When to Share Demand Information in a Simple Supply Chain?

Guillermo Gallego
Department of Industrial Engineering and Operations Research
Columbia University in the City of New York

Ying Huang
Kaan Katircioglu
and
Ying Tat Leung
T. J. Watson Research Center
Yorktown Heights, NY 10598

20 November 2000

Abstract

We assess the benefits of sharing demand information in a supply chain consisting of a single supplier and a single retailer. The retailer faces Poisson demands, economies of scale in ordering, and places orders with the supplier according to a (Q, r) policy. The supplier places orders from an outside source with ample stock. The supplier's problem is to minimize holding and backorder penalty costs with or without retail demand information. We show that the supplier's optimal policy under demand information sharing results in a cost that is independent of the order size Q. Our numerical study indicates that this policy is significantly more cost effective than the supplier's optimal base–stock policy, in the absence of demand information. The supplier, however, can do better than following a base–stock policy by delaying orders arriving from the retailer. We consider two delay base–stock policies, fixed and random, and show that random delay base–stock policies are optimal.

There are three schemes, increasingly more difficult to implement, that result in lower expected costs for both the supplier and the retailer. The first is demand information sharing. This works if the retailer's expected cost decreases as a consequence of the supplier's optimal use of demand information. The second involves demand information *selling*. This works, if the retailer's expected cost increases, but this increase is smaller than the decrease in expected cost that accrues to the supplier. Finally, if neither sharing nor selling is mutually beneficial, the supplier and the retailer can jointly minimize the sum of their expected costs and share the benefits of doing so.

1 Introduction

Since the early 1990s, leaders in many industries have realized that they must take the viewpoint of the supply chain to reduce inventory costs and improve logistics. Major retailers have initiated efforts to enable demand forecast collaboration and joint inventory management. These efforts have greatly benefited from the advent of electronic connectivity technologies and industry-wide standards (such as VAN/EDI and Internet/CPFR). One of the simplest forms of retailer/supplier collaboration is the sharing of demand–and goods movement–related data. There is a common belief that this improves the suppliers' ability to forecast demands and to produce the right amount of products at the right time while maintaining minimum inventory levels.

The most common form of demand information sharing is for retailers to send point-of-sales (POS) and inventory position information to the suppliers through EDI or other electronic means. If used properly, the suppliers can infer sales rates, demand seasonality, and promotional effects from POS data. Furthermore, suppliers may be able to deduce the inventory control policies used by retailers, and therefore be in a better position to anticipate future orders.

Without the explicit push from the retailers, many such collaboration and data sharing efforts would not have happened. Wal–Mart, the retailing industry supply chain management leader, developed a proprietary system called RetailerLink in the early 1990s. Wal–Mart requested their key suppliers to install RetailerLink on designated PCs connected to the host system at Wal–Marts IT processing center. The RetailerLink allowed the suppliers to view a wide variety of demand-related information such as POS, inventory, and forecasts at different store locations. Other retailers encouraged their suppliers to set up collaborative forecast and inventory replenishment programs such as Vendor Managed Inventory (VMI). In VMI programs, the suppliers make (sometimes, jointly with retailers) inventory replenishment decisions based on POS and inventory information transmitted from the retailers. The suppliers have the choice of deciding which products in the retailers distribution centers need to be replenished as long as certain pre-defined business measurements are met. In theory, this could allow the suppliers to optimize the coordination between replenishment decisions, production scheduling, and distribution.

Of course, retailers may have concerns about how demand information is going to be used by the supplier. First and foremost, retailers request suppliers to maintain confidentiality of demand information. In most cases, retailers want to make sure their data will not be accessible to competitors. On the other hand, retailers also have concerns about how the shared demand information might be used operationally by the supplier in ways not to their advantage. For instance, retailers usually request that suppliers in VMI programs to meet service level requirements. This is to ensure that suppliers reduction in finished goods inventories will not have an adverse effect on their ability to respond to retailers orders. This demonstrates that retailers believe suppliers can improve efficiency by reducing inventory levels while achieving the same or higher service levels through increased demand visibility. If this becomes a reality, it can translate into a cost advantage for both the retailers and the suppliers. Thus, one central question is whether sharing demand information can be mutually beneficial, and what to do if it is not.

In this paper we consider a supply chain consisting of a single supplier and a single retailer facing Poisson demands where the retailer places batch orders according to a (Q, r) policy to take advantage of economies of scale. The supplier's objective is to minimize her own holding and penalty costs. We show that the optimal policy for the supplier under demand information sharing, is to *delay* the placement of an order until the retailer's inventory position drops to $r + n$ for some $n \in \{1, \ldots, Q\}$. We also show that the expected cost under this policy is independent of Q.

We compare the cost of the optimal policy under demand information sharing to the cost of the optimal base–stock policy without demand information, and find that the supplier can derive significant benefit from timely demand information. We find instances where the retailer's lead time stochastically *decreases* by sharing demand information. These instances represent win–win situations since the expected costs of both the supplier and the retailer go down. On the other hand, we also find instances where sharing demand information stochastically *increases* the retailer's lead time, and consequently his expected cost. In these instances there is no base–stock policy the supplier can use to reduce his cost without increasing that of the retailer. In summary, there are win–win instances where both the supplier and the retailer are better off, and there are instances where it is best *not* to share demand information unless an external compensating mechanism is used.

