left(\bar{\sigma}^2\right)}{\sigma_k^2}$$

we see that

$$\begin{aligned} \frac{z'-z^*}{\sigma_l^2} &= \frac{\mu_l\left(r_l-c\right)-d_l\left(y_l\right)}{\sigma_l^2}-\frac{u\left(\bar{\sigma}^2+\sigma_k^2+\sigma_l^2\right)-u\left(\bar{\sigma}^2+\sigma_k^2\right)}{\sigma_l^2} \\ &\geq \frac{\mu_k\left(r_k-c\right)-d_k\left(y_k\right)}{\sigma_k^2}-\frac{u\left(\bar{\sigma}^2+\sigma_k^2\right)-u\left(\bar{\sigma}^2\right)}{\sigma_k^2}=\frac{z^*-z''}{\sigma_k^2}. \end{aligned}$$

It follows by $z'-z^* \geq 0$ that y' is also an optimal solution. ■

Proof of Proposition 6.5. From the preceding Propositions we can see that we only have to search sequentially over p and no other combinations can be optimal. Therefore, it suffices to search over p and choose the cut-off level p^* that maximizes the manufacturer's expected profits. ■

Proof of Proposition 6.6. The profit function is

$$\begin{aligned} E\Pi_{SC}^C &= \left(\sum_{l\leq p^*} x_l\left(r_l\mu_l-d_l\left(y_l\right)\right)\right)-C\left(\mu_{SC}\left(x(l)\right),\sigma_{SC}^2\left(x(l)\right)\right)\geq \\ &\left(\sum_{l\leq p} x_l\left(r_l\mu_l-d_l\left(y_l\right)\right)\right)-C\left(\mu_{SC}\left(x(l)\right),\sigma_{SC}^2\left(x(l)\right)\right)\ \forall p\backslash p^*. \end{aligned}$$

and the ordering is

$$\begin{aligned} \frac{\mu_1\left(r_1-c\right)-d_1\left(y_1\right)}{\sigma_1^2} &\geq \frac{\mu_2\left(r_2-c\right)-d_2\left(y_2\right)}{\sigma_2^2}\geq\cdots \\ &\geq \frac{\mu_N\left(r_N-c\right)-d_N\left(y_N\right)}{\sigma_N^2} \end{aligned}$$

The profit margin for all retailers is then zero and the manufacturer takes all supply chain profits. ■

Proof of Proposition 6.7.

The wholesale price is increasing in δ because

$$\frac{dw_i(\delta)}{d\delta} = \left(r_i - \frac{d_i(y_i)}{\mu_i}\right) - c\left(\sigma_{SC}^2(x)\right) \geq 0$$

by optimality of the centralized solution. It is still rational for the supplier to choose the retailers because

$$\begin{aligned} E\Pi_{SC}^C &= \sum_{l \leq p^*} x_l \delta\left(\left(r_l - c\left(\sigma_{SC}^2(x)\right)\right)\mu_l - d_l(y_l)\right) \\ &\geq \sum_{l \leq p} x_l \delta\left(\left(r_l - c\left(\sigma_{SC}^2(x)\right)\right)\mu_l - d_l(y_l)\right) \ \forall p \backslash p^* \end{aligned}$$

and

$$\delta\frac{\mu_1 r_1 - d_1(y_1)}{\sigma_1^2} \geq \delta\frac{\mu_2 r_2 - d_2(y_2)}{\sigma_2^2} \geq \ldots \geq \delta\frac{\mu_N r_N - d_N(y_N)}{\sigma_N^2}$$

for $0 \leq \delta \leq 1$. ■

Proof of Proposition 6.8.

