DISSERTATION

Titel der Dissertation

„Essays on Inventory Sharing, Transshipments and Coordination in Supply Chain Management“

Verfasser

Dipl.-Wirtsch.-Ing. Arkadi Seidscher

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 094 151 403
Dissertationsgebiet lt. Studienblatt:	Logistics and Operations Management
Betreuerin / Betreuer:	Prof. Dr. Stefan Minner

Acknowledgements

I would like to express my gratitude to Prof. Dr. Stefan Minner for giving me the opportunity to work on this highly interesting and challenging subject, for supervising my work, but also for the highly professional, yet constantly open and friendly support throughout our cooperation. Further I would like to thank my great colleagues at the Chair for Logistics and Supply Chain Management for their numerous ideas and their constructive feedback during the years. In addition, a special thank goes to Prof. Maqbool Dada, who gave me the opportunity for the research stay at the Johns Hopkins Carey Business School at Baltimore. Most importantly, I want to thank my parents for their constant support, motivation and encouragement during these four years as well as for their patience.

Abstract

This thesis addresses problems in supply chain management from operational and strategic points of view. On the operational level an effective inventory control is a key to successful supply chain management. Intelligent inventory management is not limited to decisions about the quantity and location of the goods in the supply network but it also includes stock transfer rules between markets and locations within a distribution network.

Beyond the operational perspective on the supply chain this thesis examines the strategic long term horizon decisions of the market participants under competition. Supply chains facing competition from a rival distribution channel have to coordinate their own decisions with the rival parties in the market.

We study different variants of stochastic inventory models with the aim to present new inventory control methods and to show how these methods can improve the performance of the supply chain. First, we examine how a distribution network can benefit from a proactive inventory sharing between the warehouse locations. Second, we suggest to use the information about the inventory in the supply pipeline into the design of stock transfer rules. We show how this method can reduce the operational cost. Third, we analyze the behavior of supply chain participants competing for the customers which differentiate the products in price and quality. Here we use game theory to derive the optimal strategic decisions regarding the price and quality of the products.

Zusammenfassung

Die vorliegende Dissertation befasst sich mit Problemenstellungen im Lieferkettenmanagement (engl.: Supply Chain Management) sowohl aus operativen als auch strategischen Gesichtspunkten. Auf der operativen Ebene ist ein effektives Bestandsmanagement (engl.: Inventory Management) der Schlüssel zum erfolgreichen Supply Chain Management. Intelligentes Inventory Management ist nicht nur beschränkt auf Entscheidungen bezüglich der Anzahl und der Position der Waren im Versorgungsnetz, sondern umfasst auch Regeln bezüglich des Transfers von Waren zwischen Märkten und Standorten innerhalb eines Vertriebsnetzes.

Jenseits der operativen Sicht auf die Lieferkette untersucht diese Arbeit die strategisch-langfristigen Entscheidungen der Marktteilnehmer unter Wettbewerb. Lieferketten die einer Konkurrenz von rivalisierenden Vertriebskanälen ausgesetzt sind, müssen ihre eigenen Entscheidungen mit den übrigen Parteien im Markt koordinieren.

Wir untersuchen verschiedene Varianten der stochastischen Bestandsmodelle mit dem Ziel, neue Methoden der Bestandskontrolle zu präsentieren und zu zeigen, wie diese Methoden die Performance der Beschaffungskette verbessern können. Zunächst untersuchen wir, wie ein Vertriebsnetz aus einer proaktiven Bestandsmitbenutzung zwischen den Lagerstandorten profitieren kann. Zweitens schlagen wir vor, die Informationen über die Lagerbestände in Transit bei der Gestaltung der Transferregeln einzubeziehen. Wir zeigen, wie dieses Verfahren die Betriebskosten reduzieren kann. Drittens analysieren wir das Verhalten der Teilnehmer der Lieferkette im Wettbewerb um die Kunden, die in Preis und Qualität des Produktes unterscheiden. Hier verwenden wir spieltheoretische Methoden zur Herleitung der optimalen strategischen Entscheidungen bezüglich Preis und Qualität der Produkte.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In order to achieve a good performance, every company selling or distributing a product or a service to its customers requires efficient supply chain management. When designing a supply chain, two levels of planning are usually distinguished, depending on the time horizon: strategic and operational. The strategic supply chain decisions highly depends on the behavior of other involved parties and require market coordination among the participants. On the operational level where inventory control and management take place, a company can be analyzed isolated from other rival parties.

In this thesis selected aspects out of wide range of inventory control problems with lateral transshipments are studied. Inventory systems are usually used in a wide range of industries, such as car manufacturing, aircraft industry, communication systems, etc. to provide service support for products whose customers are distributed over an extensive geographical region. In order to guarantee a high service level many companies use an extensive distribution network of inventory locations. The investment in inventory in such networks can be very high. Flint (1995) states that world wide spare parts inventory in the aviation industry alone amounted to $ 45 billion at that time. The recent report of Solis et al. (2008) quantifies the total value of secondary inventory (spare parts such as aircraft and ship engines and their components and accessories) of the U.S. military forces to $ 82.6 billion. Also in other industries and in other types of inventories, a large amount of money is invested and this has increased over the years. This makes it crucial to manage them effectively and efficiently. Downsizing of a few percent of necessary inventory without decreasing customer service would lead to large cost savings in absolute terms.

Inventory systems usually consist of several production or distribution echelons. The echelons e.g. local retailers, central warehouses, plants, etc. build a supply chain within an inventory system. The multi-echelon structure is used to achieve a high responsiveness within a supply chain and to benefit from centralization advantages such as economies of scale. The supply channels of most inventory systems are hierarchical, which means transportation flows from one echelon to the next. There exist a vast number of models which allow more flexibility within an inventory system. Inventory pooling is used in echelons which goes along with lateral transshipments between the members of a particular echelon in case of stockouts. Direct deliveries or so called emergency transshipment are used to fill a demand from a higher level echelon in the system. Most quantitative analysis on supply chain management issues is dominated by the framework of multi-echelon systems where relationships between a single local warehouse and a single or multiple market is considered. Typically demand can be served by several members in one echelon, due to geographical allocation or different suppliers. Hence, lateral transshipment can also be interpreted as a redirection of demand, which is another option to increase the supply chain performance. A key distinction within the literature is between reactive and proactive policies. A reactive policy performs transshipments when a stockout or potential stockout occurs, by shipping either the whole demand or the number of items short from a different location. A proactive policy performs transshipments in advance of a potential stockout to avoid high penalty cost for not satisfying the demand immediately from stock.

Aside form the operational control of inventory systems, this thesis also examines some aspects of strategic decisions under competition. When studying operational decisions, a company is considered to be isolated from other competing companies in the market or it is assumed that the operational decisions do not influence the behavior of other market participants. The assumption of no competition do not hold for long term decisions on the strategic level. A firm distributing a product, with the exception of a monopolistic setup, is rarely a single player in the market. Usually a market consists of several different players competing with each other for the customers' demand. The market players (e.g. manufacturer, retailer, distributor, etc.) might be equal or differ in their market power. The market power is influenced by the company size, technology cost advantage, or regulations. Companies with higher market power can dictate the price or quality of the product in the market or are dominant in negotiations with their rivals. An isolated view of a company from other market participants regarding the positioning of the product, market entry or supply contracts would lead to suboptimal decisions. Observing the incentives and possible actions of other players in the market, the optimal strategic decisions can be found by anticipating the behavior of business rivals.

Usually the market rivals compete for the customers not only in price but also in quality of a product. Quality refers to different product characteristics which influence the perception of the customer and demand. These aspects might be either objective or subjective. Objective quality attributes are e.g. performance, features or durability of a product, whereas subjective quality aspects are e.g. design, aesthetics or brand image. The attributes which influence the attractiveness of the product to customers might have several dimensions, which can be aggregated to a single dimension (quality) of customers' interest. An utility-based customer enjoys an utility surplus from increasing quality and decreasing price. A product which maximizes customer's utility leads to the highest demand. From the companies perspective, a higher product price increases the revenue per item but decreases the overall demand. Also the increase of quality leads to higher demand but requires cost investments for the production process or advertising. The optimal strategic decisions regarding price and quality of a product are those that optimize the profit of the company.

The dissertation is organized in three main sections, Chapter 2 to 4. Chapter 2 and 3 study design and control of inventory systems applying problem-specific solution methods. Chapter 4 investigates competition and coordination within a supply chain and between the rivals in the market based on game theory. The thesis concludes with a summary and an outlook for further research in Chapter 5. The Chapters 2 to 4 in detail:

Chapter 2 studies lateral transshipments which are an effective strategy to pool inventories. A Semi-Markov decision problem formulation for proactive and reactive transshipments in a multi-location continuous review distribution inventory system with Poisson demand and one-for-one replenishment policy is presented. For a two-location model the monotonicity of an optimal policy is stated. In a numerical study, the benefits of proactive and different reactive transshipment rules are compared. The benefits of proactive transshipments are the largest for networks with intermediate opportunities of demand pooling and the difference between alternative reactive transshipment rules is negligible. Chapter 2 is based on a paper by Seidscher and Minner (2013).

Chapter 3 investigates reactive lateral transshipments between locations of the same echelon in a two echelon inventory system. However, lateral shipments from other locations are only beneficial if the next regular replenishment order will arrive later than the delivery through transshipment. It is therefore essential to have information about the supply pipeline to track the delivery of outstanding orders. New innovative real time supply information systems allow to find optimal decisions about the cost efficient supply sources from various alternatives. In this chapter we study a two-echelon inventory system with lateral transshipments between locations of the same echelon and emergency deliveries from upper echelons. A new analytic approach for the exact analysis of a single warehouse is presented. This model is then used to analyze the performance of an aggregated model for a two-echelon distribution network. After showing the model accuracy, we compare the results to models which ignore pipeline information and identify significant cost savings. Further, the use of pipeline information influences the optimal mix of flexibility options, i.e., pipeline information increases the fraction of demand satisfied by local warehouses and decreases the fraction of demand satisfied through lateral transshipment and direct deliveries. Chapter 3 is based on a working paper by Seidscher and Minner (2012).

Chapter 4 addresses a situation when a manufacturer of a branded product may face legal or illegal competition in quality and price from a rival manufacturer. For legal activities, this type of competition may come from, for example, store brands in supermarkets, parallel imports or generic drugs in pharmacies. Illegal competition (e.g. smuggled goods or counterfeits) is associated with a certain risk of interdiction for a rival manufacturer. First, different major problem categories from illegal and legal competition are presented, discussed and classified. For one specific case of illegal competition with two manufacturers and two retailers a four stage model is developed to examine the impact of such competition with regards to the risk of interdiction. We find that in some cases, the rival manufacturer is better off carrying the risk of interdiction then passing the risk to the retailer. Further, we observe that higher risk of interdiction decreases the quality of the illegal substitute and increases the quality of the branded product. Chapter 4 is based on a working paper by Seidscher et al. (2013).

Chapter 2

A Semi-Markov Decision Problem for Proactive and Reactive Transshipments between Multiple Warehouses

2.1 Introduction

Rapid changes in demand structure driven by emerging markets force companies to redesign their distribution networks. High service levels require shorter lead times which can be met only by new distribution locations close to emerging demand nodes. Angel et al. (2006) emphasize the growing importance of European regional distribution centers supplied directly from factories and report that the customers are not concerned about the location the goods are coming from, as long as their service requirements are met. This offers the advantage that customers do not necessarily have to be served by pre-selected locations but can be supplied by the location which is best for the supplier and therefore enables further inventory pooling opportunities. An inventory system which distributes a product to customers over a large geographical region often requires several locations. Theoretically, each market can be served by each warehouse if transshipment cost and service time are negligible. Markets face time constraints from service level agreements and, given geographical distances, markets usually can be served only by a limited number of warehouses within a predetermined service time. In this case, the network is divided into pooling groups that can share their inventories.

Lateral transshipments are a strategy for inventory sharing between the members of a pooling group. They are widely used when transshipment costs are relatively low compared to high holding costs and penalty costs for failing to meet demand. This cost structure is very common for expensive spare parts. However, transshipments are also used for parts in general. Two main types of lateral transshipments can be identified in the literature, proactive and reactive transshipments. They differ as to when transshipments are performed. Proactive transshipments are used in advance of stockouts to avoid future stockouts. Reactive transshipments are triggered when a warehouse faces a stockout by transferring inventory from another warehouse with stock on hand. Applying reactive lateral transshipments, the challenge is to find optimal base stock levels and the optimal source for a transshipment. Proactive transshipments additionally require a decision on the optimal moment for the stock movement.

Alternatively, lateral transshipments can be seen as a redirection of demand, a view that we will follow in this chapter. If the base warehouse is out of stock when a customer arrives, this results in a request for a reactive transshipment that can physically be either shipped to the customer via the original base warehouse or directly from the sending warehouse. Redirecting an incoming customer to another warehouse even if the base warehouse for this customer still has on-hand stock can be seen as a proactive transshipment. We study proactive and reactive transshipments between several markets and warehouses in different network structures under continuous review. The warehouses use one-for-one inventory replenishment policies which are common for expensive spare parts, because of high prices, low demand rates, and rather negligible fixed ordering costs (see Axsäter (2003)).

Lateral transshipments have been widely addressed in the literature. Recent overviews are provided in Wong et al. (2006) and Paterson et al. (2011). In the next four paragraphs, we summarize the main literature grouped by timing of review and by pro- or reactiveness of the shipments, i.e. proactive and period review, proactive and continuous, reactive and continuous, and reactive and periodic.

The research on proactive transshipments is dominated by periodic review, because the beginning and the end of each period provide a natural occasion for stock redistribution. Reactive transshipments were analyzed both in periodic and continuous review settings. One of the first redistribution models for proactive transshipments with a single period is presented in Allen (1958). All retailers together build a single pooling group where the transshipment times for redistribution are negligible and the stock can be rebalanced subject to transportation cost. Agrawal et al. (2004) consider a system where rebalancing is done at a predetermined time before all demand has been realized and present a dynamic programming formulation to determine optimal decisions. They conclude that a dynamic policy often outperforms a static one. Tagaras and Vlachos (2002) combine proactive transshipment and replenishment decisions. Two locations are supplied periodically by a central warehouse and the redistribution point in time has a large effect on performance. Furthermore, they investigate several different types of ordering and transshipment policies and find that type and variability of demand are the key parameters for an appropriate transshipment policy. Our approach differs in that it allows for transshipment decisions being triggered by every demand event in continuous time and allowing for an arbitrary number of multiple locations.

For proactive transshipments in a continuous time review setting, Van Wijk et al. (2011)

study a two location inventory network for repairable parts where both warehouses can hold back some inventory. They derive the optimal transshipment policy for both warehouses by a Markov Decision Process and present conditions under which the optimal policy is of a simple type. The two location nature does not require a source decision, which is one of our main contributions. Paterson et al. (2012) analyze a network with multiple warehouses that follow (R,Q) policies and complete pooling. They assume a significant fixed cost for carrying out a transshipment and combine proactive and reactive elements to develop a hybrid policy for stock redistribution. Triggered by a stockout as for reactive policies, the existence of fixed transshipment costs provides an incentive to not just clear a shortage but use the opportunity for further inventory rebalancing, an element of proactive transshipments. The hybrid policy significantly improves a pure reactive policy. As opposed to our approach, transshipments are triggered by a stockout and the proactive nature is only driven by fixed costs driven economies of scale.

For reactive lateral transshipments under continuous review and backordering, Lee (1987) proposes a model for pooling groups of identical locations where replenishments are one-for-one and compares three transshipment rules for the choice of the sending source: random rule, maximum stock on hand, and smallest number of outstanding orders. Lee finds that for all rules the use of transshipments leads to significant cost savings but there is no significant performance difference between the three rules. Axsäter (1990a) extends the model to non-identical locations within a pooling group but only analyzes a random transshipment rule. Dada (1992) analyzes a pooling model under lost sales where the inventory is controlled by one-for-one replenishment policy. He uses a Markov chain approach to calculate different system performance measures. Alfredsson and Verrijdt (1999) extend this model to include emergency deliveries from upper echelons. They further find that the assumption of exponentially distributed lead times is rather robust. Kutanoglu and Mahajan (2009) extend the model by Alfredsson and Verrijdt (1999) to time-based service constraints. They assume that immediate service rather than stock availability is the key to successful service. Therefore, they analyze a situation where the warehouses are required to satisfy a certain part of overall demand within a given time window. While the above references consider complete pooling, Reijnen et al. (2009) consider lateral transshipments for expensive spare parts with delivery time requirements in partial pooling structures. In case of a stockout, the transshipment source is obtained from a predetermined, ordered list (e.g., according to transshipment time). The authors present a Markov-analysis to evaluate the system performance and present algorithms to set the initial base-stock levels. They further integrate the use of lateral transshipments into the design of inventory networks which can lead to significant cost reductions. Lien et al. (2011) investigate the benefits of transshipments depending on the network configuration and show analytically that chain configurations known from manufacturing flexibility analysis are a superior choice in transshipment settings, too. Our work differs from the above as it generalizes to proactive transshipments and further allows for partial (endogenously determined) pooling between warehouses only constrained by maximum service time requirements.

Gross (1963) examines reactive lateral transshipments in a single period, single echelon setting with multiple locations and characterizes the optimal policy. Robinson (1990) analyzes a multi-period problem and derives the optimal solution for identical locations or a maximum of two different locations. For more than two non-identical warehouses, a linear programming based heuristic is presented. Herer et al. (2006) consider the problem of Robinson for a more general cost structure and provide an approach which guarantees convergence. Tagaras (1999) studies a network with three warehouses, complete pooling and (R,Q) replenishment policy and compares three transshipment rules: random rule, risk balancing and maximum stock on hand using simulation. He finds no significant differences between the transshipment rules. While the above references study inventory shipments after all demand in the period has been observed, Archibald (2006) analyzes a situation where continuous demand may trigger a reactive transshipment.

In the following, we present a hierarchical optimization approach with centralized control and information that is capable to solve both the proactive and the reactive transshipment problem. At the lower level, we use a Semi-Markov-Decision-Process (SMDP) to find the optimal transshipment policy. For a proactive strategy, this problem optimally determines the source location for every incoming customer demand, for the reactive policy the source is either optimally chosen or determined by a simple predetermined rule. The following rules will be considered:

1. Random: The source is chosen with uniform probability among the feasible warehouses with stock on hand.

2. Highest inventory level: The warehouse with maximum stock on hand among the potential sources is chosen.

3. Minimal transshipment cost: The warehouse with minimum transshipment cost is chosen.

4. Highest run-out-time: The warehouse with maximum run-out-time (fraction of stock on hand and mean arrival rate) is chosen.

At the upper level, we determine the required order-up-to-levels by a greedy heuristic anticipating the optimal demand and sourcing management by the SMDP for any transshipment policy.

Section 2.2 introduces model assumptions and notation while also presenting the SMDP formulation. In Section 2.3, we discuss the solution methods for solving small instances to optimality. For a simple two-location network, we characterize the optimal policy and provide structural results of the optimal transshipment policy. The numerical study in Section 2.4 compares the performance of proactive and different reactive transshipment rules for various network configurations and parameter settings. We summarize the results and indicate further research for this chapter in Section 4.5.

2.2 Model formulation

2.2.1 Notation and assumptions

We consider $|R|$ warehouses $r \in R$ selling a single product to $m \in M$ markets with Poisson demand. The demand rate for market m is denoted by λ_m and unsatisfied demand is assumed to be lost at a unit penalty cost l_m. Each warehouse r continuously reviews inventory and replenishes items from an external supplier with infinite supply according to an (S-1,S)-policy with base stock S_r. The replenishment lead time of warehouse r is exponentially distributed with mean $1/\mu_r$ and the replenishment unit cost is z_r. The storage of units at warehouse r is subject to holding costs h_r per unit per unit of time.

Due to service level requirements and geographical distance, each market m can only be served by a subset of warehouses W_m, where the warehouses in the list W_m are ordered by increasing distance to market m. We assume that longer distance corresponds to higher shipment cost to fulfill the service level requirement. The closest source of market m is called the base warehouse $w(m)$ which is by definition the first warehouse in the list W_m. The base warehouse serves the respective market without additional costs, i.e., delivery costs are included in the unit replenishment cost. If a market m is served from another source r than the base warehouse, this can be seen as a transshipment with an additional shipment cost t_{mr}. Although we do not model delivery and transshipment times explicitly, the assignment of markets to warehouses through W_m accounts for such a delivery time requirement. Therefore, a warehouse which is not an element of W_m can be interpreted as a location whose delivery time does not fulfill the service time requirements of market m. Furthermore, we assume $(l_m - z_r) > 0 \; \forall r \in W_m$, i.e. a warehouse is only a candidate source for a market if the replenishment cost is smaller than the unit lost sales penalty. Table 2.1 summarizes the notation.

2.2.2 States and decisions

In the following, we model the second-stage transshipment problem for given base-stock levels S_r as a SMDP. A state represents a combination of inventory levels for all warehouses. There is a set of different decision alternatives in each state. If one alternative has been chosen, the system transfers to another state with a corresponding probability after a random period of time. The objective is to minimize expected long-run average costs under centralized information and decision. Within the general SMDP framework, we specify the differences in decision space and transition probabilities for proactive and reactive strategies. For ProTS, the optimal decisions, i.e. the sending warehouse for an incoming demand of a market, have to be found in each state. In ReTS, decisions are limited to states where the base warehouse for a customer is out of stock, i.e. a reactive transshipment is triggered. In that case, the sending source can either be determined optimally from all candidate warehouses (Rule 5) or be fixed in advance by a simple rule. In the latter cases, the SMDP formulation only serves to evaluate the long-run average

Table 2.1 List of notation

R	Set of warehouses
M	Set of markets
I	Set of states
λ_m	Customer arrival rate at market m
μ_r	Replenishment rate at warehouse r
S_r	Base stock level at warehouse r
h_r	Holding cost per unit and unit of time at warehouse r
z_r	Unit replenishment cost at warehouse r
l_m	Penalty cost per unit of lost demand at market m
t_{mr}	Transshipment cost per unit to market m from warehouse r
W_m	List of possible warehouses for market m
W_{im}	List of possible warehouses for market m in state i
$w(m)$	Base warehouse of market m
i_r	Inventory level at warehouse r in state i
q_{im}	Sending warehouse for market m in state i
Q_i	Vector of decisions in state i
T_i	Set of feasible transition states from state i
$p_{ij}(Q_i)$	Transition probability from state i into state j under decision Q_i

cost of a policy.

Let $i = (i_1, \cdots, i_{|R|}), i \in I$ be a state that contains the inventory levels of all warehouses and $0 \leq i_r \leq S_r$ where I is the set of all states. A decision $q_{im} = r, r \in W_m$ assigns a potential next demand from market m to warehouse r in state i. The vector of all decisions in state i is expressed as a list of warehouses that serve the particular markets, $Q_i = (q_{i1}, \cdots, q_{i|M|})$. The set of possible supply warehouses for every market can change in each state if some warehouses are out of stock. Warehouses which are out of stock are not considered as potential sources, except the base warehouse. Assigning a demand to the base warehouse which is out of stock is interpreted as a lost sale. For a given state i and market m we denote the list of potential sources as $W_{im} = \{r \in W_m \mid r = w(m) \vee i_r > 0\}, i \in I, m \in M$. We note that the proactive decision problem could alternatively (at the cost of a more complex exposition) be modeled by an after demand arrival state representation where we do not have to make a potential sourcing decision in advance of a demand arrival but can include the market of the next demand into another state variable.

Proactive Transshipments For ProTS, all markets in every state can be served by every warehouse among the possible sources which have stock on hand or the base warehouse.

$$q_{im} \in W_{im}, \forall m \in M,\ \forall i \in I \tag{2.1}$$

The definition in (2.1) includes decisions q_{im} which direct demand to a base warehouse with zero inventory even though other potential sources still have stock on hand if this is not economically beneficial, i.e., when the penalty cost is relatively low compared to the transshipment and replenishment cost.

Reactive Transshipments ReTS initiates transshipments only if the base warehouse is out of stock. If the base warehouse has stock on hand, the decision is $q_{im} = w(m), \forall i \in I,\ m \in M$. If the base warehouse is out of stock, either a transshipment is triggered or the demand is lost. The source for the transshipment, q_{im}, is determined by one of the following rules.

1. Random rule:
   If the base warehouse is out of stock, the source is randomly chosen from all feasible supply sources with positive stock. Market m in state i has an opportunity for a transshipment only if the list of potential sources contains more than one element, $|W_{im}| > 1$.

2. Highest inventory rule:
   If the base warehouse is out of stock, the warehouse with the highest inventory among the potential sources W_{im} is chosen as source for market m in state i.

$$q_{im} = \underset{r \in W_{im}}{\operatorname{argmax}} \{i_r\} \tag{2.2}$$

3. Minimal transshipment cost rule:
   The warehouse with minimum transshipment cost is chosen as sending location for lateral transshipment for market m in state i among the potential sources W_{im}.

$$q_{im} = \underset{r \in W_{im}}{\operatorname{argmin}} \{t_{mr}\} \tag{2.3}$$

4. Highest run-out-time rule:
   The run-out-time ρ_{ir} of warehouse r in state i is defined as current inventory level divided by the sum of the demand rates of all markets being primarily assigned to warehouse r, $\rho_{ir} = i_r / \sum_{m \in M | r = w(m)} \lambda_m$.

$$q_{im} = \underset{r \in W_{im}}{\operatorname{argmax}} \{\rho_{ir}\} \tag{2.4}$$

5. Optimal decision rule:
   This rule chooses the source so as to minimize expected cost. Therefore, the decision space comprises all possible warehouses W_m with stock on hand in state i and the base warehouse if the demand should not be satisfied.

$$q_{im} \in W_{im} \tag{2.5}$$

The transshipment rules 2, 3, 4, and 5 can have ties between the warehouses. In that case the first warehouse in the list of potential sources W_{im} is chosen (which, by assumption of ordering, has the lowest transshipment cost) but, of course, this can also be assumed differently.

