We consider a periodic-review, serial inventory system, in which each stage replenishes inventory according to an echelon-stock \((r, nQ, T)\) policy – a standard echelon-stock \((r, nQ)\) policy executed every \(T\) periods. Demands are independent, identically distributed in different periods and occur at the most downstream stage. Two types of fixed costs are considered: setup cost and replenishment cost. The former is associated with each batch ordered; the latter is incurred in each order period. We study two models. The first (second), referred to as the \(T\)-model (\(Q\)-model), assumes that batch sizes (replenishment intervals) are fixed. The objective is to determine reorder points and replenishment intervals (batch sizes) such that the average total cost per period is minimized. For both models, we establish lower and upper bounds on the average inventory costs per period for each stage. These cost bounds, which lead to an optimization algorithm and close-to-optimal heuristics, are constructed based on a novel idea of developing bounds on the average penalty cost per period to each upstream stage for holding inadequate stock. We also discuss extensions with continuous-time models and alternative fixed cost assumptions.

These analytical results are used to investigate system design issues. Among others, replenishment interval decisions should be made before batch size decisions. In addition, when setup costs are equal to replenishment costs, continuous-review \((r, Q)\) policies result in lower average total cost than periodic-review \((r, T)\) policies. However, the cost benefit decreases when setup costs at downstream stages are large.

1. Introduction

Materials are usually replenished in batches and according to a fixed schedule. For instance, most manufacturers place purchase orders of several batches with suppliers when they run their material requirements planning (MRP) systems. MRP systems often run weekly, bi-weekly or monthly, resulting in periodic ordering with suppliers. Large retail chains, such as Wal-Mart, replenish items in this fashion. The items are often replenished in batches and delivered to the sites periodically. In this paper, we consider an inventory system that resembles this cyclical replenishment practice. Specifically, we consider a serial inventory system in which each stage replenishes according to
an echelon-stock \((r, nQ, T)\) policy – at the beginning of every \(T\) periods, an echelon-stock \((r, nQ)\) policy is executed, where \(r\) is the echelon reorder point and \(Q\) the batch size. We call the moment that a stage is allowed to order as order epoch. For order coordination, we assume that both replenishment intervals and batch sizes satisfy integer-ratio constraints. Two models are considered. The first (second) assumes that batch sizes (replenishment intervals) are fixed, and a fixed cost is associated with each order epoch (batch). The objective is to determine optimal reorder points and replenishment intervals (batch sizes) such that the average total cost \(per\ period\) is minimized. We term the first as the \(T\)-model and the second as the \(Q\)-model throughout the paper.

In the deterministic demand regime, these two models are the same. When demand is stochastic, these models represent two variants that can be derived from the deterministic model. For single-stage systems, Hadley and Whitin (1963) study these two models in their classic textbook. They derive average total cost function for given policies. Since then, there has been much progress in developing theories for the \(Q\)-model. Noteworthy examples are the papers by Zipkin (1986), Federgruen and Zheng (1991), Zheng (1992), and Gallego (1998). While the \(Q\)-model has been well studied in the literature, the \(T\)-model has not drawn much attention until recently. Rao (2003) shows that the total cost of an \((R, T)\) policy is jointly convex in policy parameters \(R\) is the base-stock level). He also develops a worst-case bound on the optimal cost for the \((R, T)\) policy.

For serial systems with fixed order costs, the form of the optimal policies is difficult to characterize (Clark and Scarf 1962). Researchers therefore focus on simple and effective policies, such as \((r, nQ)\) policies. See, for example, Axśater and Rosling (1993), Chen (2000), Chen and Zheng (1998), De Bodt and Graves (1985). However, these models either consider continuous-review policies or periodic-review policies with an implicit assumption of single-period replenishment intervals. One representative work is Chen and Zheng (1998), in which they develop heuristics and an exhaustive search algorithm for finding optimal batch sizes. For the model with multi-period replenishment intervals, there are papers that consider different system configurations, such as assembly and distribution systems (e.g., Atkins and Iyogun 1988, Cachon 1999, Graves 1996, and Yano and Carlson 1998). In these papers, the replenishment intervals are given. Recent developments for serial systems with \((r, nQ, T)\) policies show that the total cost can be evaluated recursively. With fixed batch sizes and replenishment intervals, the echelon-stock \((r, nQ, T)\) policies are indeed optimal, and the optimal reorder points can be obtained sequentially (see Chen 2000, van Houtum et al. 2006, Chao and Zhou 2005).

The considered \(Q\)-model is a generalization of Chen and Zheng (1998). It is natural to ask
whether their approach of developing the optimization algorithm and heuristics can be applied to the $Q$-model or even the $T$-model. By using the cost functions introduced in §2, it is not difficult to show that their approach can be applied to the $Q$-model. However, their approach cannot be applied to the $T$-model (see a remark in §3.2 for an explanation). As a result, our focus in this paper is to find optimal replenishment intervals for the $T$-model. Indeed, determining replenishment intervals is the problem of “setting the reorder intervals and finding the required safety stocks” identified by Graves (1996, p.4) or of the “first level” decision classified by van Houtum et al. (2006). This issue is important because determining effective replenishment intervals helps to consolidate shipment activities. To our knowledge, Feng and Rao (2005) is the only paper that studies this problem. They consider a two-stage system and derive an approximate total cost function, and use Golden section search to solve for policy parameters. However, it may be difficult to extend their approach to more general systems.

In this paper, we first construct lower and upper bounds on the average penalty cost per period charged to each upstream stage for holding inadequate stock (or the average induced-penalty cost per period). Specifically, we derive these induced-penalty cost bounds by means of regulating the replenishment intervals of its downstream stages – if the downstream stages set the replenishment intervals equal to that of the upstream stage (resp., to one period), the resulting expected penalty function per period is a lower (resp., an upper) bound to the exact one. Substituting these bounds with the exact penalty cost, we establish cost bounds for the inventory holding and penalty cost per period at each stage. These stage cost bounds are used further to develop an algorithm for finding optimal solutions and two close-to-optimal heuristics. Interestingly, we can obtain parallel results for the $Q$-model by regulating downstream batch sizes. The resulting algorithm is more efficient than the one that can be derived by using Chen and Zheng’s (1998) approach. In a numerical study, the total number of feasible solutions evaluated by our algorithm is on average 70% less than that by Chen and Zheng’s algorithm. Finally, we discuss extensions with continuous-time models and alternative fixed cost assumptions.

The analytical results are used to investigate system design issues in a numerical study. First, van Houtum et al. (2003) categorize both batch sizes and replenishment intervals as “first level” decisions because these decisions have a longer term impact on system performance. We refine this statement by observing that the impact of replenishment intervals on optimal cost is larger than that of batch sizes. Thus, replenishment interval decisions should be made first before batch size decisions. Second, fixed replenishment/setup costs and echelon holding costs are highly related to
the decision of which stages should use the same replenishment intervals/batch sizes in a chain. In addition, it is more cost efficient to reduce leadtime or fixed costs at a downstream stage. Finally, some information systems can replenish inventory in a real-time fashion, which is similar to a continuous-review policy. We assess the value of such information systems by comparing a continuous-review \((r, Q)\) policy and a periodic-review \((r, T)\) policy under Poisson demand. We find that the former leads to a smaller cost. However, the cost benefit decreases when setup costs at downstream stages are large, demand rate is high, or backorder cost penalty is small.

The rest of the paper is organized as follows. §2 illustrates the dynamics of inventory variables and provides a recursion to evaluate the average total cost per period. §3 constructs lower and upper bounds on the average inventory holding and backorder costs per period for each stage. §4 develops two heuristics and an optimization algorithm. §5 summarizes the parallel results for the \(Q\)-model. §6 performs numerical studies on the heuristics and discusses system design issues. §7 summaries the paper and discusses several extensions. All proofs are in the appendix.