The value function for the retailer echelon expected profits equals

$$v = (1-\delta)\left(\left(\sum_{l \leq p^*} x_l\left(r_l\mu_l - d_l(y_l)\right)\right) - C\left(\mu_{SC}(x(l)), \sigma_{SC}^2(x(l))\right)\right)$$

from the optimal centralized solution. The basic axioms for the Shapley value are (i) $v(\emptyset) = 0$ (ii) superadditivity $v(S \cup T) \geq v(S) + v(T)$ with S and T disjoint subset (cooperation can only help and not hurt). Clearly, (i) holds since an empty supply chain does not produce any profits or costs. For (ii), the superadditivity follows from the optimality condition of the supply chain and the concave variability cost function $u\left(\sigma_{SC}^2(x)\right)$.

Then the expected profit for retailer i equals

$$\begin{aligned} E\Pi_i(y_i) &= (r_i - w)\,\mu_i - d_i\,(y_i) \\ &= (1-\delta)\,\pi_i \\ &= (1-\delta) \sum_{i \notin V \subseteq N} \frac{|V|!(n-|V|-1)!}{n!} \left(v\,(V \cup \{i\}) - v\,(V)\right) \end{aligned}$$

It follows the optimal wholesale price w_i for retailer i. ■

Proof of Proposition 6.9. The proof is analogous the Proof of Proposition 6.4. ■

Proof of Proposition 6.10. The proof is analogous the Proof of Proposition 6.5. ■

Proof of Proposition 6.11. The proof is analogous the Proof of Proposition 6.8. ■

Proof of Proposition 6.12. The retailer generates a fraction $\lambda_i = \frac{E\Pi_i^{fair}(x^*,y^*)}{E\Pi_{SC}^{CE}(x^*,y^*)}$ of the centralized supply chain profits. Then we can set

$$\begin{aligned} &\lambda_i \Bigg(\left(r_i - \frac{C\left(\mu_{SC}\,(x^*)\,,\sigma^2_{SC}\,(x^*)\right)}{\mu_{SC}(x^*)} \right) \\ &- \left(h_i + (1-\alpha)\,\frac{C\left(\mu_{SC}\,(x^*)\,,\sigma^2_{SC}\,(x^*)\right)}{\mu_{SC}(x^*)} \right) \int_{\xi=0}^{y_i} (y_i - \xi) f_i(\xi) d\xi \\ &- \left(b_i - (1-\alpha)\,\frac{C\left(\mu_{SC}\,(x^*)\,,\sigma^2_{SC}\,(x^*)\right)}{\mu_{SC}(x^*)} \right) \int_{\xi=y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi \Bigg) \\ = \;& (r_i - w_i) - \left(h_i + (1-\alpha)\,\frac{C\left(\mu_{SC}\,(x^*)\,,\sigma^2_{SC}\,(x^*)\right)}{\mu_{SC}(x^*)} + t_h \right) \cdot \\ &\int_{\xi=0}^{y_i} (y_i - \xi) f_i(\xi) d\xi \\ &- \left(b_i - (1-\alpha)\,\frac{C\left(\mu_{SC}\,(x^*)\,,\sigma^2_{SC}\,(x^*)\right)}{\mu_{SC}(x^*)} + t_b \right) \int_{\xi=y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi. \end{aligned}$$

It follows that

$$\begin{aligned} w_i &= (1-\lambda_i)\, r_i - \lambda_i \frac{C\left(\mu_{SC}\left(x^*\right), \sigma^2_{SC}\left(x^*\right)\right)}{\mu_{SC}(x^*)}, \\ t_{hi} &= (1-\alpha)\left(\lambda_i - 1\right) \frac{C\left(\mu_{SC}\left(x^*\right), \sigma^2_{SC}\left(x^*\right)\right)}{\mu_{SC}(x^*)} \\ t_{bi} &= \left(\lambda_i - 1\right)\left(b_i + h_i\right) + (1-\alpha)\left(1-\lambda_i\right) \frac{C\left(\mu_{SC}\left(x^*\right), \sigma^2_{SC}\left(x^*\right)\right)}{\mu_{SC}(x^*)} \end{aligned}$$

