

Inventory control in a decentralised two-stage make-to-stock queueing system

Yasemin Arda^{ab*} and Jean-Claude Hennet^c

^aHEC-ULg, 4000 Liège, Belgium; ^bLAAS-CNRS, Toulouse, France; ^cLSIS UMR CNRS 6168, Faculté de Saint Jérôme, Université Paul Cézanne, Marseille, France

(Received 10 December 2007; final version received 3 February 2008)

In an Enterprise network, several companies interact to produce families of goods. Each member company seeks to optimise his own production and inventory policy to maximise his profit. These objectives are generally antagonistic and can lead to contradictory choices in the context of a network with a high degree of local decisional autonomy. To avoid a global loss of economic efficiency, the network should be equipped with a coordination mechanism. The present article describes a coordination contract negotiated between a manufacturer and a supplier. The purpose of the negotiation is to determine the price of the supplied intermediate goods and the delay penalty in case of a late delivery. For a manufacturer with a dominant contracting position, the outcome of the negotiation can be computed as a Stackelberg equilibrium point. Under the resulting contract, the two-stage supply chain reaches globally optimal running conditions with the maximal possible profit obtained by the manufacturer and the smallest acceptable profit obtained by the supplier.

Keywords: inventory control; manufacturing systems; stochastic modelling; queueing network models; Stackelberg Games

1. Introduction

Production organisation in the context of networked enterprises has raised many issues and has been the subject of many research works and international projects (such as the European Coordination Project CO-DESNET) related to strategic, tactical and operational problems. A production network has an intrinsically antagonistic nature, due to the contradiction between its global functional goal – to achieve the desired customer service level – and the individual economic objectives of the partners. Negotiation is an appealing way to overcome this contradiction while respecting the decentralised nature of the network. Its main advantages are the respect of the decisional autonomy of the actors and the simplicity of the information transfer process.

This study concentrates on the tactical issue of organising the product flow between a manufacturer and a supplier, as the manufacturer, producing and selling an end-product in the final market, procures a key component from his supplier. The purpose of negotiation is thus to establish a contract setting the rules of trade and delivery between the two partners. Both the supplier and manufacturer operate in a make-to-stock manner with a limited capacity production process and control the local output inventory

according to a base-stock policy. Here, each firm is modelled as a make-to-stock single-server queue in order to combine demand and order delivery lead time uncertainties and reflect the capacitated nature of the production processes. Because of random effects, all customer demands cannot be satisfied immediately. Moreover, the service level provided by the system is influenced by the inventory control decisions of both firms. Since each firm is an individual profit maximising entity, the inventory control decisions of the firms may deviate from the system optimal solution. Carrying inventory at the supplier level creates inventory holding costs for the supplier but assures an acceptable lead time for the manufacturer and consequently an acceptable delivery delay for the final customers. However, the supplier may prefer to install a lower base-stock level since he cares less about customer backorders. On the other hand, the manufacturer with a higher inventory holding cost may prefer the supplier to keep a higher base-stock level. The purpose of the study is thus to use the ability of Game Theory in anticipating decisions, to determine the contract parameters resulting from negotiation and evaluate the global performance of the system.

Game theoretical applications in supply chain management are reviewed by Cachon and Netessine

*Corresponding author. Email: yasemin.arda@ulg.ac.be

(2004) and Leng and Parlar (2005). Cachon (2003) reviews the literature on supply chain coordination with contracts. Cachon (1999) analyses different coordination schemes that can be used in serial supply chains with inventory competition. Cachon and Zipkin (1999) analyse a two-stage serial supply chain with stationary stochastic demand and fixed transportation times, i.e. a two-stage Clark and Scarf model (Clark and Scarf 1960; Chen and Zheng 1994). Each stage incurs inventory holding costs and a backorder cost is charged at both stages whenever a customer is backordered at the second stage. The stages independently choose base-stock policies to minimise their costs. The authors show that the supply chain optimal solution is never a Nash equilibrium, so competitive selection of inventory policies decreases efficiency. Under conditions of cooperation with simple linear transfer payments it is also claimed that global supply chain optimal solution can be achieved as a Nash equilibrium. For a similar system, Lee and Whang (1999) develop a nonlinear transfer payment scheme that induces each firm to choose the system optimal base-stock policies. The non-linear transfer payments proposed by Lee and Whang (1999) correspond to the payments used in the Clark and Scarf algorithm (1960) and compensate for all the expected costs faced by the retailer (second-stage firm) because of any component late delivery. Then, the second-stage base-stock level is independent of the first-stage base-stock level. Porteus (2000) uses responsibility tokens in a similar way. Whenever the supplier is unable to fill an order, he uses responsibility tokens that are equivalent to real inventory from the retailer's perspective. Then, the retailer receives perfectly reliable supply.

Game theoretic analysis of make-to-stock queueing systems has been mainly focused on interactions between a supplier and a retailer. Caldentey and Wein (2003) analyse a two-stage system where the supplier operates as a single-server queue and controls the production rate of his manufacturing facility. On the other hand, the retailer carries finished-goods inventory and specifies a base-stock policy for replenishing his inventory from the supplier. The authors characterise Nash and Stackelberg equilibriums and show that a linear transfer payment can coordinate the system. Plambeck and Zenios (2003) analyse a make-to-stock queueing system in a principal-agent framework. The supplier (agent) dynamically controls the production rate. Information asymmetry arises since the retailer (principal) cannot observe the agent's chosen production policy. The authors show that a dynamic contract constructed by the principal can coordinate the system

if the agent is risk neutral. Cachon (1999) studies competitive inventory management of a supplier/retailer system in which the supplier operates in a make-to-stock manner with a fixed mean production rate. Unmet final demands are supposed to be lost. Jemai and Karaesmen (2007) analyse a similar system where unmet demands are backordered and both firms share the related backorder costs. Gupta and Weerawat (2006) focus on the revenue-sharing contracts that a manufacturer may use to affect his supplier's inventory decisions. In their study, the supplier works in a make-to-stock but the manufacturer in a make-to-order manner.