What should the supplier do if the retailer is not willing to share demand information? Motivated by the fact that the optimal policy for the supplier under demand information sharing delays the placement of the orders, we consider *delay* base–stock policies in the absence of demand information sharing. Interestingly, delay base–stock policies have only recently attracted the attention of researchers, see Katircioglu (1998) and Moinzadeh (1999), long after Beckman (1962) dismissed delay base–stock policies, without studying them, by claiming that they have "negligible benefits in practice." We consider two delay base–stock policies, fixed and random, show that random delays are optimal, and compare their costs to the cost of the optimal policy under demand information sharing. We find that a significant portion of the benefits of demand information sharing are recaptured by the use of delay base–stock policies. Consequently, the benefits of information sharing need to be measured against the benchmark of random delay base–stock policies. It is against this benchmark that the retailer needs to decide whether or not to share demand information. The retailer should share demand information with the supplier if the retailer's expected cost decreases as a consequence of the supplier's optimal use of demand information. On the other hand, if the retailer's expected cost increases, the retailer may be willing to *sell* demand information at a price that is high enough to cover his increase in expected costs. On the other hand, the supplier will be willing to buy this information only if his expected savings are large enough to justify the cost of acquiring the information. Finally, if neither sharing nor selling is mutually beneficial, then the supplier and the retailer can jointly minimize the sum of their expected costs and share the benefits of doing so.

1.1 Literature Review

Recent papers by Bourland, Powell and Pike (1996), Chen (1998), Gavirneri, Kapuscinski, and Tayur (1999), and Lee, So, and Tang (2000) investigate the value of sharing *current* demand information in supply chains. These authors find that information sharing is valuable when order cycles are asynchronous, when the retailer orders in batches to take advantage of economies of scales, or when the demand process is autocorrelated.

Bourland et al. consider a two stage serial system where each stage follows a periodic review base-stock policy. Each stage is assumed to have ordering periods of equal lengths. The authors investigate the benefits of sharing customer demand information through EDI when the ordering cycles are *not* synchronized. Chen considers a continuous–time serial supply chain where the demand process at the downstream stage is compound Poisson. Each stage is assumed to follow an (nQ, r) replenishment policy with given batch sizes. Reorder points are computed based on echelon (centralized) and on local (decentralized) information. The cost of the two systems is compared in a comprehensive numerical study, to determine the benefits of using centralized information and the factors that make centralized information valuable. Gavirneri et al. study the holding and penalty cost of a finite capacity supplier facing demands from a single retailer following an (s, S) policy. The authors assume that the retailer's lead time is zero and that the retailer can immediately procure units from an alternative source if the supplier is not capable of filling an order. In addition, the authors assume that the retailer is willing to freely share demand information with the supplier. Under these assumptions, the authors report significant savings from demand information sharing. Lee, So, and Tang study the benefit of demand information sharing to the manufacturer in a two stage supply chain consisting of a single manufacturer and a single retailer facing an autocorrelated demand process. The manufacturer and the retailer incur linear holding and backlogging costs, experience constant lead times, and follow base–stock policies. The authors assume that the manufacturer immediately ships retail orders and procure from an alternative source in case of a shortfall. The authors report large savings for the manufacturer when the autocorrelation coefficient $\rho \in (-1, 1)$ is high. These savings are computed under the assumption that the manufacturer knows the parameters of the demand process, observes only retail orders, and makes forecasts based only on the *last* retail order. We find, however, that the reported savings are only transient when the manufacturer makes ordering decisions based on his best estimate of the retailer's current demand.[1]

A different literature stream studies the benefits of obtaining *advance* demand information. Hariharan and Zipkin (1995) study the benefits of having advance information in the form of customer lead times and show that customer lead times offset demand lead times. Gallego and Özer (1999) study a periodic review model where advance demand information over an information horizon is observed in every period. These authors show that state–dependent (s, S) policies are optimal, where the state space consists of the inventory position net of the lead time demand that has already been observed, and observed demands beyond the lead time and up to the information horizon. Gallego and Özer (2000) study a similar model for multi-echelon serial systems and show that state–dependent base–stock policies are optimal.

Other papers related to information sharing include Chen (1999) where supply chain members are divisions of the same firm, acting on local inventories and material, and where information flows are subject to delays, and Cachon and Lariviere (1997) who study contracts that allow a manufacturer to share credible demand forecasts to a sole supplier that needs to construct capacity to meet uncertain demand.

The rest of this paper is organized as follows. §2 describes the basic setting, the cost of the optimal base–stock policy in the absence of demand information sharing and the cost of the optimal policy under demand information sharing. This section ends with a numerical comparison between these two policies. §3 introduces fixed and random delayed base–stock policies in the absence of demand information, demonstrates the optimality of random delay base–stock policies, and offers computations that show that a substantial part of the gap between the cost of straight, or naive, base–stock policies and the optimal cost under demand information sharing is accounted for by the use of delayed base–stock policies. §4 describes conditions under which the retailer is better off sharing demand information, and briefly explores situations where the retailer sells demand information to the supplier, and situations where the retailer and the supplier jointly optimize the sum of their expected

[1] It can be shown that the errors in estimating current demand go down to zero at a geometric rate for most values of ρ including $\rho \in [0, 1)$.

costs. §5 offers our conclusions and directions for future research.

2 The Benefits of Demand Information Sharing

This section describes the model, the cost of the optimal base–stock policy in the absence of demand information sharing and the cost of the optimal policy under demand information sharing. This section ends with a numerical comparison between these two policies.

