Coordinated Replenishment and Shipping Strategies in Inventory/Distribution Systems

Mustafa Cagri Gurbuz, Kamran Moinzadeh*, and Yong-Pin Zhou

University of Washington Business School

January 2005

Abstract

In this paper, we study the impact of coordinated replenishment and shipment in inventory/distribution systems. We analyze a system with multiple retailers and one outside supplier. Random demand occurs at each retailer, and the supplier replenishes all the retailers. In traditional inventory models, each retailer orders directly from the supplier whenever the need arises. We present a new, centralized ordering policy that orders for all retailers simultaneously. The new policy is equivalent to the introduction of a warehouse with no inventory, which is in charge of the ordering, allocation, and distribution of inventory to the retailers. Under such a policy, orders for some retailers may be postponed or expedited so that they can be batched with other retailers’ orders, which results in savings in ordering and shipping costs. In addition to the policy we propose for supplying inventory to the retailers, we also consider three other policies that are based on these well known policies in the literature: (a) Can-Order policy, (b) Echelon Inventory policy, and (c) Fixed Replenishment Interval policy. Furthermore, we create a framework for simultaneously making inventory and transportation decisions by incorporating the transportation costs (or limited truck capacities). We numerically compare the performance of our proposed policy with these policies to identify the settings where each policy would perform well.
1 Introduction

Effective management of distribution systems has been one of the most challenging issues facing both the practitioners and the academicians for years. This has become even more critical in recent years because the advancement in information technology has facilitated information sharing between the parties in the supply chain, and more efficient transportation systems have emerged in many supply chains. Replenishment/ordering, shipment consolidation/coordination policies, different methods to allocate inventory to locations downstream, and effective utilization of information are among the most important supply chain management issues that need to be studied.

According to a study by Delaney (1999), transportation and warehousing account for over 10% of U.S. GDP. While highly dependent on one another, inventory and transportation decisions in supply chains are tackled independently in many of the previous studies. Therefore, much improvement in the supply chain can be achieved if one considers these decisions jointly.

To reduce the logistics costs, companies are increasingly moving towards such practices as shipment coordination/consolidation, which is achieved by incorporating the information on demand, inventory and supply of all the parties in the supply chain. Shipment consolidation means combining multiple shipments into a single group through coordinated replenishment (see Pooley and Stenger 1992). There are several methods to handle the flow of goods across the supply chain such as direct shipment, merge-in-transit, and cross-docking (see Brockmann 1999 and Kopczak 2000). Among them, cross-docking is considered to be the most efficient one to facilitate consolidation. Under cross-docking, shipments from the suppliers are received at an intermediary point (cross-dock facility) that usually does not hold stock. Upon receipt, they are broken and sorted and then shipped according to customers’ orders (see Gue 2001).

In addition to reducing logistics costs due to economies of scale, coordinated replenishment also enables the supplier to better allocate inventory among retailers through postponement of allocation. Maloney (2002) reports that by delaying the allocation of the inventory among retailers, more up-to-date information on the inventory status of the retailers can be incorporated. This creates flexibility in the
distribution network and allows a faster response to the retailers (see Cheung and Lee 2002). Federgruen and Zipkin (1984a) and Schwarz (1989) consider a postponement strategy in systems where the warehouse does not hold inventory and stock allocation is performed when the order arrives at the warehouse. In contrast, in this paper, postponement occurs as a result of coordinated replenishment, as the supplier waits for a certain level of demand to accumulate in the system before placing an order to replenish all the retailers simultaneously (see Lee and Tang 1997 for further discussion of delayed differentiation through product/process redesign). The allocation is based on the retailer inventory positions at the time the order is placed for the system, not the more up-to-date retailer inventory positions at the time of shipment. Moreover, in this paper a ship-all policy is used, which means outbound shipments are dispatched as soon as items arrive at the warehouse.

In this study, we investigate the value of coordinated replenishment in distribution systems. Our goal is to devise and analyze new policies for such systems and measure the effectiveness of these policies as well as other coordinated policies. We also compare these coordinated policies with a Non-Coordinated policy. While we mainly focus on total cost rate as the measure of effectiveness, we also measure the bullwhip effect associated with each policy.

The most commonly used coordinated replenishment/shipment policies in the literature can be classified as time based, quantity based, and time and quantity based (see Higginson and Bookbinder 1994). Under a time based policy, the system is replenished at pre-determined time epochs, while under a quantity based policy, the system is replenished when the total demand from the customers reach a certain level. Finally, under a time and quantity based policy, the system is replenished when either time or quantity specifications are satisfied. In this paper, we mainly study different variations of the first two types of policies.

Ordering in batches is considered one of the five main causes of the bullwhip effect (see Chen et al. 2000), which is the increase in demand variability as one moves up the supply chain. In our model, given the exogenous demand at the retailer level, the demand rate variance at the supplier, if larger than the retailer demand rate variance, represents the bullwhip effect. The policy one uses to replenish the
retailers, which determines the order size at the supplier along with the length of the order cycle, clearly impacts the variance of demand at the supplier (thus the bullwhip effect). Lower demand rate variance will lead to lower costs for the supplier. Cachon (1999) analyzes the supply chain demand variability in a distribution system with retailers implementing scheduled ordering policies and finds that switching from synchronized ordering to balanced ordering reduces supply chain holding/backorder costs. In this paper, we also analyze the variance of the demand rate at the supplier for different coordinated policies. We examine the effect of coordination on demand variability at the supplier by comparing our results with that of a non-coordinated replenishment policy.

In this paper we consider a centralized distribution system consisting of $N$ identical retailers and a warehouse managing a single product. Demand at the retailers is random and retailers hold stock of the item. The warehouse has full information about demand/inventory status and cost structure at retailers and is in charge of replenishing the stock for the system through an outside supplier who holds ample supply of the product. Warehouse holds no inventory and functions as the transit point for the physical goods. Moreover, it is responsible for making decisions as to when to replenish the system and how to allocate goods to each retailer. We also consider the case where inbound and the outbound trucks to and from the warehouse have capacity limitations.

Since the warehouse replenishes all the retailers at the same time and does not hold any inventory, the system will benefit from reduced shipping and warehouse inventory costs at the expense of possible increase in inventory related costs at the retailers. The form of the optimal warehouse replenishment/allocation policy for managing distribution systems in general, and the system described here in particular, is unknown. Therefore, our goal is to develop and study reasonable and practical policies that incorporate both inventory and transportation costs for such systems. In this paper, we propose and analyze a replenishment/allocation policy, which will be referred to as the hybrid policy throughout the paper, in addition to three other policies that are structured after well-known inventory policies studied in the literature for managing such systems.
Our contribution is three-folded: First, we propose a model for the analysis of coordinated replenishment and ordering policies. By comparing our model with the traditional inventory models (where each retailer independently places order from the outside supplier, with no warehouse in the system), we can quantify the value of coordination. Second, our model allows joint inventory and transportation decisions. By comparing the systems with and without transportation costs, our model quantifies the impact of transportation cost on the overall system. Third, we propose a new coordinated replenishment and shipment policy. We fully characterize the system dynamics (distributions for cycle time, inventory position, inbound and outbound shipping quantities, etc.) under this policy. Then we show by numerical tests that it performs better than the other three policies structured after well-known policies in the literature. We also identify the settings where each policy would perform well.

The rest of the paper is organized as follows: Section 2 surveys the related literature. In Section 3, we introduce the model, propose the policies, and define the necessary notation. Section 4 analyzes the various policies, and is followed by the numerical analysis in Section 5. Finally, we conclude with managerial insights and suggestions for future research.