This article studies a two-stage production/inventory system where each stage carries local finished-goods inventory and operates with a fixed mean production rate. In the game theory framework, the partners play a two-stage game of the Stackelberg type. The manufacturer leads the game by setting first the terms of trade, i.e. the manufacturer possesses certain power over the supplier as being the contractor. It is shown that the manufacturer can capture the maximal global profit of the supply chain using a performance-based pricing scheme. Additionally, he can select the optimal contract parameters with respect to his own base-stock level, and even jointly optimise his base-stock level and the contract parameters.

The basic assumptions of the two-stage production/inventory model are given in Section 2. Then, Section 3 focuses on the supplier's problem, which consists of optimising the base-stock level under given contract parameters. Section 4 analyses the manufacturer's optimisation problem in terms of the contract parameters and the base-stock level. The system optimal solution is derived in Section 5 while concluding remarks are given in Section 6.

2. Model description

The supply chain under investigation consists of a component supplier and a manufacturer who sells a single end-product to the consumer market. Without loss of generality (through introduction of a proportionality coefficient), it is assumed that the manufacturer needs one component provided by the supplier to fabricate one end-product. The exogenous market demand is modelled as a stationary stochastic process: the customers arrive at the manufacturer according to a Poisson process having rate λ and each customer purchases exactly one unit of end-product. Let i index the stages of the supply chain, $i=1,2$. Stage 1 represents the supplier's plant and stage 2 the manufacturer's. Each stage carries a finished-goods inventory to serve the demand and

$S_i = I_i - B_i + K_i$ in steady-state. Then, the inventory and backorder levels are given by the identities $I_i = (S_i - K_i)^+$ and $B_i = (K_i - S_i)^+$ where $(x)^+$ stands for $\max(x, 0)$. Notice that the work-in-process inventory is physically divided between the output buffer of the supplier and the processing queue of the manufacturer as shown in Figure 1.

Under a base-stock policy, a demand arrival at the last stage instantly triggers a demand at each of the previous stages. Since there is no delay in receiving raw materials, $K_1 = N_1$ at the supplier's. Clearly, the production facility of the supplier behaves as an M/M/1 queue with the traffic intensity ρ_1 . Analysis of the second-stage production process is more complicated. The number of uncompleted orders in the manufacturer's system depends on the backorder level of the supplier: $K_2 = B_1 + N_2$. Using the well-known Little's law (Kleinrock 1975), a similar dependency relation can be established between the sojourn time of uncompleted orders in the manufacturer's system and the delivery delay of the supplier. Let W_i be the sojourn time of units and L_i the sojourn time of uncompleted orders in the fabrication system i . L_i can also be interpreted as the lead time of the finished-goods inventory of stage i . Let D_i be the delay for demands placed upon stage i . The sojourn time of uncompleted orders at the manufacturer's is then defined as the sum of the component delivery delay provoked by the supplier and the sojourn time of components at the manufacturer's: $L_2 = D_1 + W_2$. Note that $L_1 = W_1$ since the supplier instantly receives a raw material at each demand arrival.

If $S_1 = 0$, the system operates like a tandem queue. Likewise, if $S_1 \rightarrow \infty$, the production facility of the manufacturer behaves as an M/M/1 queue with the traffic intensity ρ_2 . But in the case where $0 < S_1 < \infty$, the departure distribution of units from stage 1 is not a Poisson process. Moreover, the times between successive departures are serially correlated. In other words, the sojourn times of a unit and consequently the number of units at different stages are probabilistically dependent (see Buzacott, Price and Shathikumar (1992) and Lee and Zipkin (1992) for more details). Thus, one needs to solve a complex system of balance equations to determine the exact values of the limiting probabilities. In this article, the limiting probabilities of the second-stage production process are rather approximated using the method of Lee and Zipkin (1992) so as to derive closed form results. This approximation scheme is described in Section 4.

In the decentralised setting considered, each firm manages the local production/inventory control system while attempting to maximise his own steady-state expected profit per time unit. The main operational decision at each stage is the level of the inventory

control parameter, i.e. the base-stock level S_i . Each firm incurs a production cost $c_i \geq 0$ per unit produced, and a holding cost $h_i \geq 0$ per unit of inventory per time unit. Assume $h_2 > h_1$. This assumption can be justified since the unitary holding cost of a product is generally supposed to increase with the added value. Additionally, the purchasing price per unit of end-product is p_2 ($p_2 \geq c_1 + c_2$) while the purchasing price of one unit of raw material is assumed to be included in the unit production cost c_1 . Backorder costs are incurred only at the manufacturer's for customer delays: a customer backorder generates a cost $b_2 \geq 0$ per time unit.

In the case of a centralised system, a centralised planner fixes the optimal base-stock levels of the firms which maximise the steady-state expected profit rate of the centralised system:

$$\pi_0(S_1, S_2) = \lambda(p_2 - c_1 - c_2) - h_1 E[I_1] - h_2 E[I_2] - b_2 E[B_2] \quad (1)$$

Note that the holding costs of the units in the production facility of the manufacturer are excluded since the related costs, namely $h_1 E[N_2]$, are constant using the approximation method of Lee and Zipkin (1992).

As explained above, the backorder level of the manufacturer not only depends on the base-stock level S_2 and the traffic intensity ρ_2 but also depends on the backorder level of the supplier. In fact, the second-stage production process may fall idle even when $P_2 > 0$ because of component starving. This occurs when the second server completes the unit in process and yet there is no component available in the supply system of the manufacturer, i.e. when the total number of components in the system $I_1 + N_2$ becomes zero. A centralised planner takes into account the backorder level of the manufacturer while determining the system optimal first-stage base-stock level. However, in the decentralised setting, the supplier lacks the incentive to implement this action. To create supplier incentives, the general approach is to split the customer backorder costs among the firms, with a fraction $x \in [0, 1]$ which is assumed to be exogenous (Cachon and Zipkin 1999; Caldentey and Wein 2003). These backorder costs may be interpreted as goodwill penalties which do not affect the firms equally. In this article, the aim is to define a coordination scheme based on local backorder costs representing late-delivery penalties: the manufacturer incurs backorder costs at rate b_2 and charges a penalty b_1 to the supplier per unit of component backorder and per time unit. Besides, the manufacturer offers to the supplier a purchasing price p_1 per component unit. Thus, the considered contract is a two-part linear contract (p_1, b_1) whose parameters

are the additional decision variables of the manufacturer.