2.1 Model Description

We consider a supply chain consisting of one supplier and one retailer. The retailer faces Poisson demand at rate λ and incurs stationary fixed plus linear ordering costs, and linear holding and backlogging costs. Since demands are backlogged, optimal policies for the retailer and the supplier are independent of the unit variable cost. Because of economies of scale in ordering it is optimal for the retailer to follow a (Q, r) replenishment policy. Under this policy the retailer places orders of size Q with the supplier whenever his inventory drops to r. See Federgruen and Zheng (1992) for an efficient algorithm to compute the parameters Q and r. We assume that the retailer uses a fixed Q and optimizes over r. Retailer's may fix Q to facilitate packaging and handling. Fixed values of Q with good guaranteed worst case performance are the economic order quantity, Q_d, see Zheng (1992) and Axsäter (1993) and $\sqrt{2}Q_d$, see Gallego (1998).

We assume that the supplier orders from an outside source with ample stock and incurs linear holding and backlogging costs at rates h and p respectively. Units ordered from the outside source arrive at the supplier[2] after l units of time.

2.2 Base-Stock Policies without Information Sharing

Without demand information sharing, the supplier observes only orders of size Q from the retailer. The time between orders observed by the supplier is a random variable, say T_Q, whose distribution is equal to that of the sum of Q exponential random variables, each with

[2]It is possible to think of the supplier as a manufacturer and of l as her production lead time.

parameter λ. T_Q has an Erlang distribution with parameters Q and λ, so $E[T_Q] = Q/\lambda$ and $\text{Var}[T_Q] = Q/\lambda^2$. We will assume the supplier knows the distributional form of T_Q or equivalently that the supplier knows that the retailer faces Poisson demand at rate λ. If the supplier did not know λ she could estimate it by either the method of moments, the method of maximum likelihood, or by using exponential smoothing. We will not consider the impact of errors in estimation and refer the reader to Drezner et al. (1996) for more information on this subject.

Base–stock policies are used frequently in practice and there is a widespread believe among practitioners that base–stock policies are optimal in the absence of economies of scale. Under a base–stock policy the supplier places an order, of size Q, to raise her inventory position to $b = mQ$, for some integer m, whenever the retailer places an order. The supplier immediately ships the order to the retailer if the units are available on inventory. Otherwise the shipment is delayed until the units are available. An order placed by the supplier at time t arrives at time $t + l$ and will be used to satisfy the mth order arriving from the retailer after time t. Let T_{mQ} be the time until m additional orders (of size Q) are placed by the retailer. If $t + l < t + T_{mQ}$ the order arrives early and the supplier incurs a holding cost $hQ(T_{mQ} - l)^+$. On the other hand, if $t + l > t + T_{mQ}$ the order is late and the supplier incurs a backlogging cost $pQ(l - T_{mQ})^+$. Since the retailer places orders at rate λ/Q the supplier's long run average expected holding and backlogging cost is given by $H(mQ)$ where

$$\begin{aligned} H(y) &= h\lambda E[T_y - l]^+ + p\lambda E[l - T_y]^+ \\ &= hE(y - N_l)^+ + pE(N_l - y)^+ \qquad (1) \end{aligned}$$

where N_t denotes a Poisson random variable with parameter λt. The equality in (1) is justified by Little's law, since, for example,

$$\lambda E[T_y - l]^+ = EN_{[T_y - l]^+} = E(y - N_l)^+.$$

Thus, the optimal base-stock policy for the supplier is found by minimizing $H(y)$ over

$y \in \{mQ : m \text{ integer}\}$. Since the increments

$$H((m+1)Q) - H(mQ) = (h+p) \sum_{k=mQ+1}^{(m+1)Q} P(N_l < k) - pQ$$

are increasing in m, the largest optimal solution is found by finding the smallest integer m, say m_o, such that $H((m+1)Q) - H(mQ) > 0$. Mathematically,

$$m_o = \inf\{m : \text{integer} : \frac{1}{Q} \sum_{k=mQ+1}^{(m+1)Q} P(N_l < k) > \frac{p}{h+p}.\}. \tag{2}$$

We summarize this result in the following proposition.

Proposition: 1 *The largest optimal base-stock policy for the supplier is $b_o = m_o Q$ where m_o is given by (2).*

The case $Q = 1$ is of special interest, since in this case the retail orders coincide with the demand since by the retailers. Thus, in this case, the supplier has full demand information. The optimal base stock level, say b^*, is given by the smallest integer b satisfying

$$P(N_l \leq b) > \frac{p}{h+p}. \tag{3}$$

Obviously, $H(b^*) \leq H(m_o Q)$ for all $Q > 1$. In fact $H(b^*)$ is a cost lower bound on all policies, for all Q, even if the supplier has full demand information.

2.3 Optimal Policies under Demand Information Sharing

We now assume that the retailer shares demand information with the supplier so the supplier always knows the retailer's inventory position. The following lemma shows that an optimal policy for the supplier consists of monitoring the retailer's inventory position and placing an order of size Q when the inventory position drops to a given level. To state this policy formally we need to introduce some notation. Let b^* as in (3), let m be an integer such that

$$(m-1)Q < b^* \leq mQ,$$

and let $n = b^* - (m-1)Q$. Notice m and n are unique and that $n \in \{1, \ldots, Q\}$.

