POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT
Vol. 21, No. 4, July–August 2012, pp. 698–714
ISSN 1059-1478 | EISSN 1937-5956 | 12 | 2104 | 0698

DOI 10.1111/j.1937-5956.2011.01301.x

Designing Service Level Contracts for Supply Chain Coordination

Marcel A. Sieke
Department of Supply Chain Management and Management Science, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany, marcel.sieke@uni-koeln.de

Ralf W. Seifert
College of Management of Technology, Ecole Polytechnique Fédérale de Lausanne, CH-1015, Lausanne, Switzerland,
IMD – International Institute for Management Development, Chemin de Bellerive 23, CH-1001, Lausanne, Switzerland, seifert@imd.ch

Ulrich W. Thonemann
Department of Supply Chain Management and Management Science, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany, ulrich.thonemann@uni-koeln.de

Supply contracts are used to coordinate the activities of the supply chain partners. In many industries, service level-based supply contracts are commonly used. Under such a contract, a company agrees to achieve a certain service level and to pay a financial penalty if it misses it. The service level used in our study refers to the fraction of a manufacturer's demand filled by the supplier. We analyze two types of service level-based supply contracts that are designed by a manufacturer and offered to a supplier. The first type of contract is a flat penalty contract, under which the supplier pays a fixed penalty to the manufacturer in each period in which the contract service level is not achieved. The second type of contract is a unit penalty contract, under which a penalty is due for each unit delivered fewer than specified by the parameters of the contract. We show how the supplier responds to the contracts and how the contract parameters can be chosen, such that the supply chain is coordinated. We also derive structural results about optimal values of the contract parameters, provide numerical results, and connect our service level measures to traditional service level measures. The results of our analyses can be used by decision makers to design optimal service level contracts and to provide them with a solid foundation for contract negotiations.

Key words: service level; contracts; supply chain; financial penalty
History: Received: October 2010; Accepted: June 2011, after 2 revisions.

1. Introduction

Supply contracts govern the activities of the partners of a supply chain. Various types of supply contracts are used in supply chains, such as buyback and revenue sharing contracts. Cachon (2003) provides an overview of this literature. These types of contracts can be used to coordinate supply chains and incentivize decision makers to place order quantities and hold inventories that are optimal from a supply chain perspective. Inventory systems with a variety of service level measures, such as availability or lead-time requirements, are also addressed in research. In this article, we study the coordination of supply chains using a service level contract that enforces pre-specified service levels with financial penalty payments. In this study, service level refers to the fraction of a manufacturer's demand filled by the supplier in a given period. If the fraction of demand filled by the supplier in a period is below a certain threshold, the supplier pays a penalty to the manufacturer. Such contract types are popular in practice. The objective of this study is to provide analytical results for these service level contracts and show how they can coordinate a supply chain.

Under a service level contract, the supply chain partners agree on a target service level for the supplier and on a penalty payment that is due if the target service level is not achieved. Service level contracts are a common type of contract in various industries. In the consumer goods industry, for instance, 70% of the retailers measure the service levels of their suppliers (Thonemann et al. 2003, 2005). One example of a company that has implemented a service level contract is *dm-drogeriemarkt*, one of the largest German drugstore chains (Mostberger 2006). *dm-drogeriemarkt* continuously monitors the service levels of its suppliers and requires penalty payments if the service level of a supplier drops below a pre-specified limit. In the consumer goods industry, measuring supplier service levels and enforcing penalty payments are important factors for ensuring high supply chain performance. Behrenbeck et al. (2007), for instance, analyze the supply chain performance of 33 of the largest European retailers and cluster the companies into two groups: a group that achieves high supply chain performance ("champions") and a group that does not achieve high supply chain performance ("followers"). They report that 83% of the champions measure their suppliers' service levels and 40% of the champions enforce financial penalties if a pre-specified target is not met. Out of the followers, only 59% measure their suppliers' service levels and only 5% of the followers enforce financial penalties.

Service level contracts are also used in other industries. In the process industry, Fab-tek provides titanium products for industrial use. Fab-tek has implemented a service level contract and charges penalty payments for late deliveries (Shapiro et al. 1992). In the automotive industry, missing components can result in downtimes of expensive manufacturing processes or in expensive re-work. In this industry, many original equipment manufacturers use service level contracts with their first-tier suppliers to guarantee high availability (Stratmann 2006).

One of the challenges when designing a service level contract is agreeing on the contract parameters, that is, the values of the target service level and the penalty payment. In the literature, service level-based inventory policies have been developed for a variety of settings and have been analyzed extensively (see, e.g., Zipkin 2000). The most commonly used service level measures seem to be α and β service levels. The α service level, also referred to as in-stock probability or Type I service, measures the probability that all demand of an order or period can be filled. The β service level, also referred to as fill rate or Type II service, measures the percentage of demand that can be filled. These measures can be effectively calculated if the planner has sufficient data for the service period. The time length of the measuring period has a high impact on the value of the service level (Thomas 2005). Because we focus on designing service level contracts for a single period measurement, we adopt period-specific service levels that resemble the α and β service levels. We then use these service level measures to design contracts for two-echelon supply chains. Our interest is to develop service level contracts that can coordinate the supply chain.

The remainder of the article is organized as follows: In section 2, we review the related literature. In section 3, we develop a mathematical model of a two-echelon supply chain that is governed by a service level contract. In section 4, we analyze the optimal response of the supplier. In section 5, we build on the optimal response functions and derive contracts that coordinate the supply chain. We also provide several structural results to gain managerial insight into the interdependencies of the contract parameters. At the end of this section, we will also show that there exist certain points where our service level contract parameter corresponds to the widely used α and β service levels. These *contract consistent points* clearly have a certain appeal to practitioners. In section 6, we analyze event sequences that are different from the one used in the previous sections. In section 7, we conclude. All proofs are contained in the Appendix.

2. Literature Review

Contract-based incentive schemes for coordinating supply chains with inventory-limited suppliers have been extensively analyzed in literature. One possibility to achieve coordination is by reallocating backorder costs in the supply chain, such that the supplier makes a decision that is also optimal from a supply chain perspective. For example, Lee and Whang (1999) analyze a decentralized serial supply chain where the upper echelon faces inventory holding costs, but no backorder penalty costs. They develop an optimal nonlinear incentive scheme that incentivizes all echelons to choose the supply chain optimal base stock levels. They subsidize certain echelons for holding inventory and penalize others for shortages, for example, by using a shortage reimbursement if the upstream site fails to deliver. Porteus (2000) builds on this approach and introduces responsibility tokens. These tokens are given to the downstream echelon as a replacement for real units whenever the order cannot be fully filled. The issuer of the responsibility token bears all the financial consequences of delayed deliveries at lower echelons. Porteus shows that the supply chain can be coordinated with his approach. Cachon and Zipkin (1999) show how supply chain coordination can be achieved in a serial two-echelon supply chain with a payment scheme that depends linearly on the backorders of the two echelons and the retailer's on-hand inventory.

More recently, Lutze and Özer (2008) analyzed promised lead-time contracts in a serial supply chain. The retailer orders from the supplier and in case the retailer's order request exceeds the supplier's inventory on-hand, the supplier can replenish from an alternative source, for which the supplier then has to pay a certain penalty. The supplier offers a menu of promised lead times and lump sum payments between the retailer and the supplier. With this contract, the retailer can pay the supplier for a shorter lead time and thereby reward the supplier for a better performance. The authors also study the impact of a retailer's private service level (equivalently backlogging cost) information to the end consumers on the supplier's and retailer's profits. Lutze and Özer (2008) explicitly prove that the optimal solution for their system follows an echelon base stock policy. In contrast, we penalize the supplier for inferior performance and take the base stock policy (which is not necessarily optimal in our setting) as given and optimize over base stock levels.

Using service levels in multi-echelon inventory systems has also been addressed by other authors. In a study at a manufacturer for electronic testing equipment, Cohen et al. (1999) identified the need for differentiating service levels by offering two types of service modes (normal and emergency). Wang et al. (2002) build on this work and analyze service differentiation based on demand-lead-time in a two-echelon distribution system. Service level models are also very popular in the area of call center operations where the performance of outsourced call centers is controlled closely to maintain a certain customer service level. Milner and Olsen (2008) analyze a contract setup where acceptable delay times are specified and a failure to meet the specified delay times results in financial penalties. As in our study, the penalties can vary depending on the severity of underperformance. Although the basic idea of controlling the delay of call centers comes close to our problem of ensuring the timely delivery of physical products, the focus on queuing models, which model the service delivery, prevents the direct application to our problem, in which the supplier can build up safety stock well in advance. Recent research of Kim et al. (2007) comes closer to our inventory setup: Kim et al. (2007) analyze a three-parameter contract with a fixed payment, a reimbursement of the supplier's cost, and a backorder penalty cost. They show that the supply chain can be coordinated for risk-neutral decision makers, but focus on subsystem availability instead of service levels, as in our research.