2. Problem Formulation

We consider a periodic-review, serial inventory model with \(N\) stages. Material flows from stage \(N\) to stage \(N-1\), from \(N-1\) to \(N-2\), etc. until stage 1 where customer demand occurs. We assume that the demands in different periods are continuous and independent, identically distributed with mean \(\mu\) per period. There is constant transportation leadtime \(l_j\) between stage \(j+1\) and stage \(j\), where \(l_j \in \mathbb{N}\), the set of positive integers. Time is divided into periods of equal length, say, to be one, and the time horizon is infinite. Unsatisfied demands are fully backlogged. We assume linear inventory holding and backorder costs. Echelon holding cost \(h_j\) is charged per period for each unit of inventory held in the echelon \(j\). Define the local inventory holding cost for stage \(j\), \(h'_j = \sum_{i=j}^{N} h_i\). Backorder cost \(b\) is incurred per period for each unit of unsatisfied customer demand.

Each stage replenishes inventory according to an echelon-stock \((r, nQ, T)\) policy. Specifically, stage \(j\) orders at the beginning of every \(T_j\)-th period, according to an echelon-stock \((r, nQ)\) policy: If the echelon inventory order position (= inventory on order + inventory on hand + inventory at or in transit to all downstream stages - backorders) at stage \(j\) is less than or equal to the echelon reorder point \(r_j\), the stage places an integer multiple of \(Q_j\) units to the upstream stage so as to bring the inventory order position to between the interval of \([r_j, r_j + Q_j]\). We call the moment and the period that a stage is allowed to order as order epoch and order period, respectively. For order coordination, we assume that both batch sizes and replenishment intervals satisfy integer-
ratio constraints: \( T_{j+1} = \tau_j T_j \) and \( Q_{j+1} = q_j Q_j \), where \( T_j, \tau_j, q_j \in \mathbb{N}, j = 1, ..., N - 1 \). Fixed setup cost \( k_j \) occurs for each batch \( Q_j \) ordered by stage \( j \). Also, fixed replenishment cost \( K_j \) occurs for stage \( j \) at each order epoch, regardless of whether an order is placed. (Our results can be applied to other fixed cost assumptions; see §7.) The sequence of events in each period is as follows: at the beginning of a period, each stage receives orders from its downstream stage, if any; each stage places an order to its upstream stage at each order epoch, regardless of whether an order is placed; a shipment sent from an upstream stage is received; a shipment to the downstream stage is sent. Demand occurs during a period. The costs are evaluated at the end of a period.

To see inventory dynamics under echelon-stock \((r, nQ, T)\) policies, we define the following inventory state variables:

\[
\begin{align*}
IP_j(t) &= \text{echelon inventory in-transit position of stage } j \text{ at the beginning of period } t, \\
&= \text{inventory in transit to and at all stages } i(i \leq j) - \text{stage 1 backorders} \text{ at the beginning of period } t, \\
IL_j^-(t) &= \text{net echelon inventory level at the beginning of period } t \text{ of stage } j, \\
IL_j(t) &= \text{net echelon inventory level at the end of period } t \text{ of stage } j.
\end{align*}
\]

Let \( L_j = \sum_{i=j}^N l_i \) and \( D(\tau) \) be the demand over \( \tau \) periods. Define \( \lfloor z \rfloor \) the roundoff operator, which returns the greatest integer less than or equal to \( z \), a real number. Also, let \( \text{mod}_x(y) \) be the operator that returns the remainder of \( y \) divided by \( x \), and \( x, y \in \mathbb{N} \).

Suppose stage \( N \) places an order at the beginning of period \( t \), which brings the echelon inventory order position to \( IP_N(t) \). (Since stage \( N \) has ample supply, its echelon inventory order and in-transit positions are equal.) This order, arriving at the beginning of period \( t + l_N \), directly or indirectly determines the net echelon inventory levels \( IL_j^-(t + L_j + \ell) \) and \( IL_j(t + L_j + \ell), \ell = 0, ..., T_N - 1, j = 1, ..., N \). These net echelon inventory levels \( IL_j^- \) in a cycle of \( T_N \) periods determine the average total cost in this cycle.

More specifically, for stage \( N \), \( IP_N(t) \) and the demand jointly determine \( IL_N^-(t + l_N + \ell) \) since there are no other order opportunities before period \( t + T_N \). That is, \( IL_N^-(t + l_N + \ell) = IP_N(t) - D(l_N + \ell), \ell = 0, ..., T_N \). For stage \( j, j < N \), we assume that the replenishment activities are synchronized\(^1\). That is, stage \( j \)'s first order epoch is at the beginning of period \( t + L_j + 1 \), at which stage \( j + 1 \) just receives a shipment. Clearly, stage \( j \) has another \( \tau_j - 1 \) order epochs before stage \( j + 1 \)'s next order period due to integer-ratio constraints. The \( IP_j \) are constrained by \( IL_j^+ \)

\(^1\)A synchronized policy dominates a non-synchronized one; see Chao and Zhou (2005).
Figure 1: A Two-stage Example: $l_2=2$, $l_1=1$, $T_2 = 4$, and $T_1 = 2$.

as below:

$$IP_j \left( t + L_{j+1} + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right) = O_j \left[ IL_{j+1} \left( t + L_{j+1} + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right) \right], \quad \ell = 0, 1, ..., T_N - 1,$$  \hfill (1)

where

$$O_j[x] = \begin{cases} x, & x \leq r_j, \\ x - mQ_j, & \text{otherwise}, \end{cases}$$ \hfill (2)

and $m$ is the smallest nonnegative integer so that $x - mQ_j \leq r_j + Q_j$.

Then, these $IP_j$ and the demand determine $IL_j^-$ and $IL_j$ as below: For $\ell = 0, 1, ..., T_N - 1$,

$$IL_j^- (t + L_j + \ell) = IP_j \left( t + L_{j+1} + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right) - D(l_j + \text{mod}_{T_j}(\ell)), \hfill (3)$$

$$IL_j (t + L_j + \ell) = IP_j \left( t + L_{j+1} + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right) - D(l_j + \text{mod}_{T_j}(\ell) + 1). \hfill (4)$$

Figure 1 is a two-stage example to illustrate these inventory dynamics. In this example, $l_2=2$, $l_1=1$, $T_2 = 4$, and $T_1 = 2$. A cycle for stage $j$ includes the periods $t + L_j + \ell, \ell = 0, 1, 2, 3$.

In this paper, we focus the analysis on the $T$-model by assuming the batch sizes $(Q_1, ..., Q_N)$ are fixed. Consequently, the long-run average total cost per period is

$$C(r, T) = \sum_{j=1}^N \left( \frac{K_j}{T_j} \right) + \frac{1}{T_N} \mathbb{E} \left[ \sum_{\ell=0}^{T_N-1} \left( \sum_{i=1}^N h_i [IL_i(t + L_j + \ell) + (b + h_i') [IL_i(t + L_j + \ell)]^-] \right) \right], \hfill (5)$$

where $x^- = \max\{0, -x\}$; $r$ and $T$ are the vectors of reorder points and replenishment intervals.

The expectation in the second term of (5) is the average inventory holding and backorder costs occurred in a cycle of $T_N$ periods. We divide it by $T_N$ to obtain the average cost per period.

A bottom-up procedure can be used to evaluate $C(r, T)$. For notational simplicity, let $T_j$ be the vector of $(T_1, T_2, ..., T_j)$, so $T_1 = T_1$ and $T = T_N$. The logic behind this procedure is that, at each
iteration, we evaluate a truncated $j$-stage serial system, referred to as echelon-$j$. More specifically, define $G_j(y, T_j)$ as the average inventory holding and backorder costs per period for echelon-$j$, conditioning on $IP_j(t) = y$. That is,

$$G_1(y, T_1) = \frac{1}{T_1} \left( \sum_{\ell=0}^{T_1-1} E[h_1(y - D(l_1 + \ell + 1)) + (b + h_1')((y - D(l_1 + \ell + 1))^\ell) \right].$$ (6)

For $j = 2, ..., N$, define recursively

$$G_j(y, T_j) = \frac{1}{T_j} \sum_{\ell=0}^{T_j-1} E\left[h_j(y - D(l_j + \ell + 1)) + G_{j-1}\left(y - D\left(l_j + \left\lfloor \frac{\ell}{T_{j-1}} \right\rfloor T_{j-1}\right), T_{j-1}\right)\right].$$ (7)

Since $IP_N$ is uniformly distributed (e.g., Zipkin 1986), the total average cost per period is

$$C(r, T) = \sum_{j=1}^{N} \left( \frac{K_j}{T_j} \right) + \int_0^{Q_N} G_N(r_N + x, T_N) dx.$$ (8)

It can be verified that (5) is equal to (8). Also, by setting $T_j = 1, j \leq N$, the above inventory dynamics as well as the recursion reduce to those in Chen and Zheng (1994) and Chen (2000).

