A Simple Heuristic for Serial Inventory Systems with Fixed Order Costs

Kevin H. Shang

The Fuqua School of Business, Duke University, Durham, North Carolina 27708, USA
khshang@duke.edu

July 2004; Revised July 4, 2005

Order coordination is key to improving supply chain efficiency. To coordinate orders, it is crucial to find centralized solutions for order quantities. In the deterministic demand regime, there exist easily implementable integer-ratio policies which guarantee 94% effectiveness. However, no analogous heuristic exists for the stochastic model. In this paper, we aim to propose one. Specifically, we consider serial inventory systems with stationary demand. There are fixed order costs for each order placed. We propose a heuristic that is 94% effective for the deterministic model, and demonstrate that the same approach may be used to find near-optimal order quantities for its stochastic counterpart. Our heuristic involves two steps. First, cluster the stages based on ratios of cost parameters. Second, solve a single-stage problem for each cluster. Numerical studies show that this simple heuristic outperforms the existing heuristics. Our study suggests that managing order quantities in the stochastic system may be as simple as managing those in the deterministic system. Thus, insights from the deterministic model may be carried over to the stochastic one.
(*Key Words: Simple Heuristic; Power-of-Two Policies; Fixed Order Costs; Order Quantities*)

---

1. Introduction

In many production/distribution systems, materials move in batches from one location to another. A batch can be a full container moving from one work station to another in a production system or an order taken from supplier to retailer in a distribution system. While correctly determining batch sizes between entities in a supply chain is important in order to coordinate orders, it is difficult to obtain exact optimal solutions. Therefore, most research on the topic aims to develop simple and effective heuristics. When demand is *deterministic*, there exists an easily implementable integer-ratio heuristic which guarantees 94% effectiveness. These heuristics are obtained by solving the relaxed problem in two steps: (1) Determine optimal clusters among stages based on cost parameters (the stages in a cluster have the same order quantity). (2) Solve an EOQ problem for each cluster. After this, a heuristic policy can be obtained by converting the relaxed problem's solutions into integer-ratio order quantities. One well-known integer-ratio heuristic is the "power-of-two" policy, which, as the name suggests, converts its solutions into power-of-two order quantities (Maxwell and Muckstadt 1985, Roundy 1985). This class of heuristic policies has been shown effective in many other general settings (e.g., Atkins and Sun 1995, Federgruen et al. 1992, Jackson et al. 1988, Roundy 1986). We refer the reader to Atkins (1990), Muckstadt and Roundy (1993), Silver et al. (1998), and Zipkin (2000) for a complete review.

Unfortunately, no analogous solution procedure exists for the stochastic regime. The purpose of this paper is to propose one. In particular, we consider a serial inventory system with fixed order costs at each stage. There are linear holding costs at all stages and linear backorder costs at the most downstream stage. Leadtimes are constant. The objective is to minimize the long-run average system-wide cost. The model was first introduced by Clark and Scarf (1962). They point out that characterizing optimal policies is extremely difficult. Thus, an optimal policy, even if it exists, may be difficult to implement. As a result, this paper considers stationary echelon (r, nq) policies. Here the order quantity is always a multiple of q, chosen so as to raise the echelon inventory position to the interval of $(r, r+q]$. We call q the base order quantity; the base order quantities q must satisfy an integer-ratio constraint in order to fulfill order coordination.

Most of the work on (r, nq) policies focuses on developing approximate or exact total cost expressions (see, e.g., Axsäter 1993, 2000, Axsäter and Rosling 1993, Badinelli 1992, Cachon 1995, Chen and Zheng 1994a, De Bodt and Graves 1985, Svoronos and Zipkin 1988). To our knowledge, the only paper that optimizes both reorder points and order quantities is Chen and Zheng (1998). Nevertheless, finding optimal policy parameters requires extensive numerical search, which usually requires considerable computation effort. On the other hand, when base order quantities are fixed, Chen (2000) derives recursive equations to calculate the optimal reorder points. Shang and Song (2003b) further construct single-stage solution bounds and approximations for the optimal reorder point. Recently, Doğru et al. (2005) show that the optimal reorder points satisfy newsvendor inequalities. In other words, if we can find effective base order quantities, we can use them to search for the corresponding optimal or approximated reorder points. Along these lines, our focus here is to develop a simple and effective heuristic for base order quantities.

The idea for constructing this heuristic stems from the integer-ratio heuristic in the deterministic model. In particular, we first propose an integer-ratio heuristic which also guarantees 94% effectiveness in the deterministic model. This heuristic employs a bottom-up procedure for finding the optimal order quantity by restricting the order quantity of a cluster to an integer multiple of that of the downstream cluster. Then, we show that the same solution procedure can be applied to find near-optimal base order quantities for its stochastic counterpart. The building blocks of this heuristic are derived from two theoretical findings. First, we derive a new objective function which is the sum of N independent cost functions from single-stage (r, nq) systems with the original problem data. We show that effective base order quantities can be obtained by considering this new objective function. Second, we show that the optimal order quantity for a single-stage (r, nq) system increases with the ratio of the fixed order cost to the holding cost. With these results, we may determine *near-optimal* clusters based on the ratios of cost parameters. An extensive numerical study indicates that this heuristic outperforms the existing heuristics in Chen and Zheng (1998).

The contributions of this paper include the following. First, we simplify the computation involved in obtaining effective base order quantities in the stochastic model. This result streamlines order coordination in serial supply chains. Second, our result generalizes the solution procedure for finding the effective order quantities of both deterministic and stochastic models. This extends the sections of existing studies on deterministic models in which the authors attempt to shed light on managing stochastic systems (See, e.g., Atkin and Sun 1995, Chen 1998). Third, our heuristic provides a justification for why the base quantities in the stochastic model are always larger than those obtained from the corresponding deterministic model (see Chen and Zheng 1998). Finally, our work suggests that managing order quantities in a stochastic serial inventory system may be as simple as managing those in its deterministic counterpart. Thus, many insights from the deterministic model can be carried over to the stochastic model.

The rest of the paper is organized as follows: In §2, we review the solution procedure in the deterministic model and suggest an effective integer-ratio heuristic. In §3, we show that the same solution procedure can be applied to the stochastic model. In §4, we test the heuristic for the stochastic model and summarize insights. §5 concludes the paper. Appendix A summarizes Chen's (2000) results related to ours. All proofs are presented in Appendix B.

2. A Heuristic for Deterministic Model

In this section, we briefly review the existing work related to our analysis. Then, we propose an integer-ratio heuristic which is 94% effective. As we shall show in §3, the same heuristic approach can be applied to the stochastic model.

Consider an N-stage serial production/distribution system. Materials flow from stage N to

$N-1$, $N-1$ to $N-2$, etc, until stage 1, where constant demand rate λ occurs. Define $h_j^{'}$ and h_j the installation and echelon holding costs at stage j respectively and $h_j^{'} = \sum_{i=j}^{N} h_i$, $j = 1, ..., N$. A fixed order cost k_j is charged for placing an order at stage j. The objective is to minimize the long-run average cost. Schwarz (1973) show that the optimal policy must possess both zero-inventory and nested properties. In this paper, we focus on stationary policies. Let the order quantity at stage j be q_j. Thus, the problem can be formulated as follows:

$$\begin{array}{lll} (DP) & \min_{\mathbf{q}} & \sum_{j=1}^{N} C_j^d(q_j) \\ & \text{s.t.} & q_j = n_j q_{j-1}, \quad j = 2, ..., N, \end{array}$$

where $C_j^d(q_j) = k_j\lambda/q_j + q_j h_j/2$ and n_j are positive integers. The lower bound to the optimal cost of (DP) can be obtained by solving its relaxation, i.e., solving the same objective function with the relaxed constraints $q_j \geq q_{j-1}$. Since the objective function is jointly convex in q_j, a cluster algorithm can find the optimal solution (Geoffrion 1967, Maxwell and Muckstadt 1985, Schwarz and Schrage 1975). Let $S = \{1, 2, ..., N\}$. For any $i, j \in S$ with $i \leq j$, the set $\{i, i+1, ..., j\}$ is called a cluster. For any cluster $c(k)$ with index k, define

$$\underline{Q}_k = \arg\min_{Q} \sum_{i \in c(k)} C_i^d(Q). \tag{1}$$

A partition of S is a set of disjoint clusters whose union is S. A partition $\{c(1), c(2), ..., c(m)\}$ is optimal if and only if

(i) $\underline{Q}_1 < \underline{Q}_2 < ... < \underline{Q}_m$, and

(ii) for each cluster $c(k) = \{l_1, ..., l_2\}$, there does not exist an l with $l_1 \leq l \leq l_2$ so that $\underline{Q}_{k^-} < \underline{Q}_{k^+}$, where $c(k^-) = \{l_1, ..., l\}$ and $c(k^+) = \{l+1, ..., l_2\}$.

Since $\underline{Q}_k$ in (1) is equal to $\sqrt{2(\sum_{i \in c(k)} k_i)\lambda)/(\sum_{i \in c(k)} h_i)}$, we can replace $\underline{Q}_k$ in (i) and (ii) with the corresponding cost ratio $(\sum_{i \in c(k)} k_i)/(\sum_{i \in c(k)} h_i)$ to search for an optimal partition. See Zipkin (2000) for an example of using a two-dimensional graph to perform this task.

Let $\{c(1), ..., c(m)\}$ be an optimal partition. A heuristic policy $(q_1^0, ..., q_N^0)$ for (DP) can be obtained by converting $(\underline{Q}_1, ..., \underline{Q}_m)$ into integer-ratio vector $(Q_1^0, ..., Q_m^0)$ (i.e., Q_k^0 is an integer multiple of Q_{k-1}^0, $k = 2, ..., m$) and letting $q_i^0 = Q_k^0$ for $i \in c(k)$. It is well known that the effectiveness of the resulting policy $(q_1^0, ..., q_N^0)$ is equal to

$$\min \quad \left\{ \frac{1}{2}\left(\alpha_k + \frac{1}{\alpha_k}\right), k = 1, ..., m \right\} \times 100\%,$$

where $\alpha_k = Q_k^0/\underline{Q}_k$. Roundy (1985) considers a one-warehouse, multi-retailer system and defines a general class of q-optimal integer-ratio policies. His result can be modified to accommodate the serial model.

Lemma 1 *(Roundy 1985) If α_k is in the root-two interval $[1/\sqrt{2}, \sqrt{2}]$, the policy $(q_1, ..., q_N)$ is 94% effective.*

A well-known power-of-two policy can be found by setting $Q_k^0 = 2^{x_k}, k = 1, ..., m$, where x_k is the unique integer with $2^{x_k}/\sqrt{2} \leq \underline{Q}_k \leq 2^{x_k}\sqrt{2}$. It is easy to verify that the resulting α_k ratios are within the range of $[1/\sqrt{2}, \sqrt{2}]$.

We suggest an alternative procedure to find 94%-effective Q_k^0: After an optimal partition is determined, we solve an EOQ problem for each cluster by restricting the order quantity in $c(k)$ to an integer multiple of the order quantity in $c(k-1)$, $k = 2, ..., m$. Specifically, let the resulting optimal order quantity be Q_k^s.

$$Q_1^s \quad = \quad \underline{Q}_1 = \arg\min_{Q_1} \sum_{i \in c(1)} C_i^d(Q_1) = \sqrt{\frac{2(\sum_{i\in c(1)} k_i)\lambda}{\sum_{i\in c(1)} h_i}}.$$

Set $Q_k = nQ_{k-1}^0$, where n is a positive integer, and solve

$$Q_k^s \quad = \quad \arg\min_{Q_k} \sum_{i\in c(k)} C_i^d(Q_k).$$

Denote the heuristic solution by $(q_1^s, ..., q_N^s)$, where $q_i^s = Q_k^s, i \in c(k), k = 1, ..., m$.