Service level contract models are related to backorder cost models used in supply chain contracting. For instance, consider a backorder cost model. For given base stock levels, we can compute service levels and backorder levels, that is, a backorder level can be computed for each service level. In backorder cost models, a penalty is charged per backorder, whereas in our service level model, a penalty is charged each time an agreed-on service level is not reached. We believe such service level modes are popular in practice. Hence, we study them in this article.

Our study contributes to the contracting literature by developing finite horizon service level measures that represent the fraction of a manufacturer's demand filled by the supplier in a period and are closely related to the traditional service levels. We propose the coordination of supply chains by using such service level measures in combination with periodic financial penalty payments. We show that supply chains can be coordinated with these contracts.

3. Model Description

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

The physical unit inventory holding costs of the supplier and manufacturer are h_s and h_m (with $h_s < h_m$ to avoid trivial solutions) and are charged against the inventory left over at the end of a period. The manufacturer receives r for every unit sold and encounters a backorder penalty cost b_m for each unit that is backordered at the end of a period. This cost is interpreted as usual and includes losses in customer goodwill (Porteus 1990). The supplier encounters penalty costs if she cannot deliver at least a fraction s of the manufacturer's orders in a specific period. We refer to s as the contract service level and require $0 < s \leq 1$. A detailed model of two distinct penalty payment terms will be discussed below. The lead time of the supplier is $L_s > 0$ and the lead time between the supplier and the manufacturer is $L_m > 0$.

Demand is stochastic, stationary, continuous, and independent between periods. Demand can be arbitrarily distributed as long as the p.d.f. is strongly unimodal or logconcave. For an in-depth treatment of logconcave distributions and their application to inventory control, we refer the reader to Rosling (2002). To keep our analyses concise, we will focus on distributions with infinite non-negative support. Distributions with finite support can be treated analogously, but they require the definition of feasible regions for the parameter values of the supply contracts, which makes the analysis much more complex and adds little value.

We denote the demand over t periods by D_t and the corresponding p.d.f. and c.d.f. by $f_t(\cdot)$ and $F_t(\cdot)$, respectively. Note that the logconcave property is inherited to demand convolutions (Karlin and Proschan 1960), that is, if the demand over a single period is logconcave, then the demand over the lead time is also logconcave. Note that for an infinite horizon, periodic review problem with backordering under base stock policy, both the supplier and the manufacturer observe the same demand information. We optimize with respect to the total expected profits over an infinite planning horizon. Because demand is stationary, the optimal base stock level is stationary and the inventory level at the end of period $t + L_s$ only depends on the demand over the previous $L_s + 1$ periods (Silver et al. 1998). Similarly, the available inventory at the beginning of period $t + L_s$, that is, after receiving the incoming order but before the demand for the period $t + L_s$ occurs, only depends on the demand over the previous L_s periods.

In inventory control, two commonly used service levels are the α and β service levels. The α service level (also known as in-stock probability or Type-I service) specifies the fraction of periods in which demand is completely filled. The β service level (also known as fill rate or Type II service) specifies the expected fraction of demand that is filled in a period. We can compute the infinite horizon α and β service levels as (Sobel 2004)

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

and

$$\beta = 1 - \frac{E\left(\left[D - \left(y_s^{SC^*} - D_{L_s}\right)^+\right]^+\right)}{\mu} = \frac{\int_{x=0}^{y_s^{SC^*}} [F_{L_s}(x) - F_{L_s+1}(x)]dx}{\mu}. \tag{2}$$

These measures are based on an infinite horizon analysis of the inventory systems, that is, an infinite number of periods is used for measuring the service level, whereas in our study we measure the service level in each period. The lengths of the time horizons over which the service levels are measured matter. Thomas (2005) shows that service level measures with long and short time horizons might differ significantly and that the expected finite horizon service level is always greater than the infinite horizon service level.

We evaluate the service level in every period. Our contract service level s refers to the fraction of demand filled by the supplier in a period. The manufacturer specifies the contract service level and the supplier is expected to achieve this target service level in each period. If the fraction of demand filled by the supplier in a period is below s, a penalty must be paid by the supplier to the manufacturer. We analyze two types of supply contracts, which we refer to as flat penalty and unit penalty contracts. Although the contract service s level resembles the infinite horizon β service level, the contract service level is based on a one-period measurement and non-compliance leads to an immediate cash outflow. This also can mean that a service level s does not necessarily lead to a β of similar level because the supplier treats s as another cost term in her expected profit function and not as a side constraint of her optimization problem.

Under a *flat penalty contract*, the supplier pays the manufacturer a fixed amount p for each period in which the contract service level s is not met, that is, for each period in which the supplier does not fill at least a fraction s of the manufacturer's orders. The flat penalty contract is related to the traditional α service level because the quantity of the shortage is ignored. Let D denote the demand of the current period and let D_{L_s} denote the demand over the previous L_s periods. Then, the inventory available for filling demand of the current period is $y_s - D_{L_s}$ and the flat penalty function can be written as

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

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

Under the *unit penalty contract*, the supplier is charged a penalty of p for each unit she delivers fewer than sD if she fills at least some of the demand of the current period. If she fills no demands of the current period, she is charged a penalty of p for each unit of demand of the current period. The unit penalty contract is related to the traditional β service level, because the quantity of the shortage is taken into account. The unit penalty function can be written as

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

The first case holds if no inventory is available at the beginning of the period. The second case holds if some inventory is available at the beginning of the period.

The objectives of both parties are the maximization of the individual expected profits. The supplier has information on the demand distribution, but does not need any information on the cost structure or the inventory levels of the manufacturer. The manufacturer has information on the inventory holding cost of the supplier h_s, the unit cost c of the supplier, and the lead times L_s and L_m. In industries where suppliers are audited by manufacturers, such as in the automotive, medical equipment, and electronics industries, manufacturers can estimate these costs quite accurately. In industries where such first-hand estimates are not available, manufacturers can rely on industry benchmarks that are readily available (e.g., SCORmark 2008).

The service level contract is only offered once at the beginning of the relationship and is not revised during the course of the relationship. The service level contract specifies the wholesale price w, the type of penalty function (flat or unit penalties), the penalty cost factor p, and the contract service level s. The sequence of contract-related events is similar to a Stackelberg game: The manufacturer first determines all contract parameters (w,p,s) and then offers the contract to the supplier. The supplier accepts the contract if her expected profit per period is above her reservation profit, $\hat{\Pi}_s$, and rejects the contract otherwise. In section 6 we analyze alternative event sequences where the manufacturer does not determine all contract parameters.

The overall objective is to maximize the supplier's and the manufacturer's expected profits over an infinite horizon. Clearly, this is only an approximation of a real-world setting where the contract is valid only for a given period of time.

Given the contract (w,p,s), the supplier decides on the base stock level y_s that maximizes her expected profit per period, that is,

$$E\Pi_s^*(w,p,s) = \max_{y_s} E\Big[(w-c)D - h_s(y_s - D_{L_s+1})^+ - P_\kappa(y_s,p,s,D,D_{L_s})\Big],$$

where $\kappa = \{f,u\}$ depending on which penalty contract is used. We assume that the wholesale price w does not influence the holding cost factor h_s. The decision variable of the supplier is the supplier's base stock level y_s. As discussed before, the inventory level at the end of period $t + L_s$ only depends on the demand over the previous $L_s + 1$ periods. The available inventory at the beginning of period $t + L_s$, that is, after receiving the incoming order but before the demand for the period $t + L_s$ occurs, only depends on the demand over the previous L_s periods. We show later that the supplier's problem is unimodal in y_s and hence a unique maximizer exists.

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

$$\begin{aligned} E\Pi_m^* = \max_{y_m,w,p,s} & E[(r-w)D + P_\kappa(y_s,p,s,D,D_{L_s})] \\ & - h_m I_m(y_m,y_s) - b_m B_m(y_m,y_s) \\ & s.t. \quad E\Pi_s^*(w,p,s) \geq \hat{\Pi}_s, \end{aligned} \tag{3}$$

where $\kappa = \{f,u\}$ depending on which penalty contract is used. We assume that the contract parameters are only fixed once during the course of the contract relationship and that the manufacturer follows a base stock policy. He only optimizes within this class of policies and we note that a base stock policy is not necessarily optimal.