2.2.3 State transitions

A state transition from state i to state j is either triggered by a demand or the arrival of an outstanding order. Let T_i^A denote the average time of being in state i. As the stochastic process of the time until the next event is again a Poisson process with a rate equal to the sum of the individual demand rates and the arrival rates of all outstanding orders, $T_i^A = 1/\left(\sum_{m\in M}\lambda_m + \sum_{r\in R}(S_r - i_r)\mu_r\right)$. $\lambda_m T_i^A$ is the probability that the next event is a demand at market m. Each event leads either to an increase, decrease, or no inventory change (in case of a lost sale) at one warehouse. The set of feasible transition states from state i, T_i is therefore grouped into three subsets $\{T_i^+, T_i^0, T_i^-\}$. The subset T_i^+ comprises all states triggered by the arrival of an order. The inventory at a single warehouse increases by one unit while the other warehouses remain at the same inventory level. The subsets T_i^0 and T_i^- are triggered by demand arrival. T_i^0 represents the set where a demand is not satisfied and the system remains in the same state. T_i^- is a set of states where the inventory at a single warehouse decreases by one unit and all other warehouses retain their inventory levels. Next we define those states $j \in I$ which result from a transition that is triggered by an event (demand or arrival) at warehouse r^*.

$$\begin{aligned} T_i^+(r^*) &= \{j \in I \mid j_{r^*} = i_{r^*} + 1; j_r = i_r \ \forall r \neq r^*\} && (2.6)\\ T_i^0(r^*) &= \{j \in I \mid i_{r^*} = 0; j_r = i_r \ \forall r \neq r^*\} && (2.7)\\ T_i^-(r^*) &= \{j \in I \mid j_{r^*} = i_{r^*} - 1; j_r = i_r \ \forall r \neq r^*\} && (2.8)\end{aligned}$$

Then, $T_i^+ = \cup_{r^*\in R} T_i^+(r^*)$, $T_i^0 = \cup_{r^*\in R} T_i^0(r^*)$, and $T_i^- = \cup_{r^*\in R} T_i^-(r^*)$.

By considering all events for each state under decision set Q_i, we identify the transition probabilities $p_{ij}(Q_i)$ that the system will be in state j under decision set Q_i leaving state i.

$$p_{ij}(Q_i) = T_i^A \cdot \begin{cases} \sum\limits_{r^*\in R} \sum_{m\in M | q_{im}=r^* \vee i_{r^*}>0} \lambda_m & \text{if } j \in T_i^- \\ \sum\limits_{r^*\in R} \sum_{m\in M | q_{im}=r^* \vee i_{r^*}=0} \lambda_m & \text{if } j \in T_i^- \\ \sum_{r^*\in R}(S_{r^*} - i_{r^*})\mu_{r^*} & \text{if } j \in T_i^+ \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

The first case covers all transitions with a demand that depletes inventory, the second case covers all demands resulting in a lost sale, and the last case the arrival of an outstanding order.

ReTS with random transshipment source requires a modification of the first case of above transition probabilities. Let $p_i(m) = 1/(|W_{im}| - 1)$, where $|W_{im}| - 1$ is the number of possible sources for market m which have stock on hand in state i. A market m will be served by an arbitrary non-base warehouse with probability $p_i(m)$.

$$p_{ij}(Q_i) = \left(\sum_{m \in M | w(m) = r^* \vee |W_{im}| = 1} \lambda_m + \sum_{m \in M || W_{im}| > 1} \lambda_m p_i(m) \right) \cdot T_i^A \quad \text{if} \quad T_i^-(r^*) = j \tag{2.10}$$

2.2.4 Cost function

The holding cost in state i is determined by the current inventory levels i_r for each warehouse and the time the system resides in state i which is independent of Q_i.

$$y_i = \sum_{r \in R} h_r \cdot i_r \cdot T_i^A \tag{2.11}$$

A penalty cost l_m occurs when the decision Q_i assigns the demand of market m to the base warehouse $r = w(m)$ which is out of stock ($i_r = 0$) weighted with the probability that the demand is for market m.

$$l_i(Q_i) = \sum_{m \in M | q_{im} = w(m) \wedge i_{w(m)} = 0} l_m \cdot \lambda_m \cdot T_i^A \tag{2.12}$$

The cost of placing a replenishment order is associated with satisfying a demand by a warehouse. Therefore, we consider those state transitions where the inventory level reduces.

$$z_i(Q_i) = \sum_{r \in R} z_r \cdot p_{i, T_i^-(r)}(Q_i) \tag{2.13}$$

The transshipment cost occurs only if market m is served by a source other than the base warehouse $w(m)$.

$$t_i(Q_i) = \sum_{r \in R \setminus w(m)} \sum_{m \in M | q_{im} = r} t_{mr} \cdot \lambda_m \cdot T_i^A \tag{2.14}$$

For ReTS with random rule this again requires a modification.

$$t_i(Q_i) = \sum_{m \in M || W_{im}| > 1} \lambda_m \sum_{r \in W_{im} \setminus w(m)} t_{mr} \cdot p_i(m) \cdot T_i^A \tag{2.15}$$

The expected cost for being in state i under decision Q_i is

$$e_i(Q_i) = y_i + l_i(Q_i) + z_i(Q_i) + t_i(Q_i), \ \forall i \in I. \tag{2.16}$$

2.3 Optimal solution

2.3.1 Solution algorithm

To solve an infinite horizon SMDP we must find decisions Q_i for every state $i \in I$ such that the average cost per unit of time C for given base stock levels S_r is minimized. This problem can be solved by linear programming, policy iteration or value iteration, see Tijms (2003). We use linear programming to find the minimal expected average cost per unit of time C^*. To find the optimal base stock levels S_r for each warehouse, we use a heuristic algorithm based on a greedy approach as suggested in Kranenburg and van Houtum (2009). This method proceeds iteratively and in each iteration places or displaces a unit of stock at the warehouse that provides the biggest improvement on total expected cost. This iterative procedure requires the solution of the Linear Program in each iteration step.

The number of iterations required for finding the minimal cost is already high for relatively small networks and the computation time increases rapidly for each iteration with the level of base stocks. Thus, in order to limit the required computation time it is crucial to find a good starting solution. Therefore, we separately optimize base stock levels for each warehouse as an initial solution for the greedy algorithm. Finding cost minimizing base stock levels for each warehouse can be considered as optimizing an $M|M|S|S$ queue where the steady-state probabilities can be obtained analytically. Let j be a number of units, $0 \leq j \leq S$ and p_j the steady-state probability that the warehouse has j units on hand (for brevity, we omit that all probabilities p_j are functions of S). Then,

$$\frac{1}{p_S} = \sum_{j=0}^{S} \frac{(\lambda/\mu)^j}{j!} \quad \text{and} \quad p_j = \frac{(\lambda/\mu)^{S-j}}{(S-j)!} p_S, \; j = 0, ..., S-1. \tag{2.17}$$

The total cost for a given base stock level includes all relevant costs from above except the transshipment cost. Let $\lambda_r = \sum_{m \in M | r = w(m)} \lambda_m$ denote the total demand rate of all markets with r as the base warehouse.

$$\begin{aligned} C_r(S) &= h_r \sum_{j=0}^{S} j p_j + p_0 \sum_{m \in M | r = w(m)} l_m \lambda_m + z_r \lambda_r (1 - p_0) \\ &= h_r S + \lambda_r (h_r/\mu_r - z_r) p_0 + p_0 \sum_{m \in M | r = w(m)} l_m \lambda_m + \lambda_r (z_r - h_r) \end{aligned} \tag{2.18}$$

We iteratively place an extra unit of inventory at the warehouse until no further cost reduction can be achieved. This procedure is exact as $C_r(S)$ is convex in S as $h_r S$ is linear in S and p_0 is the well known Erlang loss function being a convex function of S (see Messerli, 1972).

2.3.2 Numerical illustration and structural properties of an optimal policy

We illustrate the optimal sourcing decisions for given base stock levels in a network with two warehouses and three markets as depicted in Figure 2.1. The market to warehouse assignment is $W_1 = \{1\}$, $W_2 = \{1, 2\}$ and $W_3 = \{2\}$. The connections to the base warehouse are illustrated with a solid line and the other opportunities with a dashed line.

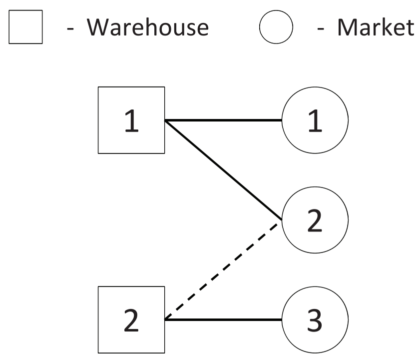


Figure 2.1 Example network

For a base case with identical warehouses and parameter values $\lambda = 0.3$, $\mu = 0.2$, $h = 0.2$, $l = 20$, $z = 1$, $t_{22} = 0.5$ and $S_1 = S_2 = 5$ (although these are not cost optimal levels), we compute the optimal sourcing policy for market 2 which is a function of the inventory levels of both retailers, as illustrated in Figure 2.2. "No transshipment" indicates that market 2 is served by the base warehouse and "Transshipment" corresponds to a shipment from warehouse 2.

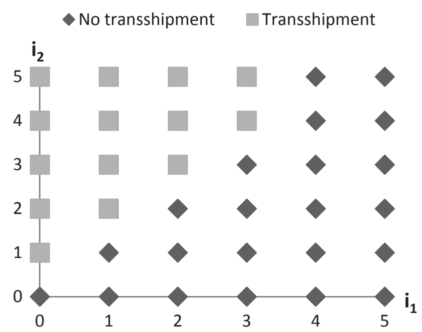


Figure 2.2 Optimal policy of market 2 for a base case with $S_1 = S_2 = 5$

The optimal policy for the considered base case is to use transshipments already when there are small differences in inventory levels in order to avoid future stockouts with additional penalty cost when demands for single-sourced markets 1 or 3 will occur. Note that a reactive policy has transshipments only if the base warehouse 1 is out of stock.

An increase of the base stock level at the base warehouse 1 for market 2 increases the separation line between the two decision regions (see Figure 2.3a). Decreasing the penalty cost, we observe that transshipments are preferably used when inventory differences are large (see Figure 2.3b). Different arrival rates $\lambda_1 = \lambda_3 = 0.7$, $\lambda_2 = 0.1$, i.e., higher stockout

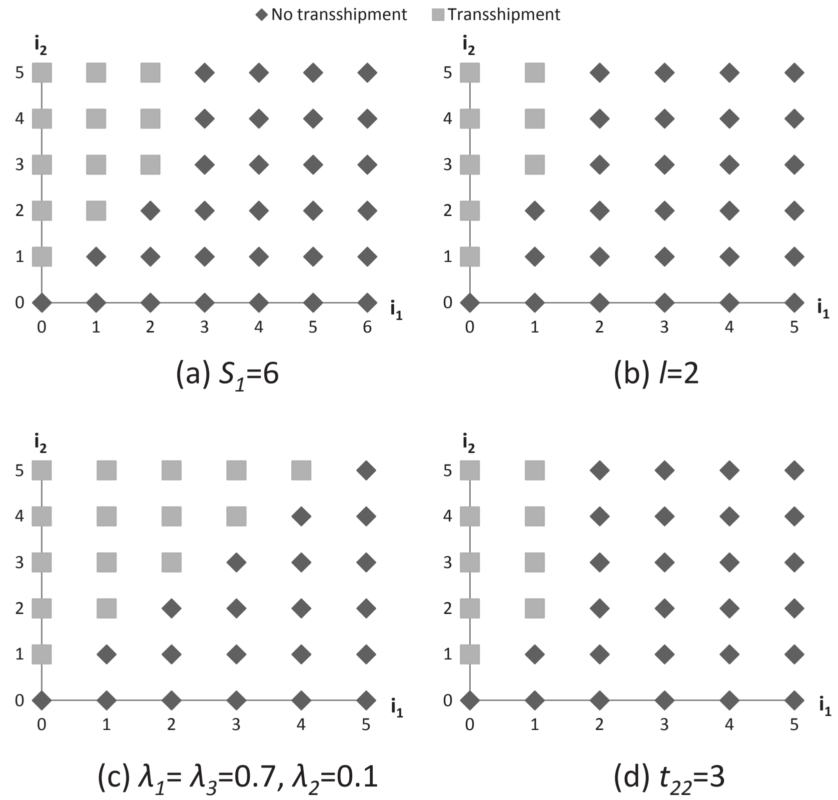


Figure 2.3 Optimal policy of market 2 varying single parameter

probability at the base warehouse, (see Figure 2.3c) show that transshipments are used more often. A higher transshipment cost (see Figure 2.3d) increases the separating line between the two decisions, i.e. reduces the attractiveness of transshipments.

For this particular network, we analyze properties of an optimal policy. The data for the decision process simplifies to the following. Let (i_1, i_2) denote the inventory state and q_2 the only decision required. Applying uniformization using the average time in state i,

$$T_i^A = \frac{1}{(S_1 - i_1)\mu_1 + (S_2 - i_2)\mu_2 + \sum_{i=1}^{3} \lambda_i} \tag{2.19}$$

we find the direct cost in state i taking decision q_2, $\tilde{e}_i(q_2)$ to be

$$\tilde{e}_i(q_2 = 1) = \begin{cases} h_1 i_1 + h_2 i_2 + z_1(\lambda_1 + \lambda_2) + z_2 \lambda_3 & i_1 > 0, i_2 > 0 \\ h_1 i_1 + z_1(\lambda_1 + \lambda_2) + l_3 \lambda_3 & i_1 > 0, i_2 = 0 \\ h_2 i_2 + z_2 \lambda_3 + l_1 \lambda_1 + l_2 \lambda_2 & i_1 = 0, i_2 > 0 \\ l_1 \lambda_1 + l_2 \lambda_2 + l_3 \lambda_3 & i_1 = i_2 = 0 \end{cases} \tag{2.20}$$

$$\tilde{e}_i(q_2 = 2) = \begin{cases} h_1 i_1 + h_2 i_2 + z_1 \lambda_1 + z_2(\lambda_2 + \lambda_3) + t_{22} \lambda_2 & i_1 > 0, i_2 > 0 \\ h_1 i_1 + z_1 \lambda_1 + l_2 \lambda_2 + l_3 \lambda_3 & i_1 > 0, i_2 = 0 \\ h_2 i_2 + z_2(\lambda_2 + \lambda_3) + l_1 \lambda_1 + t_{22} \lambda_2 & i_1 = 0, i_2 > 0 \\ l_1 \lambda_1 + l_2 \lambda_2 + l_3 \lambda_3 & i_1 = i_2 = 0 \end{cases} \tag{2.21}$$

The transition parameters $\tilde{p}_{ij}(q_2)$ are

$$\tilde{p}_{ij}(q_2=1) = \begin{cases} \lambda_1+\lambda_2 & j=((i_1-1)^+,i_2) \\ \lambda_3 & j=(i_1,(i_2-1)^+) \\ (S_1-i_1)\mu_1 & j=(i_1+1,i_2) \\ (S_2-i_2)\mu_2 & j=(i_1,i_2+1) \end{cases}, \tag{2.22}$$

$$\tilde{p}_{ij}(q_2=2) = \begin{cases} \lambda_1 & j=((i_1-1)^+,i_2) \\ \lambda_2+\lambda_3 & j=(i_1,(i_2-1)^+) \\ (S_1-i_1)\mu_1 & j=(i_1+1,i_2) \\ (S_2-i_2)\mu_2 & j=(i_1,i_2+1) \end{cases}. \tag{2.23}$$

Let $V(i)$ denote the per unit of time value of being in state i taking the optimal decision. Then,

$$V(i) = \min_{q_2\in\{1,2\}} \left\{ \tilde{e}_i(q_2) + \sum_{j\in T_i} \tilde{p}_{ij}(q_2)V(j) \right\}. \tag{2.24}$$

Let a threshold function $\bar{i}_2(i_1)$ define the minimum inventory level of warehouse 2 given the inventory level of warehouse 1 where it is optimal to source a market 2 demand from warehouse 2 (transshipment). The threshold function is characterized by

$$\bar{i}_2(i_1) = arg\min_{i_2} \left\{ z_2 + t_{22} + V(i_1,i_2-1) \le z_1 + V(i_1-1,i_2) \right\}. \tag{2.25}$$

Proposition: The optimal policy is monotone in the value of the state variables.

Proof: Monotonicity of the optimal decision (see Puterman (1994), Koole (2006)) of the policy in all the states is assured if the following conditions are satisfied.

i.) the direct cost is an increasing, superadditive function in the state, which can be observed from (2.20) and (2.21) and

ii.) the probability to leave a state to a state with more inventory (i.e. arrival of an outstanding order) is an increasing, superadditive function of the state variables. This holds because

$$\begin{aligned} &\frac{(S_1-i_1)\mu_1+(S_2-i_2)\mu_2}{(S_1-i_1)\mu_1+(S_2-i_2)\mu_2+\sum_{k=1}^{3}\lambda_k} + \frac{(S_1-i_1-1)\mu_1+(S_2-i_2-1)\mu_2}{(S_1-i_1-1)\mu_1+(S_2-i_2-1)\mu_2+\sum_{k=1}^{3}\lambda_k} \\ \le\ &\frac{(S_1-i_1-1)\mu_1+(S_2-i_2)\mu_2}{(S_1-i_1-1)\mu_1+(S_2-i_2)\mu_2+\sum_{k=1}^{3}\lambda_k} + \frac{(S_1-i_1)\mu_1+(S_2-i_2-1)\mu_2}{(S_1-i_1)\mu_1+(S_2-i_2-1)\mu_2+\sum_{k=1}^{3}\lambda_k} \end{aligned}$$

q.e.d.

As a consequence, the threshold function $\bar{i}_2(i_1)$ indicating the region without transshipments from warehouse 2, leaving all other parameters and decisions unchanged, is increasing in the inventory level i_1. Further, from our numerical investigation, we find that the threshold function

- increases with the base-stock level of the base warehouse 1, and decreases with the base-stock level of warehouse 2,
- increases with the base warehouse replenishment rate μ_1 and decreases with the replenishment rate μ_2 of warehouse 2,
- decreases with the base warehouse replenishment cost z_1 and increases with the backup warehouse replenishment cost z_2,
- decreases with the base warehouse holding cost h_1 and increases with the backup warehouse holding cost h_2,
- decreases with the lost sales penalties of markets 1 and 2 and increases with market 3's penalty l_3,
- decreases with the demand rates of markets 1 and 3 and increases with the demand rate λ_2 of market 2,
- increases with the transshipment cost t_{22}.

2.4 Numerical study

We compare proactive and reactive transshipments and investigate the performance of different transshipment rules. To evaluate the different transshipment strategies and source selection rules, we consider several different networks with complete and partial pooling. In order to quantify the cost differences, we define

$$\Delta_x = \frac{C^x_{ReTS} - C_{ProTS}}{C_{ProTS}} \cdot 100, \tag{2.26}$$

which states the percentage of total cost savings of ProTS compared to ReTS under transshipment rule x. Furthermore, we define Δ_{No} as the improvement of ProTS compared to a system with no transshipments when all markets are assigned to their base warehouse. In order to quantify the pooling level of a network, we define an inventory pooling factor ϑ as

$$\vartheta = \frac{\sum_{m=1}^{|M|}(|W_m| - 1)\lambda_m}{\sum_{m=1}^{|M|}(|R| - 1)\lambda_m}. \tag{2.27}$$

The numerator is the sum of pooled arrival rates (weighted by the number of possible sources except the base warehouse) and the denominator is the maximum pooling opportunity (sum of arrival rates multiplied by the number of all warehouses except the base warehouse).

2.4.1 Experimental design

Due to the computational burden of an exact analysis of transshipment policies and because we are mainly interested in the general benefits of proactive over reactive transshipments, we investigate rather small networks with three warehouses and three or six markets, respectively. Markets can be differently assigned to one or multiple warehouses. For each network type, we generate 5 different instances by randomly (uniformly) assigning markets to warehouses using a given number of links. Besides the randomly generated networks, we considered pre-specified networks with complete pooling, networks exhibiting a chaining structure, and a particular partial pooling network. The latter structure was used to analyze the improvements between proactive and reactive transshipments for different pooling factors. By varying the arrival rates for markets assigned to warehouses that exhibit complete pooling or no pooling, we generate a set of parameter values ϑ. Three warehouses and six markets were investigated for 9, 12, 15 and 18 links (complete pooling). Figure 2.4 shows the investigated networks with pre-specified structure.

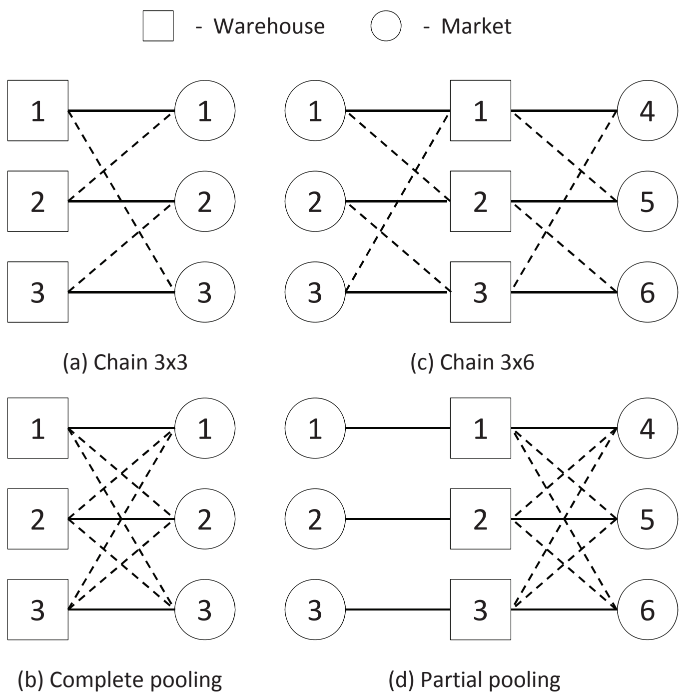


Figure 2.4 Network structures

We assume that all warehouses have identical cost parameters and expected replenishment times. The unit replenishment cost is fixed to one ($z = 1$). All other parameters are varied. Defining the transshipment cost t_{mr}, we follow the idea that higher geographical distance corresponds to higher transshipment cost. Specifically, we assume that transshipment cost doubles with each additional warehouse in the set of possible sources of a market.

$$t_{mr} = \begin{cases} t_f \cdot 2^{(y-1)} \mid r = r_y, \forall y \in |W_m| & \text{if} \quad y \geq 1 \\ 0 & \text{if} \quad y = 0 \end{cases} \tag{2.28}$$

where y is the position index of warehouse r in the list W_m and $y = 0$ denotes the base warehouse with position index zero. To avoid a bias in the definition of transshipment cost, the index positions are assigned in such a way that every warehouse appears at each position equally often. This definition limits the number of transshipment cost parameters to t_f only. Demand rates are specified by choosing an average demand rate λ and a range parameter σ such that $\lambda\sigma$ is the largest difference between arrival rates. The $|M|$ arrival rate values for the markets are equidistantly spaced over the interval $[\lambda(1-0.5\sigma), \lambda(1+0.5\sigma)]$ and randomly (uniform) assigned to the markets.

Table 2.2 Input parameter values

Parameter	Value
Arrival rate (λ)	0.1, 0.3, 0.5, 0.7
Arrival rate interval (σ)	0, 0.2, 0.5, 0.9
Replenishment rate (μ)	0.2, 0.4
Holding cost (h)	0.1, 0.2, 0.4
Penalty cost (l)	10, 20, 30, 50
Transshipment cost factor (t_f)	0.25, 0.5, 1, 2

Our numerical study used a full factorial design of the parameters shown in Table 2.2. For chain networks, we limited the number of instances by assuming identical arrival rates ($\sigma = 0$) and we only considered $\lambda \in \{0.1, 0.3, 0.5\}$. Networks with three warehouses and six markets were only considered for $\lambda \in \{0.1, 0.3, 0.5\}$.

The linear programs were solved using the solver XPRESS-MP version 7.2 on a PC with Intel(R) Core(TM) i7-2600 3.40 GHz and 16GB RAM.

2.4.2 Results

For the three warehouses and three markets (see Tables 3 and 4), we observe an average saving of ProTS compared to the situation without transshipments of 35.41% for the complete pooling network (1536 instances), 17.14 % for a network with 6 links (7680 instances) and 22.92% for the chaining network (288 instances). The average benefit only increases slightly to 23.94% for a larger chaining network with 5 rather than 3 locations and decreases slightly to 21.05% for the chaining network with six markets (Figure 4c). Because of our identical cost assumption, proactive transshipments are not beneficial compared to optimal reactive transshipments in a complete pooling network. In the chaining network, $\Delta_5 = 1.20\%$ and in the network with 6 links, $\Delta_5 = 0.77\%$. No benefit for a proactive approach in a complete pooling network is an intuitive result, since, in the setup with identical cost parameters and complete pooling, the decision about the optimal moment for transshipment is irrelevant because all warehouses can be considered as one source.