In the Stackelberg setting, the manufacturer (the Stackelberg leader) acts first and proposes a two-part linear contract (p_1, b_1) to his supplier. The supplier accepts only a contract that yields his reservation profit, which reflects the outside opportunity of the firm. Here, it is assumed that the supplier is in a totally competitive market and thus his reservation profit is zero. In the case of contract acceptance, the supplier decides the base-stock level S_1 that he will install. Besides the external demand information, all costs, parameters and rules are also common knowledge.

The contract (p_1, b_1) defines a transfer payment from the manufacturer to the supplier. The expected value of this transfer payment per time unit is given by $E[T(S_1, p_1, b_1)] = \lambda p_1 - b_1 E[B_1]$. The steady state expected profit rates of the supplier $\pi_1(S_1, p_1, b_1)$ and the manufacturer $\pi_2(S_1, S_2, p_1, b_1)$ are then

$$\pi_1 = \lambda(p_1 - c_1) - h_1 E[I_1] - b_1 E[B_1], \quad (2)$$

$$\pi_2 = \lambda(p_2 - p_1 - c_2) + b_1 E[B_1] - h_2 E[I_2] - b_2 E[B_2]. \quad (3)$$

The defined Stackelberg setting is similar to a principle-agent framework without hidden action (unobservable action which defines a moral hazard situation) (Plambeck and Zenios 2003) or hidden information (*ex ante* information asymmetry also called as adverse selection) (Corbett and de Groote 2000; Corbett, Zhou and Tang 2004). The manufacturer acts as the principal and the supplier as the agent. The manufacturer's optimisation problem is

$$\Pi_1 : \max_{p_1, b_1, S_2} \pi_2(S_1^*, S_2, p_1, b_1) \quad (4)$$

$$\text{s.t. } S_1^* = \arg \max_{S_1} \pi_1(S_1, p_1, b_1) \quad (5)$$

$$\pi_1(S_1^*, p_1, b_1) \geq 0. \quad (6)$$

The incentive compatibility constraint (5) states that, given a contract (p_1, b_1) , the supplier chooses a base-stock level that maximises his profits. In addition, the individual rationality constraint (6) bounds the supplier's optimal profit level from below since the supplier accepts only a contract that yields a non-negative profit. Since there is no information asymmetry, the manufacturer correctly anticipates the optimal action of the supplier given a contract (p_1, b_1) , i.e. the optimal first stage base-stock level which maximises $\pi_1(S_1, p_1, b_1)$. The manufacturer then takes into account the optimal action of the supplier in order to determine an acceptable contract (p_1, b_1) .

Let $P_{k_i} = \Pr\{K_i = k_i\}$. The expected inventory and backorder levels of the firms are simply

$$E[I_i] = \sum_{k_i=0}^{S_i} (S_i - k_i)P_{k_i}, \quad E[B_i] = \sum_{k_i=S_i}^{\infty} (k_i - S_i)P_{k_i}. \quad (7)$$

Since K_1 is identical to the number of items in an M/M/1 queue with the traffic intensity ρ_1 , one has $P_{k_1} = \rho_1^{k_1}(1 - \rho_1)$. Then,

$$E[B_1] = \frac{\rho_1^{S_1+1}}{1 - \rho_1}, \quad E[I_1] = S_1 - \frac{\rho_1(1 - \rho_1^{S_1})}{1 - \rho_1}. \quad (8)$$

Notice that W_1 is identical to the sojourn time of the M/M/1 queue with the traffic intensity ρ_1 and exponentially distributed with rate $(\mu_1 - \lambda)$, i.e. $E[W_1] = 1/(\mu_1 - \lambda)$. The relation between D_1 and W_1 is given by $E[D_1] = \Pr\{I_1=0\}E[W_1]$. The stock-out probability of the supplier is simply $\Pr\{I_1 = 0\} = \Pr\{K_1 \geq S_1\} = \rho_1^{S_1}$. Clearly, $E[D_1] = E[B_1]/\lambda$ and the expected rate of the transfer payment is rewritten as $E[T(S_1, p_1, b_1)] = E[\lambda(p_1 - b_1 D_1)]$. As a result, two possible techniques can be used to implement the contract (p_1, b_1) : either the price is constant but there is a penalty that depends on the amount of backorders, or the price varies as a function of the observed lateness of delivery.

In the rest of this article, S_i is treated as a continuous non-negative variable. This assumption ignores the restriction of S_i to the set of non-negative integers which are more relevant for industrial applications. On the other hand, it provides closed form results that can be used to approximate the optimal integer values.

3. Supplier's problem

The profit function of the supplier $\pi_1(S_1) = \pi_1(S_1, p_1, b_1)$ is given by

$$\pi_1(S_1) = \lambda(p_1 - c_1) - h_1 \left(S_1 - \frac{\rho_1}{1 - \rho_1} \right) - (h_1 + b_1) \frac{\rho_1^{S_1+1}}{1 - \rho_1}.$$

Let $\pi'_1(S_1)$ be the first derivative, and b_1^{\min} the value of the backorder penalty b_1 which solves $\pi'_1(0) = 0$:

$$b_1^{\min} = -h_1 \left(1 + \frac{1 - \rho_1}{\rho_1 \ln \rho_1} \right) \quad (9)$$

Profit function $\pi_1(S_1)$ is concave in S_1 . On the other hand, it is decreasing in S_1 if $b_1 < b_1^{\min}$. Then, the following proposition states the optimal action of the supplier as a response to the proposed contract.

Proposition 1: Given the contract (p_1, b_1) , the optimal base-stock level that maximises $\pi_1(S_1)$ is

$$S_1^*(b_1) = \begin{cases} \ln(\alpha(b_1)) / \ln \rho_1 & \text{if } b_1 > b_1^{\min} \\ 0 & \text{if } b_1 \leq b_1^{\min} \end{cases} \quad (10)$$

where

$$\alpha(b_1) = \frac{-h_1(1 - \rho_1)}{(h_1 + b_1)\rho_1 \ln \rho_1}. \quad (11)$$

The penalty cost b_1 is the only contract parameter that affects the optimal action of the supplier, which is increasing in b_1 for $b_1 > b_1^{\min}$. Note that $b_1^{\min} > 0$ since $\varphi_1(x) = (1-x)/x \ln x$ is increasing for $x \in (0, 1)$ with $\lim_{x \rightarrow 1^-} \varphi_1(x) = -1$. The non-negativity constraint $S_1^*(b_1) \geq 0$ is naturally satisfied since $\alpha(b_1^{\min}) = 1$.