Our objective is to solve the following problem: $\min_{r,T} C(r, T)$ subject to $T_{j+1} = \tau_j T_j$, $T_j, \tau_j \in \mathbb{N}, j = 1, ..., N-1$. This problem is difficult to solve because, in addition to integer-ratio constraints, the objective function is not jointly convex in $r$ and $T$. In the next section, we shall construct upper and lower bounds on $C(r, T)$. These results lead to an optimization algorithm and heuristics for the $T$-model.

3. Bounds

We develop a lower bound on $C(r, T)$ for any feasible $(r, T)$ policies. (As we shall see, an upper bound is a natural extension.) This bound is derived by constructing a lower-bound on the average induced-penalty cost per period. (Clark and Scarf (1960) interpret the induced-penalty cost as a penalty charged to an upstream stage if the stage cannot fulfill its downstream stage’s orders.) We first introduce a decomposition scheme, in which the average induced-penalty function per period is used to decouple $C(r, T)$. This decomposition scheme sets the stage for the construction of the lower and upper bounds.
3.1 Decomposition of Total Cost Function

We first characterize the induced-penalty function incurred at each order epoch. We start with stage 1. Let \( g_1(y, T_1) \) be the inventory holding and backorder costs for stage 1 for given \( IP_1(t) = y \) after ordering, which is equal to \( G_1(y, T_1) \). Suppose that stage 1 has ample supply, \( IP_1 \) will be uniformly distributed over \((r_1, r_1 + Q_1)\) at steady state (e.g., Zipkin 1986). The average inventory cost for stage 1 is \( I_1(r_1, T_1) = \int_0^{Q_1} g_1(r_1 + x, T_1) dx / Q_1 \). Let the optimal reorder point for any given \( T_1 \) be \( r_1(T_1) = \arg \min_y I_1(y, T_1) \). In reality, stage 1 does not have ample supply from stage 2, because \( IP_1(t) \) is constrained by \( IL_2(t) \). Now set \( IN_2(t) = y \). The penalty cost charged to stage 2 at this order period \( t \) for holding inadequate stock, \( g_{1,2}(y, T_1) \), is

\[
 g_{1,2}(y, T_1) = g_1(O_1[y, T_1], T_1) - I_1(r_1(T_1), T_1),
\]

where

\[
 O_1[y, T_1] = \begin{cases} y, & y \leq r_1(T_1), \\ y - mQ_1, & \text{otherwise}, \end{cases}
\]  

(9)

\( m \) is the smallest nonnegative integer so that \( y - mQ_1 \leq r_1(T_1) + Q_1 \), and \( g_1(O_1[y, T_1], T_1) \) is the actual inventory cost for stage 1.

For stage \( j, j = 2, ..., N - 1 \), assume that stage \( j \) has ample supply from its upstream stage. Conditioning on \( IP_j(t) = y \), the resulting average inventory cost for stage \( j, g_j(y, T_j) \), is the sum of the average inventory cost and induced-penalty cost incurred in the order cycle of \( T_j \) periods. That is,

\[
 g_j(y, T_j) = \frac{1}{T_j} \left( \sum_{\ell=0}^{T_j-1} E \left[ h_j(y - D(l_j + \ell + 1)) + g_{j-1,j} \left( y - D \left( l_j + \left\lfloor \frac{\ell}{T_j-1} \right\rfloor T_j-1 \right), T_j-1 \right) \right] \right).  
\]  

(10)

Notice that \( y - D(l_j + \left\lfloor \ell/T_j-1 \right\rfloor T_j-1) \), \( \ell = 0, ..., T_j - 1 \) represent the \( IL_j^- \) values at stage \( j - 1 \)’s order epochs. With ample supply, \( IP_j \) is uniformly distributed over \((r_j, r_j + Q_j)\) at steady state. Let the average inventory cost for stage \( j \) be \( I_j(y, T_j) = \int_0^{Q_j} g_j(y + x, T_j) dx / Q_j \) and

\[
 r_j(T_j) = \arg \min_y I_j(y, T_j).
\]  

(11)

Since stage \( j \) in fact does not have ample supply, the induced penalty cost function charged to stage \( j + 1 \) in this order period is

\[
 g_{j,j+1}(y, T_j) = g_j(O_j[y, T_j], T_j) - I_j(r_j(T_j), T_j),
\]  

(12)
where \( O_j[y, T_j] \) is the same as (9) except changing stage index from 1 to \( j \). Finally, we can compute

the inventory-related cost for stage \( N \), \( g_N(y, T_N) \), by setting \( j = N \) in (10).

Let \( C(T) \) denote the average total cost per period when \( r_j(T_j) \) are substituted for \( r_j \) in (8). Denote the average total cost per period for stage \( j \) as \( c_j(T_j) = K_j/T_j + I_j(r_j(T_j), T_j), j = 1, \ldots, N \). We have

**Proposition 1**  
(1) \( r_j(T_j), j = 1, \ldots, N \) are the optimal echelon reorder points for given \( T \). (2) \( r_j(T_j) \) increases in \( T_j \). (3) \( C(r, T) \geq C(T) = \sum_{j=1}^{N} c_j(T_j) \).

Proposition 1 Part (2) states that the optimal reorder point \( r_j(T_j) \) increases in \( T_j \). Part (3) means that we decouple the average total cost per period into each stage by using the induced-penalty cost function.

### 3.2 Lower-Bound Functions

This section establishes lower-bound functions for \( I_j(r_j(T_j), T_j), j = 1, \ldots, N \). The construction of the lower-bound functions starts from generating a series of *regulated* induced-penalty functions \( g_{j+1}(\cdot, T) \) in the order periods as follows: For any positive integer \( T \), let \( IP_j(t) = y \) after ordering and define

\[
g_j(y, T) = \frac{1}{T} \left( \sum_{\ell=0}^{T-1} \mathbb{E}[h_1(y - D(l_1 + \ell + 1)) + (b + h_1')(y - D(l_1 + \ell + 1))] \right). \tag{13}
\]

For \( j = 2, \ldots, N \), suppose \( g_{j-1}(\cdot, T) \) is known. Define the *regulated* inventory cost per period as \( L_{j-1}(y, T) = \int_0^{Q_{j-1}} g_{j-1}(y + x, T)dx/Q_{j-1} \), and \( r_{j-1}(T) = \arg\min_y L_{j-1}(y, T) \), the optimal reorder point for stage \( j - 1 \) if the replenishment interval is \( T \) periods. Define the *regulated* induced-penalty cost per period with \( IN_{j-1}(t) = y \) as

\[
g_{j-1,j}(y, T) = g_{j-1}(O_{j-1}[y, T], T) - L_{j-1}(r_{j-1}(T), T), \tag{14}
\]

where

\[
O_{j-1}[y, T] = \begin{cases} y, & y \leq r_{j-1}(T), \\ y - mQ_{j-1}, & \text{otherwise}, \end{cases} \tag{15}
\]

\( m \) is the smallest nonnegative integer so that \( y - mQ_{j-1} \leq r_{j-1}(T) + Q_{j-1} \). Conditioning on \( IP_j = y \), the regulated inventory cost per period for stage \( j \) is

\[
g_j(y, T_j) = \frac{1}{T_j} \left( \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ h_j(y - D(l_j + \ell + 1)) + g_{j-1,j}(y - D(l_j + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j), T_j) \right] \right). \tag{16}
\]
The construction is very similar to that for generating \( g_j(\cdot, T_j) \). The difference is that in each iteration, we replace \( T_j = (T_1, \ldots, T_j) \) in the \( g_j(\cdot, T_j) \) by \((T_j, T_j, \ldots, T_j)\). Consequently, the new regulated function \( g_j \) becomes a function of \( T_j \) only. This procedure is equivalent to stating that for a stage \( j \), we obtain \( g_j(\cdot, T_j) \) by regulating all of its downstream stages \( i, i < j \) to use the same replenishment interval \( T_j \).