Table 2.3 shows the average cost savings of proactive transshipments over no transshipments and over optimally chosen reactive transshipments for chaining networks with 3, 5 warehouses and 6 markets for each parameter value. We observe that the overall benefits of transshipments (Δ_{No}) decrease with the average demand rate λ. In this case, however, the relative benefit of a proactive over a reactive strategy (Δ_5) increases. For increasing average replenishment lead time, holding cost and penalty cost both the value of transshipments and the advantage of proactive transshipments increase, whereas for increasing transshipment costs, both values decrease.

Table 2.3 Average savings in chain networks with (warehouses x markets) (in percent)

		Δ_{No}			Δ_5		
Parameter	Value	(3×3)	(5×5)	(3×6)	(3×3)	(5×5)	(3×6)
λ	0.1	25.07	26.05	24.00	0.95	1.78	1.12
	0.3	22.64	23.70	20.53	1.27	2.38	1.34
	0.5	21.06	22.08	18.62	1.38	2.86	1.37
μ	0.2	23.59	24.58	21.81	1.37	2.76	1.48
	0.4	22.26	23.30	20.29	1.02	1.92	1.07
h	0.1	19.57	20.78	16.84	0.97	1.92	0.99
	0.2	23.41	24.56	21.39	1.22	2.36	1.28
	0.4	25.79	26.49	24.91	1.41	2.74	1.56
l	10	18.81	19.06	17.36	0.47	1.11	0.62
	20	22.62	23.68	20.72	1.02	2.06	1.12
	30	24.19	25.48	22.21	1.40	2.66	1.46
	50	26.08	27.52	23.90	1.90	3.52	1.92
t_f	0.25	24.98	26.41	23.03	1.90	3.75	1.94
	0.5	24.06	25.22	22.17	1.44	2.82	1.54
	1	22.56	23.50	20.71	0.96	1.84	1.06
	2	20.09	20.64	18.30	0.49	0.95	0.56
Overall		22.92	23.94	21.05	1.20	2.34	1.28

Table 2.4 shows the average savings of ProTS over different reactive rules detailed by single parameter values for a network with 6 links and the complete pooling network with three warehouses and three markets. The results illustrate that in a complete pooling network, ReTS with minimal transshipment cost selection rule provides optimal decisions. This result is evident, since transshipment cost is the only cost factor which is subject to optimization choosing the sending location if the other cost parameters are identical between the warehouses. In networks with 6 links, rule 3 can on average be improved by 1.20% if we use proactive transshipments and is inferior to rules 2, 4 and 5. However, the general differences between the reactive transshipment rules are rather negligible.

Table 2.4 Average savings in a network with three warehouses and markets (in percent)

		6 links						complete pooling					
Parameter	Value	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5
λ	0.1	20.33	0.89	0.49	0.65	0.62	0.31	45.9	1.09	0.8	0	0.97	0
	0.3	17.31	1.39	0.71	1.11	0.66	0.60	35.92	1.13	0.82	0	0.9	0
	0.5	15.81	1.59	0.93	1.33	1.03	0.85	31.41	1.06	0.72	0	0.82	0
	0.7	15.10	1.98	1.41	1.72	1.40	1.34	28.39	1.02	0.71	0	0.84	0
σ	0	16.57	1.70	1.09	1.46	1.17	0.97	37.59	1.3	0.83	0	0.83	0
	0.2	16.68	1.73	1.17	1.47	1.19	1.15	37.69	1.26	0.89	0	1.13	0
	0.5	16.92	1.58	1.03	1.33	1.08	0.83	36.23	1.14	0.85	0	0.99	0
	0.9	18.40	0.84	0.23	0.58	0.30	0.15	30.11	0.61	0.48	0	0.57	0
μ	0.2	17.14	1.32	0.76	1.08	0.82	0.74	36	1.02	0.71	0	0.82	0
	0.4	17.15	1.59	1.01	1.33	1.04	0.81	34.81	1.14	0.82	0	0.95	0
h	0.1	13.94	1.09	0.66	0.83	0.69	0.56	27.54	1.14	0.83	0	0.97	0
	0.2	17.49	1.43	0.83	1.16	0.89	0.74	36.08	1.12	0.79	0	0.91	0
	0.4	19.99	1.87	1.16	1.62	1.20	1.02	42.6	0.97	0.67	0	0.76	0
l	10	13.77	0.82	0.49	0.56	0.52	0.41	26.95	1.33	0.93	0	1.07	0
	20	16.93	1.23	0.72	0.97	0.75	0.62	34.69	1.13	0.8	0	0.92	0
	30	18.26	1.68	0.99	1.42	1.07	0.89	38.11	1	0.72	0	0.82	0
	50	19.59	2.12	1.33	1.86	1.36	1.18	41.87	0.85	0.61	0	0.71	0
t_f	0.25	19.53	2.35	1.36	2.12	1.42	1.12	39.73	0.34	0.24	0	0.27	0
	0.5	18.42	1.62	1.03	1.40	1.07	0.94	37.96	0.66	0.46	0	0.53	0
	1	16.52	1.22	0.79	1.00	0.83	0.74	34.7	1.22	0.86	0	0.99	0
	2	14.10	0.66	0.36	0.31	0.40	0.28	29.24	2.07	1.5	0	1.73	0
Overall		17.14	1.46	0.88	1.20	0.93	0.77	35.41	1.08	0.76	0	0.88	0

The results for networks with three warehouses and six markets are summarized in Table 2.5 (5760 instances for the networks with 9, 12, and 15 links, 1152 instances for the complete pooling networks). We observe that, increasing the number of links in the network, the average savings compared to no transshipments increase, $\Delta_{No} = 12.21\%$ for 9 links, $\Delta_{No} = 17.17\%$ for 12 links, $\Delta_{No} = 22.96\%$ for 15 links and $\Delta_{No} = 26.75\%$ for complete pooling. Potential savings of ProTS relative to ReTS with rule 5 first increase from $\Delta_5 = 1.17\%$ for 9 links to $\Delta_5 = 3.18\%$ for 12 links and then decrease to $\Delta_5 = 1.83\%$ for 15 links and finally to $\Delta_5 = 0.00\%$ for complete pooling network. The value of proactive transshipments increases with faster replenishment lead times, lower transshipment costs, higher lost sales penalty and inventory holding costs, and similarity in demand rates.

We observe that improvements of proactive transshipments strongly depend on the number of links in a network. We use the partial pooling network (Figure 4d) to determine several pooling parameter values ϑ. We generated in total 6144 instances using the parameters from Table 2, with the average demand rates combining 4 values for the group of pooled and 4 values for the unpooled warehouses. This results in 16 combination, 4 of them having a pooling factor of 0.5. The first line in Table 6 is trivial, the values of the last line are obtained from Table 4 for complete pooling. The results in Table 2.6 show that the savings of ProTS relative to ReTS increase until $\vartheta = 0.58$ and then decrease. The second effect is that, up until $\vartheta = 0.70$, transshipment rule 3 shows higher potential savings compared to transshipment rules 2 and 4, which corresponds to inferior rule. From $\vartheta = 0.75$, rule 3 shows lower cost savings and is superior to transshipment rules 2 and 4.

Table 2.5 Average savings in a network with three warehouses and six markets (in percent)

		9 links						12 links					
Parameter	Value	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5
λ	0.1	13.76	0.88	0.84	0.84	0.85	0.84	19.35	3.20	2.70	2.73	2.71	2.63
	0.3	11.93	1.29	1.24	1.24	1.25	1.24	16.78	3.87	3.40	3.44	3.39	3.32
	0.5	10.93	1.48	1.43	1.43	1.44	1.43	15.37	4.12	3.65	3.70	3.66	3.59
σ	0	12.19	1.21	1.18	1.19	1.18	1.16	16.24	3.93	3.55	3.57	3.54	3.53
	0.2	12.21	1.24	1.19	1.20	1.22	1.19	16.44	4.02	3.63	3.64	3.62	3.59
	0.5	12.22	1.24	1.21	1.21	1.19	1.18	16.88	3.87	3.46	3.47	3.46	3.40
	0.9	12.23	1.20	1.19	1.20	1.20	1.17	19.12	3.09	2.35	2.47	2.39	2.20
μ	0.2	12.80	1.52	1.49	1.48	1.49	1.48	18.22	3.86	3.33	3.42	3.34	3.27
	0.4	11.62	0.91	0.86	0.87	0.87	0.86	16.11	3.60	3.17	3.16	3.17	3.10
h	0.1	9.52	0.92	0.88	0.88	0.89	0.87	13.05	2.74	2.34	2.33	2.35	2.28
	0.2	12.33	1.26	1.21	1.24	1.24	1.19	17.33	3.75	3.28	3.31	3.29	3.21
	0.4	14.77	1.51	1.48	1.49	1.48	1.45	21.15	4.70	4.13	4.23	4.13	4.06
l	10	9.84	0.68	0.64	0.64	0.65	0.64	14.05	2.43	1.93	1.89	1.94	1.81
	20	11.97	1.16	1.11	1.11	1.13	1.09	16.85	3.35	2.87	2.91	2.87	2.80
	30	12.89	1.37	1.33	1.34	1.33	1.32	18.14	4.06	3.59	3.66	3.59	3.54
	50	14.13	1.68	1.64	1.66	1.65	1.63	19.62	5.06	4.60	4.68	4.60	4.56
t_f	0.25	14.06	1.72	1.69	1.69	1.69	1.68	19.89	5.33	4.83	4.97	4.82	4.81
	0.5	13.14	1.47	1.42	1.42	1.44	1.41	18.47	4.36	3.86	3.96	3.86	3.83
	1	11.76	1.06	1.01	1.03	1.02	1.00	16.48	3.17	2.69	2.72	2.70	2.63
	2	9.87	0.64	0.59	0.61	0.60	0.58	13.82	2.07	1.60	1.49	1.62	1.44
Overall		12.21	1.22	1.18	1.19	1.19	1.17	17.17	3.73	3.25	3.29	3.25	3.18

		15 links						complete pooling					
Parameter	Value	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5
λ	0.1	26.81	2.82	1.74	1.80	1.74	1.55	29.06	1.22	0.86	0.00	0.99	0.00
	0.3	22.21	3.00	1.91	1.97	1.91	1.76	23.06	1.23	0.87	0.00	0.90	0.00
	0.5	19.88	3.74	2.70	2.40	2.35	2.17	22.02	1.12	0.76	0.00	0.76	0.00
σ	0	22.07	3.48	2.51	2.47	2.37	2.29	24.82	1.26	0.86	0.00	0.86	0.00
	0.2	22.32	3.42	2.49	2.41	2.32	2.23	26.63	1.24	0.89	0.00	1.06	0.00
	0.5	22.73	3.22	2.27	2.15	2.18	1.96	29.66	1.09	0.84	0.00	0.95	0.00
	0.9	24.72	2.50	1.20	1.21	1.14	0.83	28.93	1.16	0.79	0.00	0.93	0.00
μ	0.2	23.63	2.16	1.11	0.95	0.88	0.67	27.07	1.14	0.77	0.00	0.84	0.00
	0.4	22.30	4.17	3.12	3.17	3.12	2.98	26.26	1.33	0.98	0.00	1.10	0.00
h	0.1	18.26	1.81	0.77	0.77	0.77	0.57	22.57	1.29	0.96	0.00	1.03	0.00
	0.2	23.26	2.55	1.46	1.52	1.46	1.27	27.71	1.25	0.86	0.00	0.97	0.00
	0.4	27.37	5.17	4.12	3.88	3.78	3.63	31.55	1.06	0.69	0.00	0.78	0.00
l	10	18.93	4.60	3.60	3.02	3.14	2.93	21.44	1.57	1.07	0.00	1.21	0.00
	20	22.60	2.94	1.90	1.92	1.90	1.73	26.02	1.25	0.86	0.00	0.96	0.00
	30	24.24	2.66	1.55	1.65	1.55	1.38	28.11	1.12	0.80	0.00	0.87	0.00
	50	26.08	2.65	1.43	1.65	1.43	1.26	30.25	0.98	0.72	0.00	0.78	0.00
t_f	0.25	25.38	4.05	2.79	3.05	2.79	2.70	29.40	0.40	0.29	0.00	0.31	0.00
	0.5	24.31	3.50	2.32	2.50	2.32	2.22	28.15	0.77	0.53	0.00	0.61	0.00
	1	22.53	2.78	1.72	1.75	1.72	1.55	26.33	1.35	0.96	0.00	0.97	0.00
	2	19.62	2.63	1.63	0.94	1.17	0.83	23.11	2.34	1.65	0.00	1.86	0.00
Overall		22.96	3.16	2.12	2.06	2.00	1.83	26.75	1.21	0.85	0.00	0.94	0.00

Table 2.6 Average savings for parameter ϑ in partial pooling network (in percent)

Value	Δ_{No}	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5
0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.13	3.54	1.92	1.85	1.87	1.87	1.83
0.17	4.83	2.37	2.25	2.28	2.28	2.22
0.25	7.25	3.08	2.90	2.95	2.90	2.83
0.30	8.10	4.18	3.98	4.04	4.01	3.93
0.38	10.17	4.77	4.50	4.59	4.53	4.42
0.42	10.68	5.27	4.99	5.07	5.03	4.91
0.50	13.76	5.14	4.76	4.84	4.78	4.63
0.58	14.41	5.84	5.42	5.52	5.45	5.27
0.63	16.56	5.63	5.18	5.24	5.17	4.98
0.70	17.70	5.49	5.01	5.03	5.00	4.77
0.75	22.75	4.51	3.93	3.85	3.95	3.60
0.83	23.65	3.89	3.32	3.13	3.35	2.92
0.88	23.29	3.33	2.79	2.53	2.84	2.36
1.00	35.41	1.08	0.76	0.00	0.88	0.00

Comparing the proposed transshipment rules, the random rule is significantly inferior to all other rules (two-tailed t-test, $p < 0.001$). Rule 2 performs slightly better than rule 4. This effect is caused by our definition of run-out-time which only considers the demand rates of primarily assigned markets to a warehouse and not of markets which trigger a transshipment from the respective warehouse. Rule 3 does not significantly differ from rule 4 for pooling factors $\vartheta < 0.17$ ($p = 0.20$) whereas for $0.17 < \vartheta < 0.62$ it is significantly inferior ($p < 0.001$) and it is superior for larger pooling factors $\vartheta > 0.75$ ($p < 0.001$). This implies that for very low pooling opportunities the choice of a rule is irrelevant since the potential savings are in general very low. For low to intermediate pooling opportunities, a careful deployment of inventory is more important, whereas for high pooling, transportation choice becomes more important.

Figure 2.5 visualizes the results from Table 2.6 gained from the partial pooling network and from the three warehouse and six markets (3×6) networks from Table 5 for 9, 12, 15 and 18 links which are shown by the crosses and triangles. The potential cost savings of proactive transshipments for different pooling values (compared to random or optimal choice) of the sending location first increase and then decrease. Networks with intermediate flexibility and pooling opportunities exhibit the largest savings for proactive over reactive transshipments.

2.4.3 Sensitivity to lead time distribution

The assumption of exponentially distributed lead times might appear to be rather limiting. In line with the literature, e.g., Alfredsson and Verrijdt (1999) for reactive transshipments and complete pooling and Reijnen et al. (2009) for partial pooling, we analyze the sen-

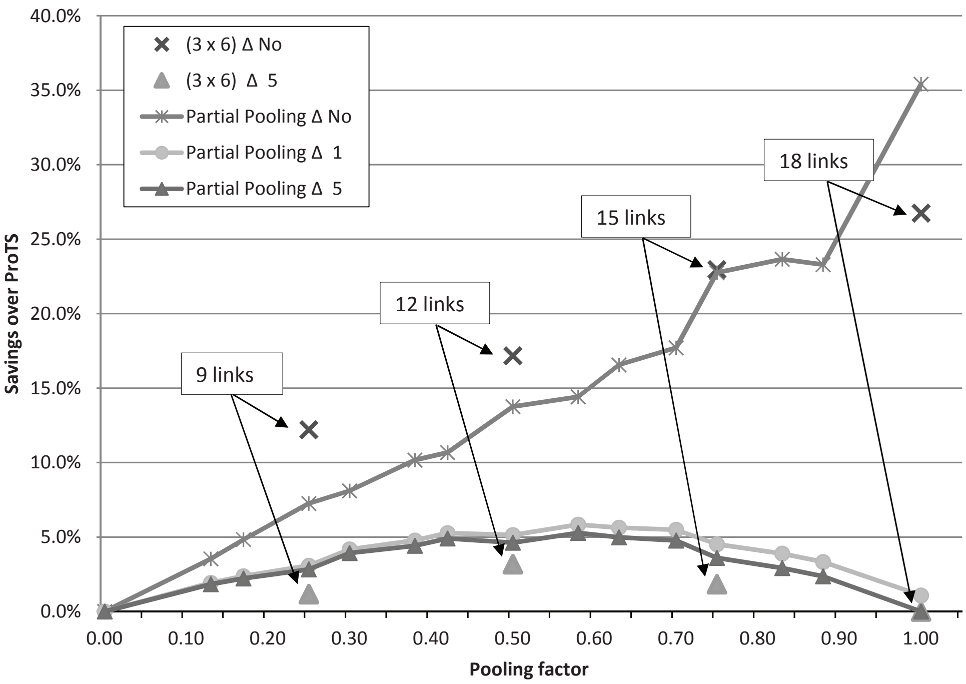


Figure 2.5 Savings by pooling factor

sitivity of the results with respect to the lead time distribution assumption. We use the network depicted in Figure 2.1 with all possible combinations of parameters shown in Table 2.2 (1536 instances). The base stock levels were optimized under the exponential lead time assumption. Using 4,000,000 arrivals, the system was simulated with deterministic (det) and gamma (gam_c) distributed lead times with a coefficient of variation of $c = 0.5$ and $c = 2$ exhibiting the same mean as the exponential distribution. The average cost difference of lead time type x is defined by $\delta_x = |C_{ProTS} - C_x^{sim}|/C_{ProTS} \cdot 100$, where C_x^{sim} are the costs obtained through simulation with x corresponding to deterministic or exponentially and gamma distributed lead times. The average difference over all instances is denoted by $\overline{\delta}_x$. The results show cost differences for exponential and for both tested gamma distributions of $\overline{\delta}_{exp} = 0.0467\%$, $\overline{\delta}_{gam_{0.5}} = 0.0418\%$, $\overline{\delta}_{gam_2} = 0.0536\%$. For deterministic lead times, the average cost differences are $\overline{\delta}_{det} = 0.0631\%$. These results support that the exponential lead time assumption is robust. Only for deterministic lead time, we observe slightly larger, but still negligible differences. This finding is supported by the fact that isolated warehouses can be considered as M/M/S/S queues which have the same resulting lead time characteristics as M/G/S/S systems.

2.5 Conclusion

We formulated an SMDP to optimize proactive transshipment decisions and evaluated different reactive transshipment rules. The contribution to the literature is an integrated continuous review approach that contains both, proactive and reactive transshipments and determines an optimal source for a transshipment in contrast to widely used predetermined assignments. While usually either a proactive or reactive approach is followed and the benefits compared to no transshipments or between different rules are reported, we present a comparison between proactive and reactive transshipments for several networks with different connectivity. Our results indicate that proactive transshipments can lead to significant cost reductions in comparison to using only reactive transshipments, especially in systems that exhibit only intermediate opportunities of inventory pooling. This result has an implication on the design of inventory pooling networks. Besides the well known fact that a clever linking of markets to warehouses might achieve benefits close to complete pooling, the full benefits can only be achieved if transshipments are optimized proactively. If reactive transshipments are used, our results confirm earlier findings that the differences between different transshipment rules are negligible.

Further research is required to overcome the curse of dimensionality when providing results for large networks as our exact analysis is limited to small problems. Here, Approximate Dynamic Programming (see Powell, 2007) and tailored proactive transshipment rules with optimized parameters are promising venues for extension.

Chapter 3

Emergency Supply Flexibility in a Two-Echelon Inventory System with Pipeline Information

3.1 Introduction

Real time supply information supported through advances of tracking and tracing technologies such as global positioning systems (GPS) or radio frequency identification (RFID) provides better visibility of orders within the supply chain. This additional information allows locations to select the fastest among several replenishment sources and can thereby reduce customer waiting times. Innovative real time replenishment policies within the concept of available-to-promise yield cost savings but require new advanced inventory management decision support.

After sales service networks for expensive spare parts in aviation, defense or communication sectors are characterized by having low volume random demand rates but fast responsiveness requirements when systems fail. This reinforces the question "What is the right supply chain for service parts" (Fisher (1997)). High stock levels guarantee low service times but make the network economically inefficient. Companies respond to this problem by introducing flexibility options for replenishment of items. Examples of such options are lateral transshipments and expedited deliveries characterized by the faster the delivery, the more expensive the shipment. For recent overviews of transshipments in inventory systems we refer to Paterson et al. (2011) and Wong et al. (2006). Because of geographically widely spread markets it is often practical to have several local warehouses which provide service to local markets. For reasons of quick response and economies of scale, local warehouses are supplied by a central warehouse which is supplied by a plant or an external supplier. In case of a stockout, lateral transshipments from other local warehouses or expedited deliveries from upper echelons can be used to fill the demand. However, it is necessary to define a strategy for serving demand when a local warehouse is out of stock. A fast but expensive lateral transshipment is only beneficial if the next free item through regular replenishment arrives later than the transshipment. Ordered items which have been sent by the central warehouse but have not yet arrived are considered as pipeline inventory. Most inventory models with transshipment and expedited delivery ignore pipeline inventory and trigger flexibility options every time a stockout occurs instead of waiting for the arrival of the next unit from regular replenishment. This is mainly due to the unrealistic assumption that transshipment lead times are negligible. Especially for slow moving items with long lead times, the pipeline can contain a considerable part of overall system stock.

The model analyzed in this chapter is an extension of classical multi-echelon inventory models with Poisson demand, one-for-one replenishment and deterministic lead times. We study a distribution network where several local warehouses are served by one central warehouse. The central warehouse is supplied by a plant. In case of a stockout, a local warehouse can either trigger a lateral transshipment from another (randomly chosen) local warehouse with stock on hand or an expedited delivery from an upper echelon (e.g. central warehouse or plant). Lateral transshipment is only beneficial if it will arrive before the next free outstanding order through the regular replenishment. We incorporate pipeline information into the decision when to initiate a lateral transshipment and when to wait until the arrival of the next outstanding order. In order to provide the fastest delivery option to the customer this decision is based on time. Therefore, the transshipment lead time, which we assume is identical for every local warehouse, is the reference time making the decision to satisfy the demand through the pipeline inventory. Satisfying the customer through pipeline inventory a local warehouse (partially) backorder the demand. We define the problem as time dependent partial backordering and analytically derive an exact solution for one local warehouse. Further, we integrate these results in a multi-echelon model to estimate the fraction of demand satisfied through different shipment options.

This chapter contributes to three research streams, lateral transshipments, partial backordering and advance information in supply chain. The first stream focuses on lateral transshipments with positive lead times and one-for-one replenishment to control inventory levels. It is a common practice in multi-echelon networks to manage low volume expensive inventory by a one-for-one replenishment control policy (for an overview see, Axsäter (2003)). Lee (1987) presents a two-echelon inventory system with lateral transshipments between identical local bases. Axsäter (1990a) extends Lee's model to non-identical bases. Lee (1987) and Axsäter (1990a) both assume positive transshipment and stock replenishment lead times but do not track the delivery status of outstanding orders, i.e., the next free resupply unit may arrive earlier than the transshipment. However, both authors recognize the potential importance to consider pipeline inventory and mention that their rule is suboptimal. Dada (1992) investigates an inventory system with several local bases and a central depot where the reaction to a stockout is either a lateral transshipment or an emergency expedited delivery from the depot. Pipeline inventory is only used if lateral transshipments and expedited deliveries are not possible.

Applying a Markov chain approach, he develops an approximation to calculate different system performance measures. Alfredsson and Verrijdt (1999) present a model with alternative emergency shipments from the depot and the manufacturing facility. They ignore pipeline inventory completely, but indicate the possible impact of pipeline flexibility on system performance and report potential cost reductions observed from simulations. Kutanoglu and Mahajan (2009) use the model of Alfredsson and Verrijdt (1999) to implement time-based service levels where the sending location for lateral transshipment is chosen according to a priority list rather than randomly. The authors mention that a more realistic operation would be that the warehouse uses regular replenishments to satisfy demands in a timelier manner. They overcome this problem by assuming that the deliveries of lateral transshipment are nearly instant. Kranenburg and van Houtum (2009) allow partial pooling and develop a new iterative approach to calculate the key figures. They show that the major part of the full pooling benefits is obtained via partial pooling if the local warehouses which provide lateral transshipment are well chosen. The main difference between our model and the papers discussed above is that we incorporate the pipeline inventory of each local warehouse into the decision when to trigger lateral transshipment and when to wait until the arrival of the regular replenishment.