Proposition: 2 *An optimal policy for the supplier under demand information sharing calls for ordering Q units to raise the inventory to mQ whenever the retailer's inventory position drops to $r+n$. Moreover, the cost of this policy is $H(b^*)$.*

Proof: We assume that the supplier can place an order only at demand epochs. Equivalently, the supplier can place an order only when there is a change in the retailer's inventory position. This restriction is without loss of generality since the interdemand epochs are exponentially distributed. The proposed policy calls for the supplier to place an order of size Q (to bring the inventory position up to mQ) whenever the retailer's inventory position "drops" to $r+n$, where $n \in \{1, \dots, Q\}$.

Suppose the retailer's inventory position drops to $r+n$ at time t. At this time the supplier places an order of size Q. This order will arrive at the retailer at time $t+l$ and is targeted to satisfy the mth order placed by the retailer, which arrives at time $t + T_{(m-1)Q+n} = t + T_{b^*}$. The supplier's order is early if $l \leq T_{b^*}$ and late otherwise. Using the same logic we did to compute the supplier's cost without demand information sharing we see that the supplier's long run average expected holding and backlogging cost is given by $H(b^*)$. Since b^* minimizes $H(\cdot)$ and $H(b^*)$ is a lower bound on all policies we conclude that proposed policy is optimal.

$\square$

Remark: Notice that $H(b^*)$ is independent of Q. Notice also that the resulting policy is not strictly a base-stock policy unless $n = Q$. Indeed, for $n < Q$ the supplier's inventory position is $(m-1)Q$ between the time the retailer places an order until the time his inventory drops to $r+n$ at which time the supplier's inventory position is raised to mQ. We can think of this policy as a state dependent base-stock policy where the state is the retailer's inventory position. If the inventory position is above $r+n$ the base-stock level is $(m-1)Q$ otherwise it is mQ. We can also view this policy as *delaying* the placement of the supplier order until $Q-n$ demands occur at the retailer. The expected delay is $(Q-n)/\lambda$.

2.4 Numerical Comparison

We have written a code to evaluate the cost with and without demand information sharing for several values of Q. For the set of parameters $\lambda = 20, l = 1$, $h = 1$, and $p = 9$, we found that H is minimized at $b^* = 26$. Thus the cost under demand information sharing is $H(26) = 8.19$ regardless of the value of Q.

Without demand information sharing the value of $b_o = m_o Q$ depends on Q. Table 1 shows the cost $H(m_o Q)$ for selected values of Q. Notice that $H(b^*)$ coincides with $H(m_o Q)$ when $Q = 1$.

Q	$m_o Q$	$H(m_o Q)$
1	26	8.19
5	25	8.31
10	30	10.32
15	30	10.32
20	20	17.77
25	25	8.31
30	30	10.32
40	40	20.00
50	50	30.00
> 200	0	180.00

Table 1: Numerical Comparison

There are two important conclusions we can draw from Table 1. First, the cost can be much lower under demand information sharing, e.g., $H(26) = 8.19$ versus $H(200) = 180$. Notice that we can select Q such $H(0) = p\lambda < H(Q)$ and $m_o = 0$ is optimal. On the other hand, the cost under information sharing is $\min_y H(y)$ where the minimization is taken over the positive integers. It is possible to select the cost parameters and the demand rate such that $\min_y H(y)$ is negligible relative to $p\lambda$, so the ratio of these two costs can be made as large as desired.

The second conclusion has to do with how information sharing affects the retailer. Suppose it takes L units of time for a shipment to get from the supplier to the retailer. If we take the view, as in Gavirneri et al., that the penalty cost is the cost of expediting late orders, then the retailer's cost is not affected by the wholesaler's policy. On the other hand, if stockouts at the wholesaler cause shipments delays to the retailer, then the retailer's total lead time is $(l - T_{b^*})^+ + L$ under demand information sharing, and $(l - T_{m_oQ})^+ + L$ in the absence of demand information sharing. Since the retailer's total lead time is stochastically decreasing in the base-stock level used by the supplier, the retailer is better off sharing demand information if $m_oQ < b^*$, and is worse off if $m_oQ > b^*$. In the above example the retailer will be willing to share demand information for values of $Q \in \{5, 20, 25\}$ and for $Q > 200$. These are win–win cases. On the other hand, for values of $Q \in \{10, 15, 30, 40\}$ or $Q \in [40, 200]$ the retailer is worse off by sharing demand information. Moreover, if $m_oQ > b^*$ the supplier cannot improve his costs without decreasing that of the retailer. As a final comment, we clarify here that although Q has been treated as exogenous, our computations indicate that the retailer's cost may rise by sharing demand information even if we endogenize Q.

3 Delay Base–Stock Policies

We have assumed that the supplier follows a base-stock policy in the absence of demand information. However, the policy under demand information *delays* the placement of orders by the supplier until the inventory position at the retailer drops to $r + n$. In the absence of demand information the supplier can also delay orders by a fixed amount of time in an attempt to mimic the policy that uses demand information. It takes, on average, $(Q - n)/\lambda$ units of time for the retailer's inventory position to drop for $r + Q$ to $r + n$ which suggest that it may be good idea to delay the order by approximately $(Q - n)/\lambda$ units of time.

The idea of delaying orders is related to just–in–time replenishment, and is best understood in a setting where demand is deterministic with rate λ. Under this setting the supplier can place her orders so they arrive just–in–time to meet orders placed by the retailer. To see how this works, write the supplier's lead time l as $l = kQ/\lambda + q$ where k is an integer and $0 \leq q < Q/\lambda$, i.e., $q = l \bmod Q/\lambda$, and let the supplier place an order $\tau_d = Q/\lambda - q$ units of time after the retailer places an order. Suppose, for example, that the retailer places an order at time t, then the supplier places an order at time $t + \tau_d$ to raise her inventory position to $b = mQ$, where $m = k + 1$. This order arrives at time $t + \tau_d + l = t + mQ/\lambda$ *just–in–time* to satisfy the mth order placed by the retailer. Under this policy the supplier's inventory is always equal to zero.