2 Literature Review

Ordering/replenishment policies in inventory/distribution systems have been a major area of research in recent years. Most of the ordering/replenishment policies are based on either the echelon inventory or the installation inventory. Axsater and Rosling (1993) show that, in a serial inventory system, echelon stock policies are optimal and dominate the installation policies. However, in general, it can be shown that neither type of policy dominates the other. In the distribution systems (i.e., one supplier and multiple retailers), which typify many real world systems, the form of the optimal replenishment/stocking policy is unknown. The only study known to the authors, which attempts to identify the settings where echelon policies are superior to installation policies is that of Axsater and Juntti (1996) who evaluate the performance of these policies via simulation. For distribution models with no information sharing, which necessitates the use of installation stock policies, see Deurmeyer and Schwarz (1981), Lee and Moinzadeh (1987), Moinzadeh and Lee (1986), and Svoronos and Zipkin...


Among the studies mentioned above, Cheung and Lee (2002), Cetinkaya and Lee (2000), and Cachon (2001) are the closest to ours. Cheung and Lee (2002) analyze a distribution system with shipment coordination and stock rebalancing at the retailers using an echelon based replenishment policy for the warehouse. Cetinkaya and Lee (2000) also consider a quantity based shipment consolidation policy and incorporate the transportation costs into the model along with inventory related costs. Cachon (2001) analyzes three policies (variations of time and quantity based policies) for a system with one retailer, one warehouse, multiple products and also limited truck capacities.

In this paper, we propose a new ordering and dispatch policy referred to as the hybrid policy and compare it with three other policies structured after well-known inventory policies in the literature. The first one, referred to throughout the paper as Policy 0, is a special type of the Can-Order type policy (see Balintfy 1964, Ozkaya et al. 2003 and Federgruen, Groenevelt, and Tijm 1984). The other two are echelon based (quantity based) and fixed replenishment interval (time based) policies (referred to as Policy 1 and Policy 2 throughout the paper). Moreover, we also compare the hybrid policy to the
traditional model where retailers order independently using a continuous review \((Q, R)\) policy to see the impact of coordination.

Traditionally, transportation costs have not been jointly considered with inventory related costs, with the following exceptions. Federgruen and Zipkin (1984b) integrate routing and inventory decisions with capacitated vehicles. Similarly, Aviv and Federgruen (1998) consider fixed cost of shipment, with a capacity limit on the quantity to be shipped. Another paper by Federgruen and Zipkin (1984a) considers linear and fixed + linear ordering costs in their numerical tests (although they model any type of ordering cost structure). Federgruen and Zheng (1992) study joint replenishment with general ordering/transportation cost structure. Cachon (2001), Lee et al. (2003) assume transportation costs that depend on the number of capacitated trucks dispatched to replenish the retailers, rather than the number of units shipped, while Cetinkaya et al. (2000) consider a variable unit shipment cost. Chan et al. (2002), consider incremental and all unit discount transportation cost structures under a cross-docking setting. In this paper, we also incorporate limited truck capacities and transportation penalty costs and make joint inventory and transportation decisions.

3 The Model

We consider a centralized distribution system consisting of an outside supplier and \(N\) identical retailers. Each retailer faces a random demand that follows Poisson distribution, holds inventory, and fully backorders excess demand. In the coordinated replenishment system, there is a centralized decision maker responsible for ordering, receiving, allocating, and dispatching shipments to retailers. For ease of exposition, we assume that a warehouse serves this function. We also assume that the warehouse has access to information about demand, inventory levels and relevant costs at the retailers. The warehouse holds no inventory and orders from an outside supplier with ample supply. The transit times from the warehouse to each retailer, \(L\), and from the outside supplier to the warehouse, \(L_0\), are both constant. The constant transit time is a reasonable assumption due to the competitive environment in logistics with the emerging third party logistics that guarantees high level of service (on time deliveries with very little variation).
The warehouse is a stylized representation of a cross-dock operation. The drawback of keeping no warehouse inventory is that, after the warehouse places an order for the retailers, it takes a constant time, $L+L_0$, for the items to reach the retailers. So effectively the retailer lead time is $(L+L_0)$ longer than the lead time had the warehouse kept inventory. The advantage of keeping no warehouse inventory is that it eliminates holding cost at the warehouse. Moreover, the coordinated replenishment by the warehouse reduces ordering and transportation costs at the warehouse: Instead of placing orders in small batches, receiving and shipping small quantities for each retailer separately, the warehouse places a large order for all the retailers at the same time and receive all the inbound shipments at once.

To quantify the benefit of replenishment/shipment coordination in a distribution system, we compare systems where each retailer places orders from the outside supplier independently of each other (Non-Coordinated policy), with those where the warehouse employs coordinated replenishment for all the retailers and holds no inventory (e.g. the hybrid policy below).

**NON-COORDINATED POLICY**: Each retailer uses a continuous review $(Q,R)$ policy independent of one another and is responsible for its own replenishment process.

**HYBRID POLICY**: The warehouse orders to raise all the retailers’ inventory position to $S$ whenever any retailer’s inventory position drops to $s$ or the total demand at all the retailers reaches $Q$.

The hybrid policy is based on both the installation and the echelon inventory positions. The system replenishment may be triggered in two ways: (1) installation trigger: whenever any retailer’s inventory position drops to $s$; or (2) echelon trigger: whenever the echelon inventory position drops to $NS-Q$. The values of $s$, $S$, and $Q$ will be optimally determined. When a retailer’s inventory position first drops to $s$, it sends a signal that other retailers may soon need to be replenished. The hybrid policy then orders to bring every retailer’s inventory position back to $S$. Hence, the installation trigger in the hybrid policy is proactive in the sense that most retailers will be replenished before their reorder point is reached so as to dampen the effect of longer lead time caused by the lack of warehouse inventory. The echelon trigger in the hybrid policy is designed to reduce the transportation penalty cost, because it reduces the
variation in the total quantity shipped to the warehouse from the outside supplier. We will verify this numerically in Section 5.

To better understand the performance of the proposed hybrid policy within the coordinated replenishment framework, we also analyze and compare the following three well-known policies:

**POLICY 0:** The warehouse orders to raise all the retailers’ inventory position to $S$ whenever any retailer’s inventory position drops to $s$.

**POLICY 1:** The warehouse orders to raise all the retailers’ inventory position to $S$ whenever the total demand at all the retailers reaches $Q$.

**POLICY 2:** The warehouse orders to raise all the retailers’ inventory position to $S$ every $T$ time units.

Policies 0 and 1 are special cases of the hybrid policy, where $Q$ and $S-s$, respectively, are sufficiently large. While for both policies the total replenishment order/transportation quantity is echelon based, the trigger for ordering is installation based for Policy 0 and echelon based for Policy 1. Furthermore, Policy 0 is also a special case of the “can-order policies”, first introduced by Balintfy (1964), where the can-order levels are equal to the order-up-to level for all the retailers. Policy 2 is a time-based policy, analogous to the periodic-review order-up-to policy. Similar to the hybrid policy, Policy 0 is proactive in reducing retailer’s shortage as whenever one retailer triggers the order, all other retailers are replenished, even though their inventory positions are above the reorder point. In contrast, Policies 1 and 2 are echelon based and opt towards reducing ordering costs.