In a simulation study Sherbrooke (1992) investigates the impact of pipeline inventory in a two-echelon system. He provides upper and lower bounds to derive approximate expressions for expected system backorders. Yang and Dekker (2010) introduce a two-echelon inventory model, explicitly taking pipeline stocks into account. They assume that customers accept a certain waiting time for which no penalty cost occurs. If the next free outstanding order arrives later than the tolerable time span, either a lateral transshipment is triggered or the demand is backordered and a penalty cost is charged. Different to our work in this chapter, they allow backorders at local warehouses when all locations are out of stock and no regular replenishment will arrive in time. Another difference is that they incorporate the time for waiting of arrival of the next delivery in the pipeline as a service level and not as a decision variable when to request the fastest delivery options. Howard et al. (2010) incorporate pipeline information for requesting an emergency shipment in a multi-echelon inventory system. They consider an inventory system similar to Axsäter et al. (2013), integrating information on orders in the replenishment pipeline. Their distribution network consists of several retailers and a support warehouse. When a retailer is out of stock and there is no unreserved replenishment order in a pipeline which will arrive within a tolerance time, then the demand is satisfied by an emergency shipment. Optimizing the base stock levels and the tolerance times for each location, they show that ignoring pipeline information can nearly double the relative cost. Compared to our work there are important differences. Firstly, we consider a multi-echelon structure where the central warehouse does not have an infinite capacity and the optimal capacity (base stock level) has to be determined. This assumption means that the regular replenishment lead times are subject to delay when the central warehouse runs out of stock. Secondly, we assume that local warehouses can share their inventory by using lateral transshipments before requesting an emergency delivery from upper echelons.

Reviewing the relevant literature on partial backordering, we focus on stochastic models with one-for-one replenishment policies as this is most related to our work. Partial backordering or lost sales add both realism and analytical complexity to the model. Das (1977) studies an $(S-1, S)$ system with a Poisson demand arrival process and exponentially distributed replenishment times. Customers are assumed to wait a certain time before canceling the order. He concludes that larger customers' willingness to wait does not necessarily reduce the system cost under all circumstances. The main difference is that in our model the demand is only backordered if the maximum tolerable delay of the customer can be guaranteed. In Moinzadeh (1989), an arriving demand facing a stockout is backordered with a known probability. Nahmias and Smith (1994) integrate the aspect of partial backordering in a two-echelon inventory system. They assume that the waiting probability for the customer is known but depends on the type of items. They test the sensitivity of optimal stocking levels to the probability of waiting for the next arriving item. Abad (1996) considers pricing and inventory control for perishable products with partially backordered demand when the backlog is cost efficient for a reseller. We model such a situation in a stochastic inventory problem that leads to partial backordering determined by the shipment time of alternative supply, lateral transshipment. The transshipment lead time, which is fixed is a reference time for maximum waiting time for arrival of the next uncommitted item through the regular replenishment in case of backorder. Our model differs from those discussed above mainly with respect to a fixed waiting time for a backorder, whereas other models assume a known backorder probability. A fixed waiting time is another performance measure to guarantee good customer service. Tempelmeier (2000) studies different optimization criteria for logistical processes and finds that standard service level constraints provide only limited information about service quality. A customer who finds an inventory location depleted does not immediately experience bad service per se, but observes a delay. Therefore, control of customer order waiting times offers another opportunity to design an inventory system in a cost minimal way.

In the research stream related to advance supply information the type of available additional information on the supplied side strongly depends on the problem setup. Song and Zipkin (1996) study a scenario where the supply system is not fixed but evolves over time. They assume that the supplied location obtains information about the current and future lead times. Their main concern is to determine how much information is really useful. Gaukler, Özer, and Hausman (2008) analyze a (Q,R) order framework combined with emergency ordering policy that uses the outstanding orders's progress information. They propose an improved inventory control policy and show the potential cost savings attributed to the order progress information in the numerical study. Chew, Lee, and Sim (2013) investigate potential cost savings in a periodic review model with variability in the transportation link with and without the asset visibility in the supply chain. An inventory system with two supply sources, a normal one and a faster but more expensive emergency source is studied by Song and Zipkin (2009). They use the real-time supply information to determine where in the order pipeline the outstanding orders are. When too much of the pipeline stock is too far away they use the emergency supply. They show that the solution for the order policies has a simple product form. Sherbrooke (1992), Yang and Dekker (2010), Howard et al. (2010) also use the real-time information to incorporate the pipeline inventory in the operational and strategic decisions. We contribute by considering the additional visibility of the pipeline inventory in the context of lateral transshipments between multiple locations and emergency supply from upper echelons.

The remainder of this chapter is organized as follows. Section 3.2 presents assumptions and notation and describes the modeling approach. Section 3.3 shows the accuracy of our model for key performance measures. Section 3.4 studies effects of pipeline information and shows how and when firms can benefit from it.

3.2 Problem description

3.2.1 Assumptions and notation

We consider a single-product continuous review two-echelon distribution network with a central warehouse $n = 0$ and $n = 1, ..., N$ local warehouses. Each local warehouse n faces a Poisson demand process with a rate of λ_n. All warehouses use a one-for-one replenishment policy with a base stock level S_n and storage is subject to holding cost h_n per item and per unit of time at warehouse n. One-for-one inventory replenishment policy is common for e.g. expensive spare parts, because of high prices, low demand rates, and rather negligible fixed ordering costs (see Axsäter (2003)). Orders with the central warehouse arrive at warehouse n after a deterministic shipping time T_n and the replenishment unit cost is r_n. Each local warehouse has full information about outstanding replenishment orders and knows the arrival time of each outstanding shipment at any time. The structure of the distribution network and the sequence of delivery options to fulfill the demand is depicted in Figure 3.1. The delivery options are numbered according to their priority of execution and are designed as follows:

Option 1 When a customer arrives at local warehouse n, the demand is satisfied from on-hand stock.

Option 2 If no items are on hand, the customer is served by the next free outstanding order in the pipeline to warehouse n which will arrive before the delivery time of lateral transshipment D_0. We set the transshipment lead time as a reference time for waiting for the arrival of the next free outstanding order. The reference time can also be set different from the lead time of lateral transshipment, but to choose the fastest delivery option is the only economically reasonable approach if the penalty cost for waiting is high compared to the shipment cost. Lateral transshipments are a common strategy for inventory sharing between the members of a pooling group. They are widely used when transshipment costs

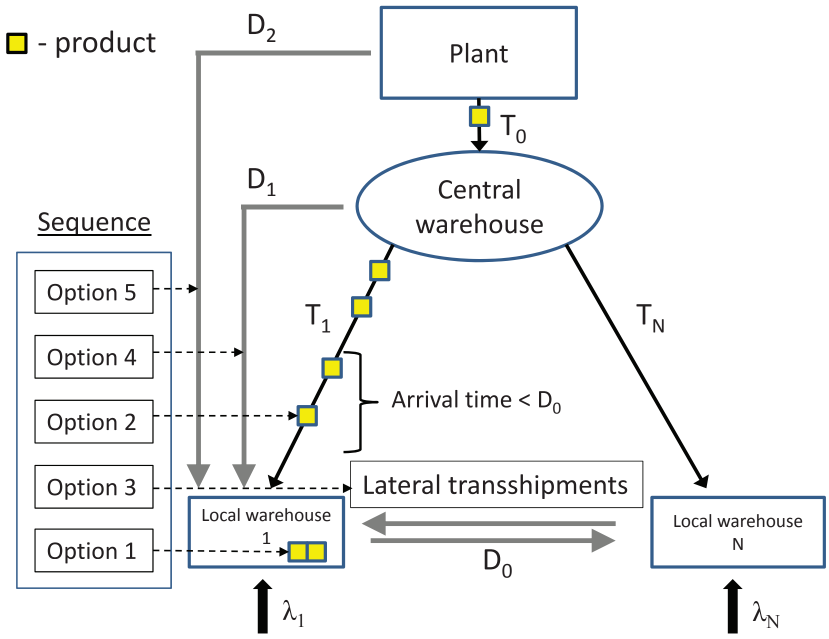


Figure 3.1 Distribution network structure and delivery options

are relatively low compared to high holding costs and penalty costs for failing to meet demand. This cost structure is very common for e.g. expensive spare parts. Within a pooling group the average transshipment times between the locations are assumed to be identical for all local warehouses. This assumption implies that the reference time for using a free outstanding item from the supply pipeline for filling the demand is also identical for every local warehouse.

If an item arrives earlier through normal replenishment than by transshipment, the demand is backordered and a new replenishment order from the depot is triggered. This implies that the customer's backorder is subject to a guaranteed maximum delay which is equal to the delivery time of the lateral transshipment D_0. Demand, including backordered demand, is satisfied first come first served (FCFS), i.e. for a backordered customer the next free outstanding order is reserved. Demand which is not satisfied from stock is subject to a penalty cost z_n per unit of time for waiting at local warehouse n.

Option 3 If the next uncommitted item in the pipeline arrives later than a lateral transshipment and another local warehouse has stock on hand, then the demand is served by a lateral transshipment for a transshipment price c_0 per item. The local warehouse which serves the request is randomly chosen among the local warehouses with stock on hand. A random rule to choose the sending location for a lateral transshipment is not a very restrictive assumption. Lee (1987) tests three transshipment rules: random rule, maximum stock on hand, and smallest number of outstanding orders and concludes that there is no significant performance differences between the rules. Seidscher and Minner (2013) test five different rules: random, highest inventory level, minimal transshipment cost, highest run-out-time and compare these rules to the optimal location choice. They show numerically that differences between the transshipment rules are negligible.

If no local warehouse has stock on hand, then the local warehouse which next receives an uncommitted item through the normal replenishment within the reference time D_0 backorders the demand. Note that it might not always be a case that a local warehouse exists which will receive an outstanding free delivery within D_0. If such a local warehouse exists, then the item arrives at the location and will be sent to the original warehouse via lateral transshipment. Then, the delay consists of the transportation time from the base, which backorders the demand, plus an additional waiting time for arrival of uncommitted item through the normal replenishment. In this case the customer's backorder is subject to the guaranteed minimum delay of D_0 and the maximum delay of $2D_0$. If there is no such local warehouse in the distribution network, then option 3 can be considered as not possible and the next option (4) applies.

This rule results from the approximation approach we use to make the model tractable. We note that it is not the fastest delivery option for filling the demand if the expedited delivery time from the central warehouse is shorter than the maximum delay of $2D_0$, but usually the central warehouse has a large geographical distance to the local warehouses which are situated in one geographical area to build an inventory sharing pooling group with short shipment times between the locations.

Option 4 If all local warehouses are out of stock and no free items in the pipelines will arrive within the delivery time of the lateral transshipment and the central warehouse has stock on hand, the customer is served by a expedited delivery from the central warehouse at a cost c_1 per unit. We let D_1 denote the delivery time from the central to the local warehouse and assume that the expedited delivery is much faster than the regular replenishment from the central warehouse.

Option 5 If option 1 to 4 are not possible, this means a customer arrives when no stock is at local warehouses, no free items in the pipelines will arrive within the delivery time of the lateral transshipment and no stock on hand is at the central warehouse, then the customer is served by expedited delivery from the plant at a cost c_2 per unit. D_2 denotes the shipment time of expedited delivery from the plant to the local warehouse. We assume that the expedited delivery from the central warehouse is much faster than from the plant. Alternatively, a lost sale an be modeled with a penalty cost of $c_2 + D_2 z_n$ per unit for warehouse n.

Note that for expedited deliveries from the central warehouse and from the plant we assume identical shipment times for each local warehouse. This assumption results from the previously assumed geographical allocation of the warehouses, where a pooling group of local warehouses has small distances between the locations but the central warehouse and the plant have a large distance to a pooling group.

α_n denotes the fraction of demand satisfied trough lateral transshipments, β_n the fraction of demand satisfied through normal replenishment. β_n includes the fraction of demand satisfied directly from stock and from backordered demand. The fraction of demand satisfied directly from the central warehouse is denoted by γ and the fraction of demand satisfied through expedited deliveries from the plant is denoted by θ. Table 3.1 summarizes the notation.

Table 3.1 List of Notation

N	Number of local warehouses, where $n = 0$ denotes the central warehouse and $n = 1, ..., N$ the local warehouses
λ_n	Customer arrival rate at local warehouse n
S_n	Base stock level at warehouse n
T_n	Replenishment time of warehouse n
D_0	Lateral transshipment delivery time
D_1	Expedited delivery time from the central warehouse
D_2	Expedited delivery time from the plant
α_n	Fraction of demand satisfied by lateral transshipments at warehouse n
β_n	Fraction of demand satisfied by local warehouse n
γ	Fraction of demand satisfied by the central warehouse
θ	Fraction of demand satisfied by the plant
h_n	Unit holding cost per time unit at warehouse n
r_n	Unit replenishment cost at warehouse n
z_n	Penalty cost per time unit for waiting at warehouse n
c_0	Lateral transshipment cost per item
c_1	Expedited delivery cost from the central warehouse per item
c_2	Expedited delivery cost from the plant per item

3.2.2 The single warehouse case

We start the analysis by focusing on a single local warehouse. We study a single local warehouse isolated from the central warehouse. This implies that for the replenishment times for a local warehouse we can ignore the possible delivery delays in case of central stock-out. When we study the complete distribution network in Chapter 3.2.3 we will relax this assumption and present a solution method to incorporate the limited supply of the central warehouse in the analysis. This single warehouse problem is identical to the single-echelon model of Howard, Reijnen, Marklund, and Tan (2010), but we present a different approach to derive the key performance measures.

When the local warehouse is out of stock the next fastest delivery option is a lateral transshipment with the shipment time D_0. In order to provide a customer a faster delivery than the lateral transshipment, we consider the free outstanding orders in the pipeline which will arrive within D_0. Consequently, the decision either to backorder a demand or not is based on time and the reference time for this decision is D_0. Each backorder implies an arrival of the next uncommitted outstanding order within the time D_0 and is faster than the alternative triggering of lateral transshipment.

Observation 1 *Any unit ordered by a warehouse becomes available D_0 time units before the physical arrival at the warehouse.*

Observation 1 is the direct consequence of a one-for-one replenishment policy, deterministic lead times and the assumption to backorder demand if any ordered unit in the pipeline will arrive within the reference time D_0. The availability of an item means, physical presence on stock or an arrival within the reference time D_0 and is visualized in Figure 3.2. Consequently, the availability replenishment rate μ_n^a for warehouse n is the fraction of time between sending an item from the central warehouse and the point in time when an item is available for the local warehouse and is define as

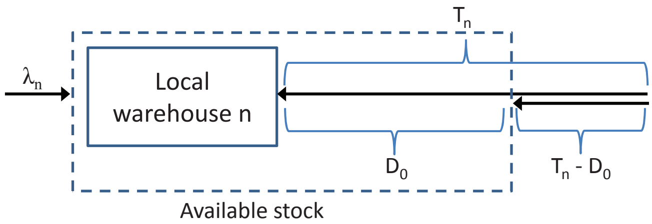


Figure 3.2 Single local warehouse n

$$\mu_n^a = 1/(T_n - D_0), \tag{3.1}$$

where the fraction is the physical delivery time T_n less the reference time D_0.

A demand is only backordered if an item is available otherwise a demand is lost. The single local warehouse with fixed transportation lead times and unlimited supply from the central warehouse can be described as an $M|D|S_n|S_n$ queue where the state represents the number of available units in the system, which become available with rate μ_n^a. Note, because we do not allow backorder of a demand if an item is not available we can model a single location as an $M|D|S_n|S_n$ queue where the number of orders in the system is limited. Since the warehouse uses one-for-one replenishment where each inventory reduction results in a new outstanding order, the maximum number of available units (on-hand stock + items which will arrive within D_0 time units) in the system is equal to the base stock level S_n. We notice that the local warehouse uses a common base stock policy where the base stock level S_n refers to the on-hand inventory, although we describe the queue with the number of available units.

The $M|D|S_n|S_n$ queue can be analyzed as a $M|M|S_n|S_n$ queue since steady-state probabilities are independent of the lead time distribution (see Gross et al. (2008)). Let j be the number of available units, $0 \leq j \leq S_n$ and p_j^n the steady-state probability that the warehouse n has j units available. The steady-state probabilities can be obtained analytically (see Gross et al. (2008)). The well-known expression for p_j^n is

$$\frac{1}{p_{S_n}^n} = \sum_{j=0}^{S_n} \frac{(\lambda_n/\mu_n^a)^j}{j!} \quad \text{and} \quad p_j^n = \frac{(\lambda_n/\mu_n^a)^{S_n-j}}{(S_n-j)!} p_{S_n}^n \tag{3.2}$$

The portion of demand which is satisfied through normal replenishment is $\lambda_n(1-p_0^n)$.

Knowing the steady-state probabilities does not directly allow to determine the holding and backordering cost, since the state represents only the number of available units. We recall available stock includes physical items on stock and the uncommitted items in the pipeline which will arrive within time D_0. We use the method of unit tracking similar to Axsäter (1990b).

Observation 2: *Any ordered unit which becomes available as an X^{th} unit after requesting the replenishment is dedicated to satisfy the X^{th} demand after the one triggering the current replenishment.*

Observation 2 is the consequence of the previously assumed order policy, the first come first serve rule and deterministic replenishment times. From the point of time when the unit is available, it becomes physical after the reference time of D_0. Before that it is still part of the pipeline inventory and is in transit to the warehouse. Therefore, holding costs only occur if the demand for a particular unit arrives after the reference time D_0. Backorder cost occurs only if the demand arrives before the reference time.

Under Poisson demand with rate λ_n is

$$v_n(t,X) = \frac{\lambda_n^X t^{X-1} e^{-\lambda_n t}}{(X-1)!} \tag{3.3}$$

the probability density function for the time until the occurrence of the X-th demand and it is an Erlang distribution. If an ordered item becomes available in the system as an X^{th} unit, then the conditional expected time the item is physically on stock is given by

$$\varepsilon_X^n = \int_{D_0}^{\infty} (t-D_0) v_n(t,X) dt \tag{3.4}$$

and the conditional expected time the item is the part of the pipeline inventory and backordered is

$$\eta_X^n = \int_0^{D_0} (D_0-t) v_n(t,X) dt. \tag{3.5}$$

The probability that an ordered item becomes available as the X^{th} unit in the long-run average is equal to the steady-state probability p_X^n that the system has X units available from (3.2).

Let q_n denote the inventory time units and w_n the backordering time units a local warehouse n has as a long-run average per one demand unit. Since every ordered item can be available as X^{th} unit in the system and the number of available units has a range of $1, ..., S_n$ we obtain the following expressions for q_n and w_n,

$$q_n = \sum_{s=1}^{S_n} p_s^n \varepsilon_s^n \quad \text{and} \quad w_n = \sum_{s=1}^{S_n} p_s^n \eta_s^n. \tag{3.6}$$

The exact analysis above allows us to determine all necessary performance measures: fraction of demand satisfied through normal replenishment and average holding and backorder cost.

Next we construct a model for aggregated local warehouse to calculate the performance measures for the central warehouse. These performance measures are equal for all the local warehouses, even if the locations are not symmetric.

3.2.3 The aggregate local warehouse

We follow the approach presented in Dada (1992) and Alfredsson and Verrijdt (1999) that from the central warehouse's point of view the local warehouses can be considered as one aggregated warehouse with a base stock level $S = \sum_{n=1}^{N} S_n$ and aggregate customer arrival rate $\lambda = \sum_{n=1}^{N} \lambda_n$. The aggregated local warehouse and the central warehouse can be modeled as a two-dimensional Markov-chain (see Alfredsson and Verrijdt (1999)). They show that their model assuming exponentially distributed lead times is not sensitive to the lead time distribution. We use this approach to approximate the deterministic lead times. Different to Alfredsson and Verrijdt (1999), we use the findings from the single warehouse analysis and model the aggregate local warehouse by the number of available units in the system instead of by the number on hand inventory.

Then, a system state is described by a tuple (i, j), where i is the number of items on stock at the central warehouse and j the number of available items at the aggregate local warehouse. The shipment time of the aggregate local warehouse is $T = \sum_{n=1}^{N} \lambda_n \cdot T_n/\lambda$. This is an approximation, since we assume that local warehouse n receives items from the central warehouse equal to the demand rate λ_n. This is not always true because of the lateral transshipments between the locations. Since we model the aggregate local warehouse with the number of available units according to the findings in Section 3.2.2, the replenishment rate has to be adjusted to the availability replenishment rate. Thus, we define the availability replenishment rate of the aggregate local warehouse as follows, $\mu^a = 1/(T - D_0)$. The state space and the associated transition rates between the states are depicted in Figure 3.3, which is similar to Alfredsson and Verrijdt (1999) for the case $S_0 = S = 2$, where $\mu_0 = 1/T_0$ is the rate at which the replenishment orders arrive at the central warehouse. A transition from state $(2, 0)$ to $(1, 0)$ and from $(1, 0)$ to $(0, 0)$ results by filling the demand by expedited delivery from the central warehouse. The states space range for the central warehouse $-S \leq i \leq S_0$ and $0 \leq j \leq S$ for the aggregate local warehouse. Note that the minimum stock level at the aggregate local warehouse is zero but the minimum stock level at the central warehouse is $-S$. This is a direct result of the one-for-one replenishment policy at local warehouses with no backorder if no item is

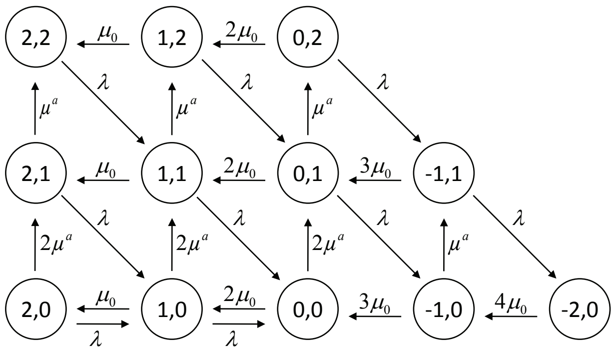


Figure 3.3 Aggregate State Space Model

available.

The two-dimensional Markov chain can be solved by defining the corresponding transition equations of the system for each state and the normalizing criterion. For given base stock levels S_0 and S the transition equations are

$$\pi_{ij} = \begin{cases} \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+(S-(j-1))\mu^a\pi_{i,j-1}}{\lambda+(S_0-i)\mu_0+(S-j)\mu^a} & \text{if} \quad i = S_0 \vee (0 < i \leq S_0; j = S) \\ \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+(S-(j-1))\mu^a\pi_{i,j-1}+\lambda\pi_{i+1,j+1}}{\lambda+(S_0-i)\mu_0+(S-j)\mu^a} & \text{if} \quad 0 \leq i < S_0; 0 < j < S \\ \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+\lambda\pi_{i+1,j}}{\lambda+(S_0-i)\mu_0+(S-j)\mu^a} & \text{if} \quad 0 < i < S_0; j = 0 \\ \frac{\mu^a\pi_{i,j-1}}{(\lambda+(S_0-i))\mu_0} & \text{if} \quad -S < i < 0; j = S - i \\ \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+\lambda(\pi_{i+1,j+1}+\pi_{i+1,j})}{(S_0-i)\mu_0+S\mu^a} & \text{if} \quad i = 0; j = 0 \\ \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+\lambda\pi_{i+1,j+1}}{(S_0-i)\mu_0+(S+i)\mu^a} & \text{if} \quad -S \leq i < 0; j = 0 \\ \frac{(S_0-(i-1))\mu_0\pi_{i-1,j}+(S+i-(j-1))\mu^a\pi_{i,j-1}+\lambda\pi_{i+1,j+1}}{\lambda+(S_0-i)\mu_0+(S+i-j)\mu^a} & \text{if} \quad -S < i < 0; 0 < j < S - i \\ \frac{\lambda\pi_{i+1,j+1}}{(S_0+S)\mu_0} & \text{if} \quad i = -S; j = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

and the normalizing criterion is

$$1 = \sum_{i=-S}^{S_0} \sum_{j=0}^{S} \pi_{ij} \tag{3.8}$$

where π_{ij} is the steady-state probability that the system has i units at the central warehouse and j available units at the aggregated local warehouse.

3.2.4 Performance measures

Having found the steady-state probabilities, γ and θ can be determined as

$$\gamma = \sum_{i=1}^{S_0} \pi_{i0} \quad \text{and} \quad \theta = \sum_{i=-S}^{0} \pi_{i0}. \tag{3.9}$$

γ is equal to the probability that all local warehouses are out of stock and the central warehouse has inventory on stock. θ is the sum of probabilities that all warehouses and the central warehouse are out of stock.

Applying Little's law that the average waiting time at the central warehouse is equal to the average number of backorders ($\sum_{i=-S}^{-1} -i \sum_{j=0}^{S+i} \pi_{ij}$) divided by the average arrival rate of orders ($(1-\theta)\lambda$), the expected time delay Δ for normal replenishment orders at the central warehouse is

$$\Delta = \frac{1}{(1-\theta)\lambda} \sum_{i=-S}^{-1} -i \sum_{j=0}^{S+i} \pi_{ij}. \tag{3.10}$$

Equations (9) and (10) are consistent with Alfredsson and Verrijdt (1999).