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

The third term of the objective function is the expected inventory holding cost per period. We assume that the holding cost h_m does not depend on the wholesale price w, for example, by pegging it to the sales price r. Then $I_m(y_m, y_s)$ denotes the average inventory level at the end of a period. Following Cachon (2003), it can be computed as

$$\begin{aligned} I_m(y_m,y_s) = & F_{L_s+1}(y_s)\int_{\delta=0}^{y_m}(y_m-\delta)f_{L_m+1}(\delta)d\delta \\ & + \int_{x=y_s}^{\infty}\int_{\delta=0}^{y_m+y_s-x}(y_m+y_s-x-\delta) \\ & \times f_{L_m+1}(\delta)f_{L_s+1}(x)d\delta dx. \end{aligned}$$

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

$$\begin{aligned} B_m(y_m,y_s) = & F_{L_s+1}(y_s)\int_{\delta=y_m}^{\infty}(\delta-y_m)f_{L_m+1}(\delta)d\delta \\ & + \int_{x=y_s}^{\infty}\int_{\delta=y_m+y_s-x}^{\infty}(\delta-(y_m+y_s-x)) \\ & \times f_{L_m+1}(\delta)f_{L_s+1}(x)d\delta dx. \end{aligned}$$

The constraint $E\Pi_s^*(w,p,s) \geq \hat{\Pi}_s$ ensures that the supplier achieves an expected profit that is greater than or equal to her reservation profit $\hat{\Pi}_s$. It can always be satisfied by choosing appropriate contract parameters.

4. Supplier Response

In this section, we analyze how the one-period expected profit function of the supplier is affected by the type and parameters of the supply contract and show that the expected profit function is quasi-concave in the supplier's base stock level. We also derive the optimality conditions for both contract types.

4.1. Flat Penalty Contract

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

$$\Pr(sD \le y_s - D_{L_s}) = \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{s}\right) dx$$

and the probability that the supplier incurs a penalty charge is

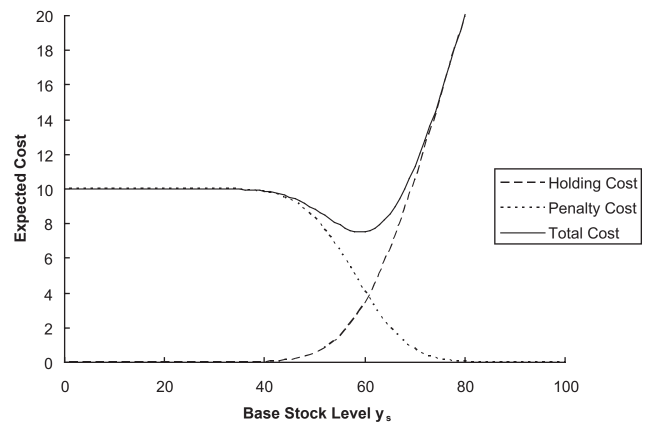
$$\Pr(sD > y_s - D_{L_s}) = 1 - \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{s}\right) dx.$$

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

$$\begin{aligned} E\Pi_s^f(y_s) &= E[(w-c)D - h_s(y_s - D_{L_s+1})^+] \\ &\quad - p\Pr(sD > y_s - D_{L_s}) \\ &= (w-c)\mu - h_s \int_{x=0}^{y_s} (y_s - x) f_{L_s+1}(x) dx \\ &\quad - p\left(1 - \int_{x=0}^{y_s} f_{L_s}(x) F\left(\frac{y_s - x}{s}\right) dx\right). \end{aligned} \tag{4}$$

Figure 2 illustrates the cost terms for truncated normally distributed demand with $\mu = 20$, $\sigma = 5$, $L_s = 2$, $h_s = 1$, $p = 10$, and $s = 0.9$. The expected cost function

Figure 2 Holding Cost, Penalty Cost, and Total Cost for a Flat Penalty Contract



is not necessarily convex in the base stock level y_s. Similarly, it can be demonstrated that the expected profit function is not concave in y_s. However, Proposition 1 states that the expected profit function $E\Pi_s^f(y_s)$ is quasi-concave in the base stock level y_s and states the optimality condition.

Proposition 1. *Under a flat penalty contract, $E\Pi_s^f(y_s)$ is quasi-concave in y_s. The optimal base stock level satisfies*

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

4.2. Unit Penalty Contract

Under a unit penalty contract, the supplier incurs a penalty charge of p for each unit of demand that is backordered and that exceeds the number of units required by the service level contract. To determine the expected penalty per period, we use the demand space partitioning of Figure 3. In area 3, all backorders and demands are filled and it suffices to analyze areas 1 and 2. In area 1, the demand over the previous L_s periods was greater than the base stock level y_s. The

Figure 1 Demand Space Partitioning

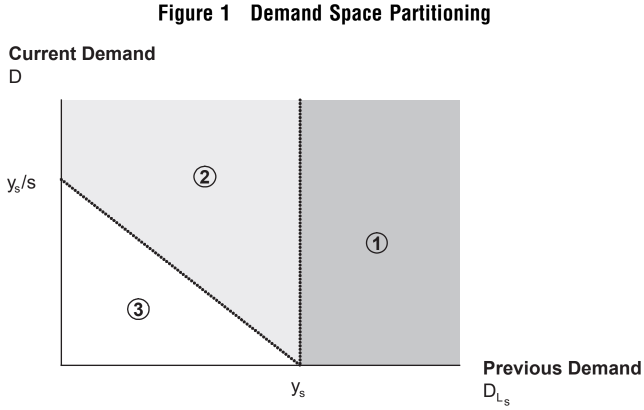
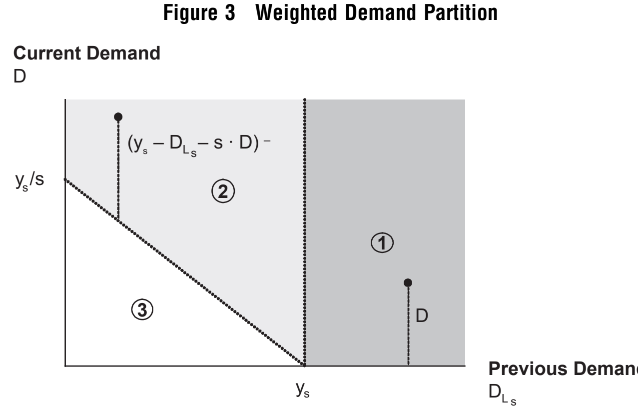


Figure 3 Weighted Demand Partition



inventory is insufficient to fill the backorders of the previous periods and the supplier incurs a penalty payment of pD. In area 2, all previous backorders are filled, but the inventory is less than sD and the supplier incurs a penalty of $p(y_s - D_{L_s} - sD)^-$. So, the expected penalty charge per period is

$$
\begin{aligned}
&E[P_u(y_s, s, D, D_{L_s})] \\
&= p\left(\int_{x=0}^{y_s}\int_{\varphi=\frac{y_s-x}{s}}^{\infty}\left(\varphi-\frac{y_s-x}{s}\right)f(\varphi)f_{L_s}(x)d\varphi dx\right. \\
&\quad \left.+\int_{x=y_s}^{\infty}\mu f_{L_s}(x)dx\right) \\
&= p\left(\int_{x=0}^{y_s} b_s\left(\frac{y_s-x}{s}\right)f_{L_s}(x)dx+(1-F_{L_s}(y_s))\mu\right),
\end{aligned}
$$

where

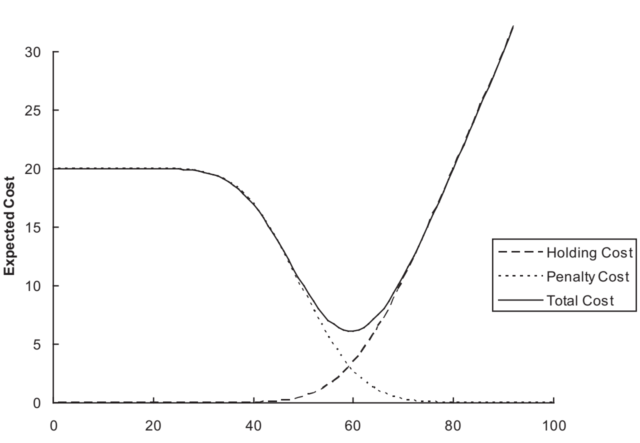
$$b_s(z) = \int_{\delta=z}^{\infty}(\delta - z)f(\delta)d\delta$$

is the expected backorder level. The expected one-period profit function is

$$
\begin{aligned}
E\Pi_s^u(y_s^u) = (w-c)\mu - h_s \int_{x=0}^{y_s}(y_s-x)f_{L_s+1}(x)dx \qquad (5)\\
-p\left(\int_{x=0}^{y_s} b_s\left(\frac{y_s-x}{s}\right)f_{L_s}(x)dx+(1-F_{L_s}(y_s))\mu\right).
\end{aligned}
$$

Figure 4 illustrates the cost terms for truncated normally distributed demand with μ = 20, σ = 5, L_s = 2, h_s = 1, p = 1, s = 0.9. As before, the expected total cost function is not convex in y_s and the expected profit function is not concave in y_s. However, Proposition 2 states that the expected profit function $E\Pi_s^u(y_s)$ is quasi-concave and states the optimality condition.