by solving for the contract parameters and substitution. Clearly the critical ratio equals the centralized solution because

$$\begin{aligned} \frac{b_i + t_b - (1-\alpha)\, w_i}{h_i + b_i + t_b + t_h} &= \frac{\lambda_i \left(b_i - (1-\alpha) \frac{C\left(\mu_{SC}(x^*), \sigma^2_{SC}(x^*)\right)}{\mu_{SC}(x^*)}\right)}{\lambda_i \left(h_i + b_i\right)} \\ &= \frac{b_i - (1-\alpha) \frac{C\left(\mu_{SC}(x^*), \sigma^2_{SC}(x^*)\right)}{\mu_{SC}(x^*)}}{h_i + b_i} \end{aligned}$$

■

Chapter 7

Conclusion and Critical Review

In this thesis, we have presented approaches for designing efficient supply chain contracts. We used an alternative formulation of the service level metric that allowed us to design contracts that are enforceable and easy to understand. Furthermore, we analyzed how demand variability can be consciously admitted to the supply chain and how this admission policy affects the pricing policy.

In this chapter, we conclude the thesis and summarize our main findings. More precisely, we highlight our contributions to current research and we critically review our models. Last, we provide an outlook into further research possibilities on the proposed contracts.

7.1 Contributions

In Chapter 4, we used a finite-horizon service level measure that was proposed by Thomas (2005). We developed two different service level contracts, a flat penalty and a unit penalty contract. We used these contracts to develop a supply chain model that is similar to Cachon and

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1_7

Zipkin (1999). Thereby, we could identify optimal contract values for these service level contracts and could compare these values with the traditional service level measures. This issue has not been analyzed in literature yet. Our analysis is based on the log-concave property of the demand distribution. We derived contract parameter values for both contract types that are able to coordinate the supply chain. We also found out that it is not always optimal for the manufacturer to offer the supplier a contract with a high service level *and* a high penalty payment at the same time. Rather, the manufacturer has to take the supplier's response into account when deciding on optimal contract parameters.

In Chapter 5, we extended current research on inventory allocation models, as for example proposed by McGavin et al. (1993) and van der Heijden (1999). We have shown that a manufacturer can realize significant cost savings from the risk pooling effect among the retailers and that an optimal allocation strategy can further enhance a manufacturer's profit. We developed the *contract balancing* approach that excludes retailers whose last period demand was exceptionally high or low from subsequent replenishments. Our analyses also show that the contract balancing approach deviates significantly from the traditional inventory balancing approach that replenishes the retailer with the lowest inventory level first and that this deviation can lead to incentive conflicts between manufacturer and retailers.

In Chapter 6, we proposed a model that selectively admits variability to the supply chain. This issue has not been analyzed in literature until now. We have shown that it is beneficial for a manufacturer to exclude some retailers or distribution channels. We derived an efficient optimization approach to find the profit maximizing channel selection and developed a wholesale pricing scheme that leads to a coordinated supply chain. We also extended our model to situations where channel opportunity costs are relevant. In addition, we use a game theoretic model - the Shapley value - to allocate profits fairly among the retailers.

7.2 Critical Review

Our thesis contributes to the existing literature on inventory management and contracting. Unavoidably, we had to make simplifying assumptions to come up with tractable models, solution approaches, and analytical results. In this section we want to review some of these critical and limiting assumptions.

Some of our assumptions are in line with current research, but nonetheless limit the practical applicability of our models. For example, we assume in Chapters 5 and 6 that retailer demands are not correlated. Clearly, this is nearly never the case in practical applications. Since this demand characteristic can be encountered frequently in the context of fashion goods and high-tech articles, an analysis that incorporates correlated customer demands would improve the applicability of our model. Relaxing this assumption would significantly complicate our models and would prevent an analytical analysis of the results. Nonetheless, our channel selection model in Chapter 6 could be extended to correlated demands, but we would have to rely on numerical approaches for the optimization and analytical insights would be hard to derive.