Having found θ and γ we want to determine α_n and β_n for each local warehouse. To disaggregate the findings of aggregate local warehouse, we use the METRIC approach from Sherbrooke (1968). METRIC is not an exact approach but it provides good results for multi-echelon inventory models when the demand rates are low. Since the central warehouse can run out of stock, the replenishment orders of local warehouses are subject to delays. We assume that the average lead time of the local warehouse n is $T_n^* = T_n + \Delta$. This is an approximation of the delivery process since the delays for each local warehouse are not identical. Besides demand, each local warehouse n faces demand for lateral transshipments y_n from other local warehouses and the adjusted demand rate is $g_n = \lambda_n + y_n$. We follow the approximation idea of Axsäter (1990a) and assume that the adjusted demand at the local warehouse n follows a Poisson process with rate g_n and is independent across the different local warehouses. y_n still needs to be determined and is determined by an approach in Alfredsson and Verrijdt (1999) which was first suggested in Axsäter (1990a). Denote f_n the fraction of total demand served by warehouse n. If the warehouse has stock on hand, then $f_n\lambda = \beta_n g_n$. Otherwise, it is zero.

$$f_n\lambda = \beta_n(\lambda_n + y_n) \tag{3.11}$$

Another interpretaion for $f_n\lambda$ is: the sum of the demand satisfied directly and the sum of demand satisfied by other warehouses through lateral transshipment times the probability that the warehouse n was the source for these lateral transshipments. Then,

$$f_n\lambda = \beta_n\lambda_n + \sum_{\substack{k=1 \\ k\neq n}}^{N} \omega_k^n \alpha_k \lambda_k \tag{3.12}$$

where ω_k^n is the conditional probability that warehouse k initiates a lateral transshipment from warehouse n and the shipment is possible. Combining both equations above, we get the following expression for the demand of lateral transshipments at warehouse n,

$$y_n = \frac{1}{\beta_n} \sum_{\substack{k=1 \\ k\neq n}}^{N} \omega_k^n \alpha_k \lambda_k. \tag{3.13}$$

For the random choice of a local warehouse to source the lateral transshipment, the expression for ω_k^n looks as follows:

$$\begin{aligned}
\omega_k^n &= Pr\{k \text{ request transshipment from } n | \text{transshipment is possible}\} \\
&= \frac{Pr\{\text{transshipment is possible and } k \text{ request transshipment from } n\}}{Pr\{\text{transshipment is possible}\}} \\
&= \frac{\beta_n \sum_{A\subseteq\{1..N\}\setminus\{n,k\}} \frac{1}{|A|+1} \prod_{\substack{j=1\\ j\neq n,k}}^{N} \{x_j\beta_j + (1-x_j)(1-\beta_j)\}}{\sum_{\substack{A\subseteq\{1..N\}\setminus\{k\}\\ A\neq\emptyset}} \prod_{\substack{j=1\\ j\neq k}}^{N}\{x_j\beta_j + (1-x_j)(1-\beta_j)\}}, \quad \begin{cases} x_j = 0 & \text{if} \quad j \notin A \\ x_j = 1 & \text{if} \quad j \in A \end{cases}
\end{aligned} \tag{3.14}$$

where A are all subsets of local warehouses of either the set $\{1..N\} \setminus \{n, k\}$ or the set $\{1..N\} \setminus \{k\}$, respectively. For $N = 2$ the probability always equals to one. For $N = 3$ the expressions for ω_2^1 and ω_3^1 look as follows (for ω_1^2, ω_3^2 and ω_1^3, ω_2^3 the expressions are analogous):

$$\omega_2^1 = \frac{\beta_1(1-\beta_3/2)}{\beta_1+\beta_3-\beta_1\beta_3} \quad \text{and} \quad \omega_3^1 = \frac{\beta_1(1-\beta_2/2)}{\beta_1+\beta_2-\beta_1\beta_2} \tag{3.15}$$

For identical local warehouses, this expression simplifies to $\omega_k^n = 1/(N-1)$. Hence, if transshipment is possible the conditional probability that the request comes from any location is identical for all locations and it follows, $y_n = \alpha_n\lambda_n/\beta_n$.

Evidently, β_n and α_n depend on y_n and g_n which vice versa depend on β_n and α_n. The parameters can be determined iteratively. Therefore, assume that g_n is known. With the arrival rate g_n, the lead time of T_n^* and applying the analysis from Subsection 3.2.2, the fraction of demand satisfied by the local warehouse n can be determined as follows: $\beta_n = 1 - p_0^n$ and $\alpha_n = 1 - \beta_n - \gamma - \theta$. The iterative computation procedure starts with $\alpha_n = y_n = 0$ for all local warehouses. The convergence is reached after a few iterations, similar to Axsäter (1990a) and Alfredsson and Verrijdt (1999).

The equations (11) to (14) are basically right out Axsäter (1990a), Alfredsson and Verrijdt (1999). Different to Alfredsson and Verrijdt (1999) the equation for expected cost TC per time unit of the system includes the penalty costs w_n for backordering the demand.

$$\begin{aligned}
TC = h_0 \sum_{i=1}^{S_0}\sum_{j=0}^{S} i\pi_{ij} + \sum_{n=1}^{N} h_n q_n \lambda_n + c_0 \sum_{n=1}^{N} \alpha_n\lambda_n + (c_1\gamma + c_2\theta)\lambda \\
+ r_0(1-\theta)\lambda + \sum_{n=1}^{N} r_n\beta_n g_n + \sum_{n=1}^{N} z_n(\alpha_n D_0 + \gamma D_1 + \theta D_2 + w_n)\lambda_n
\end{aligned} \tag{3.16}$$

where $h_0 \sum_{i=1}^{S_0}\sum_{j=0}^{S} i\pi_{ij} + \sum_{n=1}^{N} h_n q_n\lambda_n$ are the holding costs, $c_0\sum_{n=1}^{N}\alpha_n\lambda_n + (c_1\gamma + c_2\theta)\lambda$ are the costs for lateral and expedited deliveries, $r_0(1-\theta)\lambda + \sum_{n=1}^{N} r_n\beta_n g_n$ are the replenishment costs and $\sum_{n=1}^{N} z_n(\alpha_n D_0 + \gamma D_1 + \theta D_2 + w_n)\lambda_n$ are the penalty costs for waiting.

3.3 Model validation

3.3.1 Validation design

Our analysis is based on several approximations. In order to validate the accuracy of our model, we compare computation (com) and simulation (sim) results. The simulation results are obtained with a relative error of 0.01 and a confidence level of 95 % (see Law (2007)).

Similar to Alfredsson and Verrijdt (1999) we define 28 instances for identical locations where we vary λ_n, the base stock levels S_n and S_0, and the delivery time of the lateral transshipment D_0. We test the defined cases which are summarized in Table 3.2 for three different network sizes $N \in \{3, 5, 10\}$. For non-identical local warehouses, we test 38

Table 3.2 Overview of cases

Case	D_0	λ_n	S_n	S_0	Case	D_0	λ_n	S_n	S_0
1	0.5	0.02	1	1	15	1.5	0.02	1	1
2				2	16				2
3		0.06	1	1	17		0.06	1	1
4				2	18				2
5			2	1	19			2	1
6				2	20				2
7		0.1	1	2	21		0.1	1	2
8				3	22				3
9				6	23				6
10				10	24				10
11			2	2	25			2	2
12				3	26				3
13			3	2	27			3	2
14				3	28				3

different cases. Here we follow the numerical design in Alfredsson and Verrijdt (1999) and modify some of the cases from the Table 3.2 with respect to arrival rate (marked with "a") or replenishment lead time (marked with "b"). The demand rate and the replenishment lead time parameters were generated such that the average was equal to the identical case of the same case and the largest difference between the parameters of locations were equal to the average value. The generated values were randomly (uniform) assigned to the local warehouses.

In order to validate the cost accuracy, we assume the following cost parameters: $(c_0, D_0) = (10, 1.5)$, $(c_1, D_1) = (30, 3)$, $(c_2, D_2) = (100, 7)$, $r_0 = r_n = 10$. Direct deliveries from the plant are the most expensive option followed by expedited delivery from the central warehouse and lateral transshipment from other local warehouses. The regular replenishment is priced the lowest of all other delivery options. For the holding and penalty costs we consider three different cost settings $A : (z_n = 100, h_0 = h_n = 3)$,

$B : (z_n = 1000, h_0 = h_n = 3)$ and $C : (z_n = 100, h_0 = h_n = 14)$. For each instance we compare the total cost obtained by simulation with the results from our model.

3.3.2 Validation results

The accuracy of our model for γ, θ and β is excellent for identical local warehouses and it does not change with the size of the network. The results for identical locations are shown in Table 3.3. For non-identical locations, the β_n values differ between the locations. To evaluate the precision of our model when a network has asymmetric locations, we report the average (avg dif) and the maximum (max dif) difference between the computed and the simulated β_n values for all locations. Table 3.4 shows the results for asymmetric locations. For non-identical networks, the accuracy is excellent for γ and θ. The precision of parameter β_n for non-identical networks of size $N = 3$ is still well and it improves with the network size. We note that the values for α can be obtained from the relationship $\beta + \alpha + \theta + \gamma = 1$.

Table 3.5 illustrates the cost computation accuracy of the model compared to the simulation results for identical local warehouses and Table 3.6 illustrates the cost results for non-identical local warehouses. For symmetric and asymmetric networks we observe, that with increasing number of local warehouses the accuracy of our model increases in almost all the cases. As with the larger size, the costs connected with the local warehouses increase. Since our cost calculation on this level is exact, it is an intuitive observation of increasing accuracy for larger networks. To visualize the cost structure with the worst performance, we arbitrarily select 3% as threshold and highlight those cases where the difference is above in Tables 3.5 and 3.6. We find that the cost structure B with the highest difference between the holding and the penalty cost has the most highlighted examples (average difference (avg dif) = 2.84 %, standard deviation (SD)= 3.92 % for identical and avg dif = 2.32 %, SD = 3.04 % for non-identical locations), followed by the structure A (avg dif = 1.08 %, SD = 1.17 % for identical and avg dif = 0.88 %, SD = 0.89 % for non-identical locations) and the C structure (avg dif = 0.53 %, SD = 0.57 % for identical and avg dif = 0.48 %, SD = 0.47 % for non-identical locations) showing the least highlighted cases. The C structure is characterized by the smallest difference in the holding and penalty cost among the three tested cost structures. Apparently, our model achieves the highest precision for the system where the difference between the holding and penalty cost is not tremendous.

These results are in line with the findings of Alfredsson and Verrijdt (1999) in terms of the same level of accuracy of the performance measures and cost. Since our modeling on the level of local warehouses is exact as in Alfredsson and Verrijdt (1999), the errors result from the approximated shipment times from the central warehouse, the METRIC approximation and the iterative solution for α_n and β_n.

Table 3.3 Performance results for identical local warehouses, $T_0 = 15$ and $T_n = 3$

	$N = 3$						$N = 5$						$N = 10$					
	γ		θ		β_n		γ		θ		β_n		γ		θ		β_n	
Case	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim
1	0.00	0.00	0.02	0.02	0.85	0.85	0.00	0.00	0.01	0.01	0.81	0.81	0.00	0.00	0.00	0.00	0.75	0.75
2	0.00	0.00	0.00	0.00	0.92	0.92	0.00	0.00	0.00	0.00	0.89	0.89	0.00	0.00	0.00	0.00	0.83	0.83
3	0.00	0.00	0.22	0.22	0.49	0.49	0.00	0.00	0.21	0.21	0.37	0.37	0.00	0.00	0.18	0.18	0.25	0.24
4	0.00	0.00	0.13	0.12	0.63	0.63	0.00	0.00	0.14	0.14	0.48	0.48	0.00	0.00	0.14	0.14	0.30	0.30
5	0.00	0.00	0.03	0.03	0.84	0.82	0.00	0.00	0.01	0.01	0.78	0.77	0.00	0.00	0.00	0.00	0.73	0.73
6	0.00	0.00	0.01	0.01	0.92	0.90	0.00	0.00	0.00	0.00	0.86	0.84	0.00	0.00	0.00	0.00	0.78	0.77
7	0.00	0.00	0.31	0.31	0.41	0.41	0.00	0.00	0.35	0.35	0.26	0.26	0.00	0.00	0.39	0.39	0.15	0.13
8	0.00	0.00	0.21	0.21	0.51	0.51	0.00	0.00	0.28	0.28	0.32	0.32	0.00	0.00	0.34	0.34	0.16	0.15
9	0.01	0.01	0.05	0.05	0.71	0.71	0.00	0.00	0.11	0.11	0.54	0.53	0.00	0.00	0.22	0.22	0.24	0.24
10	0.03	0.03	0.00	0.00	0.76	0.76	0.00	0.00	0.02	0.02	0.71	0.71	0.00	0.00	0.10	0.10	0.40	0.40
11	0.00	0.00	0.08	0.08	0.72	0.70	0.00	0.00	0.07	0.08	0.59	0.57	0.00	0.00	0.06	0.06	0.45	0.43
12	0.00	0.00	0.05	0.05	0.82	0.79	0.00	0.00	0.05	0.05	0.67	0.65	0.00	0.00	0.04	0.04	0.49	0.48
13	0.00	0.00	0.01	0.01	0.91	0.89	0.00	0.00	0.00	0.00	0.85	0.84	0.00	0.00	0.00	0.00	0.80	0.79
14	0.00	0.00	0.00	0.00	0.95	0.94	0.00	0.00	0.00	0.00	0.90	0.88	0.00	0.00	0.00	0.00	0.83	0.82
15	0.00	0.00	0.02	0.01	0.87	0.87	0.00	0.00	0.01	0.01	0.83	0.83	0.00	0.00	0.00	0.00	0.77	0.77
16	0.00	0.00	0.00	0.00	0.94	0.94	0.00	0.00	0.00	0.00	0.91	0.91	0.00	0.00	0.00	0.00	0.85	0.85
17	0.00	0.00	0.20	0.20	0.52	0.52	0.00	0.00	0.19	0.19	0.39	0.39	0.00	0.00	0.16	0.16	0.27	0.27
18	0.00	0.00	0.11	0.11	0.67	0.67	0.00	0.00	0.12	0.12	0.51	0.51	0.00	0.00	0.12	0.12	0.33	0.32
19	0.00	0.00	0.02	0.02	0.86	0.84	0.00	0.00	0.01	0.01	0.81	0.80	0.00	0.00	0.00	0.00	0.76	0.76
20	0.00	0.00	0.01	0.01	0.93	0.92	0.00	0.00	0.00	0.00	0.88	0.86	0.00	0.00	0.00	0.00	0.81	0.80
21	0.00	0.00	0.28	0.28	0.44	0.44	0.00	0.00	0.32	0.32	0.28	0.28	0.00	0.00	0.36	0.36	0.16	0.14
22	0.00	0.00	0.19	0.19	0.56	0.55	0.00	0.00	0.25	0.25	0.35	0.35	0.00	0.00	0.31	0.31	0.18	0.17
23	0.00	0.00	0.04	0.04	0.79	0.79	0.00	0.00	0.09	0.09	0.59	0.59	0.00	0.00	0.19	0.19	0.27	0.26
24	0.01	0.01	0.00	0.00	0.85	0.85	0.00	0.00	0.01	0.01	0.80	0.80	0.00	0.00	0.07	0.07	0.46	0.46
25	0.00	0.00	0.07	0.07	0.76	0.73	0.00	0.00	0.06	0.06	0.63	0.61	0.00	0.00	0.04	0.04	0.49	0.48
26	0.00	0.00	0.04	0.04	0.85	0.82	0.00	0.00	0.04	0.04	0.71	0.69	0.00	0.00	0.03	0.03	0.54	0.53
27	0.00	0.00	0.01	0.01	0.93	0.91	0.00	0.00	0.00	0.00	0.88	0.87	0.00	0.00	0.00	0.00	0.84	0.82
28	0.00	0.00	0.00	0.00	0.97	0.95	0.00	0.00	0.00	0.00	0.92	0.90	0.00	0.00	0.00	0.00	0.87	0.85

Table 3.4 Performance results for non-identical local warehouses, $T_0 = 15$

	$N=3$						$N=5$						$N=10$					
	γ		θ		β_n		γ		θ		β_n		γ		θ		β_n	
Case	com	sim	com	sim	avg dif	max dif	com	sim	com	sim	avg dif	max dif	com	sim	com	sim	avg dif	max dif
1a	0.00	0.00	0.02	0.02	0.00	0.02	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
1b	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2a	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
2b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3a	0.00	0.00	0.22	0.22	0.00	0.02	0.00	0.00	0.21	0.21	0.00	0.01	0.00	0.00	0.18	0.18	0.00	0.01
4b	0.00	0.00	0.13	0.12	0.00	0.01	0.00	0.00	0.14	0.14	0.00	0.00	0.00	0.00	0.14	0.14	0.00	0.00
5a	0.00	0.00	0.03	0.03	0.02	0.03	0.00	0.00	0.01	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.01
6b	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01
7a	0.00	0.00	0.31	0.31	0.00	0.02	0.00	0.00	0.35	0.35	0.00	0.01	0.00	0.00	0.39	0.39	0.02	0.02
8a	0.00	0.00	0.21	0.21	0.00	0.02	0.00	0.00	0.28	0.28	0.00	0.01	0.00	0.00	0.34	0.34	0.01	0.02
8b	0.00	0.00	0.21	0.21	0.01	0.01	0.00	0.00	0.28	0.27	0.00	0.01	0.00	0.00	0.34	0.34	0.02	0.02
9a	0.01	0.01	0.05	0.05	0.00	0.01	0.00	0.00	0.11	0.11	0.00	0.02	0.00	0.00	0.22	0.22	0.00	0.01
10a	0.03	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.01	0.00	0.00	0.10	0.10	0.00	0.01
10b	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.10	0.10	0.00	0.00
11a	0.00	0.00	0.08	0.08	0.02	0.05	0.00	0.00	0.07	0.07	0.02	0.04	0.00	0.00	0.06	0.06	0.01	0.02
12a	0.00	0.00	0.05	0.05	0.03	0.05	0.00	0.00	0.05	0.05	0.02	0.04	0.00	0.00	0.04	0.04	0.01	0.02
13a	0.00	0.00	0.01	0.01	0.02	0.03	0.00	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.01	0.02
13b	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01
14a	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.02
15a	0.00	0.00	0.02	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
15b	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16a	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
16b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17a	0.00	0.00	0.20	0.20	0.00	0.02	0.00	0.00	0.19	0.19	0.00	0.01	0.00	0.00	0.16	0.16	0.00	0.00
18b	0.00	0.00	0.11	0.11	0.00	0.00	0.00	0.00	0.12	0.12	0.00	0.00	0.00	0.00	0.12	0.11	0.00	0.00
19a	0.00	0.00	0.02	0.02	0.02	0.03	0.00	0.00	0.01	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.02
20b	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01
21a	0.00	0.00	0.28	0.28	0.00	0.02	0.00	0.00	0.32	0.32	0.00	0.01	0.00	0.00	0.36	0.36	0.02	0.02
22a	0.00	0.00	0.19	0.19	0.00	0.02	0.00	0.00	0.25	0.25	0.00	0.01	0.00	0.00	0.31	0.31	0.01	0.02
22b	0.00	0.00	0.19	0.18	0.01	0.01	0.00	0.00	0.25	0.25	0.00	0.01	0.00	0.00	0.31	0.31	0.01	0.01
23a	0.00	0.00	0.04	0.04	0.00	0.01	0.00	0.00	0.09	0.09	0.00	0.03	0.00	0.00	0.19	0.19	0.00	0.01
24a	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.07	0.07	0.00	0.02
24b	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.07	0.07	0.00	0.01
25a	0.00	0.00	0.07	0.07	0.03	0.06	0.00	0.00	0.06	0.06	0.02	0.04	0.00	0.00	0.04	0.04	0.01	0.02
26a	0.00	0.00	0.04	0.04	0.03	0.06	0.00	0.00	0.04	0.04	0.02	0.04	0.00	0.00	0.03	0.03	0.01	0.03
27a	0.00	0.00	0.01	0.01	0.02	0.03	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.02	0.02
27b	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.02
28a	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.02	0.02

Table 3.5 Cost results for identical local warehouses, $T_0 = 15$ and $T_n = 3$

	$N=3$						$N=5$						$N=10$					
	A		B		C		A		B		C		A		B		C	
Case	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim
1	12.12	12.10	29.56	29.43	44.46	44.42	18.18	18.18	47.86	47.89	64.52	64.50	34.52	34.48	104.6	104.2	116.0	116.00
2	13.87	13.86	21.40	21.29	57.05	57.00	19.44	19.43	34.82	34.82	76.68	76.65	34.89	34.86	82.84	82.59	127.3	127.3
3	47.86	47.88	372.4	372.8	64.20	64.15	81.09	81.30	653.0	654.9	100.3	100.4	160.8	161.1	1,329	1,333	185.5	185.4
4	34.89	34.76	237.8	236.9	58.74	58.49	63.91	64.06	482.0	483.4	89.66	89.79	140.5	140.6	1,132	1,134	170.5	170.4
5	22.89	23.35	**86.94**	**91.43**	65.20	65.60	35.64	36.17	**141.6**	**146.7**	97.91	98.39	69.24	69.59	293.4	296.8	181.7	182.0
6	21.55	22.06	**51.15**	**56.02**	74.29	74.73	33.78	34.46	**100.0**	**106.6**	106.7	107.3	67.47	67.93	251.7	256.1	190.8	191.2
7	95.48	95.56	793.0	794.0	109.4	109.3	181.8	182.4	1,553	1,559	194.6	195.1	403.3	406.3	3,501	3,533	416.3	417.4
8	75.08	74.85	595.4	593.9	94.49	94.02	155.2	155.7	1,304	1,308	171.8	172.1	373.4	375.6	3,227	3,251	388.1	388.8
9	42.82	42.36	250.3	246.5	85.82	85.12	91.97	91.98	691.9	692.2	125.2	125.0	287.7	288.4	2,434	2,442	310.2	310.4
10	43.22	43.24	163.6	163.8	127.4	127.3	58.67	58.29	310.4	307.2	126.8	126.2	189.7	189.7	1,500	1,500	231.3	231.2
11	43.93	45.19	**280.0**	**292.1**	77.44	78.61	77.71	79.32	545.8	560.8	118.2	119.8	160.6	162.0	1,204	1,218	216.2	217.3
12	**34.98**	**36.31**	**178.3**	**191.0**	77.32	78.55	65.02	66.93	**410.5**	**428.5**	113.8	115.7	145.6	147.0	1,048	1,062	209.0	210.3
13	**29.41**	**30.41**	**82.79**	**92.37**	91.68	92.58	47.41	48.49	**159.0**	**169.2**	135.9	136.9	95.07	95.99	371.7	382.5	250.2	250.1
14	**29.04**	**29.98**	**55.33**	**64.36**	101.9	102.8	46.07	47.29	**121.8**	**133.5**	145.3	146.4	93.31	94.27	**329.8**	**341.2**	259.4	259.4
15	11.95	11.94	28.10	28.04	44.26	44.23	18.00	17.99	46.41	46.25	64.34	64.30	34.37	34.23	103.5	102.1	115.9	115.7
16	13.76	13.75	20.46	20.38	56.94	56.90	19.31	19.30	33.71	33.61	76.56	76.52	34.71	34.62	81.38	80.45	127.2	127.1
17	45.89	46.09	357.0	359.0	61.64	61.80	78.23	78.56	631.9	635.0	96.37	96.68	157.0	157.1	1,307	1,309	179.4	179.2
18	32.98	33.04	222.4	223.2	56.35	56.34	61.11	61.37	460.5	463.0	86.01	86.26	136.5	136.2	1,108	1,105	164.6	164.2
19	22.16	22.68	**80.99**	**85.95**	64.33	64.79	34.91	35.34	135.8	139.8	97.11	97.47	68.71	68.54	287.8	287.8	182.1	181.1
20	21.13	21.67	**47.64**	**52.85**	73.81	74.29	33.29	33.92	**96.04**	**102.1**	106.2	106.8	66.98	67.06	246.1	248.6	191.3	190.5
21	92.21	92.81	768.7	774.4	104.9	105.5	178.5	179.8	1,535	1,547	189.4	190.6	403.3	409.4	3,528	3,592	412.6	417.3
22	71.27	71.70	566.1	570.3	89.55	89.89	151.2	152.2	1,278	1,288	165.8	166.8	372.8	377.7	3,251	3,301	383.7	387.6
23	38.99	38.80	218.6	217.2	81.37	81.00	86.60	87.00	650.9	654.8	118.4	118.8	284.4	286.3	2,432	2,451	303.1	304.8
24	40.67	40.59	141.3	140.7	124.9	124.7	54.42	53.87	274.9	269.8	122.3	121.6	182.9	182.9	1,458	1,458	221.9	221.8
25	**41.32**	**42.97**	**259.2**	**274.9**	74.17	75.77	73.86	75.59	**516.6**	**533.0**	113.2	114.9	154.8	155.8	1,164	1,176	209.0	209.2
26	**32.94**	**34.56**	**161.6**	**177.2**	74.85	76.40	**61.72**	**63.64**	**384.9**	**403.2**	109.7	111.6	140.0	141.2	1,008	1,020	202.8	203.0
27	**28.67**	**29.71**	**76.69**	**86.72**	90.82	91.77	46.66	47.55	**152.8**	**161.5**	135.3	135.9	93.69	94.58	**346.6**	**370.6**	255.1	249.1
28	**28.60**	**29.57**	**51.70**	**61.02**	101.5	102.3	45.54	46.64	**117.3**	**128.1**	145.0	145.8	92.14	93.09	**306.4**	**331.2**	264.6	258.5