For all the policies in this paper, the objective is to minimize the cost rate of the supply chain, which consists of ordering/shipment costs, and holding and shortage costs at the retailers. We assume that holding and shortage costs are linear. For the ordering/shipment costs, we will assume a fixed cost for amounts within a pre-fixed limit (transportation capacity), and variable penalty cost for units beyond this limit. This is a reasonable assumption as carriers generally charge on the basis of full truck load (FTL) and mileage as long as all the units fit in the truck. When the truck limit is exceeded, the warehouse pays extra fee to the carrier.
Many inventory models in the literature do not consider the transportation costs. Our model, however, jointly consider the inventory and the transportation costs simultaneously. This can be used to discern the effect of transportation costs. We will do the numerical analysis in Section 5 with and without the transportation costs. We define the quantities shipped from the outside supplier to the warehouse as “inbound” and those from the warehouse to the retailers (or from the supplier to the retailers under the Non-Coordinated policy) as “outbound”. Note that the inbound quantity in every order cycle is constant under Policy 1, but random under the other three policies. This will play a critical role in the numerical comparisons with and without the transportation penalty costs in Section 5. Moreover, because the outbound quantities under Non-Coordinated policy are constant, the benefit of using coordinated shipment policies (relatively to Non-Coordinated policy) should decrease when transportation penalty costs are included. Among the coordinated shipment policies, the inclusion of transportation penalty costs will improve the relative performance of the hybrid policy and Policy 1 in most cases as $Q$ can be optimized to minimize the transportation cost penalties. As a result, whether and how the warehouse incurs additional transportation costs may have a significant impact on which policy should be employed.

We define the relevant notation in Table 1. We use subscript “0” for the warehouse related and “$i$” ($i=1,...,N$) for retailer related parameters and variables. Also, we define $\Delta=S-s$.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_0$</td>
<td>Transit time from the supplier to the warehouse under coordinated replenishment</td>
</tr>
<tr>
<td>$L_i$</td>
<td>Transit time from the warehouse to retailer $i$ under coordinated replenishment</td>
</tr>
<tr>
<td>$l_i$</td>
<td>Transit time from the supplier to retailer $i$ (under Non-Coordinated policy)</td>
</tr>
<tr>
<td>$LT$</td>
<td>Total leadtime for any retailer</td>
</tr>
<tr>
<td>$T_{\Delta}^i$</td>
<td>Time of the occurrence of $\Delta$th demand for retailer $i$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Time between two consecutive replenishments</td>
</tr>
<tr>
<td>$IP_i(t)$</td>
<td>Retailer $i$’s inventory position at time $t$</td>
</tr>
<tr>
<td>$IL_i(t)$</td>
<td>Retailer $i$’s inventory level at time $t$</td>
</tr>
<tr>
<td>$Z_0$</td>
<td>The total amount shipped to the retailers in an order cycle under coordinated replenishment (random variable)</td>
</tr>
<tr>
<td>$Z_i$</td>
<td>The amount shipped to retailer $i$ in an order cycle (random variable)</td>
</tr>
<tr>
<td>$D_i(t)$</td>
<td>Demand at retailer $i$ for a period of time $t$</td>
</tr>
<tr>
<td>$N$</td>
<td>Total number of retailers in the system</td>
</tr>
<tr>
<td>$\lambda_i$</td>
<td>Mean demand rate at retailer $i$ ($i=1,...,N$)</td>
</tr>
<tr>
<td>$K_0$</td>
<td>Fixed ordering cost per order at the warehouse</td>
</tr>
<tr>
<td>$K_i$</td>
<td>Fixed cost of a shipment to retailer $i$</td>
</tr>
</tbody>
</table>
Table 1: Notation

<table>
<thead>
<tr>
<th>notation</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_i$</td>
<td>Fixed cost of a shipment to retailer $i$ from the supplier under Non-Coordinated policy</td>
</tr>
<tr>
<td>$C_0$</td>
<td>Maximum capacity of a truck from the outside supplier to the warehouse</td>
</tr>
<tr>
<td>$C_i$</td>
<td>Maximum capacity of a truck from the warehouse to retailer $i$</td>
</tr>
<tr>
<td>$\pi_i$</td>
<td>Unit backorder cost/time at retailer $i$</td>
</tr>
<tr>
<td>$h_i$</td>
<td>Unit inventory holding cost/time at retailer $i$</td>
</tr>
<tr>
<td>$f(\cdot)$</td>
<td>Probability density function of $\tau$</td>
</tr>
<tr>
<td>$F(\cdot)$</td>
<td>Cumulative density function of $\tau$</td>
</tr>
<tr>
<td>$g(n,C_0)$</td>
<td>Penalty cost the warehouse pays to the carrier when the inbound quantity is “$n$” units</td>
</tr>
<tr>
<td>$g_i(n,C_i)$</td>
<td>Penalty cost retailer $i$ pays to the carrier when the outbound quantity is “$n$” units</td>
</tr>
<tr>
<td>$CR(\cdot,\cdot)$</td>
<td>Cost rate</td>
</tr>
<tr>
<td>$a_0$</td>
<td>Inbound unit penalty cost under coordinated replenishment</td>
</tr>
<tr>
<td>$a_i$</td>
<td>Outbound unit penalty cost for retailer $i$</td>
</tr>
</tbody>
</table>

Furthermore, we let $x^+ = \max(x,0)$, $a \wedge b = \min(a,b)$, and, as in Hadley and Whitin (1963),

$$p(j;\mu) = e^{-\mu} \frac{\mu^j}{j!} \quad \text{and} \quad P(r;\mu) = \sum_{j=r}^{\infty} p(j;\mu).$$

4 Analysis

The Non-Coordinated policy is the traditional $(Q,R)$ policy. Interested readers can find detailed analysis in Section 4-7 of Hadley and Whitin (1963). For the coordinated replenishment policies, since we assume all retailers are identical, we sometimes suppress the retailer subscript $i$. The analysis applies to heterogeneous retailers as well, but the notation is much more complex.

4.1 Analysis of the Hybrid Policy

When an order is triggered, all the retailers’ inventory positions are raised to $S$. The system as represented by all the retailers’ inventory positions, therefore, is a regenerative process, and the order epochs are the regenerative points. We denote the time between two consecutive order epochs by $\tau$, and refer to it as the “cycle time”. We first derive the probability distribution of $\tau$. The time between two consecutive orders is the minimum of the time for the first retailer to realize $\Delta$ units of demand and the time to realize a total of $Q$ units of demand at all the retailers. We define the latter as $T_Q$. Therefore, we have the following identity:

$$\tau = \min(T^{\Delta}_1,\ldots,T^{\Delta}_N,T_Q),$$

(1)
where \( T_{\Delta}^i \sim \text{Erlang}(\Delta, \lambda) \) for all \( i \) and \( Q \sim \text{Erlang}(Q, N\lambda) \). While the \( T_{\Delta}^i \)'s are i.i.d., \( T_Q \) is dependent on the \( T_{\Delta}^i \)'s. Let \( D_i \) denote the random demand at retailer \( i \) during a cycle (excluding the possible triggering demand), then we have the following relation:

\[
F(t) = \Pr(\tau \leq t) = \Pr(t > T) = 1 - \Pr(D_1(t) \leq \Delta - 1, \ldots, D_N(t) \leq \Delta - 1, \sum_{i=1}^{N} D_i(t) \leq Q - 1)
\]

(2)

where \( U_j = \min(\Delta - 1, Q - 1 - d_1 - \ldots - d_{j-1}) \), with \( U_1 = \min(\Delta - 1, Q - 1) \).

Proofs of the lemmas and propositions in this section are provided in the Appendix.

**Lemma 1:** The probability density function of \( \tau, f(.) \), is given by:

\[
f(t) = N\lambda \sum_{d_1=0}^{U_1} \sum_{d_2=0}^{U_2} \cdots \sum_{d_N=0}^{U_N} \prod_{i=1}^{N-1} p(d_i; \lambda t)p(U_N; \lambda t)
\]

(3)

From (3) we note that the distribution of \( \tau \) depends on \( Q \) and \( \Delta \) (the difference between \( S \) and \( s \)), but not on \( S \) and \( s \) individually.