In addition, the full backorder assumption limits the applicability of our model. In practice, customers do not always wait for a later delivery of a product if this product is not on stock in the store. Some customers substitute with another product or go to another store. This leads to more complex models that need to include these lost sales and demand spill-over effects. Still we feel that the backorder assumption allows us to identify the important aspects of our models.

On a final note, the distribution channel and retailer selection problem in Chapter 6 might not be applicable to all situations and is worth a short discussion. Clearly, our suggestion of consciously forgoing some demand sources and consequently some revenue crosses common intuition. Nonetheless, for companies with limited resources for serving different markets, the optimal market selection approach might be worth a thought. Especially small and medium enterprises (SME) should be actively choosing their markets instead of using an unfocused distribution strategy.

7.3 Future Research

Our work offers researchers a number of potential areas for future research. It might be worth studying the impact of different coordinating contract parameter combinations on the variability of profits in Chapter 4. Until now, we have analyzed this issue numerically and found that there exist variability minimizing parameter combinations. Unfortunately, we have not been able to derive structural results.

Another interesting aspect for extending our model from Chapter 4 is to measure the service level over more than one period. Then the supplier has an additional degree of freedom on how to fill the manufacturer's orders. Furthermore, the aspect of information asymmetries could be included into the model. In the existing model, the manufacturer has full information about the supplier's cost and reservation profit. Relaxing this assumption would lead to a higher acceptance of the model and its recommendations.

For the inventory allocation problem in Chapter 5, one can think of extending our model to a setting where the manufacturer faces a unit penalty contract where the magnitude of a stock-out influences the penalty payment. Unfortunately, the unit penalty function cannot be determined as easily as the penalty function in our model. Therefore, we leave the unit penalty contracts for future research.

We have already suggested in the last section that the channel selection model from Chapter 6 would benefit from an extension to a correlated demand environment. Unfortunately, our analytical results and solution approach cannot be easily modified to handle this case. Initial thoughts however lead to the interesting conclusion that positively correlated demand might lead to an even more restrictive channel selection because the cost-driving demand variability is even amplified compared to the uncorrelated setup. Nonetheless, analytical results are needed to support this hypothesis.

Bibliography

Aberdeen Group. 2004. *Supplier Performance Management: What Leaders Do Differently*. Benchmark Report, September, 2004.

Alptekinoglu, A., C. S. Tang. 2005. A model for analyzing multi-channel distribution systems. *European Journal of Operational Research* **163**(3) 802-824.

Anderson, E., G. S. Day, V. K. Rangan. 1997. Strategic Channel Design. *Sloan Management Review* **38**(4) 59-69.

Axsäter, S., K. Rosling. 1993. *Installation vs. Echelon Stock Policies for Multilevel Inventory Control*. Management Science **39**(10) 1274-1280.

Bagnoli, M., T. Bergstrom. 2005. Log-concave probability and its applications. *Economic Theory* **26**(2) 445-469.

Balakrishnan, A., J. Geunes, M. S. Pangurn. 2004. Coordinating Supply Chains by Controlling Upstream Variability Propagation. *Manufacturing & Service Operations Management* **6**(2) 163-183.

Balakrishnan, A., M. S. Pangburn, E. Stavrulaki. 2004b. "Stack Them High, Let 'em Fly": Lot-Sizing Policies When Inventories Stimulate Demand. *Management Science* **50**(5) 630-644.

Beamon, B. M. 1999. Measuring supply chain performance. *International Journal of Operations & Production Management* **19**(3) 275-292.

M. Sieke, *Supply Chain Contract Management*, Edition KWV,
https://doi.org/10.1007/978-3-658-24382-1

Behrenbeck, K., J. Küpper, K.-H. Magnus, U. Thonemann. 2007. *Supply Chain Champions*. McKinsey & Co. Publication.