We aim to show \( L_j(\tau_j(T_j), T_j) \leq I_j(r_j(T_j), T_j) \). The following lemma sets the stage for this result.

**Lemma 2** For all \( y \) and \( j = 2, \ldots, N \),

\[
\frac{1}{T_j} \int_0^{Q_j} \sum_{\ell=0}^{T_j-1} E \left[ g_{j-1,j} \left( y + x - D \left( l_j + \left[ \frac{\ell}{T_j} \right] T_j \right), T_j \right) \right] dx \\
\leq \frac{1}{T_j} \int_0^{Q_j} \sum_{\ell=0}^{T_j-1} E \left[ g_{j-1,j} \left( y + x - D \left( l_j + \left[ \frac{\ell}{T_{j-1}} \right] T_{j-1} \right), T_{j-1} \right) \right] dx.
\]

Clark and Scarf (1960) interpret the induced-penalty cost function as the cost charged to an upstream stage if the upstream stage cannot fulfill the downstream’s orders. Lemma 2 states that for a fixed \( IP_j = y \), the average induced-penalty cost function per period obtained from the regulated system is smaller than that of the original system. Intuitively, if downstream stages \( i, i < j \) choose the largest replenishment interval, i.e., \( T_j \), these downstream stages will have higher optimal reorder points (Proposition 1, Part 2), or equivalently, carry more inventory. More inventory holding, in turn, leads to smaller backorders at the upstream stage. Consequently, the corresponding average induced penalty cost is smaller. Figure 2 is a two-stage example with \( h_1 = h_2 = 0.5 \), \( l_1 = l_2 = 1 \), \( b = 9 \), \( Q_1 = 2 \) and \( Q_2 = 4 \). We assume demand is Poisson with mean \( \mu = 4 \) units/period. We fix \( T_2 = 4 \) and allow \( T_1 \) to be 1 and 4. Figure 2(a) represents the average inventory cost per period for stage 1, given \( IP_2 = y \). Figure 2(b) represents the average inventory cost per period if stage 1 has ample supply, i.e., \( IP_2 = \infty \). Notice that both functions are larger when \( T_1 = 4 \). However, their difference, or the average induced penalty cost function per period, becomes smaller.

From (10), (16), and Lemma 2, we have \( \int_0^{Q_j} \int_0^{T_j} g_j(y + x, T_j) dx \leq \int_0^{Q_j} g_j(y + x, T_j) dx \), for all \( y \). Recall that \( \tau_j(T_j) \) and \( r_j(T_j) \) are the optimal solutions of \( L_j(y, T_j) \) and \( I_j(y, T_j) \), or equivalently, \( \int_0^{Q_j} g_j(y + x, T_j) dx \) and \( \int_0^{Q_j} g_j(y + x, T_j) dx \), respectively. Consequently, \( L_j(\tau_j(T_j), T_j) \leq L_j(r_j(T_j), T_j) \leq I_j(r_j(T_j), T_j) \).

**Theorem 3** \( L_j(\tau_j(T_j), T_j) \leq I_j(r_j(T_j), T_j) \), for \( j = 1, \ldots, N \).

Now we are ready to construct a lower bound cost for all feasible \((r, T)\) policies. For \( j = 1, \ldots, N \), define \( \varepsilon_j(T_j) = K_j/T_j + L_j(\tau_j(T_j), T_j) \). We have
Figure 2: In all figures, the dash and solid lines represent $T_1 = 4$ and $T_1 = 1$, respectively. (a) The average inventory cost per period for stage 1, given $IP_2 = y$. (b) The average inventory cost per period, provided that stage 1 has ample supply. (c) The average induced-penalty cost per period charged to stage 2.

**Proposition 4**

1. $c_j(T_j)$ is convex in $T_j$.
2. $C(T) = \sum_{j=1}^{N} c_j(T_j) \geq \sum_{j=1}^{N} c_j(T_j)$.

Clearly, we can construct an upper-bound on the total cost function by regulating the replenishment intervals of downstream stages to one period. Substituting these upper bounds with the exact ones, we can construct an upper bound on the total cost function.

**Remark.** As stated in the introduction, one can show that Chen and Zheng’s (1998) approach of deriving bounds, optimization algorithms, and heuristics can be applied to the $Q$-model. All we have to do is to substitute our $G_j$ functions defined in (7) with theirs and follow the same steps therein. However, their approach cannot be applied to the $T$-model because the single-stage $T$-model does not have the same optimality properties (i.e., properties (i)-(iii) in p. 595 of Chen and Zheng) as those in the single-stage $Q$-model.

4. **Heuristics and Optimization Algorithm**

We propose two heuristics and construct an optimization algorithm based on the cost bound functions constructed in §3.

4.1 **Heuristics**

The Cost-Bound (CB) Heuristic
This heuristic is to solve the lower-bound cost function subject to the relaxed constraints, i.e.,
\[ \min \sum_{j=1}^{N} c_j(T_j) \text{ s.t. } T_{j+1} \geq T_j, \ j = 1, \ldots, N - 1, \] and then turn the solution into one that satisfies integer-ratio constraints. From Proposition 4 Part (1), the objective is a sum of \( N \)-convex functions, so this problem can be solved efficiently by a standard clustering algorithm (see, e.g., Maxwell and Muckstadt 1985). The output of the algorithm is an optimal partition that includes disjoint clusters \( \{c(1), c(2), \ldots, c(M)\} \) (a cluster is a set of consecutive stages) and the optimal solution \( (T_{c(1)}, T_{c(2)}, \ldots, T_{c(M)}) \), where
\[
T_{c(m)} = \arg \min_{T} \left\{ \frac{K[m]}{T} + L_{c(m)}(R_m(T), T) \right\}, \quad m = 1, \ldots, M. \tag{17}
\]
Here, \( K[m] = \sum_{i \in c(m)} K_i \), \( R_m(T) = \arg \min_y L_{c(m)}(y, T) \) and \( L_{c(m)}(y, T) = \sum_{i \in c(m)} L_i(y, T) \). These optimal intervals satisfy the following two conditions:

(i) \( T_{c(1)} < T_{c(2)} < \ldots < T_{c(M)} \).

(ii) For each cluster \( c(m) = \{v, v+1, \ldots, w\} \), there does not exist a stage \( v' \) with \( v \leq v' < w \) so that \( T_{c(m^-)} \leq T_{c(m^+)} \), where \( c(m^-) = \{v, \ldots, v'\} \) and \( c(m^+) = \{v' + 1, \ldots, w\} \).

Since \( T_{c(1)}, \ldots, T_{c(M)} \) do not necessarily satisfy the integer-ratio constraints, we propose a heuristic solution, \( \hat{T}_{c(m)} \), that satisfies the constraints. Let \( \hat{T}_{c(1)} \) be \( T_{c(1)} \). For \( m = 2, \ldots, M \), \( \hat{T}_{c(m)} \) is the solution of the problem in (17) with the constraint that it is an integer multiple of the optimal solution of the downstream cluster. That is,
\[
\hat{T}_{c(m)} = \arg \min_{T} \left\{ \frac{K[m]}{T} + L_{c(m)}(R_m(T), T) \right\} \quad \text{s.t. } T = \tau \hat{T}_{c(m-1)}, \tau \in \mathbb{N}.
\]
(Because \( L_{c(m)}(R_m(T), T) \) is convex in \( T \), we only need to compare two integer-multiple of \( \hat{T}_{c(m-1)} \) values closest to \( T_{c(m)} \).) Then, set the heuristic solution \( \hat{T}_i = \hat{T}_{c(m)} \), for \( i \in c(m) \). The heuristic reorder points \( \hat{r}_i \) are the corresponding optimal reorder points. Finally, we can compute the heuristic cost, \( C^L \), by using \( (\hat{r}_j, \hat{T}_j) \).

Similarly, we can obtain another heuristic solution by solving the upper-bound cost functions. Call the resulting heuristic cost as \( C^U \). Then, we set the heuristic cost of the CB heuristic as \( C^B = \min\{C^L, C^U\} \).