While it is impossible to duplicate this clockwork result in a stochastic setting, we can get close by having the supplier delay her orders by an appropriately chosen amount of time $\tau \geq 0$. Thus, if the retailer places an order at time t, the supplier places an order at time $t+\tau$. We now argue that this is equivalent to having a supplier lead time $l+\tau$. To see this, assume that the retailer places an order at time t. The supplier's inventory position at time t is nQ for some $n < m$, where $m - n$ is the number of orders placed by the retailer over the interval $[t-\tau, t]$. All the $m - n - 1$ retail orders placed during $[t - \tau, t)$ are placed by the supplier by time $t + \tau$. Thus, the supplier's inventory position just prior to time $t + \tau$ is equal to $(m - 1)Q$. Since an order is placed by the supplier at time $t + \tau$ the supplier's inventory position is raised to mQ at time $t+\tau$. Now, the order placed at time $t+\tau$ arrives at time $t + \tau + l$ and will be on time if $t + \tau + l < t + T_{mQ}$. Otherwise the order will be late. Thus, the system behaves as if the lead time was $l + \tau$ rather than l.

We can write the cost of delayed base-stock policies as $H(mQ, \tau)$ where

$$\begin{aligned} H(y, \tau) &= h\lambda E[T_y - l - \tau]^+ + p\lambda E[l + \tau - T_y]^+ \\ &= hE[y - N_{l+\tau}]^+ + pE[N_{l+\tau} - y]^+. \end{aligned}$$

Fixed delay base-stock policy were recently proposed by Moinzadeh where he computes the optimal delay τ for a fixed order size mQ for the case where the interorder distribution is normal and for the case where only the mean and variance of the interorder distribution is known. The computational results in Moinzadeh show that significant savings can result from using a fixed delay base–stock policy.

Here we compute the optimal delay τ and the base-stock level mQ for the actual interorder Erlang-Q distribution. To find the optimal (mQ, τ) pair for fixed Q we need to minimize $H(mQ, \tau)$ over integer values of m and nonnegative values of τ. Notice that

$$H((m+1)Q, \tau) - H(mQ, \tau) = (h+p) \sum_{k=mQ+1}^{(m+1)Q} P(N_{l+\tau} < k) - pQ$$

is increasing in m so the largest optimal m for fixed τ is the smallest integer such that $H((m+1)Q,\tau) - H(mQ,\tau) > 0$, or equivalently the smallest integer such that

$$\frac{1}{Q}\sum_{k=mQ+1}^{(m+1)Q} P(N_{l+\tau} < k) > \frac{p}{h+p}. \tag{4}$$

Similarly, if we fix m then the partial derivative with respect to τ is given by

$$H_2(m,\tau) = -(h+p)\lambda P(t_{mQ} > l+\tau) + p\lambda,$$

is increasing in τ, so the first order condition (ignoring the nonegativity constraint) is

$$P(t_{mQ} > l+\tau) = P(N_{l+\tau} < mQ) = \frac{p}{h+p}. \tag{5}$$

The above argument shows that $H(mQ,\tau)$ is convex with respect to m for fixed τ and with respect to τ for fixed m. Unfortunately, $H(mQ,\tau)$ is not jointly convex, and a countable number of stationary points satisfy the first order conditions (4) and (5). Indeed, let m_1 be the smallest integer m for which (5) has a positive solution. Mathematically,

$$m_1 = \inf\{m : \text{integer } | P(N_l < mQ) = P(t_{mQ} > l) > \frac{p}{h+p}\} \tag{6}$$

Thus, corresponding to every integer $m \geq m_1$ there is a unique $\tau_m > 0$ satisfying (5). To see that (mQ,τ_m) also satisfies (4) notice that

$$P(N_{l+\tau_m} < k) > P(N_{l+\tau_m} < mQ) = \frac{p}{h+p}$$

for all $k > mQ$ so (4) holds. In addition to the points (mQ,τ_m) the boundary point $(m_oQ,0)$ is also a candidate solution where m_o is given by (2). The existence of a countable number of stationary points lead Moinzadeh to an extensive computational search for an optimal solution. The following proposition shows that we only need to consider the points $(m_o,0)$ and (m_1,τ_1) where, for convenience, we write τ_k for τ_{m_k}.

Proposition: 3 *If $m > m_1$ then $H(mQ,\tau_m) > H(m_1Q,\tau_1)$.*

.

Proof: Clearly, if $\tau' > \tau$ then $\min_y H(y,\tau') > \min_y H(y,\tau)$. This simply tells us that it is more expensive to manage a system with a longer lead time when we have the freedom of selecting the optimal base–stock level for each system. In our problem the base–stock level is restricted to multiples of Q, and this is the key fact that make fixed delay base–stock policies attractive since it is possible to have

$$\min_m H(mQ,\tau) < \min_m H(mQ,0)$$

for some positive τ. The idea is that for a positive value of τ the unrestricted optimal base–stock level may be a multiple of Q!