Proposition 2 *The policy $(q_1^s, ..., q_N^s)$ is 94%-effective.*

We perform a numerical study to examine the effectiveness of the heuristic policy, using the power-of-two policy as a benchmark. The parameters are $N = 3$, $h_j = 0.1$, 0.5, 1, 1.5, 2.5, 5, 7.5, 8.75, 10, 12.5, j=1,2,3, $k_j = 2$, 10, 50, 70, 100, 250, 500, 750, 1250, j=1,2,3, $\lambda = 20$, 50, 100, 150, 250, 500, 750, 1000, 1500, 2500. The total number of tested cases is 7.29×10^6. On average, the percentage error compared to the lower bound cost is 2.01% for the power-of-two policies and 0.37% for our heuristic policies. Power-of-two policies outperform ours only in about 10% of instances. The maximum cost errors are about the same (6.07% for the power-of-two policy and 6.05% for our heuristic policy). Our heuristic outperforms the power-of-two one for the reason that the α_k generated from the power-of-two policy is in general larger than that from our heuristic policy since the former ranges from $\sqrt{2}$ to $1/\sqrt{2}$ while the latter from $\sqrt{(n+1)/n}$ to $\sqrt{n/(n+1)}$ $(n \geq 1)$.

3. The Stochastic Model

The stochastic model is a continuous-review inventory system, in which demands arrive at rate λ with mean demand size of μ. The demand sizes are i.i.d. and independent of demand epoches. We assume that demand is continuous, and cumulative demand is a process with stationary increments and nondecreasing sample paths. Each stage implements an echelon (r_j, nq_j) policy: whenever the echelon inventory position (inventory on order +inventory on hand + inventory at or in transit to all downstream stages - backorders) is at or below the reorder point r_j, order nq_j units where n is the minimum integer required to bring the echelon inventory position to above r_j. We call q_j the base order quantity. These base quantities satisfy the integer-ratio constraints, i.e., $q_j = n_j q_{j-1}$, n_j is an integer, $j \geq 2$. Replenishment orders to stage j are delivered after a positive leadtime L_j. All stockouts are backordered with a cost rate b. The other cost parameters defined in the deterministic model are still valid here. We concentrate on minimizing the long-run average costs per unit time.

To evaluate a given (r, nq) policy, Chen and Zheng (1994a) provide a top-down recursive procedure. Define the following inventory variables in equilibrium:

$$\begin{aligned} IP_j &= \text{echelon inventory position at stage } j, \\ IN_j &= \text{echelon inventory level at stage } j. \end{aligned}$$

With the demand process described above, IP_N is uniformly distributed over $(r_N, ..., r_N + q_N]$ (Serfozo and Stidham 1978, Zipkin 1986a). The steady-state distributions of the other inventory variables can be obtained recursively: $IN_j = IP_j - D_j$ and $IP_{j-1} = O_{j-1}[IN_j]$, $j = 2, ..., N$, where

$$O_j[x] = \begin{cases} x, & x \leq r_j + q_j, \\ x - mq_j, & \text{otherwise.} \end{cases} \tag{2}$$

Here, m is the largest integer so that $x - mq_j > r_j$. Our objective is to find the policy parameters $(\mathbf{r}^*, \mathbf{q}^*)$ such that the long-run average cost is minimized while the order quantities satisfy the integer-ratio constraint:

$$\min_{\mathbf{r},\mathbf{q}} \quad C(\mathbf{r}, \mathbf{q}) = \sum_{i=1}^{N} \frac{k_i \lambda \mu}{q_i} + \mathsf{E}[\sum_{i=1}^{N} h_j IN_j + (b + h_1^{'}) IN_1^{-}] \tag{3}$$

$$\text{s.t.} \quad q_j = n_j q_{j-1}, j = 2, ..., N, \tag{4}$$

where $x^- = \max\{-x, 0\}$.

In the subsequent sections, we demonstrate how to construct a similar clustering-minimization heuristic. In §3.1, we propose a decomposition scheme for the objective function in (3) and develop an algorithm for calculating the optimal reorder points with fixed base order quantities. With these results, in §3.2, we derive a new objective function which consists of N independent cost functions from single-stage (r, nq) systems with the original problem data. We show these single-stage cost functions can be used to search for base quantities. In §3.3, we demonstrate how to identify a near-optimal partition through the cost ratios by considering this new objective function. We propose two clustering-minimization heuristics in §3.4.

3.1 Decomposition of Total Cost Function

To construct a similar clustering-minimization heuristic in the stochastic model, the first step is to decouple the total cost function in (3) into the sum of N stage-independent cost functions. We propose a bottom-up decomposition scheme that lays the foundation for this aim.

For fixed (feasible) base order quantities $(q_1, ..., q_N)$, set

$$G_1(y) \quad = \quad \mathsf{E}[h_1(y - D_1) + (b + h_1^{'})(y - D)^-]. \tag{5}$$

Suppose $G_j(\cdot), j \geq 1$ is defined. Denote the average inventory cost

$$I_j(y, q) \quad = \quad \frac{\int_0^q G_j(y + x)dx}{q}. \tag{6}$$

Let

$$r_j(q_j) \quad = \quad \arg\min_r \{I_j(r, q_j)\} = \arg\min_r \left\{ \frac{\int_0^{q_j} G_j(r + x)dx}{q_j} \right\}. \tag{7}$$

For $j = 1, ..., N - 1$, define

$$G_{j,j+1}(y) \quad = \quad G_j(O_j[y]) - I_j(r_j(q_j), q_j), \tag{8}$$

$$G_{j+1}(y) \quad = \quad \mathsf{E}[h_{j+1}(y - D_{j+1}) + G_{j,j+1}(y - D_{j-1})], \tag{9}$$

where

$$O_j[x] \quad = \quad \begin{cases} x, & x \leq r_j(q_j) + q_j, \\ x - mq_j, & \text{otherwise}. \end{cases} \tag{10}$$

Also, define

$$C_j(r_j(q_j), q_j) \quad = \quad C_j(q_j) = (k_j \lambda \mu)/q_j + I_j(r_j(q_j), q_j), \quad j = 1, ..., N. \tag{11}$$

We have

Proposition 3

(1) $I_j(y, q_j)$ *is convex in* y *for any feasible* q_j.

(2) $(r_1(q_1), ..., r_N(q_N))$ *are the optimal reorder points with fixed base quantities* $(q_1, q_2, ..., q_N)$.

(3) $C(\mathbf{r}, \mathbf{q}) \geq C(\mathbf{r}(\mathbf{q}), \mathbf{q}) = \sum_{i=1}^{N} C_i(q_i)$.

Note that Equations (8) to (10) can be used to evaluate any (r, nq) policy by replacing $r_j(q_j)$ with any given reorder points r_j.

Proposition 3 Part (1) indicates that a simple search can find the optimal reorder point $r_j(q_j)$. Part (2) shows that the recursion in Equations (7)-(9) serves an algorithm to find optimal reorder points with fixed base quantities. It is worth noting that the distinction between the above algorithm and the one in Chen (2000). Chen proposes a different recursion to decompose the total cost function (without fixed order costs): Let $\hat{G}_1(\cdot) = G_1(\cdot)$. Suppose $\hat{G}_j(\cdot), j \geq 1$ is available. The optimal reorder point $\hat{r}_j(q_j)$ is equal to $\arg\min_r\{\int_0^{q_j} \hat{G}_j(r + x)dx\}$. Substitute $\hat{r}_j(q_j)$ with r_j in $O_j[y]$ to get $\hat{G}_{j,j+1}(\cdot)$ and $\hat{G}_{j+1}(\cdot)$ as below:

$$\begin{aligned} \hat{G}_{j,j+1}(y) &= \hat{G}_j(O_j[y]) - \min_m \hat{G}_j(y + mq_j), \quad m \text{ is an integer}, &(12)\\ \hat{G}_{j+1}(y) &= \mathsf{E}[h_{j+1}(y - D_{j+1}) + \hat{G}_{j,j+1}(y - D_{j+1})]. &(13) \end{aligned}$$

Note that our induced penalty function $G_{j,j+1}(\cdot)$ as well as the resulting $G_j(\cdot)$ are different from $\hat{G}_{j,j+1}(\cdot)$ and $\hat{G}_j(\cdot)$. As we shall see in the next section, we are able to construct a lower bound on $G_j(\cdot)$, which leads to the single-stage cost function used in the heuristic. Figure 1 illustrates the differences between these cost functions. Both $G_{2,3}(\cdot)$ and $\hat{G}_{2,3}(\cdot)$ ($G_3(\cdot)$ and $\hat{G}_3(\cdot)$) functions have "waves" with a cycle length of q_{j-1}: the former occurs on the right-hand side of $G_{2,3}(\cdot)$ ($G_3(\cdot)$) and the latter on the left-hand side of $\hat{G}_{2,3}(\cdot)$ ($\hat{G}_3(\cdot)$).

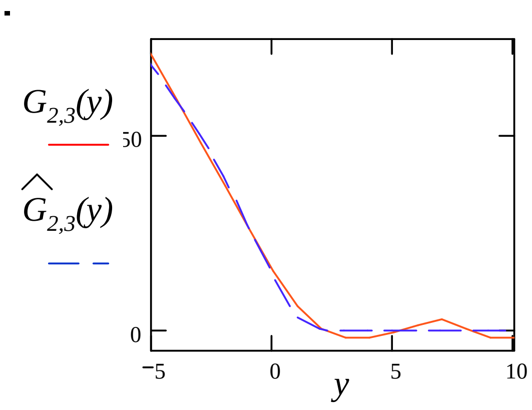
Proposition 3 Part (2) implies that the problem (3)-(4) can be reduced to

$$\begin{aligned} (SP) \qquad & \min_{\mathbf{q}} \sum_{i=1}^{N} C_i(q_i) \\ & \text{s.t.} \quad q_j = n_j q_{j-1}, j = 2, ..., N, \end{aligned}$$

and we aim to find effective base quantities to (SP).

3.2 Stage-Independent Cost Functions

It is difficult to solve (SP) directly since each $C_j(q_j)$ in the objective function depends on the downstream base quantities $q_i, i < j$ through $G_j(\cdot)$. As a result, we construct N stage-independent



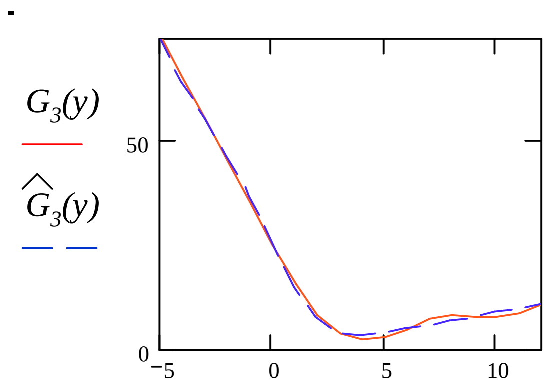


Figure 1: (a) The comparison of $G_{2,3}(y)$ and $\hat{G}_{2,3}(y)$. (b) The comparison of $G_3(y)$ and $\hat{G}_3(y)$. The system has a Poisson demand with parameters $N = 3$, $b = 10$, $\lambda = 1$, $h_j = L_j = 1$, $j = 1, 2, 3$, $k_1 = k_2 = 10$, $k_3 = 100$, $q_1 = q_2 = 6$.

functions to replace $G_j(\cdot)$, $j \geq 1$ so as to decouple the objective function. Certainly, these replacement functions must be able to generate effective order quantities for (SP).

From (7)-(9), we observe that the optimal reorder point $r_j(q_j)$ at stage j is completely independent of the decisions at its upstream stages $j+1, ..., N$. More precisely, $r_j(q_j)$ is solely determined by b, $h_1' = \sum_{i=1}^{N} h_i$, and $(q_i, r_i(q_i), D_i, h_i)$, for $i = 1, ..., j$. This motivates us to consider a truncated echelon system, termed as the *echelon-j* system, which includes the stages $j, j-1, ..., 1$ of the original system. In this truncated system, stage j is the most upstream stage with ample supply.

Now consider the *echelon-j* system. Suppose we want to determine the optimal base quantity q_j as well as the corresponding reorder point $r_j(q_j)$ at stage j with *known* downstream policy parameters $(r_i(q_i^0), q_i^0)_{i=1}^{j-1}$ and a fixed cost k_j. Proposition 4 below indicates that solving this problem is similar to solving a single-stage (r, nq) system in Zheng (1992), except for the fact that the base quantity q_j must be an integer multiple of q_{j-1}^0.