4. Manufacturer's problem

The two basic analytical approaches for performance evaluation of multi-stage make-to-stock queueing systems are the approximation methods of Buzacott et al. (1992) (called the BPS¹ method) and Lee and Zipkin (1992) (called the LZ method). The LZ and BPS¹ methods assume that the production facility at each stage behaves as an M/M/1 queue. Duri, Frein and Di Mascolo (2000) show the equivalence between these two methods and propose some extensions of the LZ method. Lee and Zipkin (1995) and Wang and Su (2007) extend the LZ method for complex base-stock systems. Buzacott et al. (1992) also propose an alternative approximation scheme (called the BPS² method) for a two-stage system. The BPS² method treats the second stage as a GI/M/1 queue (with independent inter-arrival times). Note that this is still an approximation, since the time intervals between successive departures from stage 1 are serially correlated and consequently the second stage cannot be modelled as a GI/M/1 queueing system. Gupta and Selvaraju (2006) propose an improved version of the BPS² method (referred to as the GS method) and extend the GS method for a multi-stage system. The numerical results of Gupta and Selvaraju (2006) for a two-stage system show that the mean absolute errors obtained by the LZ (or BPS¹), BPS² and GS methods are, respectively, 1.97%, 1.4%, 0.71% for $E[K_2]$ and 4.93%, 4.65%, 2.79% for $E[B_2]$. While the BPS² and GS methods may underestimate the true expected values, the LZ (or BPS¹) method always provides an upper bound for $E[K_2]$ and $E[B_2]$. Liu, Liu and Yao (2004) analyse a multi-stage system with general service time and demand inter-arrival time distributions and propose an approximation method which treats each stage as a GI/GI/1 queue. They show that their method

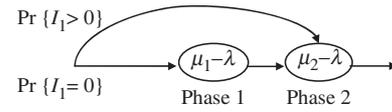


Figure 2. Phase-type approximation.

is equivalent to the LZ method for systems with exponential servers and a Poisson demand process since their method treats such a system with Poisson departure processes.

In the following, the LZ method is used to approximate the limiting probabilities of the second-stage production process for its accuracy and analytical tractability. The method is stated for the considered two-stage system using closed-form expressions of the probability distributions instead of matrix-exponential representations. Additionally, the analysis is restricted to the case of different service rates, $\mu_1 \neq \mu_2$, in order to simplify the presentation.

Lee and Zipkin (1992) approximate the arrival distribution of units into the manufacturing stage 2 as a Poisson process. Then, the sojourn time W_2 is identical to the sojourn time of an M/M/1 queue with the traffic intensity ρ_2 and exponentially distributed with rate $(\mu_2 - \lambda)$. Given the LZ approximation, the second-stage lead time $L_2 = D_1 + W_2$ has the continuous phase-type distribution given in Figure 2. The probability density function of L_2 is $f_{L_2}(t) = \rho_1^{S_1} f_{W_1+W_2}(t) + (1 - \rho_1^{S_1}) f_{W_2}(t)$. To calculate the probability P_{k_2} , the authors use the following known property: the number of outstanding orders in the system has the same distribution as the number of demands in a random time L_i . Then,

$$P_{k_2} = \left(\rho_1^{S_1+1} \frac{1 - \rho_2}{\rho_1 - \rho_2} \right) \rho_1^{k_2} (1 - \rho_1) + \left(1 - \rho_1^{S_1+1} \frac{1 - \rho_2}{\rho_1 - \rho_2} \right) \rho_2^{k_2} (1 - \rho_2).$$

Once P_{k_2} is obtained, the second-stage expected inventory and backorder levels can be calculated using Equation (7). Note that the method is exact when $S_1 = 0$.

4.1. Optimal contract parameters

Given the results of Lee and Zipkin (1992), the profit function of the manufacturer is as follows:

$$\begin{aligned} \pi_2(S_1, S_2, p_1, b_1) &= \lambda(p_2 - p_1 - c_2) - h_2 \left(S_2 - \frac{\rho_2}{1 - \rho_2} - \frac{\rho_1^{S_1+1}}{1 - \rho_1} \right) \end{aligned}$$

$$\begin{aligned}
 &+ b_1 \frac{\rho_1^{S_1+1}}{1-\rho_1} - (h_2 + b_2) \frac{\rho_2^{S_2+1}}{1-\rho_2} \\
 &- (h_2 + b_2) \left(\frac{\rho_1^{S_1+1}(1-\rho_2)}{\rho_1 - \rho_2} \right) \left(\frac{\rho_1^{S_2+1}}{1-\rho_1} - \frac{\rho_2^{S_2+1}}{1-\rho_2} \right) \\
 &- (h_2 + b_2) \frac{\rho_2^{S_2+1}}{1-\rho_2} - (h_2 + b_2) \left(\frac{\rho_1^{S_1^{*(b_1)+1}}(1-\rho_2)}{\rho_1 - \rho_2} \right) \\
 &\times \left(\frac{\rho_1^{S_2+1}}{1-\rho_1} - \frac{\rho_2^{S_2+1}}{1-\rho_2} \right).
 \end{aligned}$$

The manufacturer correctly anticipates the optimal base-stock level $S_1^*(b_1)$ which maximises $\pi_1(S_1, p_1, b_1)$. Consequently, the optimisation problem Π_1 given by Equations (4)–(6) becomes

$$\Pi_2 : \quad \max_{p_1, b_1, S_2} \pi_2(S_1^*(b_1), S_2, p_1, b_1) \quad (12)$$

$$\text{s.t. } \pi_1(S_1^*(b_1), p_1, b_1) \geq 0 \quad (13)$$

The Lagrangean of problem Π_2 is written as

$$L = \pi_2(S_1^*(b_1), S_2, p_1, b_1) + u\pi_1(S_1^*(b_1), p_1, b_1)$$

where u is the Lagrange multiplier associated with constraint (13). Setting $\partial L/\partial p_1 = -\lambda(1-u) = 0$, one obtains $u^* = 1$. According to KKT conditions, constraint (13) is then binding at the optimum and the value of the purchasing price maximising $\pi_2(S_1^*(b_1), S_2, p_1, b_1)$ for fixed b_1 and S_2 is $p_1^*(b_1)$ which satisfies $\pi_1(S_1^*(b_1), p_1^*(b_1), b_1) = 0$:

$$p_1^*(b_1) = c_1 + \frac{h_1}{\lambda} \left(S_1^*(b_1) - \frac{\rho_1}{1-\rho_1} \right) + \frac{h_1 + b_1}{\lambda} \left(\frac{\rho_1^{S_1^*(b_1)+1}}{1-\rho_1} \right)$$