We now proceed to analyze the probability distribution of the inbound quantity, \( Z_0 \). Later we will use this distribution to find the expected cycle time, the distribution of the inventory position at the retailers, the outbound quantity, and demand during a cycle. Although it is possible to obtain these measures directly, our derivations via the inbound quantity distribution significantly reduces the computational complexity. Let \( Z_i \) be the random outbound quantity to retailer \( i \). Note that \( Z_i = D_i + 1 \) if retailer \( i \) triggers the order and \( Z_i = D_i \), otherwise. Therefore, \( D_i \in [0, \Delta - I] \), \( Z_i \in [0, \Delta] \), \( \forall i \). We examine two cases: (1) \( Q > \Delta + (N-I)(\Delta-I) \) and (2) \( Q \leq \Delta + (N-I)(\Delta-I) \). Let \( \Phi \) denote the inbound quantity in the first case. Hence, \( \Phi = \sum_{i=1}^{N} Z_i \), and \( \Phi \in [\Delta, \Delta + (N-I)(\Delta-I)] \). In this case, the echelon trigger will never be initiated so we can ignore \( Q \). (In fact, the hybrid policy reduces to Policy 0, which will be analyzed in more details later.)

**Lemma 2:** When \( Q > \Delta + (N-I)(\Delta-I) \), for any retailer \( i \),
Next, we examine the second case where \( Q \leq \Delta + (N-1)(\Delta - 1) \). Let \( Z_0 \) denote the inbound quantity at the warehouse. The random variable \( Z_0 \) is in the range \([\min(Q, \Delta), Q]\). We will use a coupling argument to derive \( Z_0 \) from \( \Phi \). Let there be two systems facing the same demands. In System 1, \( Q_1 > \Delta + (N-1)(\Delta - 1) \) so \( Q_1 \) can be ignored. In System 2, \( Q_2 \leq \Delta + (N-1)(\Delta - 1) \). All other parameters are identical. For any demand realization, if it results in an inbound quantity \( \Phi > Q_2 \) in System 1, then this demand realization would certainly initiate the echelon trigger in System 2 so \( Z_0 = Q_2 \). Conversely, if a demand realization initiates an installation trigger in System 2, then it would initiate the same trigger in System 1. Hence we have the following result:

**Lemma 3:** When \( Q \leq \Delta + (N-1)(\Delta - 1) \), the inbound quantity \( Z_0 \) satisfies:

\[
Z_0 < Q \iff \Phi < Q \quad \text{and} \quad Z_0 = Q \iff \Phi \geq Q.
\]

Therefore,

\[
\Pr(Z_0 = n) = \begin{cases} 
\Pr(\Phi = n) & \text{for } \min(\Delta, Q) \leq n < Q \\
\sum_{z \geq Q} \Pr(\Phi = z) & \text{for } n = Q
\end{cases}.
\]

(5)

Now that we have the probability distribution for the inbound quantity under the hybrid policy, we use the following relation to find the expected cycle time:

**Lemma 4:**

\[
E[\tau] = \frac{E[Z_0]}{N\lambda}
\]

(6)

We will use the following recursive relation (based on \( Q \)) to compute the probability distribution of \( Z_i \) (again we use \( Z \) to simplify notation):

**Lemma 5:**
\[ P(Z \geq n | Q, \Delta, N) = \begin{cases} P(Z \geq n | Q-1, \Delta, N) + \varphi(Q, \Delta, N) & n = 0,1,\ldots,\min(\Delta, Q-1) \\ \left( \frac{1}{N} \right)^2 & n = \min(\Delta, Q) = \Delta \\ \left( \frac{1}{N} \right)^2 + \varphi(Q, \Delta, N) & n = \min(\Delta, Q) = Q \end{cases} \] (7)

where,

\[ \varphi(Q, \Delta, N) = \lambda \left( \frac{Q-1}{n-1} \right) \left( \frac{1}{N} \right)^{n-1} \left( 1 - \frac{1}{N} \right)^{Q-1-n} \left( E[\tau | Q + 1 - n, \Delta, N - 1] - E[\tau | Q - n, \Delta, N - 1] \right). \]

The recursive approach above is extremely useful for computing the probability distribution of \( Z \) for large \( Q, \Delta, \) and \( N \), because the direct method requires the consideration of all possible combinations of demands at all the retailers that would make the summation domain prohibitively large and the problem computationally intensive for large \( N, Q, \) and \( \Delta \).

The expression for \( E[\tau] \) in (6) is sufficient to derive the ordering cost at the warehouse. Now, in order to find the holding and backordering costs at the retailer, we need to find the probability distribution of the inventory level at each retailer at any point in time. As orders do not cross, we can use the following identity: \( IP(t-LT) - D(LT) = IL(t) \), where \( LT = L_0 + L \) is the lead time. Therefore, to derive the distribution of the \( IL \), it suffices to derive the probability distribution for the \( IP \). We start by conditioning on the demand at any retailer during the cycle time, \( \tau \). It is well known (Ross 1993, page 223) that for a Poisson process conditioned on \( n \) arrivals during time \((0, \tau)\), the arrival times \( S_1, S_2, \ldots, S_n \) have the same distribution as the order statistics of \( n \) independent random variables uniformly distributed on the interval \((0, \tau)\). Therefore, we have the following conditional probability distribution of the inventory position:

\[ \Pr(IP = j | D(\tau) = n) = \frac{1}{n+1} \text{ for } j = S - n, S - n + 1, \ldots, S. \] (8)

Here, \( D(\tau) \) denotes the demand faced by any retailer during the cycle time excluding the possible triggering demand (as soon as inventory position drops to \( s \), it is raised to \( S \), so the inventory position equals to \( s \) with probability 0). The following lemma gives the probability distribution of \( D(\tau) \) using (8):

**Lemma 6:** For any retailer, when \( Q = 1 \), \( P(D \geq n | Q) = \begin{cases} 1 & n = 0 \\ 0 & n \geq 1 \end{cases} \); and when \( Q \geq 2 \),
\[
P(D \geq n \mid Q) = \begin{cases} 
1 & n = 0 \\
\frac{P(Z \geq n \mid Q-1)}{n+1} & 1 \leq n \leq \min(\Delta-1, Q-1) \\
0 & n \geq \min(\Delta, Q) 
\end{cases}
\] (9)

Using Lemma 6, we find the probability distribution of the inventory position to be

\[
Pr(IP = j) = \sum_{n=S-j}^{\min(S,Q-1)} \frac{P(D = n)}{n+1}, \quad j = S - U_1, \ldots, S.
\] (10)

Lemmas 1-6 provide us with the framework to calculate the total system cost per unit time.

**Proposition 1:** The cost rate of the system under the hybrid policy is as follows:

\[
CR = \frac{K}{E[\tau]} + \frac{K + \sum_{i=2}^{N} K \Pr(Z \geq 1)}{E[\tau]} + N^* \left( (h + \pi) E[IL^*] + \pi (E[D(LT)] - E[IP]) \right)
\]

\[
+ \frac{1}{E[\tau]} \left( \sum_{n=\min(\Delta,Q)}^{Q} g(n; C_0) \Pr(Z = n) + \sum_{j=1}^{\min(\Delta,Q)} \sum_{n=0}^{\min(S,n)} g_j(n; C) \Pr(Z = n) \right).
\] (11)

**Proposition 2:**

1. For any fixed value of \(\Delta\) and \(Q\), \(CR\) is convex in \(S\).
2. For any fixed value of \(\Delta\) and \(Q\), the optimal \(S\) is the largest integer satisfying the following inequality:

\[
\sum_{j=\max(1,S-Q+1)}^{S} \Pr(IP = j)(1 - P(j; \lambda LT)) \leq \frac{\pi}{h + \pi}.
\] (12)

In the next three sub-sections, we analyze Policies 0, 1, and 2. The inbound quantity is constant \((Q)\) under Policy 1 and random under the hybrid policy, Policy 0, or Policy 2. Since Policy 2 utilizes demand pooling, its inbound quantity is less variable than those under the hybrid policy and Policy 0, even though they are also bounded from above (by \(Q\) for the hybrid policy and \(\Delta+(N-1)(\Delta-1)\) for Policy 0). As a result, one would expect Policy 1 to have more certainty and more control over its inbound transportation cost, and therefore further reduce its cost rate and start performing relatively better. For example, if it is economically viable to do so, \(Q\) can be optimized under Policy 1 so that the inbound truck capacity is never exceeded.