Proposition 4 *For any positive integer n, $C_j(q_j) = C_j(nq_{j-1}^0) = \frac{k_j \lambda \mu}{nq_{j-1}^0} + I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$ is quasi-convex in n.*

Proposition 4 indicates that the optimal order quantity at stage j, q_j^0, can be found through simple search on n, i.e., $q_j^0 = n^* q_{j-1}^0$, where $n^* = \arg\min_n \{C_j(nq_{j-1}^0)\}$. Figure 2 illustrates how the fixed order cost $k_3 \lambda \mu / (nq_2^0)$, inventory cost $I_3(r_3(nq_2^0), nq_2^0)$, and total cost $C_j(nq_2^0)$ at stage 3 change as n increases in the same example of Figure 1.

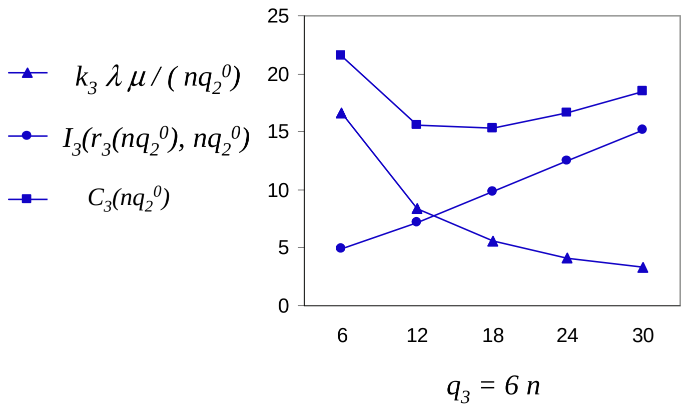


Figure 2: The functions of $k_3\lambda\mu/(nq_2^0)$, $I_3(r_3(nq_2^0), nq_2^0)$, and $C_j(nq_2^0)$ with $q_2^0 = 6$. In this example, $n^* = 3$ and $q_3^0 = 18$.

Clearly, if we can construct a parallel curve to the inventory cost $I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$, the total cost curve will shift vertically, but the minimum q_j^0 will stay the same. Below we demonstrate how to construct a stage-independent cost function (to replace $G_j(\cdot)$) such that the resulting inventory cost curve tends to be parallel to $I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$.

With fixed downstream base quantities $(q_1^0, q_2^0, ..., q_{j-1}^0)$, define

$$\tilde{G}_j^\ell(y) = G_j^\ell(y) - \sum_{i=1}^{j-1} I_i(r_i(q_i^0), q_i^0) + \sum_{i=2}^{j} h_i \mathsf{E}[\tilde{D}_{j-1}], \quad j = 1, ..., N,$$

where

$$G_j^\ell(y) = \mathsf{E}[h_j(y - \tilde{D}_j) + (b + h_j^{'})(y - \tilde{D}_j)^-] \text{ and } \tilde{D}_j = \sum_{i=1}^{j} D_i.$$

Denote

$$\tilde{I}_j^\ell(y, q) = \frac{\int_0^q \tilde{G}_j^\ell(y + x)dx}{q} \quad \text{and} \quad r_j^\ell(q_j) = \arg\min_y \tilde{I}_j^\ell(y, q_j).$$

We have

Proposition 5 *For all y, j, and positive integer n,*

(1) $G_j(y) \geq \tilde{G}_j^\ell(y)$.

(2) $I_j(y, nq_{j-1}^0)$ *and* $\tilde{I}_j^\ell(y, nq_{j-1}^0)$ *have the same asymptotic slopes, i.e.,*

$$\lim_{y\to\infty} dI_j(y, nq_{j-1}^0)/dy = \lim_{y\to\infty} d\tilde{I}_j^\ell(y, nq_{j-1}^0)/dy = h_j;$$
$$\lim_{y\to-\infty} dI_j(y, nq_{j-1}^0)/dy = \lim_{y\to-\infty} d\tilde{I}_j^\ell(y, nq_{j-1}^0)/dy = -(b + h_{j+1}^{'}).$$

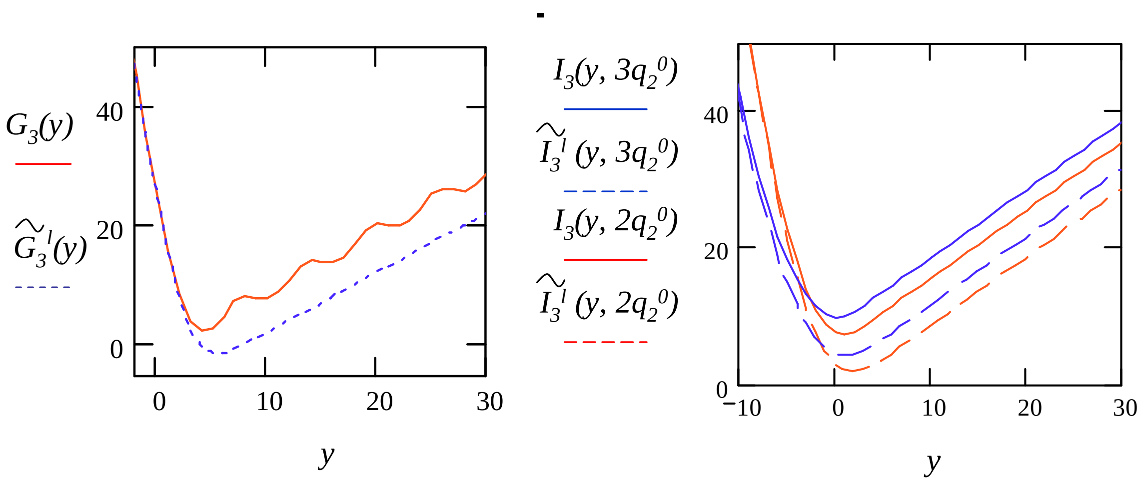


Figure 3(a) Figure 3(b)

Figure 3: (a) $G_3(y)$ is bounded by $\tilde{G}_3^\ell(y)$. (b) $I_3(y, 3q_2^0)$ and $\tilde{I}_3^\ell(y, 3q_2^0)$ have the same asymptotic slopes with 1 and -10. Also, $I_3(y, 3q_2^0) - I_3(y, 2q_2^0) = \tilde{I}_3(y, 3q_2^0) - \tilde{I}_3(y, 2q_2^0) = 3$ as $n \to \infty$.

(3) $\lim_{y\to\infty}\left(I_j(y,(n+1)q_{j-1}^0) - I_j(y, nq_{j-1}^0)\right) = \lim_{y\to\infty}\left(I_j^\ell(y,(n+1)q_{j-1}^0) - I_j^\ell(y, nq_{j-1}^0)\right)$
$= h_j q_{j-1}^0/2;$
$\lim_{y\to-\infty}\left(I_j(y,(n+1)q_{j-1}^0) - I_j(y, nq_{j-1}^0)\right) = \lim_{y\to-\infty}\left(I_j^\ell(y,(n+1)q_{j-1}^0) - I_j^\ell(y, nq_{j-1}^0)\right)$
$= -(b + h'_{j+1})q_{j-1}^0/2.$

(4) $\tilde{I}_j^\ell(r_j^\ell(nq_{j-1}^0), nq_{j-1}^0)$ *is increasing convex in* n *with asymptotic slope* $h_j(b+h'_{j+1})/(b+h'_j)(q_{j-1}^0/2)$.

Figure 3(a) shows $G_3(\cdot)$ and $\tilde{G}_3^\ell(\cdot)$ for the same example in Figure 1. (Note that the cost function $\hat{G}_j(\cdot)$ in Chen (2000) is not bounded by $\tilde{G}_j^\ell(\cdot)$.) Figure 3(b) illustrates the properties of Parts 2 and 3 in Proposition 5 in the same example. In particular, $I_3(y, 3q_2^0)$ and $\tilde{I}_3^\ell(y, 3q_2^0)$ have the same asymptotic slopes. Also, the gap between $I_3(y, 2q_2^0)$ and $I_3(y, 3q_2^0)$ (resp., $\tilde{I}_3^\ell(y, 2q_2^0)$ and $\tilde{I}_3^\ell(y, 3q_2^0)$) is equal to $h_3q_2^0/2 = 3$ as $y \to \infty$ and $-bq_2^0/2 = -30$ as $y \to -\infty$. Part 4 says that the increment of the optimal inventory costs for $\tilde{I}_j^\ell(r_j^\ell(nq_{j-1}^0), nq_{j-1}^0)$ converges to $h_j(b + h'_{j+1})/(b + h'_j)(q_{j-1}^0/2)$ as $n \to \infty$. Note that $G_j^\ell(\cdot)$ was introduced in Shang and Song (2003a, b) to construct lower bounds to the echelon cost for base-stock and batch-ordering models. We now modify it to construct a lower bound to the inventory cost for stage j.

Ideally, if we can obtain the same asymptotic slope in Part 4 for $I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$, the inventory cost functions $I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$ and $\tilde{I}_j^\ell(r_j^\ell(nq_{j-1}^0), nq_{j-1}^0)$ will be parallel as $n \to \infty$ and $r_j^\ell(nq_{j-1}^0)$ may be an effective approximation to $r_j(nq_{j-1}^0)$. Unfortunately, we fail to establish this

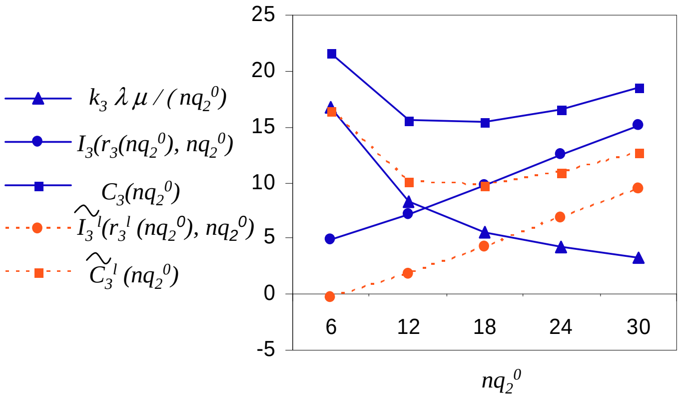


Figure 4: The comparison of $I_3(r_3(nq_2^0), nq_2^0)$ and $\tilde{I}_3^\ell(r_3^\ell(nq_2^0), nq_2^0)$ and the corresponding total costs $C_3(nq_2^0)$ and $\tilde{C}_3^\ell(nq_2^0)$.

result. However, from Parts 2 and 3 of Proposition 5, we conclude that the behaviors of $I_j(y, nq_{j-1}^0)$ and $\tilde{I}_j^\ell(y, nq_{j-1}^0)$ are very similar: Both functions have the same asymptotic slopes and the cost increments from nq_{j-1}^0 to $(n+1)q_{j-1}^0$ are the same as $y \to \infty$ or $-\infty$. Thus we conjecture that $I_j(r_j(nq_{j-1}^0), nq_{j-1}^0)$ may have a asymptotic slope close to $h_j(b+h_{j+1}^{'})/(b+h_j^{'})(q_{j-1}^0/2)$. Indeed, this conjecture is verified in the numerical observations. For instance, in the Figure 4, the increment between $I_3(r_3(6q_2^0), 6q_2^0)$ and $I_3(r_3(5q_2^0), 5q_2^0)$ is 2.724, which is very close to the value $10/(10+1) \times 6/2 = 2.727$. Note that the optimal order quantities generated from $C_3(nq_2^0)$ and $\tilde{C}_3^\ell(nq_2^0) \stackrel{def}{=} k_3\lambda\mu/(nq_2^0) + \tilde{I}_3^\ell(r_3^\ell(nq_2^0), nq_2^0)$ are the same: $q_3^0 = 18$.

Now, consider two single-stage (r, nq) systems. The inventory cost functions in the first system and the second systems are $\tilde{G}_j^\ell(\cdot)$ and $G_j^\ell(\cdot)$ respectively. With the fixed cost k_j and the known downstream base quantities $(q_1^0, ..., q_{j-1}^0)$, the optimal base order quantities and reorder points in these two systems are the same since the difference between $G_j^\ell(\cdot)$ and $\tilde{G}_j^\ell(\cdot)$ is a constant $(= \sum_{i=1}^{j-1} I_i(r_i(q_i^0), q_i^0))$. In other words, we could further replace $\tilde{G}_j^\ell(\cdot)$ with $G_j^\ell(\cdot)$ in the objective function of (SP). We formalize the result below.