The Clustering-Minimization (CM) Heuristic

This heuristic was first proposed by Shang (2004) for the Q-model with single replenishment intervals. We find that the two key observations in Shang are also true for the T-model.
First, \( T_{c(m)} \) in (17) is close to \( \tilde{T}_{c(m)} \), which is solved from the following single-stage problem:

\[
\tilde{T}_{c(m)} = \arg \min_T \left\{ \frac{K[m]}{T} + \tilde{I}_{c(m)}(\tilde{R}_m(T), T) \right\}, \tag{18}
\]

where

\[
\tilde{R}_m(T) = \arg \min_y \tilde{I}_{c(m)}(y, T), \quad \tilde{I}_{c(m)}(y, T) = \sum_{i \in c(m)} \int_0^{Q_i} \tilde{g}_i(y + x, T)dx,
\]

\[
\tilde{g}_i(y, T) = \frac{1}{T} \left( \sum_{\ell=0}^{T-1} \mathbb{E}[h_i(y - D(L_i + \ell + 1)) + (b + h'_i)(y - D(L_i + \ell + 1))^-] \right).
\]

Here, \( \tilde{I}_{c(m)}(y, T) \) is the sum of the inventory cost for the single-stage systems with the same replenishment interval \( T \), which is jointly convex in \( y \) and \( T \). Thus, solving this problem is the same as solving a single-stage problem.

Second, we observe that \( \tilde{T}_{c(m)} \) tends to increase in \( K[m]/h[m] \), where \( h[m] = \sum_{i \in c(m)} h_i \). Thus, instead of conducting the clustering algorithm, we may use cost ratios to determine an effective partition.

We formalize the CM heuristic as follows. The first step is to identify a partition by cost parameters. Let the resulting partition be \( \{c(1), \ldots, c(M)\} \). These clusters satisfy the conditions (i) and (ii) above by replacing \( T_{c(m)} \) with \( K[m]/h[m] \). (A two-dimensional diagram suggested by Zipkin (p. 128, 2000) is useful for the task.) Since the resulting solution \( \tilde{T}_{c(m)} \) does not necessarily satisfy the integer-ratio constraints, the second step is to convert \( \tilde{T}_{c(m)} \) into \( \tilde{T}_{c(m)} \) that satisfies the constraints. Similarly as in the CB heuristic, let \( \tilde{T}_{c(1)} = \tilde{T}_{c(m)} \). For \( m = 2, \ldots, M, \tilde{T}_{c(m)} \) is the solution that minimizes the right-hand side of (18), subject to \( T = \tau \tilde{T}_{c(m-1)}, \tau \in \mathbb{N} \). Finally, set the heuristic solution \( \tilde{T}_i = \tilde{T}_{c(m)} \), for \( i \in c(m) \). The heuristic reorder points \( \tilde{r}_i \) are the corresponding optimal reorder points. Denote the CM heuristic cost as \( C^M \).

### 4.2 Optimization Algorithm

The idea of the algorithm is to establish bounds for the optimal solution and then conduct a search for all feasible solutions. Let \( (T_1^*, \ldots, T_N^*) \) be an optimal solution, and \( C^h \) be any heuristic cost (say, obtained from the CB or CM heuristic). Then,

\[
C^h \geq C(T_1^*, \ldots, T_N^*) \geq \sum_{i=1}^N \xi_i(T_i^*) = \sum_{i \neq j} \xi_j(T_i^*) \geq \xi_j(T_j^*) + \sum_{i \neq j} \xi_i,
\]

where \( \xi_i \) is the optimal cost obtained by minimizing \( \xi_i(T) \). As a result, \( \xi_j(T_j^*) \) must be less than or equal to \( C^h - \sum_{i \neq j} \xi_i \). Since \( \xi_j(T) = K_j/T + L_j(\xi_j(T), T) \) is a convex function with a finite
minimizer, we can find the bounds for \( T^*_j \). Formally, define

\[
T_j = \min \{ T \in \mathbb{N} | c_j(T) \leq C^h - \sum_{i \neq j} c_i \} \quad \text{and} \quad \bar{T}_j = \max \{ T \in \mathbb{N} | c_j(T) \leq C^h - \sum_{i \neq j} c_i \}.
\]

We have

**Proposition 5** \( T_j \leq T^*_j \leq \bar{T}_j, j = 1, ..., N \).

To determine the optimal solution, it suffices to search the feasible solutions that satisfy the integer-ratio constraints.

5. The \( Q \)-Model

In the \( Q \)-model, the decision variables are \( r \) and \( Q \). The recursion in (6) and (7) can be used to evaluate any \((r, nQ)\) policies. More specifically, denote \( Q_j \) as the vector \((Q_1, ..., Q_j)\), \( Q \equiv Q_N \), \( Q_0 \equiv 0 \). So \( G_1 \) and \( G_j \) in the \( Q \)-model may be written as

\[
G_1(y, Q_0) = \frac{1}{T_1} \left( \sum_{\ell=0}^{T_1-1} \mathbb{E}[h_1(y - D(l_1 + \ell + 1)) + (b + h'_1)(y - D(l_1 + \ell + 1))] \right).
\]

\[
G_j(y, Q_{j-1}) = \frac{1}{T_j} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ h_j(y - D(l_j + \ell + 1)) 
+ G_{j-1} \left( Q_{j-1} \left[ y - D \left( l_j + \left\lfloor \frac{\ell}{T_{j-1}} \right\rfloor T_{j-1} \right) \right], Q_{j-2} \right) \right], \quad j = 2, ..., N,
\]

where \( O_{j-1} \), a function of \( Q_{j-1} \), is defined in (2). The problem is to solve \( \min_{r, Q} C(r, Q) = \sum_{i=1}^N (k_i \mu)/Q_i + \int_0^{Q_N} G_N(r_N + x, Q_{N-1}) dx/Q_N \) subject to \( Q_{j+1} = q_j Q_j, q_j \in \mathbb{N} \), \( j = 1, ..., N - 1 \).

It can be shown that all results obtained in the \( T \)-model hold for the \( Q \)-model. In particular, the decomposition scheme in (10)-(12) holds except that \( g_j(\cdot) \) depends on \( Q_{j-1} \). It can be shown that \( r_j(Q_j), j = 1, ..., N \) are optimal reorder points for given \((Q_1, ..., Q_N)\) and \( C(r, Q) \geq C(Q) = \sum_{i=1}^N c_i(Q_i) \).

For each stage \( j \), by regulating all of its downstream’s batch sizes to \( Q_j \), the resulting induced-penalty cost function per period is a lower bound to the exact one. Substituting these lower bound penalty functions with the exact ones, we can obtain the lower-bound functions parallel to Proposition 4, i.e., \( C(Q) = \sum_{i=1}^N c_i(Q_i) \geq \sum_{j=1}^N e_j(Q_j) \). Similarly, by regulating the downstream batch sizes equal to zero (one if the demand is discrete), we can obtain upper-bound functions. Then, we can construct the CB and CM heuristics and an optimization algorithm for the \( Q \)-model.

We remark that our optimization algorithm is more efficient than the one that can be derived by
using Chen and Zheng’s (1998) approach, because our algorithm searches fewer feasible solutions. We will report this result with a numerical study in §6.1.

Due to PASTA - Poisson Arrivals See Time Averages (Wolff 1982), the analytical properties for the $Q$-model with single-period replenishment intervals are essentially the same as its continuous-review counterpart with compound Poisson demand (see Chen and Zheng 1994 for a detailed discussion). Thus, all results are valid for the continuous-review, echelon-stock $(r,nQ)$ policies. (If the demand is a Poisson process, an $(r,nQ)$ policy will degenerate as an $(r,Q)$ policy.)

6. Numerical Study

In §6.1, we perform numerical studies to examine the effectiveness of the the heuristics for both models, using the optimal solution as a benchmark. In §6.2, we compare continuous-review $(r,Q)$ policies and periodic-review $(r,T)$ policies, with an emphasis on the cost benefit gained from the continuous-review $(r,Q)$ policies.