Beth, S., D. N. Burt, W. Copacino, C. Gopal, H. L. Lee, R. P. Lynch, S. Morris. 2003. Supply Chain Challenges: Building Relationships. *Harvard Business Review* **81**(7) 64-73.

Billington, C. 2005. Personal Communication at IMD Lausanne, Switzerland, November 28, 2005.

Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, United Kingdom.

Cachon, G. P. 1999. Managing Supply Chain Demand Variability with Scheduled Ordering Policies. *Management Science* **45**(6) 843-856.

Cachon, G. P. 2003. Supply Chain Coordination with Contracts. S. Graves, T. de Kok, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management.* North-Holland, Amsterdam, The Netherlands.

Cachon, G. P. 2004. The Allocation of Inventory Risk in a Supply chain: Push, Pull, and Advance-Purchase Discount Contracts. *Management Science* **50**(2) 222-238.

Cachon, G. P., A. G. Kök. 2007. Implementation of the Newsvendor Model with Clearance Pricing: How to (and How Not to) Estimate a Salvage Value. *Management Science* **9**(3) 276-290.

Cachon, G. P., M. A. Lariviere. 2001. Turning the Supply Chain into a Revenue Chain. *Harvard Business Review* **79**(3) 20-21.

Cachon, G. P., M. A. Lariviere. 2005. Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations. *Management Science* **51**(1) 30-44.

Cachon, G. P., F. Zhang. 2006. Procuring Fast Delivery: Sole Sourcing with Information Asymmetry. *Management Science* **52**(6) 881-896.

Cachon, G. P., P. H. Zipkin. 1999. Competitive and Cooperative Inventory Policies in a Two-Stage Supply Chain. *Management Science* **45**(7) 939-953.

Callioni, G., X. de Montgros, R. Slagmulder, L. N. Van Wassenhove, L. Wright. 2005. Inventory-Driven Costs. *Harvard Business Review* **83**(3) 135-141.

Carr, S., W. Lovejoy. 2000. The Inverse Newsvendor Problem: Choosing an Optimal Demand Portfolio for Capacitated Resources. *Management Science* **46**(7) 912-927.

Chen, F. 1999. Decentralized Supply Chains Subject to Information Delays. *Management Science* **45**(8) 1076-1090.

Chen, F., A. Federgruen. 2000. Mean-Variance Analysis of Basic Inventory Models. Working Paper, Columbia University, New York, NY.

Chen, F., Y. Zheng. 1993. Inventory models with general backorder costs. *European Journal of Operational Research* **65**(2) 175-186.

Chen, J., D. K. J. Lin, D. J. Thomas. 2003. On the single item fill rate for a finite horizon. *Operations Research Letters* **31**(2) 119-123.

Choi, K., J. G. Dai, J. Song. 2004. On Measuring Supplier Performance Under Vendor-Managed-Inventory Programs in Capacitated Supply Chains. *Manufacturing & Service Operations Management* **6**(1) 53-72.

Chopra, S., P. Meindl. 2004. *Supply Chain Management: Strategy, Planning, and Operations*, 2nd edition. Pearson Education, Upper Saddle River, NJ.

Clark, A. J., H. Scarf. 1960. Optimal Policies for a Multi-Echelon Inventory Problem. *Management Science* **6**(4) 475-490.

Cohen, M. A., H. L. Lee. 1988. Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods. *Operations Research* **36**(2) 216-228.

Corbett, C. J., D. Zhou, C. S. Tang. 2004. Designing Supply Contracts: Contract Type and Information Asymmetry. *Management Science* **50**(4) 550-559.

Corsten, D., T. Gruen. 2004. Stock-Outs Cause Walkouts. *Harvard Business Review* **82**(5) 26-28.

Cui, T. H., J. S. Raju, Z. Z. Zhang. 2007. Fairness and Channel Coordination. *Management Science* **53**(8) 1303-1314.