To show that the proposition holds, notice that each $m > m_1$ is such that

$$H(mQ,\tau_m) = \min_y H(y,\tau_m),$$

so mQ is the natural optimal base–stock level when the lead time is $l + \tau_m$. Consequently

$$H(mQ,\tau_m) = \min_y H(y,\tau_m) > \min_y H(y,\tau_1) = H(m_1Q,\tau_1).$$

□

3.0.1 Numerical Results

Table 2 summarizes our numerical results comparing fixed delay–base stock policies to regular base–stock policies for the same set of problems considered before. As we can see from the table, delaying orders is not always optimal, but when it is it can help capture a substantial part of the benefits of information sharing when the retailer is *unwilling* to share demand information.

It turns out that the optimal delay τ is not far, but is always smaller than $(Q-n)/\lambda$, the expected delay under demand information sharing. Indeed, for $Q \in \{10, 15, 30\}$ we have $(30-26)/\lambda = .20$ which is close to 0.1615. Similarly, for $Q = 4$, we have $(40-26)/\lambda = .70$ which is close to 0.6069, and for $Q = 50$ we have $(50-26)/\lambda = 1.20$ which is close to 1.058.

The reason the optimal delay is smaller than the expected delay under information sharing reflects the fact $h < p$, so it is best to play it safe.

Q	τ^*	m^*Q	$H(m_oQ,0)$	$H(m^*Q,\tau^*)$	$\frac{H(m_oQ,0)-H((m^*Q,\tau^*)}{H(m_oQ,0)}$
1	0	26	8.19	8.19	-
5	0	25	8.31	8.31	-
10	0.1615	30	10.32	8.83	14.44%
15	0.1615	30	10.32	8.83	14.44%
20	0.6069	40	17.77	10.33	41.87%
25	0	25	8.31	8.31	-
30	0.1615	30	10.32	8.83	14.44%
40	0.6069	40	20.00	10.33	48.35%
50	1.059	50	30.00	11.64	61.21%

3.1 Random Delay Base–Stock Policies

We now consider a more sophisticated policy that terminates the delay if the next order from the retailer arrives before the fixed delay. Suppose that the retailer places an order at time t which drops the inventory position of the supplier to $(m-1)Q$. Let τ be the fixed delay and T_Q the time until the next order from the retailer. Under the random delay base–stock policy the supplier delays her order by $\min(\tau, T_Q)$. If $\tau < T_Q$, the supplier places the order at time $t+\tau$. This order arrives at time $t+\tau+l$ and it is late when $\tau+l > T_{mQ}$. On the other hand, if $T_Q < \tau$ the supplier places the order at time $t+T_Q$. This order arrives at time $t+T_Q+l$ and it is late when $l > T_{(m-1)Q}$. In either case the inventory position at time $t+\min(\tau, T_Q)$ is $(m-1)Q$ and the ordering policy is repeated. By the reward renewal theorem the average cost of this order can be written as a function τ and m as

$$\begin{aligned} G(mQ,\tau) &= h\lambda E(T_{(m-1)Q} - l + (T_Q-\tau)^+)^+ + p\lambda E(l - T_{(m-1)Q} - (T_Q-\tau)^+)^+ \\ &= hE[(m-1)Q - N_l + (Q-N_\tau)^+]^+ + pE[N_l - (m-1)Q - (Q-N_\tau)^+]^+ \end{aligned}$$

Notice that if $\tau = 0$ then

$$G(mQ,0) = H(mQ,0) = hE[mQ - N_l]^+ + pE[N_l - mQ]^+.$$

Thus, the largest optimal m for $\tau = 0$ is m_o as defined by (2). We need to compare the cost $G(m_oQ, 0) = H(m_oQ, 0)$ to the cost of the best policy with a positive delay. Recall the definition of m_1 in equation (6). Let ν_1 be the unique solution to

$$P(N_l + N_\nu < m_1Q \mid N_\nu < Q) = \frac{p}{h+p}. \tag{7}$$

Proposition: 4 *The pair (m_1Q, ν_1) is the only strictly interior stationary point of $G(mQ, \nu)$.*

Proof: We will show that for each m, there is an optimal, possibly infinite, $\tau \geq 0$ that minimizes $G(m, \tau)$. Fixing m, and differentiating with respect to τ we find

$$\begin{aligned} G_2(m,\tau) &= -(h+p)\lambda P(T_{m-1}Q + T_Q > l+\tau, T_Q > \tau) + p\lambda P(T_Q > \tau) \\ &= \lambda P(T_Q > \tau)\left[-(h+p)P(T_{m-1}Q + T_Q > l+\tau | T_Q > \tau) + p\right] \\ &= \lambda P(N_\tau < Q)\left[-(h+p)P(N_l + N_\tau < mQ | N_\tau < Q) + p\right]. \end{aligned}$$

Notice that

$$P(N_l + N_\tau < mQ \mid N_\tau < Q) = \sum_{k=0}^{Q-1} P(N_l < mQ - k)\frac{P(N_\tau = k)}{P(N_\tau < Q)}.$$

is continuous and increasing in τ, so $H(m, \cdot)$ is convex. Moreover,

$$\lim_{\tau\to\infty} P(N_l + N_\tau < mQ \mid N_\tau < Q) = P(N_l \leq (m-1)Q).$$

Let

$$\tau_m = \inf\{\tau \geq 0 : P(N_l + N_\tau < mQ \mid N_\tau < Q) \leq \frac{p}{h+p}\} \tag{8}$$

where we set $\tau_m = \infty$ if the set is empty, i.e., if $P(N_l \leq (m-1)Q) > p/(h+p)$, and $\tau_m = 0$ if $P(N_l < mQ) \leq p/(h+p)$.