Define

$$C_j^\ell(q_j) \quad = \quad \frac{k_j\lambda\mu}{q_j} + I_j^\ell(r_j^\ell(q_j), q_j) = \frac{k_j\lambda\mu + \int_0^{q_j} G_j^\ell(r_j^\ell(q_j)+x)dx}{q_j}.$$

Replacing $C_j(q_j)$ with $C_j^\ell(q_j)$ and relaxing the constraints in (SP), we form a new problem $(SP^{'})$.

$$\begin{array}{lll} (SP^{'}) & \min & \displaystyle\sum_{j=1}^{N} C_j^\ell(q_j) \\ & \text{s.t.} & q_{j+1} \geq q_j, j = 1, ..., N-1. \end{array}$$

Since the objective function in $(SP^{'})$ is composed of N independent functions $C_j^{\ell}(q_j)$ and each of them is convex in q_j (Zheng 1992), a standard cluster technique can find the optimal partition as well as the corresponding order quantity for each cluster. Following Chen and Zheng (1998), a power-of-two-type heuristic for base quantities may be derived.

3.3 Determining an Effective Partition through Cost Ratios

The above result indicates that after order quantities for clusters are determined by executing an clustering algorithm to $(SP^{'})$, a set of feasible base quantities $(q_1^s, ..., q_N^s)$ can easily be generated. However, to make the stochastic system more transparent, i.e., to see the effects of cost parameters on optimal solutions, our goal in this paper is to identify an effective partition through cost ratios. Recall in the relaxed deterministic problem, the optimal order quantity $\underline{Q}_k$ for a cluster $c(k)$ is equal to a new EOQ solution from the model with the holding cost rate $\sum_{i \in c(k)} h_i$ and the fixed order cost $\sum_{i \in c(k)} k_i$. Thus, $(\sum_{i \in c(k)} k_i / \sum_{i \in c(k)} h_i)$ can be used to determine an optimal partition (cf. conditions (i) and (ii) in §2). In what follows, we shall show similar results for $(SP^{'})$. In particular, the optimal order quantity $\underline{Q}_k$ for any cluster $c(k)$ can be closely approximated by the optimal order quantity for a new single-stage (r, nq) system with the holding cost rate $\sum_{i \in c(k)} h_i$ and the fixed order cost $\sum_{i \in c(k)} k_i$. Second, for such a single-stage (r, nq) system, the optimal order quantity increases in $(\sum_{i \in c(k)} k_i / \sum_{i \in c(k)} h_i)$. Thus, the same two conditions in §2 may be applied to identify an effective partition. We derive this single-stage (r, nq) system below.

Consider any cluster $c(k)$ with $n(k)$ stages, i.e., $c(k) = \{v, v+1, ..., v+n(k)-1\}$. Let the optimal order quantity for this cluster be $\underline{Q}_k$, where

$$\underline{Q}_k \quad = \quad \arg\min_{Q} \sum_{i \in c(k)} C_i^{\ell}(Q) = \arg\min_{Q} \left\{ \frac{\lambda\mu \sum_{i \in c(k)} k_i}{Q} + \sum_{i \in c(k)} I_i^{\ell}(r_i^{\ell}(Q), Q) \right\}. \tag{14}$$

Applying the same idea in §3.2, if we can construct a parallel curve to $\sum_{i \in c(k)} I_i^{\ell}(r_i^{\ell}(Q), Q)$ in (14), the optimal order quantity calculated from using this curve will be still $\underline{Q}_k$. The asymptotic slope of $\sum_{i \in c(k)} I_i^{\ell}(r_i^{\ell}(Q), Q)$ is shown in the following proposition.

Proposition 6 *For any cluster $c(k)$, $\sum_{i \in c(k)} I_i^{\ell}(r_i^{\ell}(Q), Q)$ is convex increasing in Q with asymptotic slope*

$$\frac{1}{2} \sum_{t=0}^{n(k)-1} \left[h_{v+t} \left(1 + \frac{h_{v+t}}{b + h'_{v+t+1}} \right)^{-1} \right].$$

Now, consider the sum of $I_j^\ell(y,Q)$ functions in a cluster $c(k)$. Define

$$\bar{I}_k(y,Q) \quad = \quad \sum_{i\in c(k)} I_i^\ell(y,Q) = \sum_{i\in c(k)} \left(\int_0^Q G_i^\ell(y+x)dx/Q\right) = \int_0^Q \left(\sum_{i\in c(k)} G_i^\ell(y+x)dx/Q\right).$$

Note that

$$\begin{aligned}\sum_{i\in c(k)} G_i^\ell(y) &= G_v^\ell(y) + G_{v+1}^\ell(y) + ... + G_{v+n(k)-1}^\ell(y) \\ &= \mathsf{E}[h_v(y-\tilde{D}_v)^+ + (b+h'_{v+1})(y-\tilde{D}_v)^-] + ... \\ &\quad + \mathsf{E}[h_{v+n(k)-1}(y-\tilde{D}_{v+n(k)-1})^+ + (b+h'_{v+n(k)})(y-\tilde{D}_{v+n(k)-1})^-]. \end{aligned} \tag{15}$$

Thus, $\bar{I}_k(y,Q)$ is a convex function with asymptotic slopes $\sum_{t=0}^{n(k)-1} h_{v+t}$ as $y\to\infty$ and $n(k)b + \sum_{t=0}^{n(k)-1} h'_{v+t+1}$ as $y\to-\infty$. Let $R_k(Q) \overset{def}{=} \arg\min_y\{\bar{I}_k(y,Q)\}$.

Proposition 7 *For any cluster $c(k)$, $\bar{I}_k(R_k(Q),Q)$ is convex increasing in Q with asymptotic slope*

$$\frac{1}{2}\left(\sum_{t=0}^{n(k)-1} h_{v+t}\right)\left(1+\frac{\sum_{t=0}^{n(k)-1} h_{v+t}}{n(k)b+\sum_{t=0}^{n(k)-1} h'_{v+t+1}}\right)^{-1}.$$

Assume that b is sufficiently larger than h'_1. The two asymptotic slopes in Propositions 6 and 7 are actually close. (They are the same as $b\to\infty$.) For instance, consider a four-stage system with $h_j = 0.25, \forall j$ and $b = 9$. If a cluster contains stages 1, 2, and 3, the asymptotic slopes are 0.3654 and 0.3651 respectively. In other words, we may replace $\sum_{i\in c(k)} I_i^\ell(r_i^\ell(Q),Q)$ with $\bar{I}_k(R_k(Q),Q)$ in (14) to search order quantity for a cluster without distorting the value of $\underline{Q}_k$.

Finally, if we replace $\tilde{D}_i$ in $G_i^\ell(y), i\in c(k)$ with a common leadtime demand $\tilde{D}_{v+n(k)-1}$ in Equation (15) and denote these new functions as $G_i^c(\cdot)$, we can obtain a newsvendor-type cost function:

$$\begin{aligned}\mathcal{G}_k(y) &= G_v^c(y) + G_{v+1}^c(y) + ... + G_{v+n(k)-1}^c(y) \\ &= \mathsf{E}[h_v(y-\tilde{D}_{v+n(k)-1})^+ + (b+h'_{v+1})(y-\tilde{D}_{v+n(k)-1})^-] + ... \\ &\quad + \mathsf{E}[h_{v+n}(y-\tilde{D}_{v+n(k)-1})^+ + (b+h'_{v+n+1})(y-\tilde{D}_{v+n(k)-1})^-] \\ &\overset{def}{=} \mathsf{E}[(\sum_{i\in c(k)} h_i)(y-\tilde{D}_{v+n(k)-1})^+ + [n(k)b + (\sum_{i\in c(k)} h'_{i+1})](y-\tilde{D}_{v+n(k)-1})^-]. \end{aligned} \tag{16}$$

Define $\mathcal{I}_k(y,Q) = \int_0^Q \mathcal{G}_k(y+x)dx/Q$ and $\mathcal{R}_k(Q) = \arg\min_y\{\mathcal{I}_k(y,Q)\}$. It is easy to verify that $\mathcal{I}_k(\mathcal{R}_k(Q),Q)$ is convex increasing in Q with the same asymptotic slope as $\bar{I}_k(R_k(Q),Q)$. Thus, we may use $\mathcal{I}_k(\mathcal{R}_k(Q), Q)$ to replace $\sum_{i\in c(k)} I_i^\ell(r_i^\ell(Q), Q)$ in (14) to search for order quantity $\mathcal{Q}_k$ in a cluster. Namely,

$$\mathcal{Q}_k = \arg\min_Q \left\{ \frac{\lambda\mu \sum_{i\in c(k)} k_i}{Q} + \mathcal{I}_k(\mathcal{R}_k(Q), Q) \right\}. \tag{17}$$

This is a single-stage (r, nq) cost function with holding cost rate $\sum_{i\in c(k)} h_i$, backorder cost rate $n(k)b + (\sum_{i\in c(k)} h'_{i+1})$, fixed order cost $\lambda\mu \sum_{i\in c(k)} k_i$ and leadtime demand $\tilde{D}_{v+n(k)-1}$. Our remaining task is to show that $\mathcal{Q}_k$ may increase in $(\sum_{i\in c(k)} k_i / \sum_{i\in c(k)} h_i)$.

Consider a single-stage (r, nq) system, i.e., $N = 1$. We suppress the index j for simplicity. The total system cost in (3) reduces to

$$C(r, q) = \frac{\lambda\mu k + \int_0^q G(r + x)dx}{q}, \tag{18}$$

where $G(y) = \mathsf{E}[h(y - D)^+ + b(y - D)^-]$. Let $\omega \overset{def}{=} \frac{b}{b+h}$ and the optimal order quantity that minimizes (18) be q^*. Then, we have

Theorem 8 *For a single-stage (r, nq) system, the optimal order quantity q^* increases in k/h with ω fixed.*

Applying Theorem 8 to (17), $\mathcal{Q}_k$ increases in $(\sum_{i\in c(k)} k_i / \sum_{i\in c(k)} h_i)$ with $\omega_k \overset{def}{=} (n(k)b + \sum_{i\in c(k)} h'_{i+1})/(n(k)b + \sum_{i\in c(k)} h'_i)$ fixed. Since b is assumed to be sufficiently larger than h'_1, ω_k are similar across all clusters. Thus, we assume this condition holds for approximation purpose. Since $\underline{Q}_k$ is close to $\mathcal{Q}_k$, we may be able to use $(\sum_{i\in c(k)} k_i / \sum_{i\in c(k)} h_i)$ to identify an effective partition. This completes the analysis of developing the heuristic.

3.4 The Heuristics

Based on the discussion above, we construct two simple heuristics for the stochastic model backwards. These heuristics employ the same clustering step, but use different cost functions to search for order quantities.

1. *Clustering* – Group stages into clusters based on ratios of cost parameters in the same fashion as in the deterministic model. Let the resulting partition be $\{c(1), ..., c(m)\}$.

2. *Minimization* – Find the order quantity Q_k^s for each cluster $c(k)$ sequentially from $k = 1, 2, ...$ until m:

$$Q_k^s = \arg\min_Q \left\{ \lambda\mu \left(\sum_{i\in c(k)} k_i \right) /Q + A \right\}$$

$$\text{s.t. } Q = \begin{cases} Q_k^s, & k=1, \\ n_k Q_{k-1}^s, & k=2,...,m,\ n_k \text{ is some integer.} \end{cases}$$

Heuristics I and II use $\bar{I}_k(R_k(Q), Q)$ and $\mathcal{I}_k(\mathcal{R}_k(Q), Q)$ respectively in the A term.

After Q_k^s are found, the heuristic base quantities are $(q_1^s, .., q_N^s)$, where $q_i^s = Q_k^s$, for $i \in c(k)$. The heuristic reorder points $(r_1(q_1^s), ..., r_N(q_N^s))$ can be found through (7)-(9).