The outbound quantities are random for all policies (except for Non-Coordinated policy). Under the hybrid policy and Policy 0, the outbound quantities are doubly stochastic due to random inbound quantity and random disaggregation of demand, but they have a maximum of \(\Delta\). Under Policy 1 the
outbound quantities also have a maximum of \( Q \), whereas there is no upper limit on the outbound quantity under Policy 2. Detailed comparison of the four policies will be done numerically in Section 5.

### 4.1.1 Special Case of the Hybrid Policy: Policy 0

When \( Q \) is sufficiently large (i.e., \( Q > \Delta + (N-1)(\Delta - 1) \)), the hybrid policy reduces to Policy 0. As a result, the probability distributions of the demand during the cycle and the outbound quantity are given in Lemma 2, the distribution of the inbound quantity is given as \( \Phi = \sum_{i=1}^{N} Z_i \), and the probability distribution of the inventory position can be derived using (10). Moreover, we can simplify the probability distribution of \( \tau \) from (3) and the cost rate from Proposition 1.

\[
\begin{align*}
    f(t) &= N \lambda p(\Delta - 1; \lambda t) \left[ 1 - P(\Delta; \lambda t) \right]^{(N-1)}, t \geq 0. \\
    \text{Proposition 3: The cost rate of the system under Policy 0 is as follows:} \\
    CR &= \frac{K_0}{E[\tau]} + \frac{K + \sum_{i=2}^{N} K \Pr(Z_i \geq 1)}{E[\tau]} + N \left( (h + \pi)E[LT] + \pi \left( E[D(LT)] - E[IP] \right) \right) \\
    &\quad + \frac{1}{E[\tau]} \left( \sum_{n=2}^{\Delta - (N-1)(\Delta - 1)} g(n; C_0) \Pr(\Phi = n) + \sum_{n=2}^{\Delta - 1} g_i(n; C) \Pr(Z = n) + g_i(\Delta; C) \right). \\
\end{align*}
\]

### 4.1.2 Special Case of the Hybrid Policy: Policy 1

Since the retailers have independent Poisson demands with rate \( \lambda \), the total demand from all the retailers is also Poisson with rate \( N\lambda \). Therefore, the time to observe \( Q \) units of demand from the retailers has an Erlang distribution: \( \tau \sim \text{Erlang} (Q, N\lambda) \) with the expected value of \( E[\tau] = Q/(N\lambda) \). So we have:

\[
\begin{align*}
    \Pr(IP = j) &= \frac{1}{Q} \sum_{n=S-j}^{\min(S, Q)} \left( \frac{1}{N} \right)^{S-j} \left( 1 - \frac{1}{N} \right)^{n-S+j}, S - Q + 1 \leq j \leq S. \\
    \text{Lemma 7: Under Policy 1,} \\
    \Pr(Z = n) &= \binom{Q}{n} \left( \frac{1}{N} \right)^{\min(N, Q)} \left( 1 - \frac{1}{N} \right)^{Q-n}, 0 \leq n \leq Q. \\
    \text{Proposition 4: The cost rate of the system under Policy 1 is as follows:}
\end{align*}
\]
4.2 Analysis of Policy 2

Under Policy 2, the warehouse places an order with the outside supplier every $T$ time units ($\tau=T$) to bring the inventory position of every retailer to $S$. Therefore, Policy 2 resembles the periodic review, order-up-to, $<R,T>$ policy in Hadley and Whitin (1963). Under Policy 2, both the inbound and the outbound quantities ($Z_0$ and $Z$) are Poisson, with rates of $N\lambda T$ and $\lambda T$, respectively. Derivation of $B(1,S-1,T)$ can be found in Hadley and Whitin (page 260). For completeness, we also describe $B(1,S-1,T)$ in the Appendix.

Proposition 5: The cost rate of the system under Policy 2 (c.f. Hadley & Whitin 1963, p. 260) is:

$$CR = \frac{K_0 N\lambda}{Q} + \frac{K + \sum_{i=2}^{N} K \Pr(Z_i \geq 1) N\lambda}{Q} + N\pi \left[ h + \pi E[IL] + \pi E[D(LT)] - E[IP] \right]$$

$$+ \left\{ \frac{1}{E[\tau]} \left[ g(Q,C_0) + \sum_{i=1}^{N} \sum_{n=0}^{\infty} g_i(n;C) \Pr(Z = n) \right] \right\}$$

4.2 Analysis of Policy 2

Under Policy 2, the warehouse places an order with the outside supplier every $T$ time units ($\tau=T$) to bring the inventory position of every retailer to $S$. Therefore, Policy 2 resembles the periodic review, order-up-to, $<R,T>$ policy in Hadley and Whitin (1963). Under Policy 2, both the inbound and the outbound quantities ($Z_0$ and $Z$) are Poisson, with rates of $N\lambda T$ and $\lambda T$, respectively. Derivation of $B(1,S-1,T)$ can be found in Hadley and Whitin (page 260). For completeness, we also describe $B(1,S-1,T)$ in the Appendix.

Proposition 5: The cost rate of the system under Policy 2 (c.f. Hadley & Whitin 1963, p. 260) is:

$$CR = \frac{K_0}{T} P(1; N\lambda T) + N \frac{K}{T} P(1; \lambda T) + N \left\{ h \left( S - \lambda * LT - \frac{\lambda T}{2} \right) + (h + \pi) B(1,S-1,T) \right\}$$

$$+ \frac{1}{T} \left\{ \sum_{n=0}^{\infty} g(n;C_0) p(n;N\lambda T) + \sum_{i=1}^{N} \sum_{n=0}^{\infty} g_i(n;C) p(n;\lambda T) \right\}$$

5 Numerical analysis

We first investigate the performance of different coordinated replenishment policies via a numerical test. Our goal is to compare the performance of the hybrid policy against Policies 0, 1, and 2 and to identify the settings in which each policy performs well. In Section 5.1.1, we present our results for the case with no transportation penalty costs (no truck space limitations). This will reveal the effectiveness of the various coordination policies from a pure ordering/inventory perspective. Then we include the transportation penalty costs in Section 5.1.2. Clearly, the hybrid policy dominates Policies 0 and 1, which are special cases of the hybrid policy. Furthermore, as the hybrid policy and Policy 0 are both installation and echelon based, they are more responsive to the balance of inventory and shortage costs than the echelon based Policies 1 and 2. Policies 1 and 2 tend to order less frequently and reduce the ordering cost at the expense of higher shortage costs. On the other hand, one potential pitfall of the hybrid
policy and Policy 0 is that a replenishment can be triggered when only a few retailers actually need it. This can lead to higher inventory costs at the retailers. The overall cost depends heavily on the parameters. In most cases, we expect both the hybrid policy and Policy 0 to outperform the other two policies. This is verified numerically in Section 5.1.1. As we will see in Section 5.1.2, when additional transportation penalty costs are included, this may no longer hold.

In Section 5.2, we quantify the value of coordinated replenishment by comparing the hybrid policy with the Non-Coordinated policy. We conjecture that in general coordinated policy (hybrid policy) has a lower cost rate, but the limitation on truck capacities favors the Non-Coordinated policy as it has more control over its shipment size, a constant $Q$, while the hybrid policy has random shipment size. In Section 5.3, we compare the bullwhip effect under all five policies by comparing the variance of demand rate at the supplier. Finally in Section 5.4 we present insights on supply chain design issues when the hybrid policy is used.