If $\tau_m = 0$ an order should be placed immediately indicating that m is too small. On the other hand, if $\tau = \infty$, the order should not be placed before the arrival of the next order from the retailer. This is indicative that m is too large. The first argument eliminates small values of m such that $P(N_l < mQ) < p/(h+p)$. The second argument eliminates large values of m such that $P(N_l \leq (m-1)Q) > p/(h+p)$. Thus, the only integer m that results in a positive but finite value of τ_m satisfies

$$P(N_l < mQ) \geq \frac{p}{h+p} > P(N_l < (m-1)Q),$$

which is equivalent to (6). Thus $(m_1 Q, \nu_1)$ is the only stationary point G.

□

To find the optimal policy we need to compare $G(m_o Q, 0)$ and $G(m_1 Q, \nu_1)$. Let $(m * Q, \nu^*)$ denote the pair with lower cost among these two candidate solutions. Then the optimal policy for the supplier is to place an order of size Q, $\min(\nu^*, T_Q)$ units of time after receiving an order from the retailer. The supplier's inventory position after ordering is equal to $b^* = m^* Q$.

We have shown that $(m^* Q, \nu^*)$ minimizes $G(mQ, \nu)$ but we have not yet shown that this policy minimizes the supplier's long run average holding and penalty cost. We do this by showing that, appropriately interpreted, the policy $(m^* Q, \nu^*)$ minimizes the total cost over any finite horizon $[0, T]$ under any starting conditions provided that at time zero we know the inventory position and the timing of the last order, if any, placed by the retailer during the interval $(-\nu^*, 0]$. If no orders were placed by the retailer during the interval $(-\nu^*, 0]$ then an order is placed to raise the inventory position to $m^* Q$ and the policy continues as prescribed. If, on the other hand, the last order from the retailer was placed at time $t \in (-\nu^*, 0]$ then the inventory position at time zero is raised to $(m^* - 1)Q$ and the order placed at time t is delayed as prescribed by the policy. To see that this minimizes the total cost over any finite horizon, notice that this policy minimizes the expected holding and penalty cost of each order, see Katircioglu, and therefore the sum of the cost over all orders over any finite interval. Now, any admissible policy π^* that is total cost optimal over any finite horizon $[0, T]$ is necessarily long run average cost optimal since $c_{\pi^*}(T) \leq c_\pi(T)$ for all $T \geq 0$ and all admissible π implies that

$$a_{\pi^*} \equiv \liminf_{T \to \infty} \frac{c_{\pi^*}(T)}{T} \leq \liminf_{T \to \infty} \frac{c_\pi(T)}{T} \equiv a_\pi,$$

showing that π^* is also long run average cost optimal.

We now briefly compare $G(mQ,\tau)$ and $H(mQ,\tau)$. A little algebra reveals that

$$G(mQ,\tau) - H(mQ,\tau) = \{(h+p)P(N_l < (m-1)Q) - p\}\, E[N_\tau - Q]^+.$$

Notice that this allow us to compute $G(mQ,\tau)$ via

$$G(mQ,\tau) = H(mQ,\tau) + \{(h+p)P(N_l < (m-1)Q) - p\}\, E[N_\tau - Q]^+. \tag{9}$$

Notice that $G(mQ,\tau) \leq H(mQ,\tau)$ for a any positive τ if and only if

$$P(N_l < (m-1)Q) \leq \frac{p}{h+p}.$$

By the definition of m_1, it must be the case that

$$P(N_l < (m_1-1)Q) \leq \frac{p}{h+p},$$

and consequently

$$G(m_1Q,\tau) \leq H(m_1Q,\tau) \tag{10}$$

holds for all positive τ.

Although the random delay base–stock policy has the virtue of being provably optimal, our computations indicate that the savings relative to the fixed delay base–stock policy are almost always negligible. The only case where there was a small, but significant, difference in cost was for $Q = 20$ where $\nu_1 = 0.6279$ resulting in an average cost of \$10.19, a slight improvement versus the cost \$10.33 of using the fixed delay $\tau_1 = 0.6069$. A heuristic use of the fixed delay policy is to use $(m_1Q, \min(\tau_1, T_Q))$, instead of (m_1Q, τ_1), and inequality (10) guarantees that this is an improvement over the fixed delay policy. On the other hand, the computational effort of finding and evaluating the random delay base–stock policy is similar to that of the effort of finding and evaluating the fixed base–stock policy. As a final observation, note that random delay policies are unlikely to be significantly better than fixed delay policies for small and large values of Q. Indeed, for small values of Q we almost have retail demand information, whereas for large Q values T_Q becomes large and almost deterministic.

4 To Share or not to Share Demand Information

Our analysis indicates that the benefits of information sharing should be computed by comparing the cost of the random delay–base stock policy to the cost of the optimal policy under information sharing. Comparing the cost against base–stock policies greatly exaggerates the benefits of demand information.

The retailer's lead time demand under demand information sharing is

$$N_L + (N_l - b)^+,$$

where b is given by (3). On the other hand, the retailer's lead time demand without information sharing, assuming random delay, is

$$N_L + (N_l - (m-1)Q - (Q - N_\tau)^+)^+,$$

where (mQ, τ) minimizes G.