4. Numerical Studies and Insights

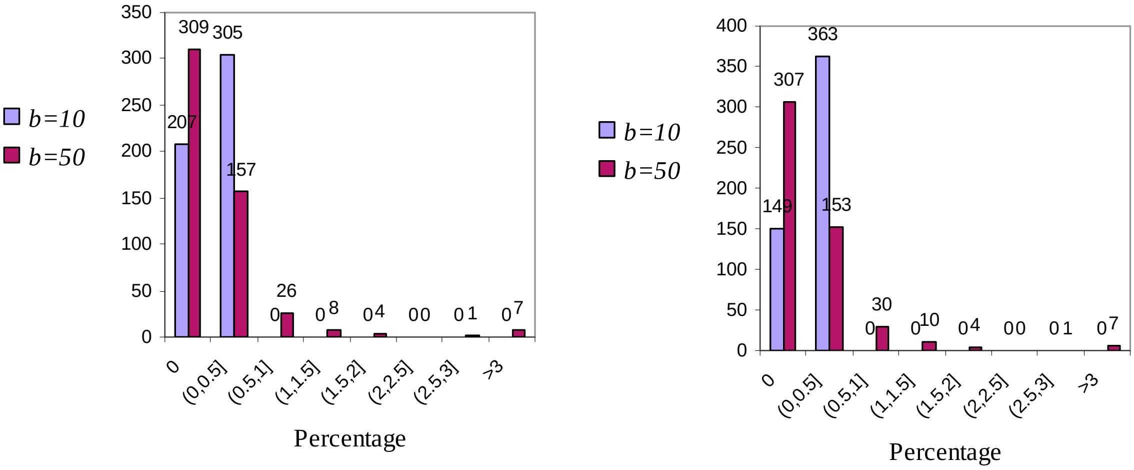
We test the effectiveness of these two heuristics by comparing heuristic costs with either the optimal cost or with the lower bound cost. The algorithms for finding optimal (r, nq) policies and the lower bound cost can be found in Chen and Zheng (1998) and (1994b) respectively.

The numerical examples are divided into four groups. The first group tests our heuristics under a wide range of system parameters, using optimal costs as a benchmark. The goal is to verify the effectiveness of these two heuristics and select a more effective heuristic for the other tests. The second group (resp., third group) aims to compare our heuristic with those in Chen and Zheng (1998). We use the same parameters as in Tables 1 and 3 (resp., Tables 2 and 4) of their paper. In the last group, we examine the effectiveness of our heuristic in conditions when the number of stages increases. We compare heuristic costs with lower bound costs.

4.1 Performance Summary

Group 1: The examples in this group have Poisson demand. We test both heuristics in a three-stage system with the following parameters: $b = 10, 50$, $\lambda = 5$, $h_j = 0.1, 1$, $L_j = 0.5, 2$, $k_j = 10, 100$, for $j = 1, 2, 3$. Note that in some instances, b may not be significantly larger than $h_1^{'}$ (e.g., $b = 10$ and $h_1^{'} = 3$). We aim to see if the heuristics still perform well in those instances. The total number of instances is 1024. Figure 5 summarizes the results for both heuristics. Remarkably, Heuristic I generates the optimal solution in 516 cases; Heuristic II generates the optimal solution in 456 cases. The average (maximum) percentage errors for Heuristic I and Heuristic II are 0.067% (3.596%) and 0.070% (3.596%) respectively. These results indicate that Heuristic I performs slightly better than Heuristic II. It is worth noting that when b is small, each heuristic generates fewer optimal solutions, but the average effectiveness increases for both heuristics.

Group 2: The demand is Poisson in this group. The examples have the following parameters in common: $N = 3, k_2 = 10, h_i = 1, L_i = 1, i = 1, 2, 3$. We varied the remaining parameters λ, k_1, k_3, and b to generate 16 examples. Table 1 reports the result. Similarly, the examples in Table 2 have



Heuristic I

Heuristic II

Figure 5: The distribution of percentage error for Group 1 examples.

the following common parameters: $b = 30$, $h_j = 1$, and $L_j = 1$. The results are listed in Table 2. Here, let $r_j^s = r_j(q_j^s)$. The average percentage error is 0.28% in Table 1 and 0.15% in Table 2. Among these total 40 examples, Heuristic I generates the optimal solution in 20 cases.

Group 3: The demand is compound Poisson in this group. The parameters are exactly the same as the examples in Group 2 except for the fact that the mean demand size is 5 for all examples. For the first 16 examples in Table 1, the average (maximum) percentage error is 0.44% (1.52%); for the rest of 24 examples in Table 2, the average (maximum) percentage error is 0.20% (1.08%). Among these total 40 examples, Heuristic I generates the optimal solution in 5 cases.

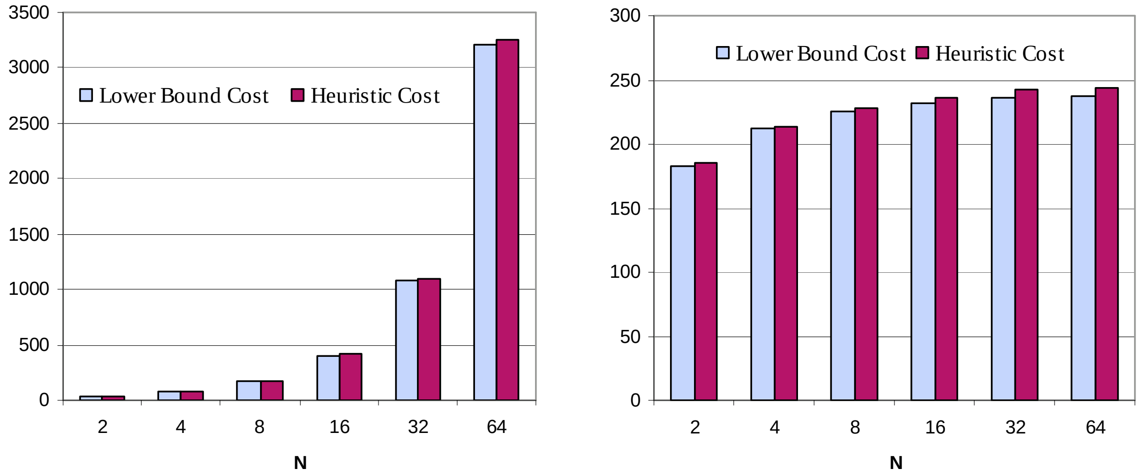
Group 4: We compare Heuristic I with the better lower-bound cost developed by Chen and Zheng when $N = 2, 4, 8, 16, 32$, and 64. We test two different scenarios. The first scenario has the following parameters: $\lambda = 1, b = 320, k_j = 100, h_j = 1, L_j = 1$. In the second scenario, we fixed total lead time equal to 16, $h_1' = 16$ with a total fixed order cost of 64 (i.e., $L_j = 16/N$, $h_j = 16/N$, and $k_j = 64/N$). The rest of the parameters are $b = 100, \lambda = 1$. In the first scenario, the average and maximum percentage errors are 0.54% and 1.48% respectively while in the second, they are 1.72% and 2.96% respectively. Figure 6 shows the percentage error with respect to the number of stages. The heuristic seems not to deteriorate much as N increases. For example, in the scenario 1 of Group 4, the percentage error grows slowly as N increases from 2 to 64 (from 0.011% to 1.476%). Since these are the results generated from comparing the heuristic cost with the lower-bound cost, the percentage error would be even smaller if we compared the heuristic with the optimal cost.

No.	λ	k_1	k_3	b	r_1^*	q_1^*	r_2^*	q_2^*	r_3^*	q_3^*	Cost	r_1^s	q_1^s	r_2^s	q_2^s	r_3^s	q_3^s	Cost	Dev(%)
1	1	10	10	10	0	6	1	6	2	6	17.7390	0	6	1	6	2	6	17.7390	0.00
2	1	10	10	20	1	5	2	5	3	5	19.9160	1	5	2	5	3	5	19.9160	0.00
3	1	10	100	10	0	7	1	7	0	14	26.4164	0	6	1	6	0	18	26.4867	0.27
4	1	10	100	20	1	5	2	5	2	15	28.6695	1	5	2	5	2	15	28.6695	0.00
5	1	100	10	10	0	11	0	11	0	11	28.2272	0	10	0	10	0	10	28.3498	0.43
6	1	100	10	20	0	10	1	10	2	10	30.9436	0	10	1	10	2	10	30.9436	0.00
7	1	100	100	10	-1	14	-1	14	-1	14	35.4643	0	12	0	12	0	12	35.8750	1.16
8	1	100	100	20	0	13	1	13	1	13	38.4416	0	12	1	12	1	12	38.6563	0.56
9	10	10	10	10	9	19	18	19	27	19	76.2283	9	19	18	19	27	19	76.2283	0.00
10	10	10	10	20	11	18	21	18	30	18	82.5368	11	18	21	18	30	18	82.5368	0.00
11	10	10	100	10	9	22	17	22	23	44	104.0117	9	18	18	18	23	54	104.1804	0.16
12	10	10	100	20	11	17	21	17	27	51	110.6577	11	17	21	17	27	51	110.6577	0.00
13	10	100	10	10	7	34	15	34	22	34	109.9492	7	35	15	35	22	35	109.9613	0.01
14	10	100	10	20	9	33	18	33	27	33	118.0225	9	33	18	33	27	33	118.0225	0.00
15	10	100	100	10	6	44	13	44	20	44	132.8222	7	37	14	37	21	37	134.5031	1.27
16	10	100	100	20	8	42	17	42	25	42	142.0274	9	37	17	37	26	37	142.9724	0.66

Table 1: Heuristic vs. Optimal Solutions

No.	λ	k_1	k_2	k_3	r_1^*	q_1^*	r_2^*	q_2^*	r_3^*	q_3^*	Cost	r_1^s	q_1^s	r_2^s	q_2^s	r_3^s	q_3^s	Cost	Dev(%)
1	0.1	10	10	10	0	2	0	2	0	2	6.1376	0	2	0	2	0	2	6.1376	0.00
2	0.1	10	10	1000	0	2	0	2	-1	14	18.0152	0	2	0	2	-1	14	18.0152	0.00
3	0.1	10	1000	10	0	2	-1	10	-1	10	21.6572	0	2	-1	10	-1	10	21.6572	0.00
4	0.1	10	1000	1000	0	2	-1	14	-1	14	29.5930	0	2	-1	14	-1	14	29.5930	0.00
5	0.1	1000	10	10	-1	9	-1	9	-1	9	23.9983	-1	9	-1	9	-1	9	23.9983	0.00
6	0.1	1000	10	1000	-1	12	-1	12	-1	12	33.5988	-1	11	-1	11	-1	11	33.7707	0.51
7	0.1	1000	1000	10	-1	12	-1	12	-1	12	33.5988	-1	12	-1	12	-1	12	33.5988	0.00
8	0.1	1000	1000	1000	-1	15	-1	15	-2	15	41.0190	-1	15	-1	15	-2	15	41.0190	0.00
9	1	10	10	10	1	6	2	6	3	6	20.9964	1	6	2	6	3	6	20.9964	0.00
10	1	10	10	1000	1	6	2	6	1	48	59.2636	1	5	2	5	1	45	59.3492	0.14
11	1	10	1000	10	1	5	0	35	0	35	71.3838	1	5	0	35	0	35	71.3838	0.00
12	1	10	1000	1000	1	6	0	48	-1	48	96.0984	1	5	0	45	0	45	96.1589	0.06
13	1	1000	10	10	0	28	0	28	0	28	79.2321	0	27	0	27	0	27	79.2778	0.06
14	1	1000	10	1000	-1	39	-1	39	-1	39	108.8846	-1	33	-1	33	-1	33	110.4091	1.40
15	1	1000	1000	10	-1	39	-1	39	-1	39	108.8846	-1	39	-1	39	-1	39	108.8846	0.00
16	1	1000	1000	1000	-1	47	-1	47	-2	47	131.9043	-1	46	-1	46	-2	46	131.9674	0.05
17	10	10	10	10	12	17	22	17	32	17	85.9392	12	17	22	17	32	17	85.9392	0.00
18	10	10	10	1000	11	18	22	18	25	144	207.3430	12	17	22	17	24	153	207.5197	0.09
19	10	10	1000	10	11	18	16	108	23	108	245.9119	12	16	16	112	22	112	246.2565	0.14
20	10	10	1000	1000	11	18	15	144	20	144	324.2378	12	16	15	144	20	144	324.2404	0.00
21	10	1000	10	10	7	88	15	88	22	88	271.2414	7	88	15	88	22	88	271.2414	0.00
22	10	1000	10	1000	6	122	12	122	19	122	365.1038	7	105	13	105	20	105	369.1794	1.11
23	10	1000	1000	10	6	122	12	122	19	122	365.1038	6	121	12	121	19	121	365.1255	0.01
24	10	1000	1000	1000	5	149	11	149	16	149	437.8022	5	148	11	148	16	148	437.8218	0.00

Table 2: Heuristic vs. Optimal Solutions



Scenario 1

Scenario 2

Figure 6: The lower-bound cost and the heuristic cost for Group 5 examples.