Figure 4 Holding Cost, Penalty Cost, and Total Cost for a Unit Penalty Contract



PROPOSITION 2. *Under a unit penalty contract, $E\Pi_s^u(y_s)$ is quasi-concave in y_s. The optimal base stock level satisfies*

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

The following Proposition 3 provides sensitivity results for flat and unit penalty contracts.

PROPOSITION 3. *Under both contract types, the optimal base stock level y_s is increasing in the penalty charge p. It is decreasing in the holding cost factor h_s. The wholesale price w does not influence the optimal base stock level, but has a direct effect on the expected one-period profit of the supplier.*

We have assumed here that the wholesale price w does not influence the holding costs h_s. It follows that w does not have an effect on the optimal base stock level. As we will show in sections 5.3 and 5.4, the reactions of the optimal base stock level y_s to changes in the contract service level s are non-trivial, such that they deserve a separate analysis.

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

5. Coordinating Contracts

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

5.1. Optimal Centralized Solution

To characterize the centralized solution, we first show how the optimal base stock levels are determined in the centralized decision-making case. We want to optimize the following joint supplier and manufacturer problem:

$$
\begin{aligned}
E\Pi_s^*(w,p,s) = \max_{y_s} E\Big[(w-c)D - h_s(y_s - D_{L_s+1})^+ \\
- P_\kappa(y_s,p,s,D,D_{L_s})\Big],
\end{aligned}
$$

where κ = {f,u} depending on which penalty contract is used, and

$$E\Pi_m^* = \max_{y_m, w, p, s} E\Big[(r-w)D + P_\kappa(y_s, p, s, D, D_{L_s})\Big] - h_m I_m(y_m, y_s) - b_m B_m(y_m, y_s)$$
$$\text{s.t.} \quad E\Pi_s^*(w,p,s) \geq \hat{\Pi}_s.$$

The centralized profit function then equals

$$\begin{aligned} E\Pi^* &= E\Pi_s^* + E\Pi_m^* \\ &= E\Big((r-c)D - h_s(y_s - D_{L_s+1})^+ \\ &\quad - h_m I_m(y_m, y_s) - b_m B_m(y_m, y_s)\Big) \end{aligned}$$
$$\text{s.t.} \quad E\Pi_s^*(w,p,s) \geq \hat{\Pi}_s.$$

The penalty payments disappear in the centralized solution, which means that the optimal solution is independent of the chosen service level contract in a centralized system. The side constraint $E\Pi_s^*(w,p,s) \geq \hat{\Pi}_s$ does not influence the operational decision concerning the base stock levels, that is, it can be satisfied by a simple transfer payment. In order to obtain the optimal base stock levels, we only have to consider the profit function $E\Pi^* = E((r-c)D - h_s(y_s - D_{L_s+1})^+ - h_m I_m(y_m, y_s) - b_m B_m(y_m, y_s))$. The solution to this problem is attributed to Gallego and Özer (2003), who were the first to show that the myopic algorithm we are using in the following yields optimal echelon base stock levels for periodic review serial systems for both infinite and finite horizon problems: The optimal centralized solution is characterized by the echelon base stock levels y_s^e and y_m^e that satisfy the manufacturer's first-order condition

$$F_{L_m+1}\big(y_m^e\big) = \frac{h_s + b_m}{h_m + h_s + b_m},$$

and the supplier's first-order condition

$$-b_m + (b_m + h_s)F_{L_s}\big(y_s^e - y_m^e\big) + (b_m + h_m + h_s) \int_{x=y_s^e - y_m^e}^{\infty} f_{L_s}(x) F_{L_M+1}\big(y_s^e - x\big)dx = 0.$$

Starting with the manufacturer, these first-order conditions yield the optimal echelon base stock levels. The optimal installation base stock levels $y_s^{SC^*}$ and $y_m^{SC^*}$ then can be derived from the optimal echelon base stock levels by setting $y_m^{SC^*} = y_m^e$ and $y_s^{SC^*} = y_s^e - y_m^e$. The corresponding expected supply chain profit per period is denoted by $E\Pi_{SC}^*(y_m^{SC^*}, y_s^{SC^*})$.

In our setting, the manufacturer's objective is to maximize his expected profit subject to a constraint that the supplier's expected profit per period is at least $\hat{\Pi}_s$. This implies that the maximum expected profit of the manufacturer is

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

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

$$\begin{aligned} E\Pi_s(y_s^{SC^*}) &= (w-c)\mu - h_s \int_{x=0}^{y_s^{SC^*}} (y_s^{SC^*} - x) f_{L_s+1}(x)dx \\ &\quad - E[P_\kappa(y_s^{SC^*}, s, p, D, D_{L_s})] \\ &= \hat{\Pi}_s. \end{aligned} \tag{6}$$

In the next subsection we show how the contract parameters are chosen in a decentralized supply chain.

5.2. Decentralized Solution

Under a decentralized solution, the manufacturer chooses (w,p,s), such that the supplier is incentivized to choose the supply chain optimal base stock level and achieves a reservation profit of $\hat{\Pi}_s$. Solving Equation (6) for the wholesale price w that satisfies the supplier's participation constraint, we obtain the optimal wholesale price of the decentralized supply chain:

$$\begin{aligned} w_m^*(y_s^{SC^*}, s, p) &= c + \frac{h_s}{\mu} \int_{x=0}^{y_s^{SC^*}} (y_s^{SC^*} - x) f_{L_s+1}(x)dx \\ &\quad + \frac{E[P_\kappa(y_s^{SC^*}, s, p, D, D_{L_s})]}{\mu} + \frac{\hat{\Pi}_s}{\mu}. \end{aligned} \tag{7}$$

The equation shows that the supply chain optimal wholesale price $w_m^*(y_s^{SC^*}, s, p)$ is equal to the sum of unit cost, expected unit inventory holding cost at the supplier, expected unit penalty cost, plus the unit reservation profit $\hat{\Pi}_s/\mu$. Thereby, $w_m^*(y_s^{SC^*}, s, p)$ ensures that the supplier's reservation profit will always be achieved.

We show next that there exist an infinite number of contracts (w,p,s) that achieve the first-best solution and show how the parameter values of these contracts can be computed.

5.3. Coordinating Flat Penalty Contract

For a flat penalty contract, Proposition 4 states for which combinations of s and p the supplier chooses the supply chain optimal base stock level.

PROPOSITION 4. *Under a flat penalty contract, the supplier's optimal base stock level is equal to* $y_s^{SC^*}$ *if*

$$p(s) = \frac{h_s F_{L_s+1}(y_s^{SC*})}{\frac{1}{s}\int_{x=0}^{y_s^{SC*}} f_{L_s}(x) f\Big(\frac{y_s^{SC*} - x}{s}\Big)dx}, \quad 0 < s \leq 1. \tag{8}$$

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

For a given contract service level s, we can compute the coordinating penalty cost $p(s)$ using Equation (8). The wholesale price w then follows from Equation (7). Therefore, there exist an infinite number of contracts that coordinate the supply chain.

Note that we present a three-parameter contract to the supplier, although a two-parameter contract would be sufficient for coordinating the supply chain. There are two main reasons for doing so. First, such three-parameter contracts are often used in practice (at the beginning of section 6, we provide details of an example from the consumer goods industry and refer to the Introduction for additional examples). Second, the over-specification allows us to choose one parameter value freely and to determine the values of the other two parameters, such that the supply chain is coordinated (e.g., we use this approach below, where we introduce contracts that are service level consistent, i.e., contracts where the contract service level is set equal to the optimal service level of the company. The wholesale price and the penalty cost are then chosen, such that the supply chain is coordinated).

Since it is well known that many contracts with two parameters can coordinate the supply chain, we want to explain in more detail the effect of our contract parameters. The contract service level s and the penalty charge p are jointly responsible for achieving a certain base stock level at the supplier, whereas the wholesale price w only ensures that the supplier earns her minimum reservation profit. Clearly, we could spare one parameter in our setting, for example, by setting the contract service level $s = 1$, which would give us a two-parameter coordinating contract, but we would lose some applicability, since many contracts in practice are negotiated with those three parameters (w,p,s).

Figure 5 shows numerical results for three examples with truncated normally distributed demand with $\mu = 20$, $\sigma = 5$, $h_s = 1$, $L_s = 2$, and $L_m = 4$. To analyze a variety of situations, we used inventory holding cost and backorder penalty cost combinations of $(h_m, b_m) = \{(1.7, 0.9), (55, 55), (1500, 1500)\}$, resulting in optimal centralized solutions of $(y_s^{SC*}, y_m^{SC*}) = \{(30, 101), (50, 100), (60, 100)\}$.