Diks, E. B., A. G. de Kok, A. G. Lagodimos. 1996. Multi-echelon systems: A service measure perspective. *European Journal of Operational Research* **95**(2) 241-263.

Emmelhainz, M. A., J. R. Stock, L. W. Emmelhainz. 1991. Consumer responses to retail stockouts. *Journal of Retailing* **67**(2) 138-147.

Eppen, G. D. 1979. Effects of Centralization on Expected Costs in a Multi-Location Newsboy Problem. *Management Science* **25**(5) 498-501.

Eppen, G. D., L. Schrage. 1981. Centralized ordering policies in a multi-warehouse inventory. L. B. Schwarz, ed. *Multi-Level Production/Inventory Control Systems: Theory and Practice*. North-Holland, Amsterdam, The Netherlands.

Federgruen, A., P. Zipkin. 1984. Approximations of Dynamic Multilocation Production and Inventory Problems. *Management Science* **30**(1) 69-84.

Federgruen, A., P. H. Zipkin. 1984b. Computational Issues in an Infinite-Horizon, Multiechelon Inventory Model. *Operations Research* **32**(4) 818-836.

Fisher, M. L. 1997. What Is the Right Supply Chain for Your Product? *Harvard Business Review* **75**(2) 105-116.

Fisher, M. L., J. Hammond, W. Obermeyer, A. Raman. 1997. Configuring a Supply Chain to Reduce the Cost of Demand Uncertainty. *Production and Operations Management* **6**(3) 211-225.

Frazier, G. L. 1999. Organizing and Managing Channels of Distribution. *Journal of the Academy of Marketing Science* **27**(2) 226-240.

Fry, M. J., R. Kapuscinski, T. L. Olsen. 2001. Coordinating Production and Delivery Under a (z, Z)-Type Vendor-Managed Inventory Contract. *Manufacturing & Service Operations Management* **3**(2) 151-173.

Gerchak, Y., D. Mossman. 1992. On the effect of demand randomness on inventories and costs. *Operations Research* **40**(4) 804-807.

Gunasekaran, A., C. Patel, E. Tirtiroglu. 2001. Performance measures and metrics in a supply chain environment. *International Journal of Operations & Production Management* **21**(1/2) 71-87.

Güllü, R., N. Erkip. 1996. Optimal allocation policies in a two-echelon inventory problem with fixed shipment costs. *International Journal of Production Economics* **46** 311-321.

Jackson, P. L. 1988. Stock Allocation in a Two-Echelon Distribution System or "What to Do Until Your Ship Comes in". *Management Science* **34**(7) 880-895.

Jackson, P. L., J. A. Muckstadt. 1989. Risk Pooling in a Two-Period, Two-Echelon Inventory Stocking and Allocation Problem. *Naval Research Logistics* **36**(1) 1-16.

Jönsson, H., E. A. Silver. 1987. Stock allocation among a central warehouse and identical regional warehouses in a particular push inventory system. *International Journal of Production Research* **25**(2) 192-205.

Johnson, M. E. 2001. Learning from Toys: Lessons in Managing Supply Chain Risk From the Toy Industry. *California Management Review* **43**(3) 106-124.

Johnson, M. E., H. L. Lee, T. Davis, R. Hall. 1995. Expressions for Item Fill Rates in Periodic Inventory Systems. *Naval Research Logistics* **42**(1) 57-80.

Kapuscinski, R., R. Q. Zhang, P. Carbonneau, R. Moore, B. Reeves. 2004. Inventory Decisions in Dell's Supply Chain. *Interfaces* **34**(3) 191-205.

Karlin, S., F. Proschan. 1960. Pólya Type Distributions of Convolutions. *The Annals of Mathematical Statistics* **31**(3) 721-736.

Kotler, P. 1980. *Marketing Management Analysis, Planning and Control*, 4th edition. Prentice-Hall, Englewood Cliffs, NJ.