4.2 Observations and Insights

(1) The effectiveness of the heuristics implies that the clustering-minimization approach actually "catches" the optimal behaviors of the considered stochastic systems. This results in interesting insights for order coordination. For instance, to reduce the order size in a location, it is not sufficient to consider the costs within that location only. Also, the fixed order costs and holding costs may play more important roles in determining effective order quantities.

(2) There are 8 examples whose percentage error are larger than 2.5% in Group 1. The parameters for these instances are $k_1 = 10, k_2 = 100, k_3 = 100, h_1 = 0.1, h_2 = 0.1, h_3 = 1$, $L_j = 0.5, 2, j = 1, 2, 3$, and $b = 50$. Consider the one with the maximum percentage error: $L_1 = 0.5, L_2 = 0.5, L_3 = 2$. Based on our clustering scheme, we have $c(1) = \{1\}$ and $c(2) = \{2, 3\}$. The heuristic order quantities are $(q_1^s, q_2^s, q_3^s) = (33, 33, 33)$. However, the optimal order quantities are $(q_1^*, q_2^*, q_3^*) = (44, 44, 44)$. If we group these three stages in the same cluster, i.e., $c(1) = \{1, 2, 3\}$, the new order quantities are $(44, 44, 44)$, which are exactly the same as the optimal solution. Similar results appear in the other seven examples. This observation suggests that there may be room to improve the cost-based clustering scheme.

(3) Scenario 2 in Group 4 indicates that it is more cost-effective to have a shorter supply chain, provided that $\sum_{j=1}^N k_j, \sum_{j=1}^N L_j$, and $h_1^{'}$ are fixed and k_j, h_j, L_j are all equally distributed. This observation is different from that in the corresponding serial system without fixed order costs or

k_1	k_2	k_3	h_1	h_2	h_3	Optimal Cost	Heuristic Cost
10	10	10	1	1	1	79.89	80.03
			1	1	0.1	48.37	48.38
			1	0.1	1	60.66	61.06
			0.1	1	1	79.89	80.03
10	10	100	1	1	1	99.48	99.48
			1	1	0.1	54.99	54.99
			1	0.1	1	78.22	78.22
			0.1	1	1	90.15	90.25
10	100	10	1	1	1	102.96	103.14
			1	1	0.1	67.76	67.76
			1	0.1	1	78.22	78.22
			0.1	1	1	92.48	92.48
100	10	10	1	1	1	104.90	104.97
			1	1	0.1	71.67	71.68
			1	0.1	1	82.37	82.37
			0.1	1	1	92.48	92.48

Table 3: The effect of echelon holding cost reduction on the optimal cost

with known order quantities (Gallego and Zipkin 1999, Shang and Song 2003a, b), where optimal cost decreases as N increases.

(4) It is more effective to reduce the echelon holding cost at an upstream stage or reduce the leadtime at a downstream stage. Table 3 and Table 4 compare several instances in Group 1. In Table 3, we fix $L_j = 2$ and reduce the echelon holding cost h_j from 1 to 0.1. Similarly, in Table 4, we fix $h_j = 1$ and reduce the leadtime L_j from 2 to 0.5. This observation is consistent with the findings in Shang and Song (2003b).

(5) Reducing fixed order cost at a downstream stage is more effective than at an upstream stage. For instance, in Table 3, consider the examples with all $h_j = 1$. If we reduce the fixed order cost from 100 to 10 for each k_j, the reduction in the optimal cost is largest when the fixed order cost 100 is located at stage 1 and smallest at stage 3.

(6) In the single-stage (r, nq) model, it is well known that the optimal order quantity is larger than the EOQ (Zheng 1992). Chen and Zheng (1998) report a similar observation in the serial model. Our study provides a justification for why the optimal order quantities in the stochastic model tend to be larger than those in its deterministic counterpart: with the same clustering scheme for both models, one solves EOQ problems and the other solves the corresponding single-stage (r, nq) models.

k_1	k_2	k_3	L_1	L_2	L_3	Optimal Cost	Heuristic Cost
10	10	10	2	2	2	79.89	80.03
			2	2	0.5	77.60	77.60
			2	0.5	2	68.85	68.86
			0.5	2	2	59.74	59.85
10	10	100	2	2	2	99.48	99.48
			2	2	0.5	97.78	97.78
			2	0.5	2	88.42	88.42
			0.5	2	2	79.31	79.31
10	100	10	2	2	2	102.96	103.14
			2	2	0.5	100.90	101.07
			2	0.5	2	92.59	92.73
			0.5	2	2	65.88	65.88
100	10	10	2	2	2	104.90	104.97
			2	2	0.5	102.76	102.76
			2	0.5	2	94.31	94.31
			0.5	2	2	85.70	85.70

Table 4: The effect of leadtime reduction on the optimal cost

(7) Our heuristic may cause "nervousness" in order quantities or batch sizes, especially when the system parameters at a downstream stage change. For instance, if the order quantity at stage 1 increases, the resulting order quantities at stage $j, j \geq 2$ will all increase. This is a disadvantage since it may not be easy to increase or reduce order quantities immediately in practice. Nevertheless, we observe that the property of insensitivity of order quantity to the optimal cost for a single-stage model is still valid for the serial model (see Zheng (1992) for the discussion of this property). In the worst-performance case in Group 1, we use the heuristic order quantity 33 instead of the optimal order quantity 44. However, the resulting percentage error is only 3.60%. In other words, in terms of implementation, we do not need to change order quantities in response to small changes in system parameters.

5. Concluding Remarks

In this paper, we consider serial inventory models with deterministic and stochastic demands, and propose a generic heuristic for both systems. This heuristic consists of one clustering step and one minimization step. For the deterministic model, we show that the heuristic policy is 94%-effective. Also, in an extensive numerical study, we show that the heuristic outperforms the power-of-two policy in about 90% of the cases. For the stochastic model, the heuristic dramatically simplifies the calculation of effective solutions in two ways. First, unlike the traditional heuristics, we can determine near-optimal clusters using cost parameters. Second, we need to solve at most N single-stage (r, nq) systems to find heuristic order quantities. In an extensive numerical study, our heuristic outperforms those proposed in Chen and Zheng (1998).

Our two-step heuristic can be carried over to assembly systems with deterministic demand. In particular, Schwarz and Schrage (1975) show that nested policies are optimal in the assembly system and that the problem formulation is virtually the same as that in the serial model. For the stochastic model, Chen (2000) shows that there exists an equivalent serial system for the assembly system, under the assumption that base quantities for the items do not decrease with their total leadtimes (the sum of the leadtimes from the current stage to stage 1). Consequently, our heuristic is also valid for such an assembly system.

We expect our results will improve the transparency of serial systems with fixed order costs. In particular, using the same effective solution approach for the deterministic and the stochastic systems, the optimality behaviors between these two systems should be similar. Future research includes extending the heuristic to distribution systems and designing simple incentive schemes to coordinate supply chain systems with scale economies.

Acknowledgments: The author thanks the associate editor and the referees for the insightful comments and suggestions that have significantly improved the paper. The author is also grateful to Robin Roundy, Jing-Sheng Song, and Paul Zipkin for useful discussions and to Sean Zhou for help with the computation.

Appendix A: Evaluation and decomposition schemes in Chen (2000)

We summarize Chen's results for later use in Appendix B. For any given echelon (r, nq) policy with fixed base quantities q_j, define $\hat{G}_1(\cdot) = G_1(\cdot)$. Suppose $\hat{G}_j(\cdot)$ is known. Let $\hat{r}_j(q_j) = \text{argmin}_y\{\int_0^{q_j} \hat{G}_j(y+x)dx\}$. Recall (12) and (13), for $j = 2, ..., N$,

$$\begin{aligned}\hat{G}_{j,j+1}(y) &= \hat{G}_j(O_j[y]) - \min_{m\in\Im} \hat{G}_j(y + mq_j), \\ \hat{G}_{j+1}(y) &= \mathsf{E}[h_{j+1}(y - D_{j+1}) + \hat{G}_{j,j+1}(y - D_{j+1})],\end{aligned}$$

where $O_j[x]$ is defined in (2) except using $\hat{r}_j(q_j)$ to replace $r_j(q_j)$. Define $\hat{I}_j(r, q) = \int_0^q \hat{G}_j(r+x)dx/q_j$. Thus, the optimal total cost is $\sum_{i=1}^N \hat{I}_i(\hat{r}_i(q_i), q_i)$.

Chen simplifies the above recursions by defining a set of echelon cost functions. Let $\hat{C}_1(\cdot) =$

$\hat{G}_1(\cdot)$. Suppose $\hat{C}_j(y)$ is known and let $\hat{r}_j(q_j) = \text{argmin}_y\{\int_0^{q_j} \hat{C}_j(y+x)dx\}$. It can be shown that

$$\begin{aligned}
\hat{C}_j(y) &= \hat{G}_j(y) + \sum_{i=1}^{j-1} \hat{I}_i(\hat{r}_i(q_i), q_i) \\
&= \mathsf{E}[h_j(y-D_j) + \hat{C}_{j-1}(O_{j-1}[y-D_j])]
\end{aligned} \tag{19}$$

and the total optimal cost $\sum_{i=1}^{N} \hat{I}_i(\hat{r}_i(q_i), q_i) = \int_0^{q_N} \hat{C}_N(r_N(q_N)+x)dx/Q_N$. Note that the above two schemes can be used to evaluate any feasible (r, nq) policies by using the given r_j in $O_j[x]$ (see Zipkin 2000, pp.321-322). That is,

$$C(\mathbf{r}, \mathbf{q}) = \sum_{i=1}^{N} k_j\lambda\mu/q_i + \sum_{i=1}^{N} \hat{I}_i(r_i, q_i) = \sum_{i=1}^{N} k_j\lambda\mu/q_i + \int_0^{q_N} \hat{C}_N(r_N+x)dx/q_N.$$

Appendix B: Proofs

Proposition 2

From Lemma 1, all we need to show is $1/\sqrt{2} \le Q_k^s/\underline{Q}_k \le \sqrt{2}$. Note that after the clustering step, we have $\underline{Q}_k < \underline{Q}_{k+1}$ for $k = 1, ..., m-1$. Consider two cases for each cluster $c(k)$.

Case 1: $\underline{Q}_k > Q_{k-1}^s$. Since $\sum_{i\in c(k)} C_i^d(Q)$ is convex, there must exist a positive integer n such that $nQ_{k-1}^s \le \underline{Q}_k \le (n+1)Q_{k-1}^s$. Within this interval, if $\underline{Q}_k > \sqrt{n(n+1)}Q_{k-1}^s$, we will choose $Q_k^s = (n+1)Q_{k-1}^s$ through our procedure. Thus, $Q_k^s/\underline{Q}_k \le \sqrt{(n+1)/n}$. On the other hand, if $\underline{Q}_k \le \sqrt{n(n+1)}Q_{k-1}^s$, we will choose $Q_k^s = nQ_{k-1}^s$ and $Q_k/\underline{Q}_k \ge \sqrt{n/(n+1)}$. Thus, $\sqrt{n/(n+1)} \le Q_k^s/\underline{Q}_k \le \sqrt{(n+1)/n}$. Since $n \ge 1$, we have $\sqrt{1/2} \le Q_k^s/\underline{Q}_k \le \sqrt{2}$.

Case 2: $\underline{Q}_k \le Q_{k-1}^s$. Through our optimization procedure, we have $Q_k^s = Q_{k-1}^s$. Thus, $Q_k^s/\underline{Q}_k = Q_{k-1}^s/\underline{Q}_k \le Q_{k-1}^s/\underline{Q}_{k-1} = 1$. Also, $Q_{k-1}^s/\underline{Q}_k = Q_k^s/\underline{Q}_k \ge 1$ in this case. Together, we have $1/\sqrt{2} \le Q_k^s/\underline{Q}_k = 1 \le \sqrt{2}$.