Figure 5 illustrates that the coordinating penalty cost $p(s)$ is increasing in the contract service level s if the base stock level at the supplier is low ($y_s^{SC*} = 30$). Similarly, $p(s)$ is decreasing in the contract service level s if the base stock level y_s^{SC*} is high ($y_s^{SC*} = 60$). For an intermediate base stock level ($y_s^{SC*} = 50$), the coordinating penalty cost $p(s)$ is decreasing–increasing in the contract service level s. This insight is important for management because it illustrates that the dependency between the contract service level s and the penalty cost p is affected by the base stock level y_s^{SC*}. In settings where the base stock level y_s^{SC*} is low, increases in the contract service level s must go along with increases in the penalty cost p. In settings where the base stock level y_s^{SC*} is high, the opposite holds. Proposition 5 states that these effects hold in general.

PROPOSITION 5. *Under a flat penalty contract, the penalty cost factor p is quasi-convex in the contract service level s. p increases in s for low base stock levels (where the marginal penalty probability is increasing in s), whereas p decreases in s for high base stock levels (where the marginal penalty probability is decreasing in s).*

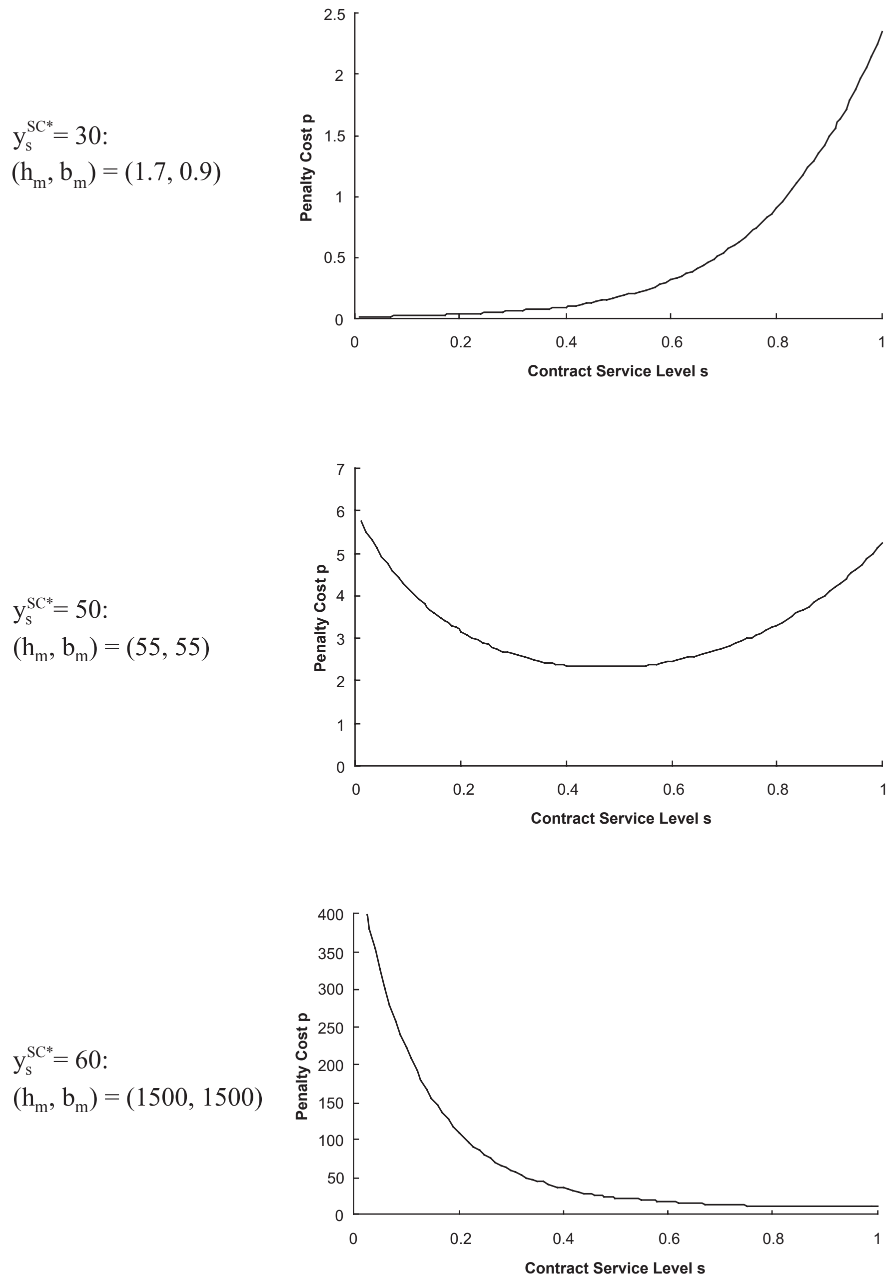
To understand the rationale behind the impact of the contract service level s on the coordinating penalty cost $p(s)$, recall that the supplier trades off marginal savings in expected inventory holding cost against marginal increases in expected penalty payments when deciding on the base stock level y_s. From Equation (4), it can be seen that in a coordinated supply chain

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

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

The question remaining is when the marginal penalty payment probability decreases in s and when it increases in s? To answer this question, we rewrite the penalty payment probability as

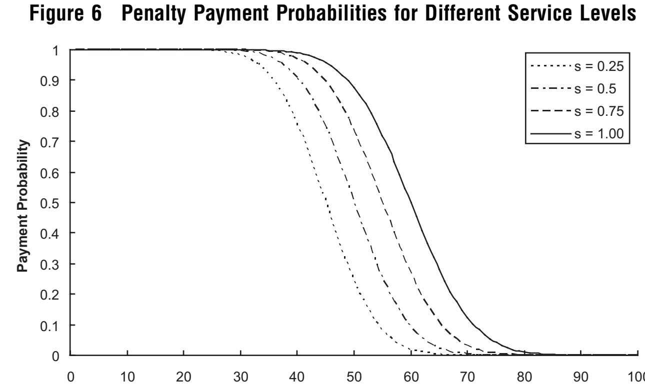
Figure 5 Coordinating Flat Penalty Contracts



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

which shows that the payment probability is a complementary cumulative distribution function (ccdf.). For our numerical example, this ccdf. is shown in Figure 6 for various contract service levels s. Since demand is logconcave distributed, the marginal penalty payment probability, that is, the derivative of the ccdf., is increasing for small base stock levels y_s and is decreasing for large y_s, which explains why the coordinating penalty cost $p(s)$ is increasing for small base stock levels y_s and decreasing for large base stock levels y_s.

Figure 6 Penalty Payment Probabilities for Different Service Levels



5.4. Coordinating Unit Penalty Contract

Analogously to Proposition 4, Proposition 6 states for which combinations of s and p the supplier chooses the supply chain optimal base stock level under a unit penalty contract.

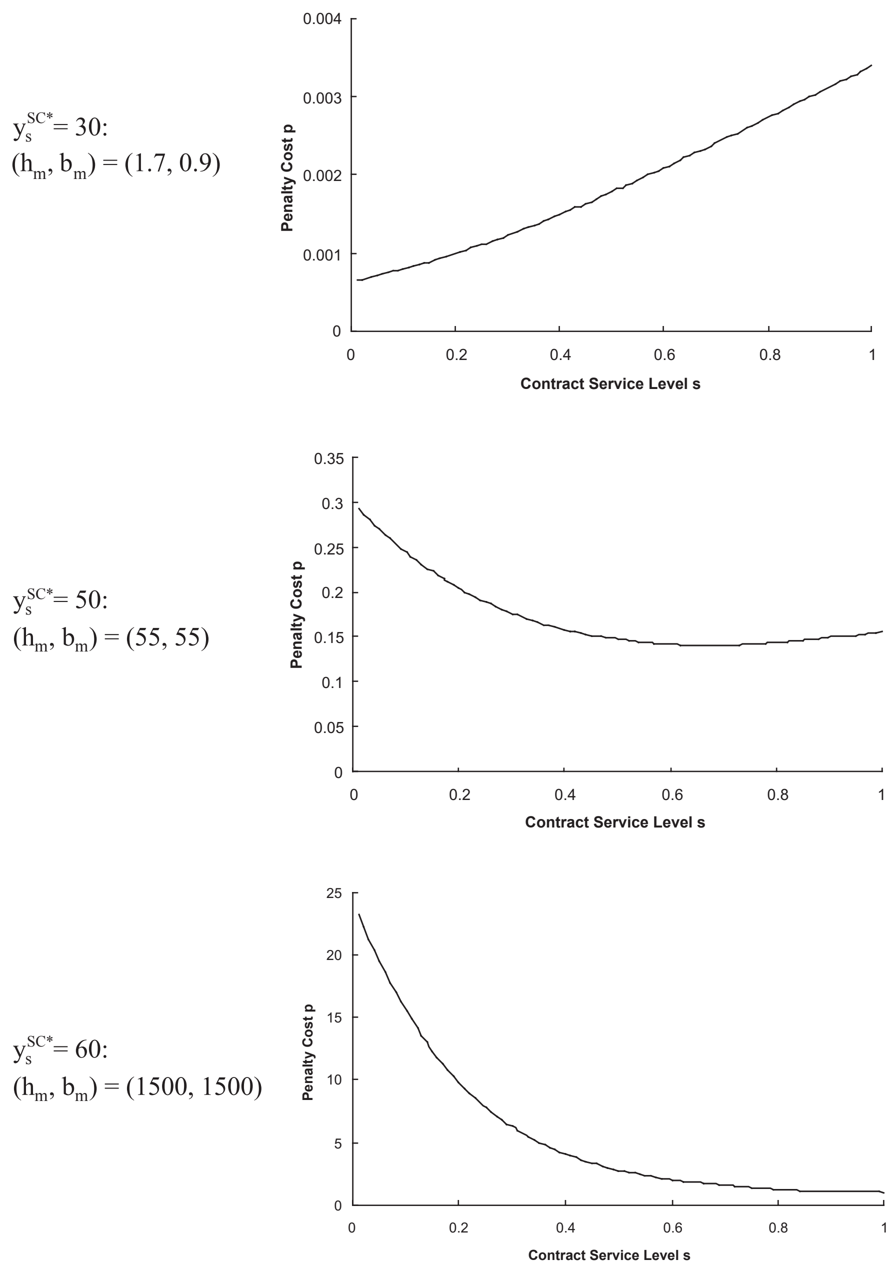
PROPOSITION 6. *Under a unit penalty contract, the supplier's optimal base stock level is equal to* $y_s^{SC^*}$ *if*