6.1 Performance of the Heuristics

We consider the following parameters for the $T$-model: $N = 3$, $h_j = 0.1, 1, K_j = 50, 200$, $l_j = 1$ for $j = 1, 2, 3$, and $b = 5, 20$. We fix $Q_3 = 8$. $Q_1$ and $Q_2$ are equal to 2, 4, or 8 that satisfy integer-ratio constraints (six combinations). Demand in each period follows a Poisson distribution with $\mu = 4$. For the $Q$-model, the parameters are the same except that $k_j = 20, 50$ for $j = 1, 2, 3$. We set $T_3 = 4$, and $T_1, T_2$ are equal to 1, 2, or 4 that satisfy the integer-ratio constraints (six combinations). The total number of instances for each of the models is 768. Denote the optimal cost $C^*$ and define the percentage error as $\epsilon^h = (C^h - C^*)/C^* \times 100\%$, $h \in \{B,M\}$. Figure 3 shows the distribution, the average, and the maximum of the percentage errors for each of the models with different $b$ values. For both heuristics, the performance is better when $b$ is large. The average percentage error for the CB heuristic is 0.54% in the $T$-model and 0.50% in the $Q$-model. The performance of the CM heuristic is similar. The average percentage error is 0.67% for the $T$-model and 0.50% for the $Q$-model. It is surprising that the CM heuristic performs slightly better than the CB heuristic for the $Q$-model with $b = 5$.

To examine the effectiveness of the heuristics when $N$ increases, we test $N = 2, 4,$ and 8. For the $T$-model, we set $K_j = 64/N$ and $Q_j = 8$, for $j = 1, \ldots, N$. For the $Q$-model, we set $k_j = 16/N$ and $T_j = 8$. The other common parameters are $h_j = l_j = 1/N$, $j = 1, \ldots, N$, $b = 9$. Table 1 reports the average percentage errors. The result suggests that neither heuristic performs worse as
Figure 3: The Performance of the Heuristics for Different Backorder costs
The CB Heuristic
The CM Heuristic

<table>
<thead>
<tr>
<th>$(N)$</th>
<th>The CB Heuristic</th>
<th>The CM Heuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$T$-Model ($\epsilon %$)</td>
<td>0.00</td>
<td>0.92</td>
</tr>
<tr>
<td>$Q$-Model ($\epsilon %$)</td>
<td>0.00</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 1: Summary of Heuristic Performance for Different $N$

As we mentioned in §5, Chen and Zheng’s (1998) approach can be used to derive lower bound functions, from which an exhaustive search algorithm can be constructed for finding an optimal solution. Here, we compare the number of feasible solutions reviewed by our and their algorithms. For a fair comparison, we use the heuristic cost $C^M$ in both algorithms. Among the 768 instances tested for the $Q$-model above, our algorithm evaluates fewer feasible solutions in all instances. The total number of feasible solutions evaluated by our algorithm is 70% less than that of Chen and Zheng’s algorithm. Consequently, the computational time of finding optimal solutions is significantly reduced.

Some numerical observations regarding optimal solutions and costs are summarized below.

(1) Both heuristics perform very well on average. In most cases, the cost-parameter-based clustering scheme in the CM heuristic generates the optimal partition. However, we observe that both heuristics perform less effectively in the following cost parameters: $(K_1, K_2, K_3) = (50, 50, 50)$ and $(h_1, h_2, h_3) = (1, 0.1, 1)$. Under these parameters, the CM heuristic suggests two clusters: $c(1) = \{1\}$ and $c(2) = \{2, 3\}$, but the optimal partition is $c(1) = \{1, 2, 3\}$. Nevertheless, using the single-stage cost function $\tilde{g}_j(\cdot, T)$ to generate solutions seems to be very effective. In the above worst-performance instances, the CM heuristic generates the optimal solution if the optimal partition is provided.

(2) Since the CM heuristic is in general very effective, it helps to visualize the optimal behaviors for the system. For example, fixed costs and echelon holding costs play an important role for determining effective replenishment intervals or batch sizes. In particular, reducing fixed cost for a stage may affect the optimal replenishment intervals or batch sizes of the other stages when the system operates optimally.

(3) We observe that the optimal batch sizes are more sensitive to the change of cost parameters than the optimal replenishment intervals; however, the optimal cost is more sensitive to the change of replenishment intervals. In addition, changing the replenishment intervals in the
Q-model often leads to changes of optimal batch sizes. van Houtum et al. (2006) categorize both “determining batch sizes and replenishment intervals” as the “first level” decision. Our study refines this statement by suggesting that replenishment interval decisions should be made before batch size decisions.

(4) For both models, it is more cost efficient to reduce the echelon holding cost at an upstream stage, reduce the leadtime at a downstream stage, and reduce the fixed cost at a downstream stage. These findings are consistent with those in Shang (2004).

### 6.2 Comparison of the \(Q\)- and \(T\)-Model

In the single-stage system with Poisson demand, Rao (2003) shows that a continuous-review \((r, Q)\) policy has cost advantage over a periodic-review \((s, T)\) policy, where \(s\) is the base-stock level (i.e., batch size is one). The cost benefit comes from the fact that an \((r, Q)\) policy is more responsive to the demand. It is foreseeable that the same result holds for the considered serial systems. Our focus here is to examine the cost benefit under different system parameters.

To examine this issue, we consider a three-stage system with Poisson demand. To compare both systems, we assume that \(k_j = K_j\) for all \(j\). For the \(T\)-model, we assume \(Q_j = 1\) for all \(j\), i.e., each stage implements an echelon base-stock policy. For the \(Q\)-model, we consider continuous-review echelon \((r, Q)\) policies. Let the optimal cost obtained from the above \(Q\)-model and \(T\)-model be \(C^*_Q\) and \(C^*_T\), respectively. Define the cost benefit as \((C^*_T - C^*_Q) / C^*_Q \times 100\%\). Fix leadtimes and echelon holding costs as \(l_j = 1\), \(h_j = 1\), \(j = 1, 2, 3\). Table 2 shows nine different fixed cost combinations (case no. 1 to 9). For each of the combinations, we further consider three sets of \(b\) and \(\lambda\). There are a total of 27 instances.

When \(K_j\) are the same for all stages (Cases No. 1 - 3), we observe that the cost benefit is larger when \(K_j\) is small. This is intuitive because when \(K_j\) is small, the \((r, Q)\) policies, which tend to

<table>
<thead>
<tr>
<th>Case No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K_1 = k_1)</td>
<td>150</td>
<td>30</td>
<td>5</td>
<td>150</td>
<td>150</td>
<td>30</td>
<td>30</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(K_2 = k_2)</td>
<td>150</td>
<td>30</td>
<td>5</td>
<td>30</td>
<td>5</td>
<td>150</td>
<td>5</td>
<td>150</td>
<td>30</td>
</tr>
<tr>
<td>(K_3 = k_3)</td>
<td>150</td>
<td>30</td>
<td>5</td>
<td>30</td>
<td>5</td>
<td>150</td>
<td>150</td>
<td>30</td>
<td>150</td>
</tr>
<tr>
<td>(b = 9, \lambda = 4)</td>
<td>4.61%</td>
<td>8.72%</td>
<td>11.68%</td>
<td>6.25%</td>
<td>6.87%</td>
<td>6.87%</td>
<td>10.26%</td>
<td>9.17%</td>
<td>11.84%</td>
</tr>
<tr>
<td>(b = 9, \lambda = 10)</td>
<td>2.48%</td>
<td>3.94%</td>
<td>4.50%</td>
<td>3.39%</td>
<td>3.39%</td>
<td>6.33%</td>
<td>5.75%</td>
<td>5.53%</td>
<td>4.43%</td>
</tr>
<tr>
<td>(b = 39, \lambda = 4)</td>
<td>11.45%</td>
<td>15.31%</td>
<td>17.86%</td>
<td>14.10%</td>
<td>14.10%</td>
<td>16.91%</td>
<td>17.20%</td>
<td>16.71%</td>
<td>17.40%</td>
</tr>
</tbody>
</table>

Table 2: Cost Benefit Comparison under Different System Parameters
order more frequently, can take the advantage of smaller fixed order costs.

Cases No. 4 - 9 provide the cost benefits when fixed costs are not identical. We observe that when \( K_1 \) is large (i.e., equal to 150), the cost benefit is smallest. Examining the optimal solutions in Cases No. 4 and 5, we find that \( T_1^* = T_2^* = T_3^* \) and \( Q_1^* = Q_2^* = Q_3^* \). This implies that when all stages belong to a single cluster, the cost benefit is smaller. This may be explained as follows: If all stages belong to a cluster, the optimal batch sizes at downstream stages tend to be larger than those otherwise. Larger batch sizes at downstream stages may hurt the responsiveness of \((r, Q)\) policies. Consequently, the cost benefit becomes smaller.