Table 3.6 Cost results for non-identical local warehouses, $T_0 = 15$

	$N=3$						$N=5$						$N=10$					
	A		B		C		A		B		C		A		B		C	
Case	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim	com	sim
1a	12.27	12.21	30.99	30.47	44.61	44.52	18.41	18.36	50.02	49.62	64.75	64.68	34.95	34.88	108.7	108.0	116.4	116.3
1b	12.12	12.09	29.49	29.29	44.49	44.42	18.18	18.17	47.79	47.77	64.55	64.52	34.52	34.48	104.5	104.2	116.0	116.0
2a	13.97	13.94	22.32	22.03	57.14	57.08	19.60	19.57	36.36	36.07	76.84	76.78	35.26	35.18	86.39	85.60	127.7	127.6
2b	13.87	13.85	21.33	21.18	57.07	57.02	19.44	19.44	34.75	34.77	76.71	76.68	34.88	34.85	82.73	82.42	127.3	127.3
3a	48.32	48.20	376.8	375.8	64.67	64.49	81.51	81.44	657.0	656.4	100.7	100.6	161.2	161.4	1,333	1,335	185.9	185.6
4b	34.79	34.43	236.4	233.6	58.81	58.23	63.80	63.72	480.5	480.1	89.73	89.55	140.3	140.0	1,130	1,128	170.6	167.0
5a	23.73	24.09	**94.86**	**98.32**	66.04	66.36	36.84	37.20	152.9	156.3	99.11	99.48	71.37	71.59	313.5	315.7	183.8	184.0
6b	21.57	22.07	**51.21**	**56.00**	74.38	74.79	33.79	34.47	**99.97**	**106.4**	106.8	107.4	67.47	67.92	251.4	255.6	191.0	191.4
7a	96.14	96.00	799.2	798.1	110.0	109.7	182.3	182.8	1,558	1,563	195.0	195.4	403.6	406.3	3,503	3,534	416.6	417.4
8a	75.92	75.46	603.4	599.7	95.34	94.63	155.9	156.2	1,310	1,313	172.4	172.6	373.7	375.7	3,229	3,253	388.5	389.0
8b	74.66	73.95	590.7	585.4	94.42	93.27	154.8	154.7	1,300	1,299	171.7	171.2	372.9	374.4	3,221	3,240	388.0	387.8
9a	43.79	43.26	259.4	255.0	86.79	86.02	93.13	92.69	702.8	699.2	126.4	125.7	288.3	288.7	2,440	2,445	310.9	310.8
10a	44.15	44.16	172.3	172.5	128.3	128.2	59.97	59.41	322.7	317.9	128.1	127.3	191.2	190.7	1,514	1,510	232.8	232.3
10b	42.50	41.96	155.3	150.7	127.3	126.5	57.84	57.02	300.9	294.2	126.6	125.4	188.7	187.7	1,488	1,481	231.1	229.7
11a	45.43	46.31	294.2	302.6	78.94	79.74	79.41	80.61	561.8	573.3	119.9	121.0	162.7	163.8	1,224	1,235	218.2	219.1
12a	36.46	37.39	**192.2**	**201.2**	78.81	79.65	66.96	68.25	428.9	441.2	115.8	116.9	148.0	149.1	1,071	1,083	211.4	212.4
13a	30.95	31.79	**97.27**	**105.4**	93.22	93.97	49.88	50.75	**182.2**	**190.7**	138.4	139.1	99.66	100.4	415.0	423.6	254.8	254.6
13b	**29.45**	**30.42**	**82.92**	**92.20**	91.84	92.68	47.45	48.50	**158.9**	**169.0**	136.1	137.0	95.09	96.02	371.3	382.3	250.5	250.4
14a	30.12	31.00	**65.54**	**74.00**	103.0	103.8	48.25	49.23	**142.5**	151.9	147.5	148.4	97.80	98.48	372.3	380.9	263.9	263.7
15a	12.10	12.06	29.56	29.17	44.42	44.35	18.24	18.17	48.64	48.01	64.58	64.49	34.83	34.63	107.7	105.8	116.3	116.1
15b	11.95	11.93	28.03	27.93	44.29	44.24	18.00	17.98	46.34	46.15	64.37	64.32	34.37	34.24	103.3	102.0	115.9	115.8
16a	13.86	13.83	21.35	21.11	57.03	56.98	19.47	19.42	35.24	34.78	76.72	76.65	35.10	34.93	85.01	83.44	127.6	127.4
16b	13.76	13.75	20.38	20.30	56.96	56.92	19.31	19.29	33.63	33.50	76.59	76.55	34.71	34.61	81.27	80.26	127.2	127.1
17a	46.42	46.42	362.0	362.1	62.17	62.14	78.73	78.77	636.6	637.0	96.87	96.91	157.4	157.4	1,312	1,312	179.9	179.5
18b	32.86	32.74	220.9	220.2	56.41	56.13	60.97	61.03	458.8	459.6	86.06	86.02	136.2	135.6	1,105	1,099	164.6	163.6
19a	23.06	23.44	**89.49**	**93.16**	65.24	65.58	36.22	36.43	148.1	150.0	98.43	98.63	71.07	70.63	310.1	307.6	184.5	183.2
20b	21.15	21.68	**47.71**	**52.85**	73.90	74.34	33.30	33.95	**96.00**	**102.2**	106.3	106.9	66.98	67.05	245.8	248.2	191.5	190.7
21a	93.00	93.23	776.1	778.5	105.7	105.9	179.1	180.2	1,540	1,551	190.0	191.0	403.5	409.5	3,531	3,593	412.9	417.5
22a	72.29	72.33	575.7	576.4	90.58	90.51	151.9	152.8	1,285	1,293	166.6	167.4	373.2	377.8	3,254	3,301	384.1	387.6
22b	70.75	70.58	560.3	559.8	89.40	88.91	150.59	151.14	1,272	1,278	165.6	165.8	371.9	376.2	3,241	3,287	383.2	386.1
23a	40.14	39.78	229.5	226.5	82.55	82.02	88.06	87.89	664.8	663.4	119.9	119.6	285.2	286.6	2,440	2,455	304.0	305.2
24a	41.73	41.62	151.4	150.5	126.0	125.7	55.96	55.11	289.6	281.7	123.9	122.8	184.9	184.0	1,476	1,468	223.9	222.9
24b	39.80	39.20	**131.5**	**126.4**	124.7	123.9	53.47	52.66	264.2	257.1	122.1	121.0	181.6	180.8	1,442	1,437	221.4	220.2
25a	43.03	44.18	**275.4**	**286.4**	75.91	77.00	75.89	77.13	535.9	547.7	115.3	116.4	157.5	157.8	1,190	1,194	211.7	211.2
26a	**34.59**	**35.77**	**177.3**	**188.6**	76.53	77.64	64.01	65.28	406.6	418.8	112.1	113.2	143.1	143.5	1,037	1,042	206.0	205.3
27a	30.32	31.19	**92.39**	**100.8**	92.51	93.28	49.38	50.03	**178.5**	**185.1**	138.0	138.4	98.78	99.24	**394.8**	**414.6**	260.3	253.8
27b	**28.71**	**29.71**	**76.86**	**86.53**	90.97	91.86	46.70	47.57	**152.8**	**161.4**	135.5	136.1	93.73	94.61	**346.4**	**370.3**	255.4	249.4
28a	29.73	30.66	**62.40**	**71.31**	102.6	103.4	47.90	48.71	**139.6**	**147.8**	147.4	147.9	97.02	97.65	**352.6**	**374.3**	269.5	263.2

3.4 Effect of pipeline inventory

3.4.1 Experimental design

To evaluate the cost reduction effect caused by including pipeline inventory compared to a situation without pipeline information, let TC^* denote the minimal expected average cost per unit of time and S_n^* the optimal base stock levels. We use simulation and local search to find the optimal base stock levels S_n^* and to obtain the optimal cost TC^* for each instance. Define

$$\delta = \frac{TC^*_{No} - TC^*}{TC^*} \tag{3.17}$$

as the potential cost savings where real-time supply information is used TC^* compared to the optimal total cost TC^*_{No} without pipeline information.

We define the emergency structure factor (e) in Table 3.7, which states the portion of delivery times of lateral transshipment compared to normal replenishment. The emergency

Table 3.7 Emergency delivery structure

From	Delivery time	Delivery cost structure (M)		
		$M = 1$	$M = 2$	$M = 3$
Local warehouse	$D_0 = T_n \cdot e$	$c_0 = r_n/e$	$c_0 = r_n/e$	$c_0 = r_n/e$
Central warehouse	$D_1 = 2 \cdot T_n \cdot e$	$c_1 = 1.1 \cdot r_n/e$	$c_1 = 1.5 \cdot r_n/e$	$c_1 = 2 \cdot r_n/e$
Plant	$D_2 = D_1 + T_0 \cdot e$	$c_2 = 1.2 \cdot r_n/e$	$c_2 = 2 \cdot r_n/e$	$c_2 = 5 \cdot r_n/e$

factor e can be seen as an acceleration factor of emergency deliveries compared to the normal replenishments. Note that the delivery costs are increasing and delivery times are decreasing in e. Furthermore, we study different delivery cost structures (M). For $M = 1$ we assume a slow shipment cost increase for expedited deliveries. The expedited delivery from the central warehouse has 10 percent higher shipment cost and the shipment from the plant is 20 percent more expensive than the lateral transshipment. $M = 2$ cost structure is moderate and the extreme case $M = 3$ where the expedited delivery from the central warehouse is twice and the expedited delivery from the plant is five times more costly than the lateral transshipment.

We assume that all warehouses have identical unit holding cost (h_n) per unit of time and identical replenishment cost (r_n) per unit. We fixed the replenishment time of the central warehouse ($T_0 = 15$) and the replenishment cost to ($r_n = 10$). All other varied cost parameters are chosen relative to the fixed cost parameters. Since we are mainly interested in the general benefits of supply chain information, we focus on identical local warehouses. Our numerical study is based on 3888 experiments that were generated using all possible combinations of parameters shown in Table 3.8.

Table 3.8 Input parameter values

Parameter	Values
Number of warehouses (N)	3, 5, 10
Delivery cost structure (M)	1, 2, 3
Arrival rate (λ_n)	0.02, 0.06, 0.10, 0.15
Replenishment time (T_n)	3, 5, 7
Emergency structure factor (e)	1/4, 1/3, 1/2
Unit holding cost per time unit (h_n)	1, 3, 8, 14
Penalty cost per time unit for waiting (z_n)	100, 500, 1000

3.4.2 Numerical results

The average cost saving from using pipeline information is $\bar{\delta} = 4.25\%$ (SD $= 2.05\%$) in a network with three identical warehouses. These savings slightly increase with the size of the network to $\bar{\delta} = 4.40\%$ (SD $= 2.26\%$) for $N = 5$ and $\bar{\delta} = 4.68\%$ (SD $= 2.51\%$) for $N = 10$. Table 3.9 shows the average potential cost savings by single parameter values. We observe that the average cost savings increase with the demand rate λ_n and the delivery time T_n. We find that holding cost (h_n) increases the advantage of the pipeline information. Higher holding cost makes it more economically attractive to use the flexibility options like emergency deliveries more frequently. In this case real time supply information provides the fastest replenishment option. The savings decrease in z_n, which is not an intuitive result since one would expect the opposite. Since we simultaneously optimize the base stock levels of the network, the increasing penalty cost has two overlapping effects. Firstly, higher penalty cost makes it expensive to wait for the regular arrival of outstanding orders which increases the benefit from pipeline information. Secondly, penalty cost leads to higher optimal base stock levels and the holding cost ratio on the total cost increases compared to the ratio of the penalty cost. Observing the e parameter, the savings increase in emergency structure factor. The intuition behind this fact is that the faster emergency delivery structure allows less average savings because shorter lateral transshipment times imply a smaller probability that the demand can be backordered and served by the next outstanding order. Consequently, the nearly instant emergency delivery structure where lateral transshipments have nearly zero lead time would lead to no potential cost saving. From Table 3.9 we can conclude that the cost for expedited deliveries from upper echelons (M) have a positive impact on savings, while the effect is very small. The presented cost savings using pipeline inventory are considerable, since the model without pipeline information is already optimized and accounts for all regular flexibility benefits and parameter optimization.

Studying the optimal base stock levels, we observe that pipeline information in general leads to lower optimal base stock levels at the central and at the aggregated local warehouse. Table 3.10 summarizes the average optimal total base stock in the network and the split between the base stock levels at the central warehouse S_0 and the aggregated local

Table 3.9 Cost reductions $\bar{\delta}$ (in percent)

Parameter	Value	$N = 3$	$N = 5$	$N = 10$
M	1	4.22	4.39	4.67
	2	4.24	4.40	4.68
	3	4.29	4.43	4.70
λ_n	0.02	3.18	3.45	3.59
	0.06	4.29	4.34	4.71
	0.10	4.66	4.84	5.08
	0.15	4.86	4.98	5.36
e	1/4	3.13	3.21	3.46
	1/3	3.96	4.16	4.32
	1/2	5.66	5.84	6.27
T_n	3	3.50	3.60	3.92
	5	4.31	4.45	4.73
	7	4.94	5.16	5.40
h_n	1	2.95	3.11	3.47
	3	4.07	4.07	4.30
	8	4.65	4.94	5.31
	14	5.31	5.49	5.66
z_n	100	4.88	5.10	5.29
	500	4.07	4.12	4.48
	1000	3.80	3.99	4.28
Average		4.25	4.40	4.68

warehouse S over the conducted numerical study for a model without pipeline information (No Info) and with pipeline information (With Info). It is also interesting to note that

Table 3.10 Average optimal base stock levels

	No Info			With Info		
N	Total Stock	S_0 (in %)	S (in %)	Total Stock	S_0 (in %)	S (in %)
3	12.17	34.3	65.7	11.89	33.7	66.3
5	19.14	35.0	65.0	18.75	34.3	65.7
10	36.51	35.1	64.9	35.93	34.7	65.3

the use of pipeline information moves the inventory of the system downstream to the local warehouses, which can be seen on the higher S values. This effect is small but significantly positive.

We also find that the pipeline information changes the supply mix of the network. Table 3.11 summarizes the changes of the ratios satisfied by different delivery options for studied network sizes. We observe that with information the average ratio of demand satisfied by plant (θ) and through lateral transshipments (α) decrease but the average ratio of demand satisfied directly by the local warehouse (β) increases. The average ratio of demand satisfied by the central warehouse with and without information is ($\gamma < 0.000$) for all three tested network sizes. Since the ratio β with information also comprises the part of demand satisfied through pipeline inventory (backorders), we find an increase of this parameter and a decrease of other delivery options in all studied examples. Apparently, the use of pipeline information reduces the importance of lateral transshipments and expensive expedited deliveries from upper echelons in the supply mix of the inventory network.

Table 3.11 Average ratios of the supply options

	θ		β		α	
N	No Info	With Info	No Info	With Info	No Info	With Info
3	0.016	0.013	0.910	0.922	0.075	0.064
5	0.009	0.008	0.893	0.909	0.098	0.084
10	0.005	0.004	0.876	0.897	0.120	0.099

3.5 Conclusion

We incorporate pipeline information into the decision when to initiate a lateral transshipment and when to wait until the arrival of the next outstanding order using a Markov-chain approach. In a numerical study for symmetric and asymmetric local warehouses, we show that the service and cost performance is close to the results obtained by simulation. The results indicate that incorporating pipeline information can lead to significant cost savings. Increasing the arrival rate, replenishment time and the holding cost increases the average cost savings of pipeline information whereas the cost saving decreases with higher penalty cost. We also find that a faster but more expensive emergency delivery structure allows less average savings because shorter lateral transshipment times imply a smaller probability for the demand to be backordered and served by the next outstanding order. Furthermore, we observe that pipeline information increases the fraction of demand satisfied by local warehouses and decreases the fraction of demand satisfied through lateral transshipment and expedited deliveries. Apparently, pipeline information reduces the importance of direct deliveries from upper echelons. Another observation is that real time supply information can lead to smaller optimal base stock levels of the total inventory system. However, at the same time we can observe a trend that pipeline information moves the inventory downstream to the local warehouses.

We have made a number of assumptions to make our model tractable. These include deterministic replenishment times and identical waiting times within a pooling group for backordering the demand. The relaxation of these assumptions and the implementation of pipeline inventory of the upper echelons (e.g. central warehouse) in the decision when to request an emergency delivery may be interesting areas for future research. In order to increase the precision of the model, another research path may be to model the local warehouses exactly (e.g. Graves (1985)) instead as aggregated. The model here also does not consider an optimal transshipment rule because it chooses the fastest delivery option. For a cost structure where the penalty cost is low compared to the delivery cost of emergency supply the cost optimal policy may have a significant impact on the inventory network performance.

Chapter 4

The Impact of Cost and Quality on Illegal Goods in the Supply Chain

4.1 Introduction

The positioning strategy of a product in a competitive market has an important influence on the business success. Usually a supply channel of a product consists of firms with manufacturing or distributing functions. Manufacturers optimize the profit of their own products and designated markets. The retailers profit optimization problem might include products and markets which differ from those of manufacturers. Coordination between firms with manufacturing or distributing functions is crucial for the business performance of each involved firm.

This chapter is motivated by several manuscripts which study the concept of profit-maximizing firms competing on two dimensions: product quality and product price. This concept was widely addressed in several research streams which have different problem settings but very similar analysis methods. There are various ways to source the products from the manufacturer through retailers to the markets. Figure 4.1 is general enough to embrace several major problem categories in one flow structure. We identify several research problems which emanate from the general structure illustrated in Figure 4.1. We first present the general structure and address the literature which is related to our work, then we present specific cases providing consecutively examples for each one. We explain how each case is related to the general structure and present the literature which investigates this particular research problem.

In the general setting we consider two products L and F, which compete in quality and price on the market L. The product L is produced by a manufacturer M_L and distributed to the market L by the retailer R_L. The manufacturer M_F which produces the product F can supply the market L either directly trough the retailer R_F or through the retailer R_L. Product F can also be offered on the domestic market D. In the market L is a two product

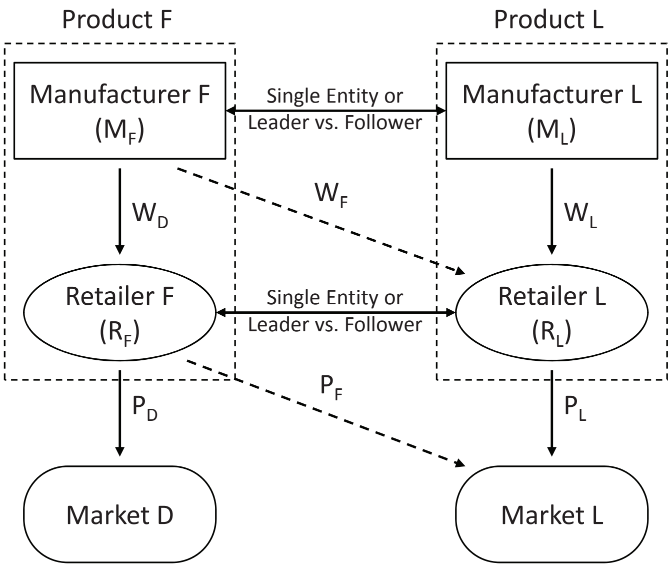


Figure 4.1 General structure of the players in the supply chain

competition whereas the market D is a single-product market without competition. When the product F is illegal (e.g. counterfeit of a branded product), the supply to the market L has a certain risk of interdiction. Two competing products in the market L do not necessarily imply that all the players in both supply chains are competing with each other. The manufacturers and the retailers might also build a joint firm. This cooperation might be horizontal, between manufacturers (or retailers) or vertical between manufacturer and retailer. When two manufacturers (or retailers) build a joint firm then they jointly optimize their profit even if the entities are geographically separated, single entities (SE). Otherwise, we assume that there exists a leader vs. follower ($L\, vs.\, F$) setting where the leader has the power to set standards in the market. In case of vertical integration we assume that the manufacturer offers the retailer the product to the production cost.

The firm's price decisions of horizontally differentiated products were addressed by McGuire and Staelin (1983). They study two distribution channels with producers and intermediate sellers. They distinguish between three different structures of the players in the market. Coughlan (1985) generalizes these results to cases of demand concavity or convexity with respect to own price. Lus and Muriel (2009) study different linear modeling approaches for two substitutable products on pricing. They find that one leads to unrealistic results when the products are more substitutable. In these cited studies the authors do not consider the quality level of the products or assume that the quality is fixed. The firms decisions for vertically differentiated products were studied one of the first by Dorfman and Steiner (1954). They analyze the optimal firm's decision in price-quantity-advertising and price-quantity-quality constellations. Mussa and Rosen (1978) provide background on the monopoly model where the consumers valuate the product in price and quality. A model for two identical firms, competing on product quality and price is studied by Moorthy (1988). They study different market scenarios: market entry, monopoly and competition, concluding for each scenario the optimal quality decisions for the firms. Banker et al. (1998) study differences in quality levels under different competition intensities. They conclude that the quantity changes strongly depend on the definition of increased competition.

We identify seven different research streams (but there might be more) which are related to the general supply chain structure presented in Figure 4.1. By setting the prices P or the wholesale prices W to infinite we close the supply of a market or a retailer, respectively. Contrary, setting $(0 < P)$ or $(W < \infty)$ we endorse the supply of a market or a retailer, respectively. With regards to structure between both manufacturers or retailers we distinguish between SE and $L\,vs.\,F$ settings. Modifying the parameters according to

Table 4.1 Parameters settings for particular problem

Case		M_L vs M_F	R_L vs R_F	W_D	W_F	P_D	P_F
1	Parallel Imports	SE	$L\,vs.\,F$	$< W_L$	∞	> 0	$< P_L$
2	National and Generic Brand	$L\,vs.\,F$	SE	$< W_L$	∞	∞	$< P_L$
3	National and Store Brand	$L\,vs.\,F$	SE	C_F	∞	∞	$< P_L$
4	Two National Brands	SE	SE	$< W_L$	∞	∞	$< P_L$
5	Non Deceptive Counterfeit	$L\,vs.\,F$	$L\,vs.\,F$	$< W_L$	∞	∞	$< P_L$
6	Deceptive Counterfeit	$L\,vs.\,F$	$L\,vs.\,F$	∞	$< W_L$	∞	∞
7	Identical Smuggled Good	SE	$L\,vs.\,F$	$< W_L$	$< W_L$	> 0	∞

$W_L > 0$ and $P_L > 0$ in each case

Single Entity - (SE), Leader vs Follower - ($L\,vs.\,F$)

Table 4.1 the general supply flow can be adjusted to each particular case. These cases are:

Case 1: Parallel Imports are unauthorized flows of products between two regions with price differences, which compete with the authorized distribution channel. Both channels distribute an identical product, which is generally produced by a single manufacturer. Despite the fact that the authorized channel and the grey importer are selling identical products, the products might differ in the quality observed by the customer. This might be because of lack of warranty or additional cost for adjustment to country specifications, etc. This practice is also known as "grey market" or "grey importer". Usually a parallel importer sells the same product to the higher-priced region. Gerstner and Holthausen (1986) study a situation when a monopolist serves a high- and low-price market. The high-value customers have also access to the cheaper market for addition transaction cost. Ahmadi and Yang (2000) study a model of parallel imports between two international markets where the parallel imports are caused by price discriminations in the markets. They provide theoretical insights on pricing strategies to prevent the market entry for the unauthorized channel. Since they do not explicitly model the difference between the product offered by unauthorized channel, they mention that grey imports are usually assessed with a discount in consumer valuation.

Case 2: National and Generic Brand In the pharmacy industry it is a common practice that after the expiration of a patent for a product another manufacturer enters the market. The second manufacturer distributes the generic product through the same retailer. The generic brand has generally the same or slightly lower quality than the product of the national brand but is offered for considerably lower price than the national brand. This practice can be also observed in the electronic and computer industry where the generic manufacturer copies the technology of the "innovator" branded product.

Case 3: National and Store Brand The discount retailer of food and non-food grocery products decides very often to produce and to distribute in her supermarkets her own low quality brand (store) to compete against the national brand. Usually the store brand is priced lower than the national brand. The supply chain structure of two products consists in this case of two manufacturers and one retailer. The difference is that the second manufacturer and the retailer build one firm. The retailer might choose the introduction of the store brand, rather than offer the generic brand, in order to reduce the risk and the dependency to the manufacturer. Raju et al. (1995) present a model which describes the characteristics on store brand introduction in a market with a number of national brands in the same product category. Each brand is produced and distributed by her own channel. Sayman et al. (2002) extend this model by making the cross-price sensitivity an endogenous decision parameter for introduction of a store brand.

Case 4: Two National Brands In the food and drug industry there is a tendency that the manufacturers introduce parallel to one highly advertised brand, a second unadvertised brand. The quality of the unadvertised brand product is generally equal or slightly lower than the quality of the national brand but is priced considerably lower than the product of the national brand. This is a common practice particularly for products which have low variable production cost and the most quality observed by the customers is gained from advertising. Because of economies of scale, the manufacturer wants to increase the production volume. The rebranding of the product assures that the buyer of the private brand cannot see the similarity to the advertised brand and does not switch to the low priced brand.