$$p(s) = \frac{h_s F_{L_s+1}(y_s^{SC*})}{\int_{x=0}^{y_s^{SC*}} \frac{\left(1-F\left(\frac{y_s^{SC*}-x}{s}\right)\right)}{s} f_{L_s}(x)dx}, \quad 0<s\leq 1. \qquad (9)$$

Similar to the flat penalty contracts, there exists a unit penalty contract for all contract service levels $0 < s \leq 1$ that coordinate the supply chain, if the underlying distribution function has infinite non-negative support.

For a given contract service level s, we can compute the coordinating penalty cost $p(s)$ using Equation (9). Figure 7 shows numerical results for our three numerical examples that are similar to the results of the flat penalty contract. For a low base stock level ($y_s^{SC*} = 30$), the coordinating penalty cost $p(s)$ is increasing in the contract service level s, for a medium base stock level ($y_s^{SC*} = 50$) it is decreasing–increasing,

Figure 7 Coordinating Unit Penalty Contracts



and for a large base stock level ($y_s^{SC*} = 60$) it is decreasing. As for the flat penalty contracts, we can state general results on the effect of the contract service level s on the coordinating penalty cost $p(s)$ in Proposition 7.

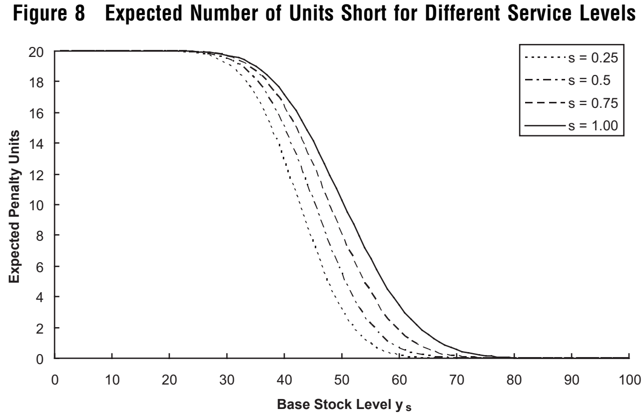
PROPOSITION 7. *Under a unit penalty contract, the penalty cost factor p is convex in the contract service level s. p increases for low base stock levels (where the marginal expected number of units short is increasing in s), whereas p is decreasing for high base stock levels (where the marginal expected number of units short is decreasing in s).*

As before, we can explain the effect of the contract service level s on the coordinating penalty cost $p(s)$ by recalling that the supplier trades off marginal savings in expected inventory holding cost against marginal increases in expected penalty payments when deciding on the base stock level y_s. From Equation (5), it can be seen that in a coordinated supply chain

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

Using the same arguments as in subsection 5.3, we see that in situations in which the *marginal expected number of units short*, that is, $d/dy_s(\int_{x=0}^{y_s} b_s(\frac{y_s - x}{s}) f_{L_s}(x)dx + (1 - F_{L_s}(y_s))\mu)$, is increasing in the contract service level s, the coordinating penalty cost $p(s)$ is increasing in s. In situations in which the marginal expected number of units short is decreasing in s, the coordinating penalty cost $p(s)$ is decreasing in the contract service level s. Figure 8 shows that the marginal expected number of units short is increasing in the contract service level s for small base stock levels y_s and is decreasing for large y_s. As before, this observation explains the effect that, for a unit penalty contract, the coordinating penalty cost $p(s)$ is increasing for small base stock levels y_s and decreasing for large base stock levels y_s.

Figure 8 Expected Number of Units Short for Different Service Levels



5.5. Contract Consistent Points

For both contract types, the manufacturer must decide which combination of s and p to specify in the contract. Since any point on the curves of Figures 5 and 7 coordinates the supply chain, the manufacturer has infinitely many combinations to choose from. However, there exists one point on each curve that is particularly attractive. We refer to these points as *contract consistent points*, because at these points the contract service level is equal to the traditional α or β service levels.

From a managerial perspective, it would be attractive to use supply contracts where the contract service level s is equal to the traditional α or β service level, because the supplier can then focus on achieving the service level by using well-known methods from inventory management, that is, by using Equations (1) and (2).

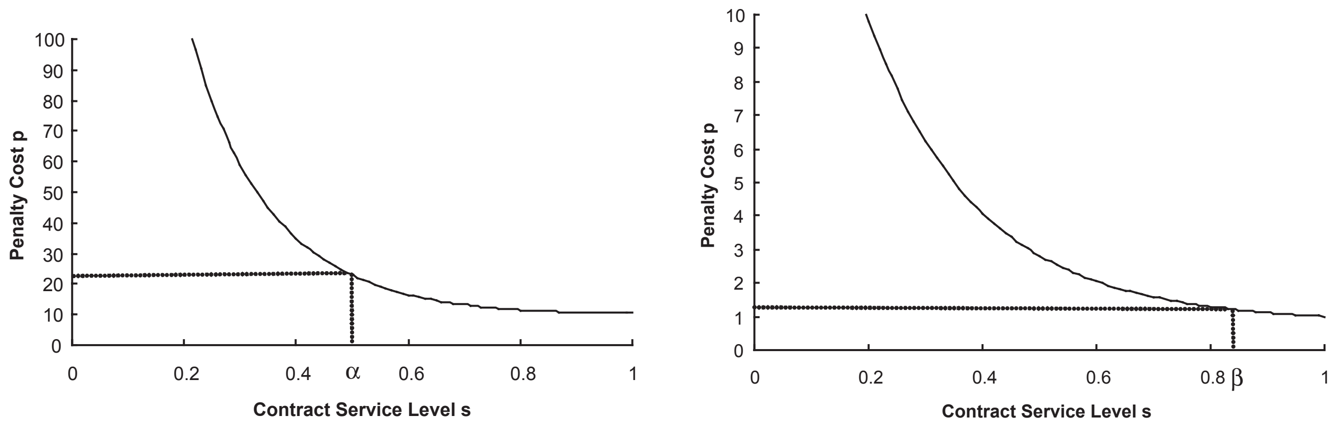
Figure 9 shows how we can design such contracts for our example with $y_s^{SC^*} = 60$, $\alpha = 50\%$, and $\beta = 82.75\%$. If we choose $s = \alpha = 50\%$ and $p = 22.86$ for a flat penalty contract or $s = \beta = 82.75\%$ and $p = 1.24$ for a unit penalty contract, the supply chain is coordinated. We have shown before that such service level consistent contracts always exist for flat and unit penalty contracts.

6. Alternative Event Sequences

So far, we have assumed that the manufacturer offers the supplier a contract and that the manufacturer specifies all contract parameters, that is, specifies s, p, and w. Our approach can be easily adapted to other event sequences. We illustrate how the approach can be adapted for a setting where the manufacturer specifies the service level s and the penalty payment p and the supplier offers a wholesale price w.