The retailer needs to compute his cost under each of these lead time demands to decide whether or not it is in his interest to share demand information. For example, when $Q = 20$, the retailer would be willing to share demand information. This is a win–win situation since both the retailer and the wholesaler benefit from sharing demand information. On the other hand, for $Q = 10$ the retailer will be unwilling to share demand information. A similar situation arrises when $Q = 15$ and when $Q = 30$.

4.1 Information Selling

If the retailer is worse off sharing demand information, the supplier may be willing to buy these information from the retailer if the expected gain to the supplier exceeds the expected cost to the retailer. Since the retailer must at least insure that his expected cost does not go up, a lower bound on the value of demand information sharing is the incremental cost to the retailer. On the other hand, the supplier will never pay more than his expected gain from demand information sharing. Thus, we can obtain, computationally, a lower an upper bound on the price at which the retailer will sell the information to the supplier. Although we are not aware of any practical instance where the retailer sells demand information for a fee, there may be an implicit cost paid by the supplier in the form of lower unit prices. Moreover, as retailers become more powerful, they may soon be in a position of actually demanding payment for demand information.

4.2 Jointly Optimizing Costs

When the retailer is unwilling to share demand information, and the value of the information to the supplier is less than the cost of the information to the retailer, the only win–win option is for the supplier and the retailer to jointly optimize the sum of their costs and then find a way to share the savings. Moving from distributed to centralized decision making is nontrivial since it entails sharing cost structures, in addition to demand information. Also, there may be incentives for misrepresenting costs. Nevertheless, joint optimization, is one instance that can benefit from the emerging use of electronic contracts that use Third Party Agreement Markup Language (TPAML) .

5 Conclusions and Directions for Future Research

Results from the literature, e.g., Gavirneni et al., indicate that a supplier using a base–stock policy can significantly reduce her *own* holding and backorder penalty costs if the retailer shares his demand information. This is confirmed in the model studied in this paper. However sharing demand information may not be in the best interest of the retailer. We have also shown that delay base–stock policies proposed by Katircioglu and by Moinzadeh can be effectively used by the supplier to reduce his costs even when the retailer is unwilling to share demand information. We plan to conduct future research in the area of information sharing. In particular, we plan to consider models where the retailer faces compound Poisson demand, and models where there are multiple retailers. Finally, we plan to consider a Markov chain driven demand model where the retailer has knowledge of the state of the Markov chain and investigate the value of sharing this knowledge with the supplier. On the other hand, there is a whole body of literature on continuous time multi-echelon inventory models, e.g., serial, assembly, and distribution systems that can benefit from either fixed or random ordering delays. We are in fact already vigorously working on some of these extensions.

References

[1] Axsäter, S. (1993) "Using the Deterministic EOQ Formula in Stochastic Inventory Control," *Management Science*, 831-834.

[2] Beckman, M., "An Inventory Model for Arbitrary Interval and Quantity Distributions of Demand." *Management Science*, Vol. 8, No. 1, Oct. 1961, pp. 35-37.

[3] Bourland, K.E., S.G. Powell, and D.F. Pyke (1996) "Exploiting Timely Demand Information to Reduce Inventories." *European Journal of Operations Research*, 92, 239-253.

[4] Cachon, G. and M. A. Lariviere (1997) "Contracting to Assure Supply: How to Share Demand Forecasts in a Supply Chain."

[5] Chen, F. (1998) "Echelon Reorder Points, Installation Reorder Points, and Value of Centralized Demand Information." *Management Science*, 44, 12, part 2 of 2, s221-s234.

[6] Chen, F. (1999) "Decentralized Supply Chains Subject to Information Delays." *Management Science*, 45, 1076–1090.

[7] Drezner, Z, J. K. Ryan, and D. Simchi-Levi (1996) "Quantifying the Bullwhip Effect: The Impact of Forecasting, Leadtime and Information." Working paper, Northwestern University. To appear in *Management Science*.

[8] Federgruen, A., and Y.S. Zheng. (1992) "An Efficient Algorithm for Computing an Optimal (Q, r) Policy in Continuous Review Stochastic Inventory Systems." *Operations Research*, 40, 808–813.

[9] Gavirneni, S., Kapuscinski, R., and Tayur, S. (1999) "Value of Information in Capacitated Supply Chains." *Management Science*, 16-24.

[10] Gallego, G. (1998) "New Bounds and Heuristics for (Q, r) policies." *Management Science*, 44, 219–233.

[11] Gallego, G., and O. Özer (1999) "Integrating Replenishment Decisions with Advance Demand Information." Working paper, Department of Industrial Engineering and Operations Research, Columbia University in the City of New York.

[12] Gallego, G., and O. Özer (1999) "Optimal Replenishment Policies for Multi–Echelon Inventory Problems under Advance Demand Information." Working paper, Department of Industrial Engineering and Operations Research, Columbia University in the City of New York.

[13] Katircioglu, K. and Atkins, D. (1998) "New Optimal Policies for A Unit Demand Inventory Problem." Working paper, IBM Research Division, Yorktown Heights, NY.

[14] Lee, H. L., So K., and Tang, C. (2000) "Information Sharing in a Two-Level Supply Chain." *Management Science*, 46, 626–643.

[15] Moinzadeh, K. (1999) "An Improved Ordering Policy for Continuous Review Inventory Systems with Arbitrary Inter-Demand Time Distributions." Working paper. School of Business. University of Seattle, WA 98195.

[16] Zheng, Y. S. (1992) "On Properties of Stochastic Inventory Systems." *Management Science*, 38, 87–103.