Kraiselburd, S., V. G. Narayanan, A. Raman. 2004. Contracting in a Supply Chain with Stochastic Demand and Substitute Products. *Production and Operations Management* **13**(1) 46-62.

Lariviere, M. A., E. L. Porteus. 2001. Selling to the newsvendor: an analysis of price-only contracts. *Manufacturing & Service Operations Management* **3**(4) 293-305.

Lee, H. L. 2004. The Triple-A Supply Chain. *Harvard Business Review* **82**(10) 102-112.

Lee, H. L., C. Billington. 1992. Managing Supply Chain Inventory: Pitfalls and Opportunities. *Sloan Management Review* **33**(3) 65-73.

Lee, H. L., C. Billington. 1993. Material Management in Decentralized Supply Chains. *Operations Research* **41**(5) 835-847.

Lee, H. L., S. Whang. 1999. Decentralized Multi-Echelon Supply Chains: Incentives and Information. *Management Science* **45**(5) 633-640.

Lee, H. L., V. Padmanabhan, T. A. Taylor, S. Whang. 2000. Price Protection in the Personal Computer Industry. *Management Science* **46**(4) 467-482.

McGavin, E. J., L. B. Schwarz, J. E. Ward. 1993. Two-interval Inventory-allocation Policies in a One-warehouse N-identical-retailer Distribution System. *Management Science* **39**(9) 1092-1107.

Metty, T., R. Harlan, Q. Samelson, T. Moore, T. Morris, R. Sorensen, A. Schneur, O. Raskina, R. Schneur, J. Kanner, K. Potts, J. Robbins. 2005. Reinventing the Supplier Negotiation Process at Motorola. *Interfaces* **35**(1) 7-23.

Meyerson, R. B. 2000. *Game Theory: Analysis of Conflict.* Harvard University Press, Harvard, MA.

Nahmias, S. 2005. *Production and Operations Analysis*, 5th edition. McGraw-Hill, Boston, MA.

Narayanan, V. B., A. Raman. 2004. Aligning Incentives in Supply Chains. *Harvard Business Review* **82**(11) 94-102.

Neale, J. J. 2003. The Role of Inventory in Superior Supply Chain Performance. T. P. Harrison, H. L. Lee, J. J. Neale, eds. *The Practice of Supply Chain Management: Where Theory and Application Converge.* Springer, New York, NY.

Pasternack, B. A. 1985. Optimal Pricing and Return Policies for Perishable Commodities. *Marketing Science* **4**(2) 166-176.

Porteus, E. L. 1990. Stochastic Inventory Theory. D. P. Heyman, M. Sobel, eds. *Handbooks in OR & MS, Vol. 2.* North-Holland, Elsevier Publishers B.V., New York, NY.

Porteus, E. L. 2000. Responsibility Tokens in Supply Chain Management. *Manufacturing & Service Operations Management* **2**(2) 203-219.

Rosling, K. 2002. Inventory Cost Rate Functions With Nonlinear Shortage Costs. *Operations Research* **50**(6) 1007-1117.

Scarf, H. 1959. The Optimality of (S, s) Policies in the Dynamic Inventory Problem. K. Arrow, S. Karlin, H. Scarf, eds. *Mathematical Methods in the Social Sciences*. Stanford, CA.

Schweitzer, M. E., G. P. Cachon. 2000. Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. *Management Science* **46**(3) 404-420.

Shapiro, B., R. McOriary, C. Cline. 1992. Fabtek (A). *Harvard Business School Case Study,* 9-592-095, Boston, MA.

Shen, Z. M., C. Coullard, M. S. Daskin. 2003. A Joint Location-Inventory Model. *Transportation Science* **37**(1) 40-55.

Silver, E. A., D. F. Pyke, R. Peterson. 1998. *Inventory Management and Production Planning and Scheduling*, JohnWiley and Sons, New York, NY.