This specific setup can be found, for example, in the consumer goods industry. In this industry, retailers often specify a service level that they expect all of their suppliers to achieve and specify penalties for not achieving them. The values of the contract service level and the penalties are based on operational issues, such as cost of an out-of-stock situation in the store and administrative cost of handling stockouts, and are communicated to the supplier and are not negotiable. Various suppliers are then asked to quote a wholesale price. Among other factors such as competitors' prices or prices they charge to other customers, suppliers typically also take into account the profit or contribution margin they want to achieve with their products, which we model by using a reservation

Figure 9 Consistent Contracts with α=50% and β=82.75%



Flat Penalty Contract   Unit Penalty Contract

profit. This reservation profit is set by the supplier, such that the minimum profitability target of the supplier is met. If a reservation profit that is greater than the minimum reservation profit is accepted by the retailer, the supplier can set a reservation profit that is greater than the minimum reservation profit. We do not analyze reservation profit optimization in this study, but consider a model where the supplier has set her reservation profit and is interested in the wholesale price she must charge in order to achieve it.

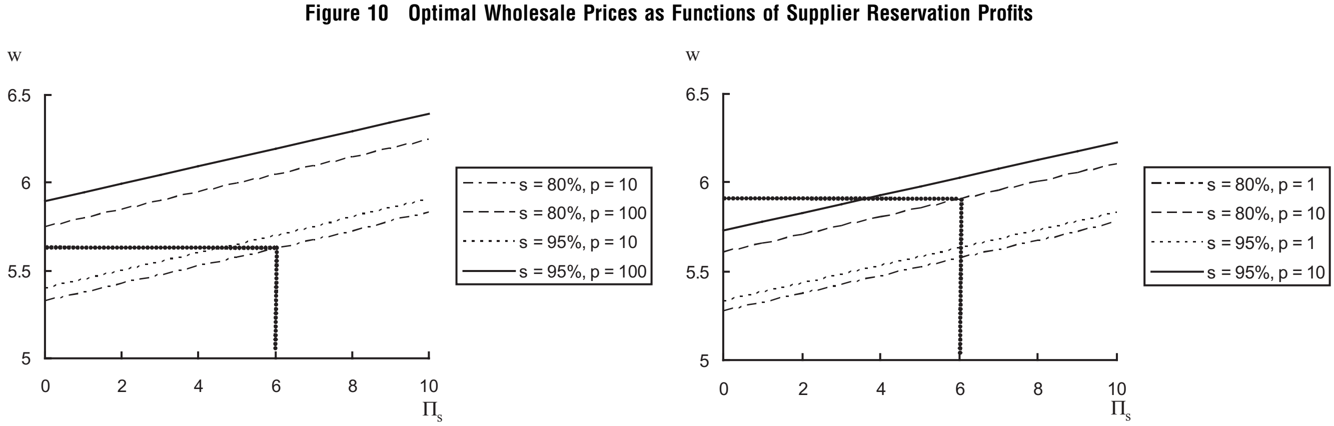
We model this setting by considering a situation where the manufacturer (the drugstore in the discussion above) sets the service level s and the penalty payment p and then the supplier (consumer goods manufacturer in the discussion above) determines the wholesale price w that allow her achieving her target reservation profit. As before, our interest is only in solutions that coordinate the supply chain, that is, in solutions where the contract parameters ensure that the supplier and the manufacturer use first-best inventory levels.

Under this event sequence, the manufacturer can find s and p by computing the optimal $y_s^{SC^*}$ as before. We have shown before that such a contract always exists. Then, the supplier computes the wholesale price w, such that she achieves her reservation profit. We obtain the optimal wholesale price of the coordinated supply chain by solving Equation (6) for the wholesale price w:

$$w_s^*(y_s^{SC^*}, s, p) = c + \frac{h_s}{\mu} \int_{x=0}^{y_s^{SC^*}} (y_s^{SC^*} - x) f_{L_s+1}(x) dx + \frac{E[P_\kappa(y_s^{SC^*}, s, p, D, D_{L_s})]}{\mu} + \frac{\hat{\Pi}_s}{\mu}. \tag{10}$$

For our example with μ = 20, σ = 5, $L_s = 2$, c = 5, and $h_s = 1$, Figure 10 shows solutions for various parameter value combinations for flat penalty and unit penalty contracts. For instance, if the manufacturer requires a service level of s = 95% and requests a penalty payment of p = 10, then the supplier can achieve a reservation profit of $\hat{\Pi}_s = 6$ under a flat penalty contract by charging a wholesale price w = 5.6. Under a unit penalty contract the supplier would charge w = 5.9.

Figure 10 Optimal Wholesale Prices as Functions of Supplier Reservation Profits



Flat Penalty Contract   Unit Penalty Contract

Various changes of the model setup can be incorporated quite easily. We can analyze a setup where the supplier's objective is maximizing her own profit by determining all contract parameters s, p, and w, but at the same time guaranteeing a reservation profit of $\hat{\Pi}_m$ at the manufacturer. To analyze this situation, we can simply replace $\hat{\Pi}_s$ in Equation (10) with $E\Pi^*_{SC} - \hat{\Pi}_m$.

We can also analyze a setup where the manufacturer sets the contract type, the supplier offers the parameters s and p, and the manufacturer quotes a maximum wholesale price w that he is willing to pay. To analyze this situation, we first determine the centralized solution and find the coordinating contract parameters s and p from Proposition 4 or 6. Then the manufacturer can solve Equation (3) for w, such that $E\Pi^*_m = \hat{\Pi}_m$.

The discussion illustrates that there exists a variety of event sequences that can be dealt with by our approach. The approach for analyzing a certain event sequence is essentially the same for all event sequences: First, the Stackelberg leader computes the first-best solution using the centralized model. Then, the Stackelberg leader determines some or all contract parameters, such that the expected profit of the Stackelberg leader is maximized, taking into account that all reservation profit constraints can be met. If not all contract parameters are specified by the Stackelberg leader, the Stackelberg follower responds to the incompletely specified contract by specifying the remaining contract parameters, such that the expected profit of the Stackelberg follower is maximized, taking into account that all reservation profit constraints are met.

7. Conclusion

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

In line with the majority of the literature on supply chain contracting, we have assumed that both supply chain partners have information on all relevant parameters. However, this might not always be the case and some authors have started analyzing settings with asymmetric information (see section 10 in Cachon [2003] for an overview). In many situations, the manufacturer might not know the exact supplier reservation profit. If the expected profit of the supplier is below the reservation profit, she will not accept the contract. Therefore, the manufacturer has to take this into account and his uncertainty on the level of the reservation profit plays an important role. One means of addressing this and other information asymmetries could be the use of contract menus. Contract menus have been used by Corbett et al. (2004) and Cachon and Zhang (2006) to allow manufacturers revealing their supplier's true costs. However, analyzing information asymmetry is beyond the scope of this study and we leave it to future research.

Appendix

PROOF OF PROPOSITION 1. We want to show that $E\Pi^f_s(y_s)$ is quasi-concave in y_s and a unique maximum y^f_s exists. Consider the derivative

$$\frac{dE\Pi^f_s(y_s)}{dy_s} = -h_s F_{L_s+1}(y_s) + \frac{p}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\Big(\frac{y_s - x}{s}\Big)dx.$$

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

$$-h_s F_{L_s+1}(y_s) + \frac{p}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\Big(\frac{y_s - x}{s}\Big)dx \geq 0$$

$$\Leftrightarrow \frac{F_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\big(\frac{y_s-x}{s}\big)dx} \leq \frac{p}{h_s}.$$

We can rewrite the last line as

$$\begin{aligned}&\frac{F_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\big(\frac{y_s-x}{s}\big)dx}\\ &\quad= \frac{F_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\big(\frac{y_s-x}{s}\big)dx}\cdot\frac{f_{L_s+1}(y_s)}{f_{L_s+1}(y_s)}\\ &\quad= \frac{F_{L_s+1}(y_s)}{f_{L_s+1}(y_s)}\cdot\frac{f_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f\big(\frac{y_s-x}{s}\big)dx} \leq \frac{p}{h_s}.\end{aligned} \tag{11}$$

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

From $(g(x)h(x))' = g(x)h'(x) + g'(x)h(x)$ the LHS of Equation (11) is non-decreasing in y_s. It follows that there exists only one positive area of $\frac{dE\Pi_s^f(y_s)}{dy_s}$ in at most one subset of y_s and the sign of $\frac{dE\Pi_s^f(y_s)}{dy_s}$ changes at most once from + to − and thus the objective function $E\Pi_s^f(y_s)$ is quasi-concave in y_s. Then, an optimal base stock level y_s^f is unique. Hence it is sufficient to set the first derivative to zero. □

Proof of Proposition 2. We want to show that $E\Pi_s^u(y_s)$ is quasi-concave in y_s and a unique maximum y_s^u exists. Consider the derivative