Finally, we find that the cost benefit tends to increase in \( b \) and decrease \( \lambda \). Thus, companies should consider installing advanced inventory tracking systems that can automatically generate replenishment orders in a real-time fashion if backlogged penalty cost is high or demand rate is low.

7. Concluding Remarks

This paper studies serial inventory systems with batch ordering and nested replenishment schedule. We consider two models by assuming either batch sizes or replenishment intervals are fixed. We develop lower and upper bounds on the exact total cost per period. These bounds are derived by regulating replenishment intervals or batch sizes at downstream stages. These bounds lead to two effective heuristics and an optimization algorithm. The algorithm generates optimal solutions faster than the existing one for the model with fixed replenishment intervals. Numerical observations suggest that, from the system-design perspective, replenishment interval decisions should receive more attention than batch size decisions. In addition, we assess the cost benefit of the continuous-review, batch-ordering policy over the periodic-review policy. We find that the cost benefit decreases if downstream stages have higher fixed replenishment/setup costs or larger demand rate, and increases in backorder cost.

Our analysis can be extended to other fixed cost assumptions. For instance, we can assume that the setup cost incurred once for multiple order batches in the \( Q \)-model. In such a setting, we need to multiply \( k_j \) by a probability that an order is triggered at a replenishment period in steady state (see Zheng and Chen (1992) for the derivation of this probability in the \( Q \)-model with single-period replenishment intervals). Since the revised fixed cost term is nonincreasing in batch size, all results derived for the \( Q \)-model hold. For the \( T \)-model, we can also assume that \( K_j \) occurs only if an order is placed. Similarly, we need to multiply \( K_j \) by the same probability. Again, we can show that the
fixed cost term is nondecreasing in replenishment interval, so all results derived for the T-model
hold. Finally, our analysis can be extended to assembly systems, assuming that base quantities
and replenishment cycles do not decrease with their total leadtimes (see Chen (2000) and Chao
and Zhou (2005) for detailed descriptions of such assembly systems).

Acknowledgment: The authors wish to thank the referees for their helpful comments.

Appendix: Proofs

Throughout the proofs, for convenience, we assume T is a continuous variable. All results hold for
the discrete-time case with additional technicality.

Proposition 1

We first review an existing result. For fixed T, Chao and Zhou (2005) show that the optimal
reorder points r_j(T_j) can be obtained recursively by minimizing a sequence of convex functions.
Their procedure is essentially the same as minimizing our G_j(·, T_j) functions defined in (6) and
(7). (Note that in each iteration, the cost function in Chao and Zhou is the total cost in an order
cycle of echelon j, while ours is the average total cost per period for echelon j.) More specifically,
r_1(T_1) = \arg \min_y \int_0^{Q_1} G_1(y + x, T_1) dx. For j = 2, ..., N, r_j(T_j) = \arg \min_y \int_0^{Q_j} G_j(y + x, T_j) dx,
where G_j(·, T_j) is defined in (7) with r_{j-1} replaced by r_{j-1}(T_{j-1}) in the O_{j-1} function.

For notational simplicity, let r_j = r_j(T_j). Also, let G_j(·, T_j) be the function that the optimal
reorder points r_i(T_i), i = 1, ..., j-1 are in place. The following lemma sets the stage for the proofs
of parts 1 and 3.

Lemma 6 G_j(y, T_j) = g_j(y, T_j) + \sum_{i=1}^{j-1} I_i(r_i, T_i) for j = 1, ..., N.

Proof. We show this result by induction. It holds for j = 1 by definition. For j = 2,

\[ G_2(y, T_2) = \frac{1}{T_2} \sum_{\ell=0}^{T_2-1} \mathbb{E} \left[ h_2(y - D(l_2 + \ell + 1) + G_1 \left( O_1 \left[ y - D \left( l_2 + \left\lfloor \frac{\ell}{T_1} \right\rfloor T_1 \right] \right), T_1 \right) \right] \]

\[ = \frac{1}{T_2} \sum_{\ell=0}^{T_2-1} \mathbb{E} \left[ h_2(y - D(l_2 + \ell + 1) + g_{1,2} \left( y - D \left( l_2 + \left\lfloor \frac{\ell}{T_1} \right\rfloor T_1 \right) \right), T_1 \right] + I_1(r_1, T_1) \]

\[ = g_2(y, T_2) + I_1(r_1, T_1), \]

where the second equality follows from that g_1(y, T_1) = G_1(y, T_1) and the definition of g_{1,2}(y, T_1).
Suppose G_j(y, T_j) = g_j(y, T_j) + \sum_{i=1}^{j-1} I_i(r_i, T_i), then for j + 1,
Let $y$ be any point in $\mathbb{R}^3$, from Lemma 6, By definition,

$$T - 1 \geq 0$$

The last inequality holds because $G_j$ and $G_{j+1}$ are non-negative.

For part 2, to show $C(T)$ is increasing in $T$, it is sufficient to show that $g_j(y, T)$ is submodular in $y$ and $T_j$. Taking derivative with respect to $y$, we have

$$\frac{1}{T_j} \int_0^{T_j} \int_0^{Q_j} \int_{z=0}^{q_{j-1}} h_j + E g'_{j-1} \left( \min \left\{ y + zQ_j - D \left( l_j + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right), r_j-1(T_j-1) \right\} + x, T_j \right) dx dz d\ell.$$

Then, taking derivative of the above function with respect to $T_j$, we have

$$\frac{1}{T_j^2} \left( T_j \int_0^{Q_j-1} \int_{z=0}^{q_{j-1}} h_j + E g'_{j-1} \left( \min \left\{ y + zQ_j - D (l_j + (r_j-1-T_j)T_j-1), r_j-1(T_j-1) \right\} + x, T_j \right) dx dz \right) - \int_0^{T_j} \int_0^{Q_j-1} \int_{z=0}^{q_{j-1}} h_j + E g'_{j-1} \left( \min \left\{ y + zQ_j - D \left( l_j + \left\lfloor \frac{\ell}{T_j} \right\rfloor T_j \right), r_j-1(T_j-1) \right\} + x, T_j \right) dx dz d\ell \leq 0.$$

The last inequality holds because $\int_0^{Q_j-1} g_{j-1}(y+x) dx$ is convex in $y$ and $\int_0^{Q_j-1} g_{j-1}(\min\{y+x, R_{j-1}(T_j-1)\} + x) dx$ is decreasing in $y$.

**Lemma 2**

By definition,

$$\int_0^{T_j} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1,\ell} (y + x - D(l_j), T_j) \right] dx$$
\[ = \int_{x=0}^{Q_j-1} \int_{z=0}^{n_j-1-1} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1} \left( \min \left\{ y + zQ_{j-1} - D(l_j), z_{j-1}(T_j) \right\} + x, T_j \right) \right] dz \] 

\[ -Q_j L_{j-1}(r_{j-1}(T_j), T_j) \]

and

\[ \int_{x=0}^{Q_j-1} \int_{z=0}^{n_j-1-1} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1} \left( \min \left\{ y + zQ_{j-1} - D(l_j + \frac{\ell}{T_j-1} T_{j-1}), z_{j-1}(T_j) \right\} + x, T_j \right) \right] dz \]

\[ = \int_{x=0}^{Q_j-1} \int_{z=0}^{n_j-1-1} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1} \left( \min \left\{ y + zQ_{j-1} - D \left( l_j + \frac{\ell}{T_j-1} T_{j-1} \right), r_{j-1}(T_{j-1}) \right\} + x, T_{j-1} \right) \right] dz \]

\[ -Q_j L_{j-1}(r_{j-1}(T_{j-1}), T_{j-1}). \]

We need to prove the following three results.

(a) Both \( g_j(y, T_{j-1}) \) and \( g_j(y, T) \) are convex in \( y \), for \( j = 1, \ldots, N \).