The condition $\pi_1(S_1, p_1, b_1) = 0$ implies $E[T(S_1, p_1, b_1)] = \lambda p_1 - b_1 E[B_1] = \lambda c_1 + h_1 E[I_1]$ and consequently

$$\pi_2 = \lambda(p_1 - c_1 - c_2) - h_1 E[I_1] - h_2 E[I_2] - b_2 E[B_2]. \quad (14)$$

This expression is identical to the one given in Equation (1). That is, the contract (p_1, b_1) transfers all the operating costs and benefits of the supplier to the manufacturer letting the supplier with zero profit. The manufacturer captures the steady-state expected profit rate of the centralised system. It is shown formally in Section 5 that this full transfer forces the manufacturer to choose the system optimal first-stage base-stock level.

Let $\pi_2(S_2, b_1)$ denote $\pi_2(S_1^*(b_1), S_2, p_1^*(b_1), b_1)$, for simplicity:

$$\begin{aligned}
 &\pi_2(S_2, b_1) \\
 &= \lambda(p_2 - c_1 - c_2) - h_1 \left(S_1^*(b_1) - \frac{\rho_1}{1-\rho_1} + \frac{\rho_1^{S_1^*(b_1)+1}}{1-\rho_1} \right) \\
 &- h_2 \left(S_2 - \frac{\rho_2}{1-\rho_2} - \frac{\rho_1^{S_1^*(b_1)+1}}{1-\rho_1} \right)
 \end{aligned}$$

The first derivative of $\pi_2(S_2, b_1)$ with respect to b_1 for $b_1 > b_1^{\min}$ can be obtained as

$$\frac{\partial \pi_2(S_2, b_1)}{\partial b_1} = \frac{h_1(b_1 + h_2 - (h_2 + b_2)\tau_0(S_2))}{(h_1 + b_1)^2 \ln \rho_1}, \quad (15)$$

where

$$\tau_0(S_2) = \frac{(1-\rho_1)(1-\rho_2)}{\rho_1 - \rho_2} \left(\frac{\rho_1^{S_2+1}}{1-\rho_1} - \frac{\rho_2^{S_2+1}}{1-\rho_2} \right).$$

Let $\varphi_2(x) = x^{S_2+1} \ln x / (1-x)$. One obtains that $\varphi_2(x)$ is decreasing for $x \in (0, 1)$ when $S_2 \geq 0$. Consequently, $\tau_0'(S_2) < 0$ and $\tau_0(S_2)$ is decreasing for $S_2 \geq 0$. Note also that $\tau_0(S_2) : [0, \infty) \rightarrow (0, 1]$. Then, if $b_2 \geq b_1^{\min}$, there exists a value S_2^{\max} which satisfies the equality $\tau_0(S_2) = (h_2 + b_1^{\min}) / (h_2 + b_2)$. The following proposition characterises the optimal value of the backorder penalty b_1 for fixed S_2 using the above observations.

Proposition 2: Given S_2 , profit function $\pi_2(S_2, b_1)$ is strictly quasi-concave for $b_1 > b_1^{\min}$ and constant for $b_1 \leq b_1^{\min}$. Then, the optimal value of the backorder penalty which maximises $\pi_2(S_2, b_1)$ is

$$b_1^*(S_2) = \begin{cases} b_1^{\text{opt}}(S_2) & \text{if } \tau_0(S_2) > (h_2 + b_1^{\min}) / (h_2 + b_2) \\ b_1^{\min} & \text{if } \tau_0(S_2) \leq (h_2 + b_1^{\min}) / (h_2 + b_2) \end{cases}$$

where $b_1^{\text{opt}}(S_2) = (h_2 + b_2)\tau_0(S_2) - h_2$.

Proof: If $\tau_0(S_2) > (h_2 + b_1^{\min}) / (h_2 + b_2)$, $b_1^{\text{opt}} > b_1^{\min}$. Then, Equation (15) implies that $\pi_2(S_2, b_1)$ is strictly increasing for $b_1^{\min} < b_1 < b_1^{\text{opt}}$ and strictly decreasing for $b_1 > b_1^{\text{opt}}$. If $\tau_0(S_2) \leq (h_2 + b_1^{\min}) / (h_2 + b_2)$, Equation (15) implies that $\pi_2(S_2, b_1)$ is strictly decreasing for $b_1 > b_1^{\min}$. \square

Receiving the contract $(p_1^*(b_1^*), b_1^*)$, the supplier installs $S_1^*(b_1^*) > 0$ if $b_1^* = b_1^{\text{opt}}$ and $S_1^*(b_1^*) = 0$ if $b_1^* = b_1^{\min}$. Clearly, if $b_2 \geq b_1^{\min}$, then $b_1^{\min} \leq b_1^* \leq b_2$, i.e. the manufacturer never proposes a backorder penalty greater than b_2 .