5.1 Numerical Analysis within Coordinated Replenishment Framework

5.1.1 Numerical Analysis without the Transportation Penalty Costs with Coordinated Replenishment

There are situations where the transportation capacity is rarely exceeded and the transportation penalty costs can be ignored. One example is Wal-Mart, who has its own fleet and is responsible for its own transportation. Then the relevant costs are ordering, inventory holding and backorders. In our numerical analysis, we have observed that the probability of having at least one order from any retailer ($P(Z \geq 1)$) is almost always 1 at optimality. Therefore, we let $K=0$ and increase $K_0$ by $NK$. The following set of parameters is used in the analysis:

$\lambda = 1, L = 1, \frac{L}{T} \in \{0.5, 1, 2\}, h = 1, \frac{h}{\pi h} \in \{0.9, 0.95, 0.99\}, K_0 \in \{2, 5, 10, 100\}, N \in \{5, 10, 20\}$.

For all policies, the optimal value of $S$ is calculated given other variables ($Q$ and $\Delta$ for the hybrid policy, $\Delta$ for Policy 0, $Q$ for Policy 1, and $T$ for Policy 2). Then numerical search is carried out over
possible values of $Q$, $\Delta$, and $T$. While the convexity of the cost function over $\Delta$ (or $Q$, $T$) is not clear, our numerical experience indicates that it is satisfied for all the cases considered.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$S^*$</th>
<th>$\Delta^*$</th>
<th>$Q_{\text{hybrid}}^*$</th>
<th>$Q_{\text{policy}1}^*$</th>
<th>$T^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_0$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
</tr>
<tr>
<td>$L_0/L$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
</tr>
<tr>
<td>$\pi/(\pi+h)$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
<td>$\nabla$</td>
</tr>
</tbody>
</table>

Table 2: Trends in average optimal values of decision variables

In total, we consider 81 cases. In Table 2, we present the behavior of the optimal decisions for each policy. To compare the effectiveness of our proposed hybrid policy with the other three policies, we compute the deviation of Policies 0, 1 and 2 from the hybrid policy as follows:

$$% \text{deviation for Policy } i = \frac{\text{CR(Policy } i\text{)} - \text{CR(hybrid)}}{\text{CR(hybrid)}} \cdot i = 0, 1, 2.$$ 

The optimal decisions reflect the tradeoff between inventory costs and the fixed ordering cost, so it is not surprising that the supplier tends to order less frequently and carry more inventory at the retailer level (higher $S^*$, $\Delta^*$, $T^*$) in systems with few retailers, large transit time and/or fixed ordering cost. Moreover, order frequency/inventory levels go up as shortages become more expensive. On average, optimal $Q^*$ in Policy 1 exhibits the same behavior as $\Delta^*$ except that it increases in systems with more retailers (higher $N$) as the average total demand rate faced by the system is increased. However, the average $Q^*$ exhibits a different behavior (first decreases and then increases) under the hybrid policy with respect to $L_0$ (we should note that $Q^*$ increases in most cases (60%) when $L_0$ increases) and $\pi$. This could be due to the interaction between $\Delta^*$ and $Q^*$. The overall joint effect of $\Delta^*$ and $Q^*$ is that the average cycle time, $E[\tau]$, increases (decreases) for higher $L_0(\pi)$, as expected.

It is known that echelon based joint replenishment policies are more effective when a supplier is responsible for replenishing many retailers (Cheung and Lee 2002). Our numerical results in Figure 1 verify this. That is, as $N$ increases, the average % deviation from the hybrid policy decreases for Policies 1 and 2, but increases for Policy 0. Recall that the major difference between the hybrid policy and Policies 1 and 2 is the installation-based trigger $\Delta$ in the hybrid policy, which is more protective of retailer
shortages. Therefore, for smaller values of \( \pi \), the relative advantage of the hybrid policy over Policies 1 and 2 diminishes. On the other hand, since order triggering is based solely on installation inventory position under Policy 0, we expect that Policy 0 is more responsive to shortages as the hybrid policy also focuses on echelon inventory position with the additional variable, \( Q \). This is reflected by the observation (see Figure 2) that \( \Delta^* \) for Policy 0 is no greater than that of the hybrid policy (Policy 0’s relative performance is better for higher \( \pi \)).

In accordance with Axsater and Juntti’s (1996) claim that echelon based policies dominate installation policies in systems where \( L_0 \) is large relative to \( L \), we also observe that average % deviations are smaller for Policies 1 and 2 for higher \( L_0 \) (see Figure 3). This could be due to the increase in the inventory levels/reorder point to mitigate the effect of longer lead time, which has more pronounced impact on the hybrid policy and Policy 0. The change in the relative performance of Policy 0 compared to the hybrid policy due to an increase in \( L_0/L \) does not seem to be significant. In Figure 4, we note that as \( K_0 \) increases, \( Q^* \) under the hybrid policy increases substantially (to balance the ordering cost), and Policy 0’s deviation from the hybrid policy is small (since when \( Q^* \) is relatively large, the hybrid policy and Policy 0 perform very closely). Furthermore, in systems with large ordering costs, order quantity/cycle time
increases making Policies 1 and 2 even more vulnerable to shortages, and therefore Policies 1 and 2 perform relatively worse (see Figure 4).

In summary, on average the hybrid policy has the most improvement over Policy 0 for large $N$, $K_0$, and small $\pi$ (the effect of $L_0$ is insignificant); and the hybrid policy has the most improvement over Policies 1 and 2 for large $K_0$, $\pi$ and small $L_0$, $N$. On a case-by-case basis, we observe that the hybrid policy dominated all other three in all cases. Policy 0 dominates Policy 1 in all but 10 cases. Moreover, Policy 2 is dominated by Policies 0 and 1 in all cases.

5.1.2 Numerical Analysis with Transportation Penalty Costs and Coordinated Replenishment

In our analysis in Section 4, we allow $g(\cdot)$, the transportation penalty cost, to have a general form. As long as this function is convex in $\Delta$, $Q$, or $T$, its specific form does not change the optimization procedure. In this section only, we assume a linear penalty cost function due to its simplicity (see Federgruen and Zipkin 1984a for fixed+linear ordering costs). Therefore, the transportation/ordering cost function is a piecewise linear convex one. It is reasonable when the warehouse uses an outside carrier who charges a fixed fee for quantity below a certain limit, and additional per-unit fee for quantity in excess of the limit; or when the warehouse uses its own fleet with a fixed transportation capacity and uses an outside carrier for additional units, who charges on a per-unit basis. Thus,

$$g(n;C_0) = (n - C_0)^+ \alpha_0 \quad \text{and} \quad g_i(n;C) = (n - C)^+ \alpha$$ (19)

For the inbound truck capacity, we let $C_0 = Q^* \gamma_0$, where $Q^*$ is the optimal value for Policy 1 that we calculated in Section 5.1.1 for each case, and $\gamma_0$ is the “inbound capacity factor”. Similarly, the outbound truck capacity is set at $C = C_0 \gamma$, where $\gamma$ is the “outbound capacity factor”. Below is a summary of all the parameter values:

$$h = 1, \lambda = 1, L = 1, \frac{h \alpha}{\lambda} \in \{0.5, 1, 2\}, N \in \{5, 10, 20\}, \frac{\pi}{\alpha} \in \{0.95, 0.99\},$$
$$K_0 \in \{2, 50, 100\}, \alpha \in \{0.5, 2\}, \frac{\pi}{\alpha} \in \{0.5, 1, 2\}, \gamma_0 \in \{0.5, 1, 1.5\}, \gamma \in \{1/N, 1.2/N\}.$$
In all, a total of 1944 cases are considered. The optimization procedure is almost identical to that in 5.1.1. The only difference is in the computation of the total cost (due to additional transportation cost).