Case 5: Non Deceptive Counterfeit Counterfeit goods are illegal lower quality copies of branded products. The products which are offered as counterfeit are in a range from branded apparel products, luxury watches to more sophisticated like spare parts or high tech products. In non deceptive form of counterfeiting the customer knows that the product is not authentic and can observe the lower quality. But because of the lower price or the inferior quality of the product itself his utility is higher to obtain the counterfeit than the genuine brand product. The non deceptive counterfeit is usually distributed through a different retailer (e.g. bazaar, street vendors) and produced by a different manufacturer than the genuine high brand product. Additionally, the counterfeit channel has a probability of confiscation by the authorities. Grossman and Shapiro (1988b) present a model for non-deceptive and genuine products with a exogenously given quality levels and conclude that counterfeits increase the consumers welfare due to their lower prices. Cho et al. (2011) relax the assumption that the quality is exogenous and study different anti-counterfeiting strategies. They show analytically, that the effectiveness of anti-counterfeiting strategies depends on the strategic response of a counterfeiter and could even decrease the consumer welfare under certain conditions.

Case 6: Deceptive Counterfeit When the customers cannot distinguish between the genuine and the counterfeit products (e.g. pharmaceuticals, chemical products) it is a deceptive type of counterfeiting. Then the supply chain structure consists of two different manufacturers and one retailer which sells both type of products to the customer. A deceptive counterfeit is usually sold at the same price as the branded product. Also here the counterfeit products might be confiscated with a certain probability from authorities. Additionally one has take into account that the counterfeit product can be also discovered by the customers. Grossman and Shapiro (1988a) study the welfare effects on the market with the deceptive counterfeits. Assuming endogenous quality level and detection likelihood which is increasing with the number consumers of the deceptive counterfeit Cho et al. (2011) show that the deceptive counterfeiter will always choose the lowest functional quality.

Case 7: Identical Smuggled Good Smuggling describes an illegal flow of products between two regions with differences in taxes or laws. The supply chain structure in this case is similar to the case of parallel imports. Different to the case of parallel import, the product can be confiscated with a certain probability and the smuggler loses the profit. The product distributed by the smuggler and the authority channel is produced by one manufacturer and the quality is generally identical (e.g. cigarettes, alcohol, gasoline, etc.).

The plan for the rest of Chapter 4 is as follows. In Section 4.2 we introduce the market setting which we want to investigate and explain the coordination method between the product distribution channels. We start our analysis of a single product channel in Section 4.3. Given the results of a single channel, two product distribution channels in competition are considered in Section 4.4. We conclude in Section 4.5.

4.2 General formulation

We begin by studying a situation where a manufacturer M_L produces a branded product L and sells it to the retailer R_L which offers it on her domestic market L. A rival manufacturer M_F decides to produce an illicit substitute F and offers it through the retailer R_F on the own domestic market D and on the market L where the branded product is offered. Thus, the market L is a two product market for the branded product L and the illicit substitute F with some product substitution effects. The domestic market D is a single product market where only the product F is offered. The supply of the illicit substitute into the market L is subject risk of interdiction γ. We assume that products can be differentiated by customers in price and quality. In this model we assume complete information, which implies that each player in the market knows the cost and the demand structure of all other actors in the market. Figure 4.2 illustrates the structure of two

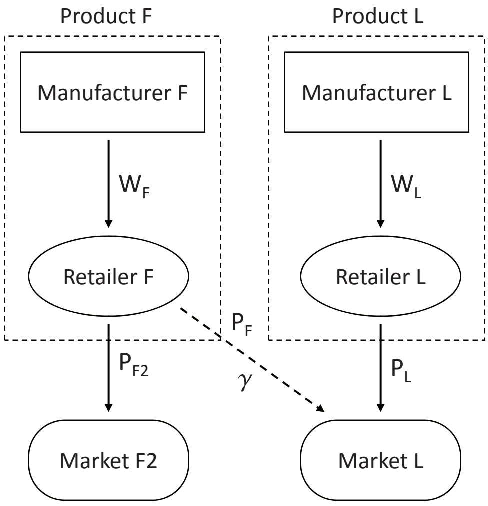


Figure 4.2 Supply chain structure

competing distribution channels. We define the following basic notation for our model:

$i = L, F$: index identifying the products L and F in the market. The manufacturer and the retailer which produce and distribute product L are assumed to have the first mover advantage or equivalently to have a leader position compared to the players associated with the product F.

$n = M, R$: index identifying the manufacturer M and the retailer R within a distribution channel of a product. We assume that within both product channels the manufacturer has a higher market power than the retailer.

P_i: per-unit price of the retailer for the product i in the market L.

P_D: per-unit price of the retailer for the product F in the market D.

W_i: per-unit wholesale price of the manufacturer of the product i.

Q_i: quantity of the product i distributed by the retailer R_i in the market L.

Q_D: quantity of the product F distributed by the retailer R_F in the market D.

X_i: level of the quality of the product i. By quality we mean any aspect of a product, including advertising, services and customer warranty which influence the demand level of the product.

e_i: initial size of the market L for the product i.

e_D: initial size of the market D for the product F.

b: measures the price substitutability among the products offered in market L. When the products are substitutes, $b > 0$; independent, $b = 0$; and complements, $b < 0$. The demand for a product increases with an increase in price of the other product for substitutes and decreases for complements.

β_i: demand sensitivity to the changes in price of a product i in the market L.

β_D: demand sensitivity to the changes in price of product F in the market D.

a: measures the degree of substitutability between two products in market L in terms of quality. It ranges from $a > 0$ when the products are substitutes, to $a = 0$ (independent) and $a < 0$ when these are complementary goods. Note that different to price, the quality substitutability implies that the demand for a product increases in quality of the other product for complementarity and vice versa for substitutes.

α_i: elasticity of the demand curve to the changes in product quality of the product i in the market L.

α_D: elasticity of the demand curve to the changes in product quality of the product F in the market D.

γ: probability of no supply or for cases of counterfeiting and smuggling goods the probability of confiscation of the product F production offered in the market L. If illicit goods are caught by authorities, the production cost is lost. The financial risk of confiscation of the illicit product F is either on the manufacturer or the retailer side which depends on the risk contract between the two players.

π_i^n: maximum profit of the manufacturer or the retailer associated with the product i.

Following the literature on utility maximizing consumers in Moorthy (1988) and Tirole (1988), we assume for the one product market D a linear demand function:

$$Q_D = e_D + \alpha_D X_F - \beta_D P_D \tag{4.1}$$

The linear demand function for two products in the market L is introduced as follows,

$$Q_i = e_i - aX_j + bP_j + \alpha_i X_i - \beta_i P_i, \quad i, j = L, F \text{ and } i \neq j \tag{4.2}$$

The above parameterization of the linear demand function does not imply necessarily $a = b$. This means that products which are identical, substitute or complementary in price are not necessarily the same in quality. The proposed linear two product demand model is consistent with the standard duopolistic case for two firms competing on price and quality level (see. Banker et al. (1998) and Balasubramanian and Bhardwaj (2004)). We make further assumption that the price in one market does not effect the demand in the other market and vice versa.

For the cost side of our model we assume that both manufacturer have the same type of cost function which can differ only in the parameterization. The production cost of the manufacturer of the product i is given by

$$C_i(X_i, Q_i) = k_i Q_i + y_i X^2, \quad i, j = L, F \tag{4.3}$$

where $k_i > 0$ is the variable production cost per unit and $y_i > 0$ is a constant of the fixed part of the production cost of manufacturers i. The quality aspect affects the total cost only in the constant cost part which are sunk once the production begins. The variable cost of the product depends only on the output quantity. The cost function describes a production process where the most value of the product is gained not from the physical or material value but either from the advertising, production equipment, organizational training, labor skills or from the innovation level of the product. The assumption that the quality affects only the fix cost related to production, enables closed-form solutions without the loss of generality and is in line with previous works (e.g. Moorthy (1988) and Balasubramanian and Bhardwaj (2004)).

We propose a Stackelberg leader-follow game for the pricing and quality decisions of the firms under unlimited capacity. For the cases we are studying (parallel imports, counterfeiting, smuggling and national brand) typically there exists one dominant player with an established product distribution channel and a second channel that wants to enter the market either on the retailer or manufacturer level. The second product channel as the follower observes the quality and the price of the leader's channel in the market as given. Thus the Stackelberg decision behavior seems to us a reasonable assumption. In our case the distribution channel of the product L is a leader and the distribution channel of the product F is a follower which wants to enter with the own product the market L as a second mover. In the general case with four players the game consists of four stages where each player makes her own decisions in the following order:

Stage 1: The manufacturer of the leader channel M_L sets simultaneously the price and the quality of the product anticipating all other decisions in her own distribution channel of the product and the channel of the competitor manufacturer as the reaction of his decisions.

Stage 2: Next the retailer of the leader channel R_L sets the price P_L by taking the quality and the price of the manufacturer M_L as given and by incorporating the anticipated reaction of the follower channel in the objective function.

Stage 3: The manufacturer of the follower channel M_F sets simultaneously the price and the quality considering the decisions in the leader supply channel as given and anticipating the demand of the retailer R_F as a function of its quality and wholesale price.

Stage 4: The retailer of the follower channel R_F chooses a prize based on the previous decisions of all players optimizing her own profit objective function.

We solve the four-stage problem by first finding the optimal response of the complete follower channel with respect to the prices and quality for given decision in the leader's channel. Here we first solve the stage four and then substitute the results in the third-stage problem. Subsequently we substitute the optimal response function of the players in the follower channel in the second-stage and first-stage problem respectively. We derive the sub-game perfect Nash equilibrium using the backward induction.

4.3 Single channel supply chain analysis

For a decentralized supply chain with a manufacturer and a retailer we want to start our analysis by asking the question: "Who should bear the risk in the distribution channel of the product F (retailer or manufacturer) if there is a certain probability of no supply to the market L?" For illicit distribution scenarios the product can be confiscated or interdicted by authorities with a certain probability. For legal products the reasons for no supply are disruptions in the supply chain (e.g. natural disasters, pirates).

4.3.1 Equilibrium for different risk scenarios

According to our previous assumption the manufacturer has the dominant market power compared to the retailer and therefore the manufacturer will dictate the risk contract in the supply channel maximizing its profit in both scenarios. Thus we calculate the optimal profit of the manufacturer for the case when the retailer bears the complete risk and for the second case when the manufacturer bears the risk. We assume that the manufacturer either completely carries the risk of confiscation or completely avoids the risk by passing the risk to the retailer.

Retailer bears the risk

In the scenario when the risk is on the retailer side the optimum profit of the retailer can be obtained by solving

$$\begin{aligned}\pi_F^R = \max_{P_F,P_D} \quad & (e'_F + \alpha_F X_F - \beta_F P_F)(P_F(1-\gamma) - W_D) \\ & +(e_D + \alpha_D X_F - \beta_D P_D)(P_D - W_D)\end{aligned} \tag{4.4}$$

where $e'_F = e_F - aX_L + bP_L$ is the size of the market L for the product F which is left after the channel distributing product L has made decisions concerning the price and quality of her own product. Since we want to study the channel of the product F independently we assume the quality and the price of the product L as given and rescale the initial market size for the product F in the market L. Analyzing both products simultaneously we will relax this assumption. We obtain the optimal price P_F^* and the quantity Q_F^* as the response functions of the wholesale price of the manufacturer W_D and X_F as follows

$$P_F^* = \frac{e'_F + \alpha_F X_F}{2\beta_F} + \frac{W_D}{2(1-\gamma)} \quad ; \quad P_D^* = \frac{e_D + \alpha_D X_F + \beta_D W_D}{2\beta_D} \tag{4.5}$$

$$Q_F^* = \frac{e'_F + \alpha_F X_F}{2} - \frac{\beta_F W_D}{2(1-\gamma)} \quad ; \quad Q_D^* = \frac{e_D + \alpha_D X_F - \beta_D W_D}{2} \tag{4.6}$$

The manufacturer optimizes her profit observing the demand of the retailer and optimizes simultaneously the wholesale price and quality of the product according to her cost function

$$\pi_F^M = \max_{W_D, X_F} \frac{1}{2}\left(e'_F + e_D + (\alpha_F + \alpha_D)X_F - (\beta_F + (1-\gamma)\beta_D)W_D\right)(W_D - k_F) - y_F X_F^2 \tag{4.7}$$

The closed form equilibrium solution results are summarized in Table 4.2, where A^R is

$$A^R = \frac{e'(1-\gamma) - (\beta' - \beta_D\gamma)k_F}{8y_F(\beta' - \beta_D\gamma) - \alpha'^2(1-\gamma)} \tag{4.8}$$

To ensure that X_F^* is positive for $e_i, \alpha_i, \beta_i, \gamma \geq 0$, we need $A^R > 0$ for which we require the following assumptions:

$$e'(1-\gamma) \geq (\beta' - \beta_D\gamma)k_F \tag{4.9}$$

and

$$8y_F(\beta' - \beta_D\gamma) \geq \alpha'^2(1-\gamma) \tag{4.10}$$

which lead to one joint condition, which has to be satisfied in order to guarantee a positive optimal quality.

$$\frac{e'(1-\gamma)}{k_F} \geq (\beta' - \beta_D\gamma) \geq \frac{\alpha'^2(1-\gamma)}{8y_F} \tag{4.11}$$

Manufacturer bears the risk

When the manufacturer carries the financial risk the retailer's profit function simplifies to the profit function of two independent markets without confiscation probability,

$$\begin{aligned} \pi_F^R = \max_{P_F, P_D} \quad & (e'_F + \alpha_F X_F - \beta_F P_F)(P_F - W_D) \\ & + (e_D + \alpha_D X_F - \beta_D P_D)(P_D - W_D) \end{aligned} \tag{4.12}$$

Table 4.2 Results when the retailer bears the risk

$X_F^* = \alpha' A^R$
$W_D^* = 4y_F A^R + k_F$
$Q_F^* = \frac{1}{2}(e'_F - \frac{\beta_F k_F}{1-\gamma} + (\alpha_F \alpha' - \frac{4y_F\beta_F}{1-\gamma})A^R)$
$Q_D^* = \frac{1}{2}(e_D - \beta_D k_F + (\alpha_D \alpha' - 4y_F\beta_D)A^R)$
$\pi_F^M = \frac{y_F(8y_F(\beta' - \beta_D\gamma) - \alpha'^2(1-\gamma))}{1-\gamma}(A^R)^2$
$P_F^* = \frac{1}{2\beta_F}(e'_F + \frac{\beta_F k_F}{1-\gamma} + (\alpha_F \alpha' + \frac{4y_F\beta_F}{1-\gamma})A^R)$
$P_D^* = \frac{1}{2\beta_D}(e_D + \beta_D k_F + (\alpha_D \alpha' + 4y_F\beta_D)A^R)$
$\pi_F^R = \frac{1-\gamma}{4\beta_F}(e'_F - \frac{\beta_F k_F}{1-\gamma} + (\alpha_F \alpha' - \frac{4y_F\beta_F}{1-\gamma})A^R)^2 + \frac{1}{4\beta_D}(e_D - \beta_D k_F + (\alpha_D \alpha' - 4y_F\beta_D)A^R)^2$
with $e' = e'_F + e_D$, $\alpha' = \alpha_F + \alpha_D$ and $\beta' = \beta_F + \beta_D$

The optimal price P_F^* and the quantity Q_F^* as the response functions of the wholesale price of the manufacturer W_D and X_F are given as (the values for P_D^* and Q_D^* are analogous),

$$P_F^* = \frac{e'_F + \alpha_F X_F + \beta_F W_D}{2\beta_F} \quad ; \quad Q_F^* = \frac{e'_F + \alpha_F X_F - \beta_F W_D}{2} \tag{4.13}$$

In this scenario the manufacturer's profit function involves the risk of no supply. This means that the quantity delivered to the retailer which was destined for the market L is lost with the probability γ. The profit optimization function is

$$\begin{aligned}\pi_F^M = \max_{W_D, X_F} \quad & \tfrac{1}{2}(e'_F + e_D + (\alpha_F + \alpha_D)X_F - (\beta_F + \beta_D)W_D)(W_D - k_F) - y_F X_F^2 \\ & -\tfrac{1}{2}(e_D + \alpha_D X_F - \beta_D W_D)W_D\gamma\end{aligned} \tag{4.14}$$

The closed form equilibrium solution results are summarized in Table 4.3, where A^M is

$$A^M = \frac{4y_F(e' - e_F\gamma) + (4y_F\beta' + \alpha_F\alpha'\gamma - \alpha'^2)k_F}{8y_F(\beta' - \beta_F\gamma) - (\alpha' - \alpha_F\gamma)^2} \tag{4.15}$$

Similar to the scenario when the retailer bears the risk, to ensure that optimal quality is $X_F^* = \frac{\alpha' - \alpha_F\gamma}{4y_F}A^M - \frac{\alpha' k_F}{4y_F} > 0$ for $e_i, \alpha_i, \beta_i, \gamma \geq 0$, we require the following assumption:

$$(\alpha' - \alpha_F\gamma)(e' - e_F\gamma) \geq (2\alpha'(\beta' - \beta_F\gamma) - (\alpha' - \alpha_F\gamma)\beta')k_F \tag{4.16}$$

Table 4.3 Results when the manufacturer bears the risk

$X_F^* = \frac{\alpha'-\alpha_F\gamma}{4y_F}A^M - \frac{\alpha' k_F}{4y_F}$
$W_D^* = A^M$
$Q_F^* = \frac{1}{2}(e_F' - \frac{\alpha'\alpha_F k_F}{4y_F} + (\frac{\alpha_F(\alpha'-\alpha_F\gamma)}{4y_F} - \beta_F)A^M)$
$Q_D^* = \frac{1}{2}(e_D' - \frac{\alpha'\alpha_D k_F}{4y_F} + (\frac{\alpha_D(\alpha'-\alpha_F\gamma)}{4y_F} - \beta_D)A^M$
$\pi_F^M = \frac{8y_F(\beta'-\beta_F\gamma)-(\alpha'-\alpha_F\gamma)^2}{16y_F}(A^M)^2 - \frac{e'k_F}{2} + \frac{\alpha'^2 k_F^2}{16y_F}$
$P_F^* = \frac{1}{2\beta_F}(e_F' - \frac{\alpha'\alpha_F k_F}{4y_F} + (\frac{\alpha_F(\alpha'-\alpha_F\gamma)}{y_F} - 4\beta_F)A^M)$
$P_D^* = \frac{1}{2\beta_D}(e_D' - \frac{\alpha'\alpha_D k_F}{4y_F} + (\frac{\alpha_D(\alpha'-\alpha_F\gamma)}{y_F} - 4\beta_D)A^M$
$\pi_F^M = \sum_{i\in\{F,D\}} \frac{1}{4\beta_i}(e_i' - \frac{\alpha'\alpha_i k_F}{4y_F} + (\frac{\alpha_i(\alpha'-\alpha_F\gamma)}{y_F} - \beta_i)A^M)^2$
with $e' = e_F' + e_D$, $\alpha' = \alpha_F + \alpha_D$ and $\beta' = \beta_F + \beta_D$

4.3.2 Optimal risk scenarios

Intuitively in a setup of perfect information in the supply chain with risk neutral players it should not make any difference between both risk scenarios for the profit of the players in the supply chain. The manufacturer would balance the decision about the risk contract simply by adjusting the wholesale price of the product. In practice we usually find that the manufacturer passes the risk to the retailer and reduces the wholesale price.

We do numerical experiments to demonstrate that there are cases, where it is optimal for the manufacture to take over the risk from the retailer. In the studied examples all the parameters of the markets except the initial size of both markets are identical and fixed to one. We vary the confiscation risk (γ) between 0 and 0.99 and the ratio of the sizes between the markets ($\Omega = e_F/e_D$) between 0.1 and 10. For instance, a market ratio of 0.1 means that the market D is ten times larger than the market L where with the market ratio of 10 it is vice versa. In all the tested market ratio combination the overall size of both markets is fixed to 200. Since we assume that the profit from one market is not used to subsidize the lost of the other market, we consequently close the market which contributes negatively to the overall profit.

Figure 4.3 shows the optimal risk decisions of the manufacturer for different market ratio and confiscation risk combinations for a base case with $\alpha_F = \alpha_D = \beta_F = \beta_D = k_F = y_F = 1$. When the confiscation risk is $\gamma = 0$ the profits in both risk scenarios are equal. But if the supply to the market L can be confiscated or get lost, the optimal risk contract depends on the market ratio. We find that for small confiscation risks ($\gamma < 0.1$) the manufacturer would take over the risk of the supply if the illicit market L is smaller then the licit market D ($\Omega < 1$) and it would avoid the risk when the ($\Omega > 1$). With increasing γ, the separation line between both risk decision increases. The intuition behind this result

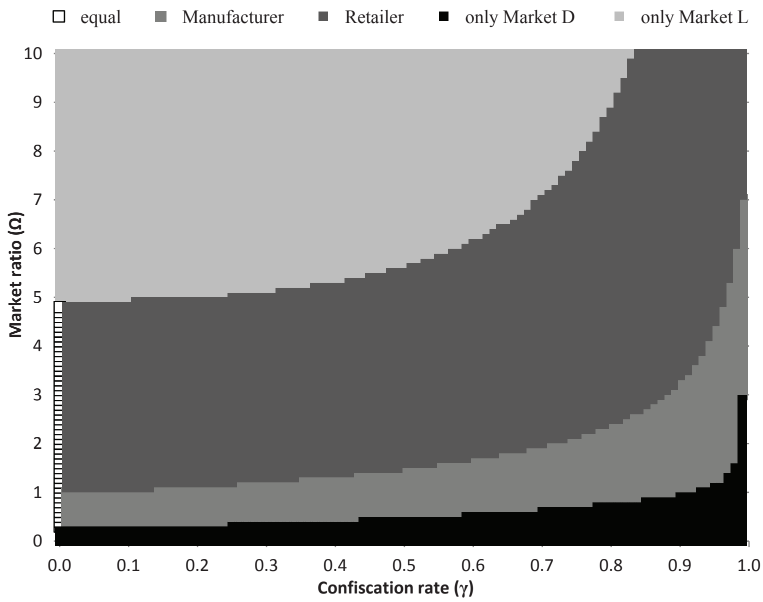


Figure 4.3 Optimal risk decisions of a manufacturer for a base case

is, that because of economies of scale the manufacturer is interested to increase the volume of the production in order to decrease the fix cost. Thus, the manufacturer tries to push the retailer to supply the illicit market L by taking over the risk and increasing her revenue. In extreme cases when one market is too large compared to another it is economically optimal to consider only a single market optimizing the price and the quality of the product F. These extreme cases are visualized in Figure 4.3 with (only Market L) and (only Market D) respectively.

Figure 4.4 shows the optimal risk decisions of the manufacturer for four different cases. In case (a) we assume that customers of the product F in the market L are more price-conscious than in the market D, whereas case (b) describes the opposite situation where the customers of the product F in the market L are more quality-conscious than in the market D. We observe that different customers behavior either shifts the separation line up (case a) or down (case b), but does not influence its convex form. Similar results we observe varying the production cost parameters in cases (c) and (d).

In all computed numerical examples we find that the wholesale price is always higher when the risk is on the side of the manufacturer. This is a plausible result, since this is how the manufacturer compensates the risk.

4.4 Two channel supply chain analysis

4.4.1 Quality and price equilibrium

We assume that the leader's market L is much larger than the domestic market D of the follower and the market L is dominated by the quality-conscious customers. This scenario is typical for a situation when a rival manufacturer makes an illegal substitute in

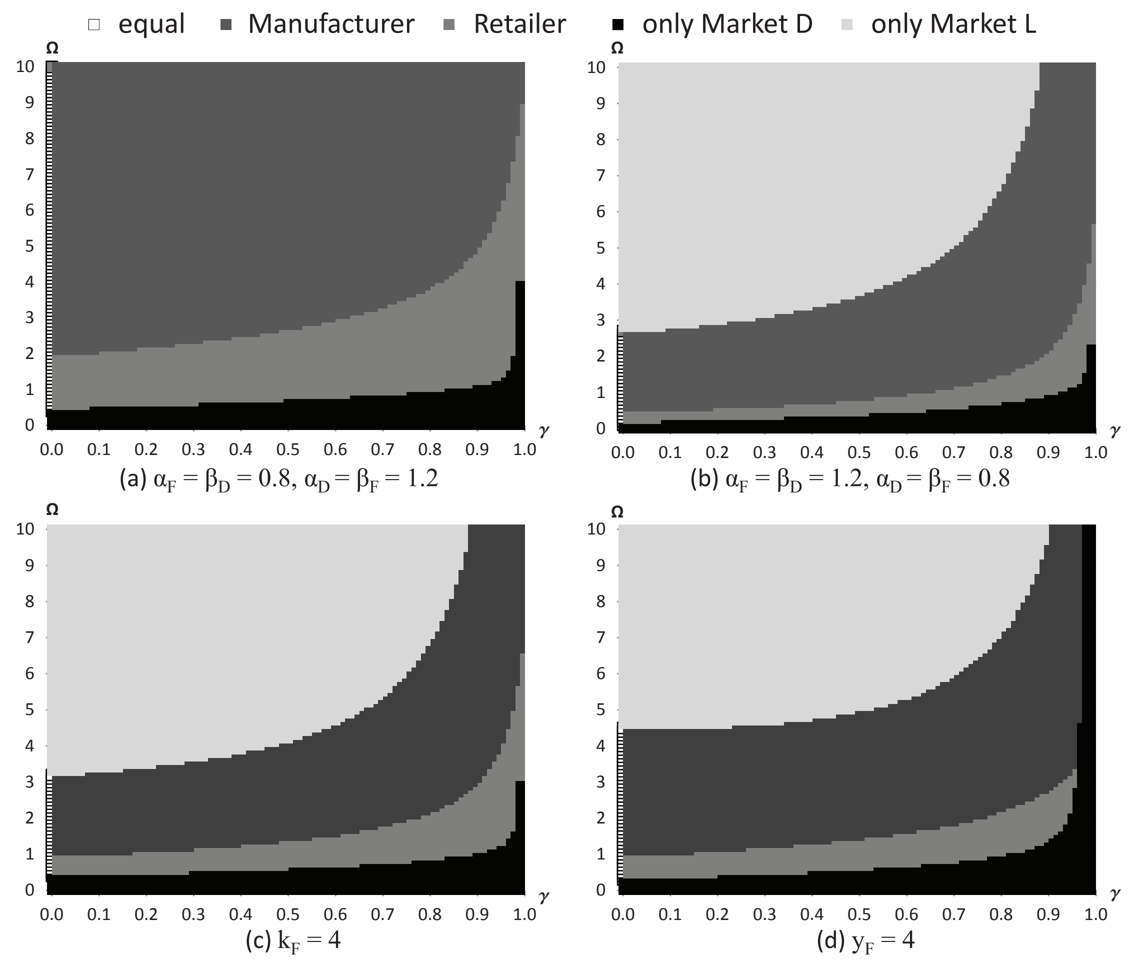


Figure 4.4 Optimal risk decisions for a manufacturer varying single parameter

a low wage country and exports the product to the industrialized country. In the previous chapter we have found that under such conditions the retailer would take over the risk. Hence, studying the quality and price settings in a two channels setup we rather focus our analyzes on the scenario when in the distribution channel of the product F the retailer bears the risk of confiscation.