4.2. Optimal base-stock level

It can be observed that the backorder penalty $b_1^{\text{opt}}(S_2)$ fixed by the manufacturer if $b_2 \geq b_1^{\min}$ and $S_2 < S_2^{\max}$ is a decreasing function of S_2 . On the other hand, the manufacturer proposes b_1^{\min} either if $b_2 \geq b_1^{\min}$ and

$S_2 \geq S_2^{\max}$, or if $b_2 < b_1^{\min}$. Since the manufacturer covers the inventory holding costs of the supplier, he prefers to offer a lower backorder penalty leading to a lower first-stage base-stock level as S_2 gets higher. When the base-stock level S_2 is sufficiently high or the backorder cost b_2 is sufficiently low, the manufacturer offers the minimum value of the backorder penalty allowing the supplier to install a zero base-stock level. The question then is how the base-stock levels of the firms must be balanced. Let $\alpha(S_2) = \alpha(b_1^{\text{opt}}(S_2))$ and $\pi_2(S_2) = \pi_2(S_2, b_1^*(S_2))$ with the first derivative

$$\pi_2'(S_2) = \begin{cases} -h_2 - (h_2 + b_2)\tau_1(S_2) & \text{for } \tau_0(S_2) > (h_2 + b_1^{\min})/(h_2 + b_2) \\ -h_2 - (h_2 + b_2)\tau_2(S_2) & \text{for } \tau_0(S_2) \leq (h_2 + b_1^{\min})/(h_2 + b_2) \end{cases}$$

where

$$\tau_1(S_2) = \left(\frac{\alpha(S_2)\rho_1(1-\rho_2)}{\rho_1-\rho_2} \right) \frac{\rho_1^{S_2+1} \ln \rho_1}{1-\rho_1} + \left(1 - \frac{\alpha(S_2)\rho_1(1-\rho_2)}{\rho_1-\rho_2} \right) \frac{\rho_2^{S_2+1} \ln \rho_2}{1-\rho_2}, \quad (16)$$

$$\tau_2(S_2) = \left(\frac{\rho_1(1-\rho_2)}{\rho_1-\rho_2} \right) \frac{\rho_1^{S_2+1} \ln \rho_1}{1-\rho_1} + \left(1 - \frac{\rho_1(1-\rho_2)}{\rho_1-\rho_2} \right) \frac{\rho_2^{S_2+1} \ln \rho_2}{1-\rho_2}. \quad (17)$$

Lemma 1: $\tau_2(S_2)$ is increasing for $S_2 \geq 0$.

Proof: $\varphi_3(x) = x \ln x / (1-x)$ is decreasing for $x \in (0, 1)$. It follows that, $\{[\rho_j(1-\rho_i) \ln \rho_j] / [\rho_i(1-\rho_j) \ln \rho_i]\}^2 < 1$ for $\rho_i > \rho_j$. As a result, for $S_2 \geq 0$,

$$\tau_2'(S_2) = \frac{(1-\rho_2)\rho_1^2 \rho_2^{S_2} (\ln \rho_1)^2}{(\rho_1-\rho_2)(1-\rho_1)} \times \left(\left(\frac{\rho_1}{\rho_2} \right)^{S_2} - \left(\frac{\rho_2(1-\rho_1) \ln \rho_2}{\rho_1(1-\rho_2) \ln \rho_1} \right)^2 \right) > 0. \quad \square$$

Lemma 1 shows that when $b_2 < b_1^{\min}$, the profit function $\pi_2(S_2)$ is concave for $S_2 \geq 0$. Then, the optimal value of the base-stock level which maximises $\pi_2(S_2)$ is $S_2^* = S_2^{\text{opt}_2}$, which satisfies $\tau_2(S_2^{\text{opt}_2}) = -h_2/(h_2 + b_2)$. Increasing nature of $\tau_1(S_2)$, defined for $S_2 \in [0, S_2^{\max}]$, is not guaranteed in general. However, a large number of numerical examples have shown that $\pi_2(S_2)$ is a strictly quasi-concave function when $b_2 \geq b_1^{\min}$. Note that, for $S_2 < S_2^{\max}$, $\alpha(S_2) < 1$ and so $\tau_1(S_2) > \tau_2(S_2)$ with $\lim_{S_2 \rightarrow S_2^{\max}} \tau_1(S_2) = \tau_2(S_2^{\max})$. Using these observations, the following proposition can be stated.

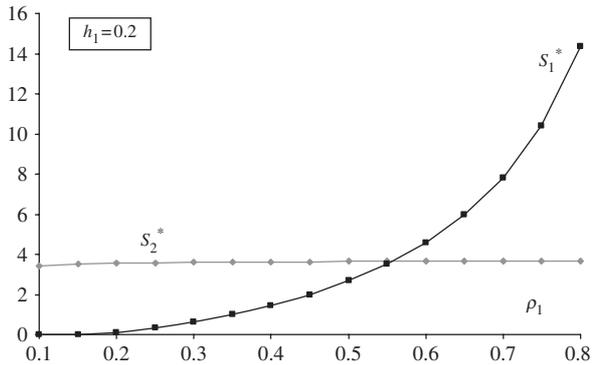
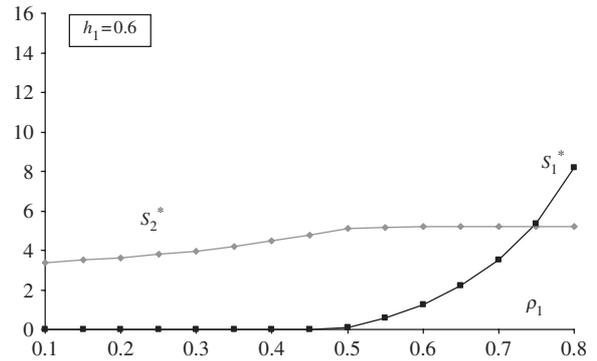


Figure 3. Optimal base-stock levels S_2^* and $S_1^*(b_1^*(S_2^*))$ for $h_2=0.8, b_2=10, \rho_2=0.5$.

Proposition 3: When $b_2 \geq b_1^{\min}$, the profit function $\pi_2(S_2)$ is strictly quasi-concave for $S_2 \geq 0$ under the condition that $\tau_1(S_2) = -h_2/(h_2 + b_2)$ has at most one solution when $\tau_2(S_2^{\max}) > -h_2/(h_2 + b_2)$, and $\tau_1(S_2)$ is increasing in S_2 when $\tau_2(S_2^{\max}) \leq -h_2/(h_2 + b_2)$. Then, the optimal value of the base-stock level which maximises $\pi_2(S_2)$ is given by

$$S_2^* = \begin{cases} S_2^{\text{opt}_1} & \text{if } \tau_2(S_2^{\max}) > -h_2/(h_2 + b_2) \\ S_2^{\text{opt}_2} & \text{if } \tau_2(S_2^{\max}) \leq -h_2/(h_2 + b_2) \end{cases}$$

where $\tau_1(S_2^{\text{opt}_1}) = [-h_2/(h_2 + b_2)]$ and $\tau_2(S_2^{\text{opt}_2}) = [-h_2/(h_2 + b_2)]$.