Proposition 3

We first show Part 2. It suffices to show that $\int_0^{nq_{j-1}} G_j(y+x)dx = \int_0^{nq_{j-1}} \hat{G}_j(y+x)dx$, for all y and for all positive integer n. Since $G_j(y) = \mathsf{E}[h_j(y-D_j) + G_{j-1,j}(y-D_j)]$ and $\hat{G}_j(y) = \mathsf{E}[h_j(y-D_j) + \hat{G}_{j-1,j}(y-D_j)]$. It is sufficient to show

$$\mathsf{E}\left[\int_0^{nq_{j-1}} G_{j-1,j}(y-D_j+x)dx\right] = \mathsf{E}\left[\int_0^{nq_{j-1}} \hat{G}_{j-1,j}(y-D_j+x)dx\right].$$

We show this by induction. When $j = 2$,

$$\begin{aligned}
&\mathsf{E}\left[\int_0^{nq_1} G_{1,2}(y-D_2+x)dx\right] = \mathsf{E}\left[\int_0^{nq_1} (G_1(O_1[y-D_2+x])dx - I_1(r_1(q_1), q_1))\right] \\
&= \mathsf{E}\left[\int_0^{nq_1} G_1(O_1[y-D_2+x])dx\right] - n_2\int_0^{q_1} G_1(r_1(q_1)+x)dx.
\end{aligned}$$

On the other hand, let $\Im$ denote the integer set.

$$\begin{aligned}
&\mathsf{E}\left[\int_0^{nq_1}\hat{G}_{1,2}(y-D_2+x)dx\right]=\mathsf{E}\left[\int_0^{nq_1}\left(\hat{G}_1(O_1[y-D_2+x])-\min_{m\in\Im}\hat{G}_1(y-D_2+x+mq_1)\right)dx\right]\\
&=\mathsf{E}\left[\int_0^{nq_1}G_1(O_1[y-D_2+x]dx\right]-n_2\left(\int_0^{q_1}\hat{G}_1(r_1+x)dx\right)\int_0^{\infty}\mathsf{P}(D_2=\xi)d\xi\\
&=\mathsf{E}\left[\int_0^{nq_1}G_1(O_1[y-D_2+x])dx\right]-n_2\int_0^{q_1}G_1(r_1(q_1)+x)dx.
\end{aligned}$$

Thus, $j=2$ is true. Suppose $j=k-1$ is true, i.e., $\int_0^{nq_{k-2}}G_{k-1}(y+x)dx=\int_0^{nq_{k-2}}\hat{G}_{k-1}(y+x)dx$. When $j=k$,

$$\begin{aligned}
\mathsf{E}\left[\int_0^{nq_{k-1}}G_{k-1,k}(y-D_k+x)dx\right]&=\mathsf{E}\left[\int_0^{nq_{k-1}}\left(G_{k-1}(O_{k-1}[y-D_k+x])-I_{k-1}(r_{k-1}(q_{k-1}),q_{k-1})\right)dx\right]\\
&=\mathsf{E}\left[\sum_0^{nq_{k-1}}G_{k-1}(O_{k-1}[y-D_k+x])dx\right]-n_k\int_0^{q_{k-1}}G_{k-1}(r_{k-1}+x)dx,
\end{aligned}$$

and

$$\begin{aligned}
&\mathsf{E}\left[\int_0^{n_kq_{k-1}}\hat{G}_{k-1,j}(y-D_k+x)dx\right]\\
&=\mathsf{E}\left[\int_0^{n_kq_{k-1}}\left(\hat{G}_{k-1}(O_{k-1}[y-D_k+x])-\min_{m\in\Im}\hat{G}_{k-1}(y-D_k+x+mq_{k-1})\right)dx\right]\\
&=\mathsf{E}\left[\int_0^{n_kq_{k-1}}\hat{G}_{k-1}(O_{k-1}[y-D_k+x])dx\right]-n_k\left(\int_0^{q_{k-1}}\hat{G}_{k-1}(r_{k-1}+x)dx\right)\int_0^{\infty}\mathsf{P}(D_k=\xi)d\xi\\
&=\mathsf{E}\left[\int_0^{n_kq_{k-1}}G_{k-1}(O_{k-1}[y-D_k+x])dx\right]\\
&-n_k\int_0^{q_{k-1}}G_{k-1}(r_{k-1}+x)dx.\qquad\text{(by induction assumption)}
\end{aligned}$$

Therefore, $j=k$ is true and the result follows.

We next show Part 1. Recall $I_j(y,q_j)=\left(\int_0^{q_j}G_j(y+x)dx\right)/q_j$. It is sufficient to show $\int_0^{q_j}G_j(y+x)dx$ is convex in q_j. Let $q_j=n_jq_{j-1}$. Note that

$$\begin{aligned}
\int_0^{q_j}G_j(y+x)dx &= \int_0^{q_j}\hat{G}_j(y+x)dx\qquad\text{(from Part 2 proof)}\\
&= \int_0^{q_j}h_j(y+x-\mathsf{E}[D_j])dx+\int_0^{q_j}\mathsf{E}[\hat{C}_{j-1}(O_{j-1}[y+x-D_j])]\\
&\quad -\sum_{i=1}^{j-1}\hat{I}_i(\hat{r}_i(q_i),q_i).\qquad\text{(from Equation (19))}
\end{aligned}$$

The first term is linear, the last term is a constant, and the second term is convex from Lemma 3 of Chen (2000).

For part (3), the first inequality is obvious. We show the equality. Let $r_j(q_j)=r_j$ for simplicity. Conditioning on $IP_N=y$, from Equation (19), we have

$$\hat{C}_N(y) = \mathsf{E}[h_N(IN_N)+\hat{C}_{N-1}(O_{N-1}[IN_N])|IP_N=y]$$

$$
\begin{aligned}
&= \mathsf{E}[h_N(IN_N) + C_{N-1}(IP_{N-1})|IP_N = y] \\
&= \vdots \\
&= \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + h_1(IN_1) + (b + h_1^{'})IN_1^{-}|IP_N = y] \\
&= \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + G_1(IP_1)|IP_N = y] \\
&= \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + G_1(O_1[IN_2])|IP_N = y] \\
&= \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + G_1(O_1[IN_2])|IP_N = y] \\
&\quad - \sum_{j=1}^{N-1} I_j(r_j, q_j) + \sum_{j=1}^{N-1} I_j(r_j, q_j).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&C_N(y) + \sum_{j=1}^{N} \frac{k_j \lambda \mu}{q_j} = \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + G_1(O_1[IN_2]) - I_1(r_1, q_1)|IP_N = y] \\
&- \sum_{j=2}^{N-1} I_j(r_j, q_j) + \sum_{j=1}^{N-1} I_j(r_j, q_j) + \sum_{j=1}^{N-1} \frac{k_j \lambda \mu}{q_j} + \frac{k_N \lambda \mu}{q_N} \\
&= \mathsf{E}[h_N(IN_N) + ... + h_2(IN_2) + G_{1,2}(IN_2) - \sum_{j=2}^{N-1} I_j(r_j, q_j)|IP_N = y] \\
&+ \sum_{j=1}^{N-1} C_j(r_j, q_j) + \frac{k_N \lambda \mu}{q_N} \\
&= \mathsf{E}[h_N(IN_N) + ... + h_3(IN_3) + G_2(IP_2) - I_2(r_2, q_2) - \sum_{j=3}^{N-1} I_j(r_j, q_j)|IP_N = y] \\
&+ \sum_{j=1}^{N-1} C_j(r_j, q_j) + \frac{k_N \lambda \mu}{q_N} \\
&= \mathsf{E}[h_N(IN_N) + ... + h_3(IN_3) + G_2(O_2[IN_3]) - I_2(r_2, q_2) - \sum_{j=3}^{N-1} I_j(r_j, q_j)|IP_N = y] \\
&+ \sum_{j=1}^{N-1} C_j(r_j, q_j) + \frac{k_N \lambda \mu}{q_N} \\
&= \mathsf{E}[h_N(IN_N) + ... + h_3(IN_3) + G_{2,3}(IN_3) - \sum_{j=3}^{N-1} I_j(r_j, q_j)|IP_N = y] + \sum_{j=1}^{N-1} C_j(r_j, q_j) + \frac{k_N \lambda \mu}{q_N} \\
&= \qquad \vdots \\
&= \mathsf{E}[G_N(y)|IP_N = y] + \frac{k_N \lambda \mu}{q_N} + \sum_{j=1}^{N-1} C_j(r_j, q_j).
\end{aligned}
$$

Note that IP_N is uniformly distributed over $(r_N, r_N + q_N]$, $\mathsf{E}[G_N(y)] = \left(\int_0^{q_N} G_N(r_N + x)dx\right)/q_N$.

Thus,

$$C(\mathbf{r}(\mathbf{q}),\mathbf{q}) \quad = \quad \frac{\int_0^{q_N} G_N(r_N+x)dx}{q_N} + \frac{k_N\lambda\mu}{q_N} + \sum_{j=1}^{N-1} C_j(r_j,q_j) = \sum_{j=1}^{N} C_j(r_j,q_j).$$

Proposition 4

For $n = 1, 2, ...,$ define

$$H_j(n) = \int_{r_j(nq^0_{j-1})}^{r_j((n-1)q^0_{j-1})} G_j(x)dx + \int_{r_j((n-1)q^0_{j-1})+(n-1)q^0_{j-1}}^{r_j(nq^0_{j-1})+nq^0_{j-1}} G_j(x)dx,$$

where $r_j(0) \stackrel{def}{=} \arg\min_y G_j(y)$. That is, $H_j(n)$ is the nth smallest q^0_{j-1} on $G_j(y)$. Thus, by definition, we have $H_j(n) \le H_j(n+1)$.

Denote

$$\begin{aligned}
\Delta_j(n) &= C_j((n+1)q^0_{j-1}) - C_j(nq^0_{j-1}) \\
&= \frac{-k_j\lambda\mu + n\int_0^{(n+1)q^0_{j-1}} G_j(r_j((n+1)q^0_{j-1})+x)dx - (n+1)\int_0^{nq^0_{j-1}} G_j(r_j(nq^0_{j-1})+x)dx}{n(n+1)q^0_{j-1}} \\
&= \frac{-k_j\lambda\mu + n\left(\sum_{i=1}^{n+1} H_j(i)\right) - (n+1)\left(\sum_{i=1}^{n} H_j(i)\right)}{n(n+1)q^0_{j-1}} \\
&= \frac{-k_j\lambda\mu + nH_j(n+1) - \sum_{i=1}^{n} H_j(i)}{n(n+1)q^0_{j-1}}.
\end{aligned}$$

Note that $-k_j\lambda\mu$ is a negative constant and $n(n+1)q^0_{j-1}$ is positive. To show $C_j(nq^0_{j-1})$ is quasi-convex in n, it suffices to show that $L_j(n) \stackrel{def}{=} nH_j(n+1) - (\sum_{i=1}^{n} H_j(i))$ is increasing in n. This is true since $L_j(n+1) - L_j(n) = (n+1)(H_j(n+2) - H_j(n+1)) \ge 0$. This completes the proof.

Proposition 5

We first show Part 1. Shang and Song (2003b) show that $\hat{C}_j(y) \ge G^{\ell}_j + \sum_{i=2}^{j} h_i \mathsf{E}[\tilde{D}_{j-1}]$. Thus, it suffices to show $\hat{C}_j(y) - \sum_{i=1}^{j-1} I_i(r_i(q^0_i), q^0_i) = G_j(y)$. We prove this by induction. When $j = 1$, $\hat{C}_1(y) = G_1(y)$. Thus the proposition holds for $j = 1$. Now take any $j \ge 1$ and suppose the proposition holds for j, i.e., $\hat{C}_j(y) - \sum_{i=1}^{j-1} I_i(r_i(q^0_i), q^0_i) = G_j(y)$. We next show that it also holds for $j + 1$.

$$\begin{aligned}
&\hat{C}_{j+1}(y) - \sum_{i=1}^{j} I_i(r_i(q^0_i), q^0_i) \\
&= \mathsf{E}\left[h_{j+1}(y - D_{j+1}) + \hat{C}_j(O_j[y - D_{j+1}]) - \sum_{i=1}^{j-1} I_i(r_i(q^0_i), q^0_i)\right] - I_j(r_j(q^0_j), q_j)
\end{aligned}$$

$$= \mathsf{E}[h_{j+1}(y - D_{j+1}) + G_j(O_j(y - D_{j+1})] - I_j(r_j(q_j^0), q_j) \qquad \text{(by induction assumption)}$$
$$= \mathsf{E}[h_{j+1}(y - D_{j+1}) + G_{j,j+1}(y - D_{j+1})]$$
$$= G_{j+1}(y).$$

The result follows immediately.