The impact of \( N, K_0, L_0, \) and \( \pi \) on the average values of \( S^*, \Delta^*, Q^*, \) and \( T^* \) is mostly the same as in Section 5.1.1. However, with the introduction of transportation penalty, at optima all policies tend to order more frequently (decreasing \( \Delta^*, Q^*, \) and \( T^* \)) to reduce the likelihood of exceeding the truck capacities. If it is economically preferable to decrease order frequency in Section 5.1.1 due to a change in a particular parameter, then the increase in \( \Delta^*, Q^*, \) and \( T^* \) will be smaller here, in the case of penalty. Furthermore, for all policies, the tendency is to reduce the inbound/outbound quantities by replenishing more frequently (decreasing \( \Delta^*, Q^*, \) and \( T^* \)) in systems with tight transportation capacity constraints and high unit penalty cost (i.e., small \( \gamma_0, \gamma, \) and large \( \alpha_0, \alpha \)). As a result of the increase in order frequency, average \( S^* \) is also reduced to avoid high inventory holding cost. The only exception is that, under the hybrid policy, \( \Delta^* \) first decreases and then increases when \( \gamma_0 \) and \( \alpha_0 \) increase (due to interaction between \( Q \) and \( \Delta \)).

Previously we conjecture that including the limited transportation capacity and adding the transportation penalty to the total cost function would improve the relative performance of Policy 1 (which has a fixed inbound quantity \( Q \)), and degrade the relative performance of Policies 0 and 2 (which have random inbound quantities), with the hybrid policy in between (which has a random inbound quantity, but upper-limited by \( Q \)). Our numerical results confirm our intuition: we observe that Policy 0 has the highest variance in inbound quantity, followed by Policy 2, the hybrid policy and Policy 1 (which has zero variance). As a result, average deviation of Policy 1 is lower here than in Section 5.1.1, while those for Policies 0 and 2 are higher (the randomness in the inbound quantity hurts Policy 0 more than Policy 2). As Table 3 also suggests, relative performances of Policy 1 and the hybrid policy improve when inbound capacity is tight. Although, average deviations do not differ significantly, maximum deviations can change drastically with the inclusion of transportation penalty. Clearly, Policy 1’s performance shows less variation than the other two policies. Moreover, we observe that in the presence
of penalty costs, Policies 1 and 2 outperform Policy 0 in 37% and 8% of the cases, respectively, a large increase from 12% and 0% in Section 5.1.1.

<table>
<thead>
<tr>
<th>% Deviations from hybrid policy</th>
<th>Policy 0</th>
<th>Policy 1</th>
<th>Policy 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>without penalty</td>
<td>Min. 0</td>
<td>0.17</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Avg. 0.23</td>
<td>1.77</td>
<td>4.48</td>
</tr>
<tr>
<td></td>
<td>Max. 1.04</td>
<td>6.53</td>
<td>13.32</td>
</tr>
<tr>
<td>with penalty</td>
<td>Min. 0</td>
<td>0</td>
<td>0.04</td>
</tr>
<tr>
<td>(Section 5.2)</td>
<td>Avg. 1.21</td>
<td>1.45</td>
<td>4.74</td>
</tr>
<tr>
<td></td>
<td>Max. 25.48</td>
<td>6.18</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3: The minimum, average, maximum deviations with/without penalty costs

Figures 5 and 6 illustrate the average deviations from the hybrid policy with respect to inbound capacity factor and unit penalty (the trend is similar for outbound unit penalty). As expected, when the inbound capacity becomes a critical factor (small \( \gamma_0 \) and large \( \alpha_0 \)), Policy 1 (0) starts to perform relatively better (worse). Numerically, Policy 1 chooses an optimal \( Q^* \) that’s less than the inbound capacity about 75% of the time to completely avoid inbound penalty (63% for the hybrid policy). Now with respect to the outbound capacity, both the hybrid policy and Policy 0 have an upper limit \( \Delta^* \). Therefore, when \( \gamma \) is low, both policies take advantage of this upper bound to reduce outbound penalty cost (however, the change in performance is negligible). Considering the impact of different parameters, we observe that \( N, \pi, K_0 \) and \( L_0 \) have almost the same effect as in Section 5.1.1 on the relative performances of policies. However, the effect of \( K_0 \) is more significant in here: when \( K_0=2 \), the hybrid policy and Policies 1 and 2 seem to have less variation, so they pay very little transportation penalty (due to very small \( Q^* \) and \( T^* \)). Therefore, Policy 0’s performance gets significantly worse in systems with very low fixed ordering cost.

![Figure 5](image1.png)

![Figure 6](image2.png)

Clearly, the hybrid policy dominates the other three policies. Even though in some cases the benefit of employing the hybrid policy is marginal, in other cases it makes significant difference (see the
maximum deviations in Table 3). Because all policies require the same information infrastructure, the hybrid policy is preferable to the other three policies, unless the warehouse management is restricted to taking action only at certain points in time (which favors Policy 2), or the contract with the carrier stipulates that inbound quantity must be multiples of a fixed integer (which favors Policy 1).

5.2 Numerical Analysis: Coordinated Policy (Hybrid Policy) vs. Non-Coordinated Policy

In this section, we investigate the value of order coordination by comparing the performance of the hybrid policy with Non-Coordinated policy via a numerical experiment. We use the same set of parameters as in Section 5.1 for \( N, K_0, L_0, \pi \) (or target service level), \( \alpha \). To see the effect of the Non-Coordinated policy’s lead time and fixed ordering costs (denoted by \( l \) and \( k \)), we vary the values of \( k \) and \( l \) (see Figures 8 and 9). In Figures 7-10, we present the results for some representative cases. The transportation penalty cost is set at zero (i.e., there is ample transportation capacity) except for Figure 10.

![Figure 7](image1.png)  ![Figure 8](image2.png)

In Figure 7, we observe that the hybrid policy enjoys the benefit of coordinated replenishment even more as \( K_0 \) increases, by placing fewer orders per unit time while providing more or less the same service level as the Non-Coordinated policy. Moreover, Figure 7 also suggests that joint replenishment is more valuable in systems with large number of retailers as the cost under the hybrid policy do not increase linearly in \( N \) (cost per retailer decreases as \( N \) goes up), while it is directly proportional to \( N \) under Non-Coordinated policy.

In Figure 8, we observe that shorter total lead time in the Non-Coordinated policy (which occurs if the handling at warehouse introduces additional time or distance in shipping) will reduce Non-Coordinated policy’s costs compared to the hybrid policy. In addition, our computational tests suggest
that Non-Coordinated policy performs relatively better as \( \pi \) increases, because Non-Coordinated policy keeps the order quantity the same most of the time while increasing the reorder point.

In reality, fixed ordering/transportation cost in the Non-Coordinated policy might be less than that of the coordinated replenishment policy, which could be due to sending less than truck-load instead of full truck-load, or additional handling/sorting charges at the warehouse. Thus, we consider the relative performance of these two policies under different ratios of fixed ordering costs. Intuitively, as \( k/K_0 \) decreases for a given \( K_0 \), the performance of Non-Coordinated policy improves. Interestingly enough, from Figure 9, there is a threshold level of “\( k/K_0 \)” below (above) which the relative performance of the Non-Coordinated policy improves when compared to the hybrid policy as \( K_0 \) increases (decreases). This is because the relative performances of these two policies mainly depend on the difference between the fixed ordering costs, and the difference between \( k \) and \( K_0 \) gets higher as \( K_0 \) increases, for any given \( k/K_0 \).