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

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

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

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

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

We know that $\frac{F_{L_s}(y_s)}{F_{L_s+1}(y_s)}$ is non-increasing in y_s and ≥ 0. Also $\left(1 - \frac{\int_{x=0}^{y_s} F(\frac{y_s - x}{s})f_{L_s}(x)dx}{F_{L_s}(y_s)}\right)$ is non-increasing in y_s and $\frac{\int_{x=0}^{y_s} F(\frac{y_s - x}{s})f_{L_s}(x)dx}{F_{L_s}(y_s)} \leq 1$ (Rosling 2002). From $(g(x)h(x))' = g(x)h'(x) + g'(x)h(x)$ the term $\frac{F_{L_s}(y_s)}{F_{L_s+1}(y_s)} \left(1 - \frac{\int_{x=0}^{y_s} F(\frac{y_s - x}{s})f_{L_s}(x)dx}{F_{L_s}(y_s)}\right)$ is also non-increasing in y_s and $\frac{d}{dy_s} \frac{\int_{x=0}^{y_s}(1 - F(\frac{y_s - x}{s}))f_{L_s}(x)dx}{F_{L_s+1}(y_s)} \leq 0$. It follows that $\frac{\int_{x=0}^{y_s} \frac{(1 - F(\frac{y_s - x}{s}))}{s} f_{L_s}(x)dx}{F_{L_s+1}(y_s)}$ is non-increasing in y_s. Then the derivative is positive in at most one subset of y_s and the sign of $\frac{dE\Pi_s^u(y_s)}{dy_s}$ changes at most once from + to − and hence the objective function is quasi-concave in y_s. Thus, an optimal base stock level y_s^u is unique and it is sufficient to set the first derivative to zero. □

Proof of Proposition 3. Flat penalty contract: The first part of the proof follows from considering the optimality criterion $\frac{F_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f(\frac{y_s - x}{s})dx} \leq \frac{p}{h_s}$ from above. For a higher penalty cost p, the base stock level y_s has to increase in order to achieve a change of signs for $\frac{dE\Pi_s^f(y_s)}{dy_s}$ since $\frac{F_{L_s+1}(y_s)}{\frac{1}{s}\int_{x=0}^{y_s} f_{L_s}(x) f(\frac{y_s - x}{s})dx}$ is non-decreasing in y_s. Similarly, the base stock level y_s has to decrease for an increasing h_s.

Unit penalty contract: The first part of the proof follows from considering the optimality criterion $\frac{\int_{x=0}^{y_s} \frac{(1 - F(\frac{y_s - x}{s}))}{s} f_{L_s}(x)dx}{F_{L_s+1}(y_s)} \geq \frac{h_s}{p}$ from above. For a higher penalty cost p, the base stock level y_s has to increase in order to achieve a change of signs for $\frac{dE\Pi_s^u(y_s)}{dy_s}$ since $\frac{\int_{x=0}^{y_s} \frac{(1 - F(\frac{y_s - x}{s}))}{s} f_{L_s}(x)dx}{F_{L_s+1}(y_s)}$ is non-increasing in y_s. Similarly, the base stock level y_s has to decrease for an increasing h_s. □

Proof of Proposition 4. The supplier solves

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

Differentiation of the expected profit function with respect to y_s yields

$$\frac{d}{dy_s} E\Pi_s(y_s) = -h_s F_{L_s+1}(y_s) + \frac{p}{s} \int_{x=0}^{y_s} f_{L_s}(x) f\left(\frac{y_s - x}{s}\right)dx. \tag{12}$$

With Equation (8) in Equation (12) we see that $d/dy_s E\Pi_s(y_s) = 0$ for $y_s = y_s^{SC^*}$. Since the expected profit function is quasi-concave, this corresponds to the optimal solution.

Proof of Proposition 5. The proof follows by the unimodal property of the demand distribution. Consider the derivative

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

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

Proof of Proposition 6. The supplier solves

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

Differentiation of the expected profit function with respect to y_s yields

$$\frac{d}{dy_s}E\Pi_s(y_s) = -h_s F_{L_s+1}(y_s) - p\int_{x=0}^{y_s} \frac{\left(F(\frac{y_s-x}{s}) - 1\right)}{s} f_{L_s}(x)dx. \quad (13)$$

With Equation (9) in Equation (13), we see that $d/dy_s E\Pi_s(y_s) = 0$ for $y_s = y_s^{SC^*}$. Since the expected profit function is quasi-concave, this corresponds to the optimal solution. □

Proof of Proposition 7. In the coordinated solution, the penalty cost factor equals

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

First, we will analyze the denominator

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

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

$$\int_{z=\frac{y_s^{SC*}}{s}}^{0} (F(z) - 1) f_{L_s}(y_s^{SC^*} - zs)dz = \int_{z=0}^{\frac{y_s^{SC*}}{s}} (1 - F(z)) f_{L_s}(y_s^{SC^*} - zs)dz = \hat{F}\left(\frac{y_s^{SC^*}}{s}\right).$$

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

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

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

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

From Bagnoli and Bergstrom (2005), Theorem 7, we know that if $\bar{F}$ is logconvex and t is a convex function, the composition $\bar{F}(t(x))$ is logconvex. From Boyd and Vandenberghe (2004), we see that logconvexity implies convexity. Scaling with a constant factor preserves convexity. Thus, $p(s)$ is convex in the contract service level s. □

References

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

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

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

Cachon, G. P. 2003. Supply chain coordination with contracts. S. Graves, T. de Kok, eds. *Handbook in Operations Research and Management Science: Supply Chain Management*. Elsevier Publishers B.V., North-Holland, The Netherlands, 229–339.

Cachon, G. P., F. Zhang. 2006. Procuring fast delivery: Sole sourcing with information asymmetry. *Manage. Sci.* **52**(6): 881–896.

Cachon, G. P., P. H. Zipkin. 1999. Competitive and cooperative inventory policies in a two-stage supply chain. *Manage. Sci.* **45**(7): 939–953.

Cohen, M. A., Y.-S. Zheng, Y. Wang. 1999. Identifying opportunities for improving Teradyne's service-parts logistics system. *Interfaces* **29**(4): 1–18.

Corbett, C. J., D. Zhou, C. S. Tang. 2004. Designing supply contracts: Contract type and information asymmetry. *Manage. Sci.* **50**(4): 550–559.

Gallego, G., Ö. Özer. 2003. Optimal replenishment policies for multiechelon inventory problems under advance demand information. *Manuf. Serv. Oper. Manag.* **5**(2): 157–175.

Karlin, S., F. Proschan. 1960. Pólya type distributions of convolutions. *Annals Math. Statist.* **31**(3): 721–736.

Kim, S., M. A. Cohen, S. Netessine. 2007. Performance contracting in after-sales service supply chains. *Manage. Sci.* **53**(12): 1843–1858.

Lee, H. L., S. Whang. 1999. Decentralized multi-echelon supply chains: Incentives and information. *Manage. Sci.* **45**(5): 633–640.

Lutze, H., Ö. Özer. 2008. Promised lead-time contracts under asymmetric information. *Oper. Res.* **56**(4): 898–915.

Milner, J. M., T. L. Olsen. 2008. Service-level agreements in call centers: Perils and prescriptions. *Manage. Sci.* **54**(2): 238–252.

Mostberger, P. 2006. *Supply Chain Controlling – Basis für gemeinsamen Erfolg (in German)*. Supply Chain Summit, September 24–26, 2006, Montreux, Switzerland.

Porteus, E. L. 1990. Stochastic inventory theory. D. P. Heyman, M. Sobel, eds. *Handbook in OR & MS*, Vol. **2**. Elsevier Publishers B.V., North-Holland, The Netherlands, 605–652.

Porteus, E. L. 2000. Responsibility tokens in supply chain management. *Manuf. & Serv. Oper. Manage.* **2**(2): 203–219.

Rosling, K. 2002. Inventory cost rate functions with nonlinear shortage costs. *Oper. Res.* **50**(6): 1007–1117.

SCORmark. 2008. *Supply-Chain Council's Benchmarking Portal*. Available at http://www.supply-chain.org/cs/benchmarking (accessed June 17, 2008).

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

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

Sobel, M. J. 2004. Fill rates of single-stage and multistage supply systems. *Manuf. & Serv. Oper. Manag.* **6**(1): 41–52.

Stratmann, U. 2006. *Gemeinsam im Takt – Lean Manufacturing (in German)*. Management Forum Seminar, June 30, 2006, Cologne, Germany.

Thomas, D. J. 2005. Measuring item fill-rate performance in a finite horizon. *Manuf. & Serv. Oper. Managem.* **7**(1): 74–80.

Thonemann, U., K. Behrenbeck, R. Diederichs, J. Großpietsch. 2003. *Supply Chain Champions: Was sie tun und wie Sie einer werden (in German)*. Financial Times Deutschland/Gabler, Wiesbaden, Germany.

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

Wang, Y., M.A. Cohen, Y.-S. Zheng. 2002. Differentiating customer service on the bases of delivery lead-times. *IIE Trans.* **34**(11): 979–989.

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