(b) Both \( g_{j-1,j}(y, T_{j-1}) \) and \( g_{j-1,j}(y, T) \) are convex decreasing in \( y \), for \( j = 2, \ldots, N \).

(c) For \( j = 2, \ldots, N \),

\[ \int_{x=0}^{Q_j-1} \int_{z=0}^{n_j-1-1} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1} \left( \min \left\{ y + zQ_{j-1} - D \left( l_j + \frac{\ell}{T_j-1} T_{j-1} \right), z_{j-1}(T_j) \right\} + x, T_j \right) \right] dz \]

\[ -\int_{x=0}^{Q_j-1} \int_{z=0}^{n_j-1-1} \sum_{\ell=0}^{T_j-1} \mathbb{E} \left[ g_{j-1} \left( \min \left\{ y + zQ_{j-1} - D \left( l_j + \frac{\ell}{T_j-1} T_{j-1} \right), r_{j-1}(T_{j-1}) \right\} + x, T_{j-1} \right) \right] dz \]

\[ \leq \sum_{\ell=0}^{T_j-1} Q_j L_{j}(r_{j-1}(T_j), T_j) - \sum_{\ell=0}^{T_{j-1}-1} Q_j I_{j}(r_{j-1}(T_{j-1}), T_{j-1}). \]  

For (a), we have already shown in Proposition 1 for \( j = 1 \). Using induction, we can easily prove both \( g_j(y) \) and \( g_j(y, T) \) is convex in \( y \) for \( j = 2, \ldots, N \). We omit the detailed steps here.

For (b), because \( r_{j-1}(T_{j-1}) \) and \( z_{j-1}(T_j) \) are the minimizer of \( g_{j-1}(y) \) and \( g_{j-1,j}(y, T) \), respectively, it can be shown that \( g_{j-1,j}(y) \) and \( g_{j-1,j}(y, T) \) are convex decreasing in \( y \) by definition.

For (c), notice that the right-hand side of the inequality is a constant and independent of \( y \). It suffices to show the difference between the left-hand side functions is less than or equal to the right-hand side constant. We use the two-stage system with \( T_1 = 1, T_2 = 2 \) and \( Q_1 = Q_2 = 1 \) to illustrate the idea. For convenience, let \( r_1 \) and \( z_1 \) denote \( r_1(T_1) \) and \( z_1(T_2) \), respectively. Recall that when \( T_1 = 1 \),

\[ G_1(y) = \mathbb{E}[h_1(y - D(l_1 + 1)) + (b + h_1'(y - D(l_1 + 1))^\|]. \]

The inequality (19) can be written as

\[ \mathbb{E}G_1(\min\{y - D(l_2, R_1)\}) - \mathbb{E}G_1(\min\{y - D(l_2, r_1)\}) \]
We first show that $\mathbb{E}G_1(\min\{y - D(l_2) - D, \tau_1 - D\}) - \mathbb{E}G_1(\min\{y - D(l_2) - D, r_1\}) \leq G_1(\tau_1) - G_1(r_1) + \mathbb{E}G_1(\tau_1 - D) - G_1(r_1). \tag{20}$

We next show that $\mathbb{E}G_1(\min\{y - D(l_2) - D, \tau_1 - D\}) - \mathbb{E}G_1(\min\{y - D(l_2) - D, r_1\}) \leq G_1(\tau_1) - G_1(r_1).$ For any $D = d$, we consider the following four different cases.

Case 1. If $y - D(l_2) - d \leq \min\{r_1 - d, r_1\}$,

$$\mathbb{E}G_1(\min\{y - D(l_2) - d, \tau_1 - d\}) - \mathbb{E}G_1(\min\{y - D(l_2) - d, r_1\}) = 0 \leq G_1(\tau_1 - d) - G_1(r_1).$$

The inequality follows from that $r_1$ minimizes $G_1(y)$.

Case 2. If $y - D(l_2) - d > \max\{\tau_1 - d, r_1\}$,

$$\mathbb{E}G_1(\min\{y - D(l_2) - d, \tau_1 - d\}) - \mathbb{E}G_1(\min\{y - D(l_2) - d, r_1\}) = G_1(\tau_1 - d) - G_1(r_1).$$

Case 3. If $\tau_1 - d \leq y - D(l_2) - d \leq r_1,$

$$\mathbb{E}G_1(\min\{y - D(l_2) - d, \tau_1 - d\}) - \mathbb{E}G_1(\min\{y - D(l_2) - d, r_1\}) = G_1(\tau_1 - d) - \mathbb{E}G_1(y - D(l_2) - d) \leq G_1(\tau_1 - d) - G_1(r_1).$$

The inequality follows from that $r_1$ minimizes $G_1(y)$.

Case 4. If $r_1 \leq y - D(l_2) - d \leq \tau_1 - d,$

$$\mathbb{E}G_1(\min\{y - D(l_2) - d, \tau_1 - d\}) - \mathbb{E}G_1(\min\{y - D(l_2) - d, r_1\}) = \mathbb{E}G_1(y - D(l_2) - d) - G_1(r_1) \leq G_1(\tau_1 - d) - G_1(r_1).$$

The inequality follows from that $r_1 \leq y - D(l_2) - d \leq \tau_1 - d$. Thus, $\mathbb{E}G_1(\min\{y - D(l_2) - D, \tau_1 - D\}) - \mathbb{E}G_1(\min\{y - D(l_2) - D, r_1\}) \leq \mathbb{E}G_1(\tau_1 - D) - G_1(r_1)$ is true. Consequently, (20) holds for the two-stage system with $T_1 = 1$, $T_2 = 2$ and $Q_1 = Q_2 = 1$. Thus, we have shown that the result is true for this special two-stage system. To prove the case with general $Q_1$ and $Q_2$, we only need to define $\bar{G}_1(y) = \int_0^Q G_1(y + x)dx$ and replace $G_1(y)$ in previous derivation with $\bar{G}_1(y)$. For general $T_1$ and $T_2$, we can follow the similar procedure but have to validate $T_2$ inequalities.

An induction based on a similar analysis can show the general $j$ case. Detailed proof is available from the authors upon request.
Proposition 4

We first prove part 1 by induction. When \(j = 1\), it is clear that \(L_1(r_1(T), T)\) is convex in \(T\). Suppose that part 1 is true for some \(j\). For \(j + 1\), by definition

\[
L_{j+1}(y, T) = \frac{\int_0^{Q_{j+1}} g_{j+1}(y + x, T)dx}{Q_{j+1}}
\]

\[
= \frac{1}{Q_{j+1}T} \left( \int_0^{Q_{j+1}} \int_{0}^{T} \mathbb{E} \left[ h_{j+1}(y + x - D(l_{j+1} + \ell + 1)) + g_{j,j+1}(y + x - D(l_{j+1}), T) \right] dxd\ell \right).
\]

Note that it is sufficient to show that \(\int_0^{Q_{j+1}} \mathbb{E} g_{j,j+1}(y + x - D(l_{j+1}), T)dx\) is jointly convex in \(y\) and \(T\).

\[
\int_0^{Q_{j+1}} \mathbb{E} g_{j,j+1}(y + x - D(l_{j+1}), T)dx
\]

\[
= \int_0^{Q_{j+1}} \mathbb{E} g_j(O_j[y + x - D(l_{j+1})], T)dx - Q_{j+1}L_j(r_j(T), T)
\]

\[
= \int_0^{Q_{j+1}} \int_{0}^{Q_j} \mathbb{E} g_j(\min\{y + zQ_j - D(l_{j+1}), r_j(T)\} + x, T)dxdz - Q_{j+1}L_j(r_j(T), T),
\]

which is convex by induction. Therefore, \(L_{j+1}(y, T)\) is jointly convex in \(y\) and \(T\). As a result, \(L_{j+1}(\ell_{j+1}(T), T)\) is convex in \(T\).

For part 2,

\[
C(T) = \sum_{j=1}^{N} c_j(T_j) = \sum_{j=1}^{N} \left( \frac{K_j}{T_j} + I_j(r_j(T_j), T_j) \right) \geq \sum_{j=1}^{N} \left( \frac{K_j}{T_j} + L_j(\ell_j(T_j), T_j) \right) = \sum_{j=1}^{N} c_j(T_j).
\]

References


24
References