Due to the economy of scale in ordering and shipping, the hybrid policy tends to order in larger batches than the Non-Coordinated policy. Moreover, the inbound quantity is random under the hybrid policy while it is constant under the Non-Coordinated policy. Therefore, the Non-Coordinated policy performs relatively better in systems with tight inbound capacity constraints (see Figure 10). Moreover, as the unit penalty cost decreases, both policies tend to increase the order quantity given that the transportation capacity is tight. We observe that the outbound quantity under the Non-Coordinated policy is already larger than the truck capacity even when \( \alpha=2 \) (i.e., transportation capacity is tight). Hence, the Non-Coordinated policy takes advantage of this constraint and increases \( Q \) as \( \alpha \) decreases, to lower the ordering cost. In contrast, the inbound truck capacity is sufficient when \( C=3 \) or 4.5, which does not provide the hybrid policy with enough incentive to adjust \( Q \) accordingly in response to a reduction in \( \alpha \).
Thus, a reduction in \( \alpha \) will benefit the Non-Coordinated policy more when \( C=3 \text{ or } 4.5 \) (see Figure 10). However, when \( C=1.5 \), the transportation capacity is tight and the hybrid policy also takes advantage of the lower unit penalty cost by increasing \( Q \) significantly, which results in better performance relative to the Non-Coordinated policy.

In summary, our numerical experiment indicates that in most cases, coordinated replenishment results in savings in supply chain costs. The value of replenishment coordination seems to be the highest in systems with large fixed ordering/transportation cost and truck capacities, many retailers, and lower service level.

### 5.3 Numerical Analysis of Demand Variability at the Supplier

The supplier’s demand rate variance is often used as a measure of the bullwhip effect, for all the four policies we have studied. In this section we examine the impact on bullwhip effect by all the policies we have considered. The results are given in Table 4. Detailed derivations can be found in the Appendix.

<table>
<thead>
<tr>
<th>Supplier’s demand rate</th>
<th>Hybrid policy</th>
<th>Policy 0</th>
<th>Policy 1</th>
<th>Policy 2</th>
<th>Non-Coordinated policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
</tr>
<tr>
<td>Variance</td>
<td>( N\lambda \left(1 + 2 \frac{Var(Z_0)}{E(Z_0)}\right))</td>
<td>( N\lambda \left(1 + 2 \frac{Var(Z_0)}{E(Z_0)}\right))</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
<td>( N\lambda )</td>
</tr>
</tbody>
</table>

**Table 4: Supplier’s demand rate**

In Table 4 we observe that there is no bullwhip effect under the Non-Coordinated policy, Policy 1 and Policy 2, because their demand rate variance at the supplier is the same as the end customer demand rate variance, \( N\lambda \). This is due to the constant batch size or replenishment interval in these policies and the fact that the end customer demand is Poisson. Specifically, for Policy 1 and the Non-Coordinated policy, the order size is constant, \( Q \). An increase (decrease) in \( Q \) is offset by a decrease (increase) in the variance of order process, and in balance, the supplier’s demand rate variance remains a constant \( N\lambda \). The same holds true for Policy 2 where the order cycle is constant.

In Table 4 we also observe that the bullwhip effect does exist for the hybrid policy and Policy 0, because the demand rate variance is higher at the supplier. Bullwhip effect has an impact on the total cost
rate when there is limitation on transportation capacity in our paper. Other than that, since we assumed ample supply at the supplier (supplier’s cost is not modeled in this paper), the bullwhip effect does not have an impact on the model considered in this paper. However, our observation of the bullwhip effect under the hybrid policy and Policy 0 does have important implications for other models that study the whole supply chain: our analysis show that the use of the hybrid policy and Policy 0 results in savings for the retailers/warehouse, but it will hurt the supplier due to the bullwhip effect.

We define the ratio of the demand variance under the hybrid policy or Policy 0 to that of Non-Coordinated policy (or Policies 1 and 2) as the *bullwhip factor*. Figures 11-13 illustrate the *bullwhip factor* of the hybrid policy and Policy 0 as the number of retailers, fixed setup cost, and outbound capacity are varied. The following set of values is used: \( \{N=10, K_0=50, \pi=19, L_0=1, \alpha_0=\alpha=2, C_0=10*C_i\} \).

The hybrid policy performs better than Policy 0, since it limits the range of the total demand from retailers with an additional control variable, \( Q \). Figures 11 and 12 suggest that the variance under the hybrid policy shows a concave behavior in \( N \) and \( K_0 \), because an increase in either parameter leads to more echelon triggers than installation triggers (*i.e.* the chances of realizing constant inbound quantity under the hybrid policy increases) as the increase in \( Q \) is not as significant as the increase in \( \Delta \). On the other hand, the demand rate variance under Policy 0 increases when \( N, K_0, \) or \( C_i \) increases, as higher \( \Delta \) or \( N \) leads to a wider range for the random inbound quantity.

![Figure 11](image1)

![Figure 12](image2)

![Figure 13](image3)

5.4 Designing the Supply Chain Using Hybrid Policy

Our earlier numerical results suggest that the hybrid policy should be used to replenish the retailers. In this section we study the design of the distribution systems that use the hybrid policy. In
Figure 14, we see that the average cost rate per retailer is decreasing in \( N \). This implies that a regional warehouse, rather than multiple local warehouses, should be used to serve a large number of retailers.

Furthermore, from Figure 15 we infer it is economically viable to locate the warehouse closer to the outside supplier. This would also lead to more bargaining power to reduce the fixed ordering cost \( K_0 \) that will in turn decrease the cost rate at a faster rate (Figure 16). Moreover, it could be beneficial if the management use its own transportation fleet, since increasing the truck capacities (increasing inbound capacity factor from 0.5 to 1) while keeping the fixed ordering cost the same would reduce the average cost rate by almost 10%.

6 Conclusion

In this paper, we analyze a distribution system with multiple retailers, a single outside supplier, and one warehouse that holds no inventory. To the best of our knowledge, this is also one of the few studies that use coordinated replenishment policies to consider both inventory and transportation issues in a supply chain. We propose a new policy, the hybrid policy, which is the mixture of a traditional echelon policy and a special type of can-order type policy. We also compare the hybrid policy with three other coordinated replenishment policies. We derive the expected cost rate function and present the procedure for obtaining the optimal decision variables for all four policies. To examine the value of replenishment coordination, we also compare the hybrid policy with the traditional decentralized \((Q,R)\) policy, which is a Non-Coordinated policy. Finally, we analyze the variance of demand rate at the supplier and observe that the two coordination policies that have low total cost rate generate the most bullwhip effect. This
raises interesting question regarding the analysis of the entire supply chain, which should include the
supplier’s costs as well.

Our numerical results suggest that the hybrid policy always outperforms the other policies, both
with and without the transportation penalty cost. We observe that most of the parameters have significant
effect on the relative performance of the hybrid policy over the other three policies. However, the impact
of the transit time from the supplier to the warehouse ($L_0$) and the outbound capacity factor ($\gamma$) on the
deviations is not noticeable. Moreover, replenishment coordination seems to significantly lower the
system cost and the value of coordination increases in the cases of large fixed ordering/transportation cost
and truck capacities, many retailers, and lower service levels. There are a number of directions for future
research. Those include the case with multiple products, systems with non-identical retailers and finally
models where the warehouse would hold inventory.

References:

Inventory Systems. Management Science. 34 (6) 791-796.

Aviv, Y., L. Federgruen. 1998. The Operational Benefits of Information Sharing and Vendor Managed
Inventory (VMI) Programs. Working paper.

Axsater, S., L. Juntti. 1996. Comparison of echelon stock and installation stock policies for two-level

Management Science. 39 1274-1280.


Cachon, G. 2001. Managing a retailer’s shelf space, inventory, and transportation. Manufacturing &
Service Operations Management. 3(3) 211-229.