Sobel, M. J. 2004. Fill Rates of Single-Stage and Multistage Supply Systems. *Manufacturing & Service Operations Management* **6**(1) 41-52.

Spengler, J. 1950. Vertical Integration and Antitrust Policy. *Journal of Political Economy* **58**(4) 347-352.

Tan, B., S. Karabati. 2004. Can the desired service level be achieved when the demand and lost sales are unobserved? *IIE Transactions* **36**(4) 345-358.

Tapiero, C. S. 2005. Value at risk and inventory control. *European Journal of Operational Research* **163**(3) 769-775.

Tempelmeier, H. 2006. *Inventory Management in Supply Networks. Problems, Models, Solutions.* Books on Demand, Norderstedt, Germany.

Thomas, D. J. 2005. Measuring Item Fill-Rate Performance in a Finite Horizon. *Manufacturing & Service Operations Management* **7**(1) 74-80.

Thomas, D. J. 2007. *10 Trends That Will Change Your Supply Chain*, World Trade Magazine, January Online Edition.

Thonemann, U. W. 2005. *Operations Management: Konzepte, Methoden und Anwendungen (in German)*. Pearson Studium, Munich, Germany.

Thonemann, U. W., A. O. Brown, W. H. Hausman. 2002. Easy Quantification of Improved Spare Parts Inventory Policies. *Management Science* **48**(9) 1213-1225.

Thonemann, U. W., K. Behrenbeck, J. Küpper, K.-H. Magnus. 2005. *Supply Chain Excellence im Handel (in German)*. Financial Times Deutschland/Gabler, Wiesbaden, Germany.

Tomlin, B. 2003. Capacity Investments in Supply Chains: Sharing the Gain Rather Than Sharing the Pain. *Manufacturing & Service Operations Management* **5**(4) 317-333.

Tsay, A. A. 1999. The Quantity Flexibility Contract and Supplier-Customer Incentives. *Management Science* **45**(10) 1339-1358.

Tsay, A. A., S. Nahmias, N. Agrawal. 1999. Modeling Supply Chain Contracts: A Review. S. Tayur, R. Ganeshan, M. Magazine, eds. *Handbooks in Operations Research and Management Science: Quantitative Models For Supply Chain Management.* North-Holland, Amsterdam, The Netherlands.

U.S. Census Bureau. 2008. *Manufacturing and Trade Inventories and Sales*, March 2008.

van der Heijden, M. C. 1999. Multi-echelon inventory control in divergent systems with shipping frequencies. *European Journal of Operational Research* **116**(2) 331-351.

Van Houtum, G. J., K. Inderfurth, W. H. M. Zijm. 1996. Materials coordination in stochastic multi-echelon systems. *European Journal of Operational Research* **95**(1) 1-23.

Verrijdt, J. H. C. M., A. G. de Kok. 1996. Distribution planning for a divergent depotless two-echelon network under service constraints. *European Journal of Operational Research* **89**(2) 341-354.

Volkswagen AG. 2005. *Volkswagen Banker's Meeting May 19, 2005.* Retrieved June 6, 2006, http://www.volkswagen-ir.de/fileadmin/vw-ir2/dokumente/inv_presentations/2005/Q2/2005-05-19_Bankers__Meeting.pdf.

Wang, Y., L. Jiang, Z. Shen. 2004. Channel Performance Under Consignment Contract with Revenue Sharing. *Management Science* **50**(1) 34-47.

Whang, S. 1995. Coordination in Operations: A Taxonomy. *Journal of Operations Management* **12**(3) 413-422.

Zipkin, P. H. 1984. On the Imbalance of Inventories in Multi-Echelon Systems. *Mathematics of Operations Research* **9**(3) 402-423.

Zipkin, P. H. 2000. *Foundations of Inventory Management.* McGraw-Hill, Boston, MA.