The channel of the product L is observing the demand curve anticipating the optimal decisions from the players distributing product F.

$$Q_L = e_L - aX_F^* + bP_F^* + \alpha_L X_L - \beta_L P_L \tag{4.17}$$

In the single channel analysis in Chapter 4.3 we have shown that the optimal reaction functions X_F^* and P_F^* depend on X_L and P_L since the market size for the product F in the market L, $e'_F = e_F - aX_L + bP_L$ depends on the price and quality decisions of the product L. Replacing the optimal decisions X_F^* and P_F^* with the optimal reactions of the channel distributing product F from Table 4.2 the demand curve of the product L reduces to the optimization of decisions X_L and P_L.

In order to simplify the analysis we rearrange the terms which do not depend on the decisions of the leader to e_L^{new}, the terms which depend on X_L to α_L^{new} and the terms which depend on P_L to β_L^{new}, respectively. The expressions for the new readjusted parameters for the market intercept, quality elasticity and price sensitivity are

$$e_L^{new} = e_L + \frac{be_F}{2\beta_F} + \frac{bk_F}{2(1-\gamma)} + \left(-a\alpha' + \frac{b\alpha_F\alpha'}{2\beta_F} + \frac{2by_F}{1-\gamma}\right) \frac{(e_F+e_D)(1-\gamma) - (\beta' - \beta_D\gamma)k_F}{8y_F(\beta' - \beta_D\gamma) - \alpha'^2(1-\gamma)} \tag{4.18}$$

$$\alpha_L^{new} = \alpha_L - \frac{ab}{2\beta_F} - a\left(-a\alpha' + \frac{b\alpha_F\alpha'}{2\beta_F} + \frac{2by_F}{1-\gamma}\right) \frac{1-\gamma}{8y_F(\beta' - \beta_D\gamma) - \alpha'^2(1-\gamma)} \tag{4.19}$$

$$\beta_L^{new} = \beta_L - \frac{b^2}{2\beta_F} - b\left(-a\alpha' + \frac{b\alpha_F\alpha'}{2\beta_F} + \frac{2by_F}{1-\gamma}\right) \frac{1-\gamma}{8y_F(\beta' - \beta_D\gamma) - \alpha'^2(1-\gamma)} \tag{4.20}$$

With the rearranged parameters the profit function of the retailer R_L is

$$\pi_L^R = \max_{P_L} (e_L^{new} + \alpha_L^{new} X_L - \beta_L^{new} P_L)(P_L - W_L) \tag{4.21}$$

We obtain for optimal price P_L^* and optimal quality Q_L as the response functions of the wholesale price of the manufacturer W_L and X_L as follows

$$P_L^* = \frac{e_L^{new} + \alpha_L^{new} X_L + \beta_L^{new} W_L}{2\beta_L^{new}} \quad ; \quad Q_L^* = \frac{e_L^{new} + \alpha_L^{new} X_L - \beta_L^{new} W_L}{2} \tag{4.22}$$

The profit function of the manufacturer M_L is given by

$$\pi_L^M = \max_{W_L, X_L} \frac{1}{2}(e_L^{new} + \alpha_F^{new} X_F - \beta_L^{new} W_L)(W_L - k_L) - y_L X_L^2 \tag{4.23}$$

where the wholesale price and the quality are optimized simultaneously. The closed form solutions for the channel of the product L are summarized in Table 4.4, where A^L is

$$A^L = \frac{e_L^{new} - \beta_L^{new} k_L}{8y_L\beta_L^{new} - (\alpha_L^{new})^2} \tag{4.24}$$

Similar to the results for the channel of the product F, we need $A^L > 0$ for which we require a joint condition to ensure a positive quality X_L^*. This leads to a necessary joint condition for the distribution channel of the product L

$$\frac{e_L^{new}}{k_L} \geq \beta_L^{new} \geq \frac{(\alpha_L^{new})^2}{8y_L} \tag{4.25}$$

which is similar to the condition of the distribution channel of the product F and differs only in the rescaled parameters e_L^{new}, β_L^{new} and α_L^{new}. The optimal response decisions of the channel which supplies product F are obtained by inserting the results of the product L in the results summarized in Table 4.2.

Table 4.4 Results for the channel of the product L

$X_L^* = \alpha_L^{new} A^L$
$W_L^* = 4y_L A^L + k_L$
$Q_L^* = 2y_L \beta_L^{new} A^L$
$\pi_L^M = y_L(8y_L\beta_L^{new} - (\alpha_L^{new})^2)(A^L)^2$
$P_L^* = 6y_L A^L + k_L$
$\pi_L^R = 4y_L^2 \beta_L^{new} (A^L)^2$

4.4.2 Impact of risk

The question we address in this section is how the risk of interdiction influences the price and quality decisions of both supply channels. Figure 4.5 illustrates the best-response curves as a function of the confiscation risk γ for the base case where the markets are equal, in size ($e_L = e_F = e_D = 100$), production cost ($k_L = k_F = y_L = y_F = 1$) and demand function ($\alpha_L = \alpha_F = \beta_L = \beta_F = 1$). Further we assume for the base case that between the products exist substitution effects in quality and price ($a = b = 0.5$). We observe that the product quality (dotted line) and wholesale price (solid line) of the

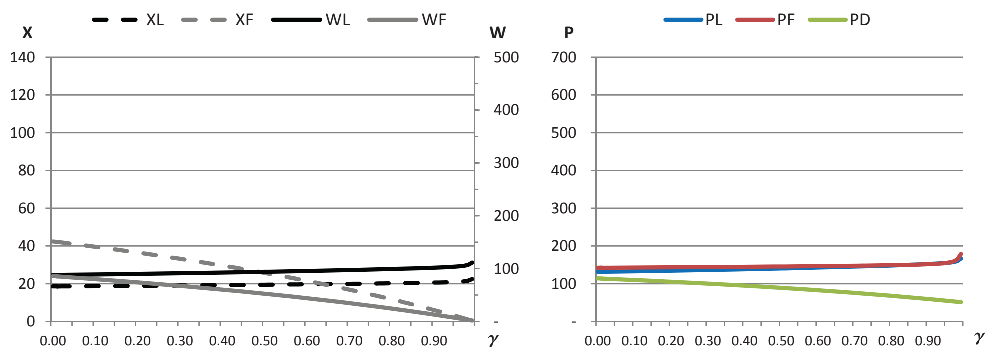
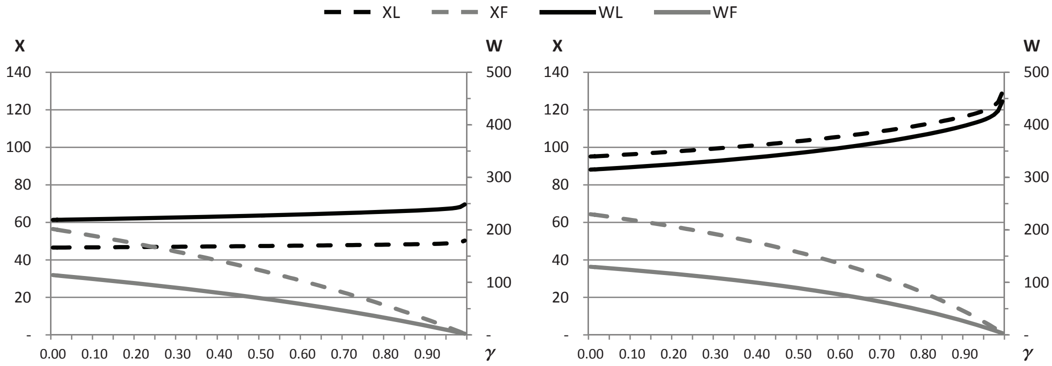


Figure 4.5 Optimal quality, wholesale and retail price for a base case

product F are decreasing in γ whereas the quality and the wholesale price of the product L is increasing in γ. For the retail prices P we find that an increase in risk of confiscation leads to lower price at the domestic market D, whereas the retail prices of both products at the market L slightly increase with a higher risk of interdiction. The protection of branded products from counterfeits or illicit substitutes leads to higher quality of original product but to higher prices, whereas quality and price of the illegal substitute are decreasing.

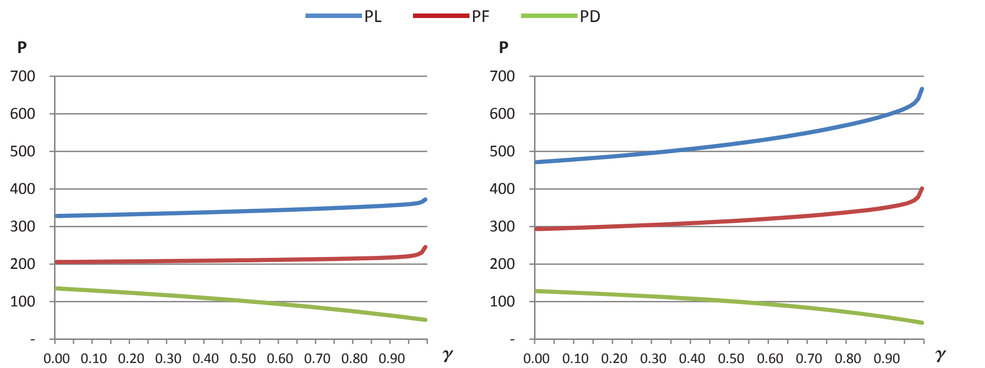
Figure 4.6 plots the quality and wholesale price functions of two different numerical examples (a) and (b). In example (a) the initial size of market L is three times larger than in the base case. In numerical example (b) we assume different behavior of the customers in each market segment. The market L is dominated by the quality-conscious customers and the market D by the price-conscious customers. In both numerical examples we



(a) $e_L = 300$ (b) $\alpha_F = \beta_D = 1.2$, $\alpha_D = \beta_F = 0.8$, $\alpha_L = 1.4$, $\beta_L = 0.6$

Figure 4.6 Optimal quality and wholesale price varying single parameters

observe the same effect of decreasing quality and wholesale price of the product F and increasing quality and wholesale price of the product L in the risk probability. The best-response curves differ only in the level and slope from the curves in the base case. Figure 4.7 illustrates the optimal price functions of two numerical examples described in Figure 4.6. Also here we indicate the same effect as in the base case of the risk on the price



(a) $e_L = 300$ (b) $\alpha_F = \beta_D = 1.2$, $\alpha_D = \beta_F = 0.8$, $\alpha_L = 1.4$, $\beta_L = 0.6$

Figure 4.7 Optimal retail price varying single parameters

of the products in different markets. The size of the market and the customers behavior influence only the level and the slope of the prices but do not change the findings from the base case.

4.5 Conclusion

Positioning of a product in a competitive market is crucial for the business success of an enterprise. In this chapter, we study two channels competing in products which are differentiated in quality and price. First we frame different related problems in one general base structure. Then we develop a four stage model to investigate the impact of competition on quality and price.

We derive close form solutions for different risk scenarios in the follower channel and visualize scenarios when the rival manufacturer is better off to bear the risk. We observe that when the leaders market is relatively small compared to the domestic market of the follower, the manufacturer of illicit substitute is better off to take over the risk of interdiction to push the retailer to supply into the leader's market. Studying price and quality decisions of two competing channels our analysis indicates that the higher risk of interdiction leads to higher quality of the branded and lower quality of the illicit substitute. Higher market protection against the illegal products leads to higher prices for the customers in the leaders market but to lower prices at the domestic market of the follower.

Further extension of the work in this chapter is to prove the findings analytically and to address general structural results. Another research direction would be to consider profit sharing between the manufacturer and the retailer in the follower supply chain which is a realistic assumption when the contracts are used to declare the risk of the supply.

Chapter 5

Summary and outlook

This thesis contributes in the field of supply chain management from two important perspectives. The major focus of the thesis lies on the optimization of an inventory network from the operational point of view but it also addresses the question of how to coordinate enterprises decisions under competition.

Chapter 2 investigates the advantages of proactive inventory sharing in a network of several local warehouses supplying a different number of markets. The most important contributions of this Chapter are: an exact model for proactive lateral transshipments and the investigation of potential cost savings compared to several reactive rules in different network structures.

The issue of inventory sharing between local warehouses in a two echelon distribution network is also examined in Chapter 3. Here the idea of using the inventory in the supply pipeline for the decision when to trigger a lateral transshipment is incorporated into the inventory model. A model to analyze the performance of such a network is presented and the major findings are: that consideration of the pipeline inventory can lead to significant cost savings and to reduction of the inventory in the network.

A supply chain of a branded product facing a competition from a illegal substitute of the rival manufacturer is investigated in Chapter 4. For two manufacturers and two retailers a four stage model of competition is developed to examine the impact of such competition with regards to the risk of interdiction of the supply of illegal product. The results highlight the cases when risk of interdiction would be taken by the manufacturer and when by the retailer. Further, the results hold how the risk level of interdiction influences the quality and prices of both products.

For the future research an interesting path may be to combine operational and strategic decisions in the design of the inventory networks (Integrated Network Design). In inventory models it is usually assumed that the set of locations which can serve a particular market is given. This assumption is determined by the given network of the local warehouses and markets. To integrate the transshipment decision in the network design would influence the decision about the number and position of the locations in a network. One possible solution approach for this problem could be an iterative process in which a location algorithm first, determines the number of required locations and the assignment of locations to the given markets that should be served. Second, an inventory model determines the optimal inventory levels of warehouses with corresponding minimal cost for the underlying network. Another approach might be analytic which incorporates the strategic decision and operational decision in one model and simultaneously derives the solution. The integration of strategic and operational decisions should lead to better solutions compared to making decisions separately.

References

Abad, P. L. (1996). Optimal pricing and lot-sizing under conditions of perishability and partial backordering. *Management Science 42*(8), 1093–1104.

Agrawal, V., X. Chao, and S. Seshadri (2004). Dynamic balancing of inventory in supply chains. *European Journal of Operational Research 159*(2), 296–317.

Ahmadi, R. and B. R. Yang (2000). Parallel imports: Challenges from unauthorized distribution channels. *Marketing Science 19*(3), 279–294.

Alfredsson, P. and J. Verrijdt (1999). Modeling emergency supply flexibility in a two-echelon inventory system. *Management Science 45*(10), 1416–1431.

Allen, S. (1958). Redistribution of total stock over several user locations. *Naval Research Logistics Quarterly 5*(4), 337–345.

Angel, B. W. F., D. A. van Damme, A. Ivanovskaia, R. J. M. Lenders, and R. S. Veldhuijzen (2006). Warehousing space in Europe: meeting tomorrow's demand. A pan-European warehousing trends study by Capgemini and ProLogis.

Archibald, T. (2006). Modelling replenishment and transshipment decisions in periodic review multilocation inventory systems. *Journal of the Operational Research Society 58*(7), 948–956.

Axsäter, S. (2003). *Supply chain operations: Serial and distribution inventory systems*, Volume 11, pp. 525–559. Elsevier.

Axsäter, S., C. Howard, and J. Marklund (2013). A distribution inventory model with transshipments from a support warehouse. *IIE Transactions 45*(3), 309–322.

Axsäter, S. (1990a). Modelling emergency lateral transshipments in inventory systems. *Management Science 36*(11), 1329–1338.

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

Balasubramanian, S. and P. Bhardwaj (2004). When not all conflict is bad: Manufacturing-marketing conflict and strategic incentive design. *Management Science 50*(4), 489–502.

Banker, R., I. Khosla, and K. Sinha (1998). Quality and competition. *Management Science 44*(9), 1179–1192.

Chew, E., L. Lee, and C.-K. Sim (2013). The impact of supply chain visibility when lead time is random. *OR Spectrum 35*, 163–190.

Cho, S., X. Fang, and S. Tayur (2011). Combating strategic counterfeiters in licit and illicit supply chains. *Working paper, Tepper School of Business*.

Coughlan, A. T. (1985). Competition and cooperation in marketing channel choice: Theory and application. *Marketing Science 4*(2), 110–129.

Dada, M. (1992). A two-echelon inventory system with priority shipments. *Management Science 38*(8), 1140–1153.

Das, C. (1977). The (S-1, S) inventory model under time limit on backorders. *Operations Research 25*(5), 835–850.

Dorfman, R. and P. O. Steiner (1954). Optimal advertising and optimal quality. *The American Economic Review 44*(5), 826–836.

Fisher, M. (1997). What is the right supply chain for your product? *Harvard Business Review 75*, 105–117.

Flint, P. (1995). Too much of a good thing, better inventory management could save the industry millions while improving reliability. *Air Transport World 32*, 103–107.

Gaukler, G. M., Ö. Özer, and W. H. Hausman (2008). Order progress information: Improved dynamic emergency ordering policies. *Production and Operations Management 17*(6), 599–613.

Gerstner, E. and D. Holthausen (1986). Profitable pricing when market segments overlap. *Marketing Science 5*(1), 55–69.

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

Gross, D. (1963). Centralized inventory control in multilocation supply systems. *Multi-stage Inventory Models and Techniques*, 47–84.

Gross, D., J. F. Shortle, J. M. Thompson, and C. M. Harris (2008). *Fundamentals of Queueing Theory*, Volume Fourth Edition. New Jersey, USA: Wiley.

Grossman, G. and C. Shapiro (1988a). Counterfeit-product trade. *The American Economic Review*, 59–75.

Grossman, G. M. and C. Shapiro (1988b). Foreign counterfeiting of status goods. *The Quarterly Journal of Economics 103*(1), 79–100.

Herer, Y. T., M. Tzur, and E. Yücesan (2006). The multilocation transshipment problem. *IIE Transactions 38(3)*, 185–200.

Howard, C., I. Reijnen, J. Marklund, and T. Tan (2010). Using pipeline information in a multi-echelon spare parts inventory system. *Working paper, Eindhoven University of Technology*.

Koole, G. (2006). Monotonicity in markov reward and decision chains: Theory and applications. *Foundations and Trends in Stochastic Systems 1*(1), 1–76.

Kranenburg, A. and G. J. van Houtum (2009). A new partial pooling structure for spare parts networks. *European Journal of Operational Research 199*(3), 908–921.

Kutanoglu, E. and M. Mahajan (2009). An inventory sharing and allocation method for a multi-location service parts logistics network with time-based service levels. *European Journal of Operational Research 194*(3), 728–742.

Law, A. M. (2007). *Simulation Modeling and Analysis* (4th ed.). McGraw-Hill New York.

Lee, H. L. (1987). A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science 33*(10), 1302–1316.

Lien, R. W., S. M. R. Iravani, K. Smilowitz, and M. Tzur (2011). An efficient and robust design for transshipment networks. *Production and Operations Management 20*(5), 699–713.

Lus, B. and A. Muriel (2009). Measuring the impact of increased product substitution on pricing and capacity decisions under linear demand models. *Production and Operations Management 18*(1), 95–113.

McGuire, T. W. and R. Staelin (1983). An industry equilibrium analysis of downstream vertical integration. *Marketing Science 2*(2), 161–191.

Messerli, E. J. (1972). Proof of a convexity property of the Erlang B formula. *The Bell System Technical Journal 51*, 951–953.

Moinzadeh, K. (1989). Operating characteristics of the (S-1, S) inventory system with partial backorders and constant resupply times. *Management Science 35*(4), 472–477.

Moorthy, K. S. (1988). Product and price competition in a duopoly. *Marketing Science 7*(2), 141–168.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic theory 18*(2), 301–317.

Nahmias, S. and S. A. Smith (1994). Optimizing inventory levels in a two-echelon retailer system with partial lost sales. *Management Science 40*(5), 582–596.

Paterson, C., G. Kiesmüller, R. Teunter, and K. Glazebrook (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research 210*(2), 125–136.

Paterson, C., R. Teunter, and K. Glazebrook (2012, September). Enhanced lateral transshipments in a multi-location inventory system. *European Journal of Operational Research 221*(2), 317–327.

Powell, W. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*, Volume 2nd Edition. Wiley-Blackwell.

Puterman, M. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, Hoboken.

Raju, J. S., R. Sethuraman, and S. K. Dhar (1995). The introduction and performance of store brands. *Management Science 41*(6), 957–978.

Reijnen, I. C., T. Tan, and G. J. van Houtum (2009). Inventory planning for spare parts netwoks with delivery time requirements. *Beta Working Paper 280, Eindhoven University of Technology*.

Robinson, L. W. (1990). Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Operations Research 38*(2), 278–295.

Sayman, S., S. J. Hoch, and J. S. Raju (2002). Positioning of store brands. *Marketing Science 21*(4), 378–397.

Seidscher, A., M. Dada, and S. Minner (2013). The impact of cost and quality on illegal goods in the supply chain. *Working paper, University of Vienna*.

Seidscher, A. and S. Minner (2012). Emergency supply flexibility in a two-echelon inventory system with pipeline information. *Working paper, University of Vienna*.

Seidscher, A. and S. Minner (2013). A semi-markov decision problem for proactive and reactive transshipments between multiple warehouses. *European Journal of Operational Research forthcoming*.

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

Sherbrooke, C. C. (1992). Multiechelon inventory systems with lateral supply. *Naval Research Logistics (NRL) 39*(1), 29–40.

Solis, W. M., T. Gosling, C. Barden, L. C. Cooper, Jr., F. Kerrison, C. Ramirez, M. Richardson, S. Pruitt, and C. Weissman (2008). Defense inventory: Management actions needed to improve the cost efficiency of the navy's spare parts inventory. Technical Report ADA491588, Government Accountability Office Washington DC.

Song, J.-S. and P. H. Zipkin (1996, October). Inventory control with information about supply conditions. *Management Science 42*(10), 1409–1419.

Song, J.-S. and P. H. Zipkin (2009). Inventories with multiple supply sources and networks of queues with overflow bypasses. *Management Science 55*(3), 362–372.

Tagaras, G. (1999). Pooling in multi-location periodic inventory distribution systems. *Omega 27*(1), 39–59.

Tagaras, G. and D. Vlachos (2002). Effectiveness of stock transshipment under various demand distributions and nonnegligible transshipment times. *Production and Operations Management 11*(2), 183–198.

Tempelmeier, H. (2000). Inventory service-levels in the customer supply chain. *OR Spectrum 22*(3), 361–380.

Tijms, H. C. (2003). *A First Course in Stochastic Models*. Chichester, England: Wiley.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

Van Wijk, A. C. C., I. J. B. F. Adan, and G. J. van Houtum (2011). Optimal lateral transshipment policy for a two location inventory problem. *EURANDOM Working Paper 27, Eindhoven University of Technology*.

Wong, H., G. van Houtum, D. Cattrysse, and D. V. Oudheusden (2006). Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research 171*(3), 1071–1093.

Yang, G. and R. Dekker (2010). Service parts inventory control with lateral transshipment that takes time. *Working paper, Erasmus University Rotterdam*.

Curriculum Vitae

Personal Data

Name	Arkadi Seidscher
Date of Birth	23.06.1982
Nationality	German

Education and Academic Experience

2009 - present	PhD Candidate in Logistics and Operations Management and Research Assistant, Chair of Logistics and Supply Chain Management at the University of Vienna, Austria
Feb - Jun 2012	Research stay at the Johns Hopkins Carey Business School in Baltimore, USA (Prof. Maqbool Dada)
2002 - 2008	Study of Industrial Engineering at the Leibniz University of Hanover, Germany

Papers

In Press	Seidscher, A. and S. Minner (2013). *A semi-markov decision problem for proactive and reactive transshipments between multiple warehouses.* European Journal of Operational Research forthcoming
Working paper	Seidscher, A. and S. Minner (2012). *Emergency supply flexibility in a two-echelon inventory system with pipeline information.*
	Seidscher, A., M. Dada and S. Minner (2013). *The impact of cost and quality on illegal goods in the supply chain.*

Talks and Presentations

Nov 2011	Seidscher, A. and S. Minner (2011). *A semi-markov decision problem for proactive and reactive transshipments between multiple warehouses.* INFORMS Annual Meeting in Charlotte, USA
Aug 2011	Seidscher, A. and S. Minner (2011). *A semi-markov decision problem for proactive and reactive transshipments between multiple warehouses.* 10th ISIR Summer School in Istanbul, Turkey