Evolution of optimal base-stock levels with ρ_1 and h_1 is given in Figure 3. In each problem, $b_2 > b_1^{\min}$ and the conditions of quasi-concavity given in Proposition 3 are satisfied. For $h_1=0.6, S_1^* > 0$ when $\mu_1 < \mu_2$. However, for $h_1=0.2$, installing a positive base-stock level at the supplier's becomes profitable even when $\mu_1 > \mu_2$.

5. System optimal solution

The steady-state expected profit rate of the centralised system is simply the sum of the expected profit rates of

the partners given by Equation (1). The first derivative of $\pi_0(S_1, S_2)$ with respect to S_1 is:

$$\frac{\partial \pi_0(S_1, S_2)}{\partial S_1} = -h_1 - (h_1 - h_2 + (h_2 + b_2)\tau_0(S_2)) \frac{\rho_1^{S_1+1} \ln \rho_1}{1 - \rho_1}.$$

Given S_2 , $\pi_0(S_1, S_2)$ is concave in S_1 if $\tau_0(S_2) > (h_2 + b_1^{\min})/(h_2 + b_2)$, and decreasing for $S_1 > 0$ otherwise. Then, the optimal value of the first-stage base-stock level maximising the system total profit is

$$S_1^0(S_2) = \begin{cases} \ln(\alpha(S_2))/\ln \rho_1 & \text{if } \tau_0(S_2) > (h_2 + b_1^{\min})/(h_2 + b_2), \\ 0 & \text{if } \tau_0(S_2) \leq (h_2 + b_1^{\min})/(h_2 + b_2). \end{cases}$$

It can be observed that $S_1^0(S_2) = S_1^*(b_1^*(S_2))$. Consequently, $S_2^0 = S_2^*$. Thus, using the defined contracting scheme, the system optimal solution is obtained as the Stackelberg equilibrium.

Since the contract $(p_1^*(b_1^*), b_1^*)$ transfers all the operating costs and benefits of the supplier to the manufacturer, the manufacturer chooses the optimal contract parameters so as to maximise the integrated system profit. Offering $(p_1^*(b_1^*), b_1^*)$ the manufacturer obtains the integrated system profit, $\pi_2^*(S_1^*, S_2^*, p_1^*, b_1^*) = \pi_0^*(S_1^0, S_2^0)$, and lets the supplier with his reservation profit, $\pi_1^*(S_1^*, p_1^*, b_1^*) = 0$. The contract $(p_1^*(b_1^*), b_1^*)$ creates a positive profit for the manufacturer only if $\pi_0^*(S_1^0, S_2^0) > 0$.

For the principal-agent problem, it is known that, under full information and if the agent (supplier) is risk-neutral (i.e. expected profit maximiser), the principal (manufacturer) can design a contract that achieves the centralised solution while leaving the supplier only with his reservation profit (Corbett and de Groote 2000; Plambeck and Zenios 2003; Corbett et al. 2004). Here, the manufacturer uses the backorder penalty b_1 in order to force the supplier to install the system optimal first stage base-stock level. Recall that the corresponding transfer payment is $E[T(S_1, p_1, b_1)] = \lambda(p_1 - b_1 E[D_1]) = \lambda c_1 + h_1 E[I_1]$. Then, the purchasing price p_1 compensates all the operational costs of the supplier for any b_1 : $p_1 = c_1 + b_1 E[D_1] + h_1 E[I_1]/\lambda$. The proposed contract can also be interpreted as a revenue sharing contract (Cachon and Lariviere 2005). For each item sold, the manufacturer pays a portion of the sales revenue to the supplier. The supplier's share is then $p_1 - b_1 E[D_1]$.

These results can also be generalised to the case where the supplier's reservation profit is greater than zero. Let $\pi_1^{\min} > 0$ be the supplier's reservation profit level. Then, the individual rationality constraint (6) becomes $\pi_1(S_1^*, p_1, b_1) \geq \pi_1^{\min}$. However, a positive reservation profit level imposed by the supplier decreases the negotiation power of the manufacturer

and can even create negative profits for the manufacturer. In order to balance the negotiation powers of the firms, consider that the manufacturer also has a reservation level $\pi_2^{\min} > 0$ which reflects the profit that he could earn by contracting another supplier. If $\pi_1^{\min} + \pi_2^{\min} \leq \pi_0^*(S_1^0, S_2^0)$, then it can be shown that there exists a Stackelberg equilibrium in which the manufacturer proposes a contract $(p_1^{**}(b_1^*), b_1^*)$. Here, $p_1^{**}(b_1^*)$ is the value which satisfies $\pi_1(S_1^*(b_1^*), p_1^{**}(b_1^*), b_1^*) = \pi_1^{\min}$: $p_1^{**}(b_1^*) = p_1^*(b_1^*) + \pi_1^{\min}/\lambda$. The corresponding transfer payment can be written as $E[T(S_1, p_1, b_1)] = \lambda c_1 + h_1 E[I_1] + \pi_1^{\min}$. Receiving the contract $(p_1^{**}(b_1^*), b_1^*)$, the supplier installs $S_1^*(b_1^*)$. The manufacturer's optimal base-stock level is S_2^* . Then, offering the contract $(p_1^{**}(b_1^*), b_1^*)$, the manufacturer obtains $\pi_2^*(S_1^*, S_2^*, p_1^{**}, b_1^*) = \pi_0^*(S_1^0, S_2^0) - \pi_1^{\min}$ while letting the supplier with his reservation profit, $\pi_1^*(S_1^*, p_1^{**}, b_1^*) = \pi_1^{\min}$. Note that the manufacturer refuses to trade with the supplier if $\pi_0^*(S_1^0, S_2^0) - \pi_1^{\min} < \pi_2^{\min}$.

6. Conclusions

In the type of supply chains considered, the manufacturer can anticipate the inventory capacity investment of his supplier through the contract parameters that he offers. The optimal value of these contract parameters can then be determined by the manufacturer jointly with his own optimal inventory capacity investment. The considered contract can be seen as a consignment vendor managed inventory contract, since the supplier owns the first-stage stock keeping responsibility for it. However, it is the manufacturer who influences the first-stage inventory capacity investment in choosing the backorder penalty. Under such a contract, the whole system is globally optimised with respect to the total inventory capacity and the global performance is achieved.

The proposed contract is not restricted to systems with exponential make-to-stock queues. For any base-stock system, the base-stock level is increasing in the backorder cost. That is, the manufacturer can always propose an appropriate backorder penalty in order to force the supplier to install the system optimal first-stage base-stock level. Moreover, through revenue sharing, the operational costs of the supplier can always be compensated. The manufacturer then obtains the integrated system profit reduced by the supplier's reservation profit and lets the supplier with his reservation profit. Using a similar approach, it would be interesting to analyse the optimal contract parameters and the optimal base-stock levels of the firms for the case of non-exponential processing times.

References

- Arda, Y., and Hennet, J.C. (2006), "Inventory Control in a Multi-supplier System," *International Journal of Production Economics*, 104(2), 249–259.
- Buzacott, J.A. Price S.M. and Shanthikumar, J.G. (1992), "Service Level in Multistage MRP and Base Stock Controlled Production Systems," in *New Directions for Operations Research in a Manufacturing System*, ed. G. Fandel, T. Gullledge, A. Jones, Springer, Berlin, pp. 445–463.
- Cachon, G.P. (1999), "Competitive and Cooperative Inventory Management in a Two-echelon Supply Chain with Lost Sales," Working paper, University of Pennsylvania.
- Cachon, G.P. (1999), "Competitive supply chain inventory management," in *Quantitative Models for Supply Chain Management*, ed. S. Tayur, R. Ganeshan, & M. Magazine, Boston: Kluwer, pp. 111–146.
- Cachon, G.P. (2003), "Supply Chain Coordination with Contracts," in *Supply Chain Management: Design, Coordination and Operation*, ed. A.G. de Kok, & S.C. Graves, Amsterdam: Elsevier, pp. 229–340.
- Cachon, G.P., and Lariviere, M.A. (2005), "Supply Chain Coordination with Revenue-sharing Contracts: Strengths and Limitations," *Management Science*, 51(1), 30–44.
- Cachon, G.P., and Netessine, S. (2004), "Game Theory in Supply Chain Analysis," in *Handbook of Quantitative Supply Chain Analysis: Modeling in the eBusiness Era*, ed. D. Smichi-Levi, S.D. Wu, & Z. Shen, Boston: Kluwer, pp. 13–66.
- Cachon, G.P., and Zipkin, P.H. (1999), "Competitive and Cooperative Inventory Policies in a Two-stage Supply Chain," *Management Science*, 45(7), 936–953.
- Caldentey, R., and Wein, L.M. (2003), "Analysis of a Decentralized Production-inventory System," *Manufacturing and Service Operations*, 5(1), 1–17.
- Chen, F., and Zheng, Y.S. (1994), "Lower Bounds for Multi-echelon Stochastic Inventory Systems," *Management Science*, 40(11), 1426–1443.
- Clark A.J. and Scarf, H. (1960), "Optimal Policies for a Multi-echelon Inventory Problem," *Management Science*, Vol. 6(4) 475–490, [50(12), 1782–1790] (2004).
- Corbett, C.J., and de Groot, X. (2000), "A Supplier's Optimal Discount Policy Under Asymmetric Information," *Management Science*, 46(3), 444–450.
- Corbett, C.J., Zhou, D., and Tang, C.S. (2004), "Designing Supply Contracts: Contract Type and Information Asymmetry," *Management Science*, 50(4), 550–559.
- Duri, C., Frein, Y., and Di Mascolo, M. (2000), "Performance Evaluation and Design of Base Stock Systems," *European Journal of Operational Research*, 127, 172–188.
- Frein, Y., Di Mascolo, M., and Dallery, Y. (1995), "On the Design of Generalized Kanban Control Systems," *International Journal of Operations & Production Management*, 15(9), 158–184.
- Gupta, D., and Weerawat, W. (2006), "Supplier-manufacturer Coordination in Capacitated Two-stage Supply Chains," *European Journal of Operational Research*, 175, 67–89.
- Gupta, D., and Selvaraju, N. (2006), "Performance Evaluation and Stock Allocation in Capacitated Serial Supply Systems," *Manufacturing & Service Operations Management*, 8(2), 169–191.
- Jemai, Z., and Karaesmen, F. (2007), "Decentralized Inventory Control in a Two-stage Capacitated Supply Chain," *IIE Transactions*, 39, 501–512.
- Karaesmen, F., and Dallery, Y. (2000), "A Performance Comparison of Pull Type Control Mechanisms for Multistage Manufacturing," *International Journal of Production Economics*, 68, 59–71.
- Kleinrock, L. (1975), *Queueing systems Volume 1: Theory*, John Wiley & Sons, New York.
- Lee, H., and Whang, S. (1999), "Decentralized Multi-echelon Supply Chains: Incentives and Information," *Management Science*, 45(5), 633–640.
- Lee, Y.J., and Zipkin, P. (1992), "Tandem Queues with Planned Inventories," *Operations Research*, 40(5), 936–947.
- Lee, Y.J., and Zipkin, P. (1995), "Processing Networks with Inventories: Sequential Refinement Systems," *Operations Research*, 43(6), 1025–1036.
- Leng, M., and Parlar, M. (2005), "Game theoretic applications in supply chain management: a review," *Information Systems and Operations Research*, 43(3), 187–230.
- Liu, L., Liu, X., and Yao, D. (2004), "Analysis and Optimization of a Multistage Inventory-queue System," *Management Science*, 50(3), 365–380.
- Plambeck, E., and Zenios, S. (2003), "Incentive Efficient Control of a Make-to-stock Production System," *Operations Research*, 51(3), 371–386.
- Porteus, E.L. (2000), "Responsibility Tokens in Supply Chain Management," *Manufacturing and Service Operations Management*, 2(2), 203–219.
- Veatch, M.H., and Wein, L.M. (1994), "Optimal Control of a Two-station Tandem Production/Inventory system," *Operations Research*, 42(2), 337–350.
- Wang, F.F., and Su, C.T. (2007), "Performance Evaluation of a Multi-echelon Production, Transportation and Distribution System: A Matrix-analytical Approach," *European Journal of Operational Research*, 176, 1066–1083.
- Zipkin, P.H. (2000), *Foundations of Inventory Management*, USA: McGraw-Hill, Boston.