We next prove Part 2.

$$\lim_{y\to\infty}\left(\frac{dI_j^\ell(y, nq_{j-1}^0)}{dy}\right) = \lim_{y\to\infty}\left(\frac{d}{dy}\left(\frac{\int_0^{nq_{j-1}^0} G_j^\ell(y+x)dx}{nq_{j-1}^0}\right)\right)$$
$$= \frac{1}{nq_{j-1}^0}\lim_{y\to\infty}\left(\int_0^{nq_{j-1}^0}\left(\frac{d}{dy}G_j^\ell(y+x)\right)dx\right)$$
$$= \frac{1}{nq_{j-1}^0}\left(\int_0^{nq_{j-1}^0} h_j dx\right) = h_j.$$

The swap of integral and derivative is allowed since $G_j^\ell(\cdot)$ is continuous over the entire region $[0, nq_{j-1}^0]$ and $\frac{d}{dy}G_j^\ell(y+\cdot)$ is bounded by h_j (Dominance Convergence Theorem). The second to the last equation holds since $\frac{d}{dy}G_j^\ell(y+x)$ converges uniformly to h_j as $y \to \infty$ on the interval $[0, nq_{j-1}^0]$. Similarly, we can show $\lim_{y\to-\infty} d\tilde{I}_j^\ell(y, nq_{j-1}^0)/dy = -(b + h'_{j+1})$.

On the other hand,

$$\lim_{y\to\infty}\left(\frac{dI_j(y, nq_{j-1}^0)}{dy}\right) = \frac{1}{nq_{j-1}^0}\lim_{y\to\infty}\left(\frac{d}{dy}\int_0^{nq_{j-1}^0} G_j(y+x)dx\right)$$
$$= \frac{1}{nq_{j-1}^0}\lim_{y\to\infty}\Big(G_j(y + nq_{j-1}^0) - G_j(y)\Big)$$
$$= \frac{1}{nq_{j-1}^0}\lim_{y\to\infty}(h_j nq_{j-1}^0 + \mathsf{E}[G_{j-1}(O_{j-1}[y + nq_{j-1}^0 - D_j]) - G_{j-1}(O_{j-1}[y - D_j])])$$
$$= h_j + \frac{1}{nq_{j-1}^0}\mathsf{E}[\lim_{y\to\infty}(G_{j-1}(O_{j-1}[y + nq_{j-1}^0 - D_j]) - G_{j-1}(O_{j-1}[y - D_j]))] = h_j.$$

The second to the last equation is due to Monotone Convergence Theorem. $\lim_{y\to-\infty} d\tilde{I}_j(y, nq_{j-1}^0)/dy = -(b + h'_{j+1})$ can be shown by noting that $\frac{d}{dy}G_j(y+x)$ converges uniformly to $-(b + h'_{j+1})$ as $y \to -\infty$ on the interval $[0, nq_{j-1}^0]$. Using the similar technique as Part 1, the result follows immediately.

We proceed to show Part 3.

$$\lim_{y\to\infty}\left(\tilde{I}_j^\ell(y, (n+1)q_{j-1}^0) - \tilde{I}_j^\ell(y, nq_{j-1}^0)\right)$$
$$= \lim_{y\to\infty}\left(\frac{\int_0^{(n+1)q_{j-1}^0}\tilde{G}_j^\ell(y+x)dx}{(n+1)q_{j-1}^0} - \frac{\int_0^{nq_{j-1}^0}\tilde{G}_j^\ell(y+x)dx}{nq_{j-1}^0}\right)$$

$$= \lim_{y\to\infty}\left(\frac{n\int_0^{(n+1)q^0_{j-1}} \mathsf{E}[h_j(y+x-D_j)+(b+h_j^{'})(y+x-D_j)^-]dx}{n(n+1)q^0_{j-1}}\right)$$

$$- \lim_{y\to\infty}\left(\frac{(n+1)\int_0^{nq^0_{j-1}} \mathsf{E}[h_j(y+x-D_j)+(b+h_j^{'})(y+x-D_j)^-]dx}{n(n+1)q^0_{j-1}}\right)$$

$$= \frac{nh_j((n+1)q^0_{j-1})^2/2-(n+1)h_j(nq^0_{j-1})^2/2}{n(n+1)q^0_{j-1}} = h_jq^0_{j-1}/2.$$

Similarly, we can show $\lim_{y\to-\infty}\left(\tilde{I}^\ell_j(y,(n+1)q^0_{j-1})-\tilde{I}^\ell_j(y,nq^0_{j-1})\right) = -(b+h_{j+1}^{'})q^0_{j-1}/2$.

The proof of $\lim_{y\to\infty}\left(I_j(y,(n+1)q^0_{j-1})-I_j(y,nq^0_{j-1})\right) = h_jq^0_{j-1}/2$ is similar as above by noting that

$$\lim_{y\to\infty}\left(n\int_0^{(n+1)q^0_{j-1}} \mathsf{E}[G_{j-1,j}(y+x-D_j)]dx-(n+1)\int_0^{nq^0_{j-1}} \mathsf{E}[G_{j-1,j}(y+x-D_j)]dx\right) = 0.$$

Also, note that $\lim_{y\to-\infty} G_j(y) = -(b+h_{j+1}^{'})y$. Thus,

$$\lim_{y\to-\infty}\left(I_j(y,(n+1)q^0_{j-1})-I_j(y,nq^0_{j-1})\right)$$

$$= \frac{\int_0^{(n+1)q^0_{j-1}} -(b+h_{j+1}^{'})(y+x)dx}{(n+1)q^0_{j-1}} - \frac{\int_0^{nq^0_{j-1}} -(b+h_{j+1}^{'})(y+x)xdx}{nq^0_{j-1}} = -(b+h_{j+1}^{'})q_{j-1}/2.$$

Finally, we prove Part 4. We ignore the constant terms for brevity, i.e., we show the properties hold for $I^\ell_j(r^\ell_j(nq^0_{j-1}),nq^0_{j-1})$. Note that $I^\ell_j(r,q)$ is jointly convex (Zipkin 1986b). It follows from the convexity preservation theorem, see, e.g., Proposition B-4 in Heyman and Sobel (1984, p. 525) that $I^\ell_j(r^\ell_j(q_j),q_j)$ is convex in q_j, or equivalently in n.

Define

$$H^\ell_j(q) \quad = \quad G^\ell_j(r^\ell_j(q)), \qquad H^\ell_j(0) = G^\ell_j(y^0_j),$$

where $y^0_j = \text{argmin}_y G^\ell_j(y)$. Following Zheng (1992), it can be shown that $\int_{r^\ell_j(q)}^{r^\ell_j(q)+q} G^\ell_j(y)dy = \int_0^q H^\ell_j(y)dy$ and $\lim_{q\to\infty} H^{\ell\,'}_j(q) = h_j(b+h_{j+1}^{'})/(b+h_j^{'})$. We now take derivative of $I^\ell_j(r^\ell_j(q),q)$ with respect to q:

$$\lim_{q\to\infty} I^{\ell\,'}_j(r^\ell_j(q),q) = \lim_{q\to\infty}\left(\frac{\int_{r^\ell_j(q)}^{r^\ell_j(q)+q} G^\ell_j(y)dy}{q}\right)^{'} = \lim_{q\to\infty}\left(\frac{\int_0^q H^\ell_j(y)dy}{q}\right)^{'}$$

$$= \lim_{q\to\infty}\frac{H^\ell_j(q)q-\int_0^q H^\ell_j(y)dy}{q^2} \tag{20}$$

$$= \lim_{q\to\infty} \left(\frac{H_j^\ell(q)q - \int_0^q \left(\int_0^y H_j^{\ell\,\prime}(x)dx + H_j^\ell(0) \right) dy}{Q^2} \right)$$
$$= \lim_{q\to\infty} \left(\frac{H_j^\ell(q)q - H_j^\ell(0)q - \int_0^q \int_0^y H_j^{\ell\,\prime}(x)dxdy}{q^2} \right)$$
$$= \lim_{q\to\infty} \left(\frac{q\int_0^q H_j^{\ell\,\prime}(x)dx - \int_0^q \int_0^y H_j^{\ell\,\prime}(x)dxdy}{q^2} \right) = \frac{1}{2} \left(\frac{h_j(b + h_{j+1}^{'})}{(b + h_j^{'})} \right).$$

From (20), the derivative of $I_j^\ell(r_j^\ell(q), q)$ is positive. It is clear that $I_j^\ell(r_j^\ell(q), q)$ is increasing in q, or equivalently in n. Since $q_j = nq_{j-1}^0$, the asymptotic slope for n is $(q_{j-1}^0/2)\left(h_j(b + h_{j+1}^{'})/(b + h_j^{'})\right)$.

Proposition 6

Following the proof of Part 4 in Proposition 5, for each stage i, $I_i^\ell(r_i^\ell(Q), Q)$ is convex increasing in Q with the asymptotic slope $(1/2)\left(h_j(b + h_{j+1}^{'})/(b + h_j^{'})\right)$. The result immediately follows.

Proposition 7

The proof is similar to that of Part 4 in Proposition 5. One needs to note that (1) $\bar{I}_k(y, Q)$ is jointly convex, and $\sum_{i\in c(k)} G_i^\ell(y)$ is convex with the asymptotic slope $\sum_{t=0}^{n(k)-1} h_{v+t}$ as $y \to \infty$ and the $n(k)b + \sum_{t=0}^{n(k)-1} h_{v+t+1}^{'}$ as $y \to -\infty$. The detailed proof is omitted.

Theorem 8

For simplicity, we assume the demand is continuous. The objective function in (18) can be rewritten as

$$C(r, q, \omega) \quad = \quad h\left(\frac{(\frac{k}{h})\lambda\mu}{q} + \frac{1}{q}(\frac{1}{1-\omega}) \int_r^{r+q} f(y, \omega)dy \right),$$

where

$$f(y, \omega) \quad = \quad (y - m)(1 - \omega) + F^1(y).$$

Here, $m = \mathsf{E}[D]$ and $F^1(y)$ is the loss function, which is defined as $F^1(y) = \mathsf{E}[(y - D)^-]$.

At optimality, we have $\frac{\partial C}{\partial r} = 0$ and $\frac{\partial C}{\partial q} = 0$. These equations can be reduced, respectively, to

$$(1 - \omega)(-q) + (F^1(r) - F^1(r + q)) = 0, \tag{21}$$
$$q(\frac{1}{1-\omega})f(r + q, \omega) - (\frac{k}{h})\lambda\mu - (\frac{1}{1-\omega}) \int_r^{r+q} f(y, w)dy = 0. \tag{22}$$

For a given q, let $r = r(q)$ be the optimal reorder point. Replace $r = r(q)$ in (21) and take the derivative with respect to q, we have

$$(\omega - 1) - F^0(r(q))r^{'}(q) + F^0(r(q) + q)(r^{'}(q) + 1) = 0,$$

where $F^0(\cdot) = \mathsf{P}(D > \cdot)$.

Thus,

$$\begin{aligned} r^{'}(q) &= \frac{(1-\omega) - F^0(r(q)+q)}{F^0(r(q)+q) - F^0(r(q))} \\ &= \frac{F^0(s^*) - F^0(r(q)+q)}{F^0(r(q)+q) - F^0(r(q))}. \end{aligned}$$

Note that $1 > F^0(s^*) - F^0(r(q)+q) > 0$, $-1 < F^0(r(q)+q) - F^0(r(q)) < 0$, and $|F^0(s^*) - F^0(r(q)+q)| < |F^0(r(q)+q) - F^0(r(q))|$. Thus, $-1 < r^{'}(q) < 0$.

In (22), define $\upsilon = \frac{k}{h}$, $r = r(q)$, and

$$G(q, \omega, \upsilon) = q(\frac{1}{1-\omega})f(r(q)+q, \omega) - \upsilon\lambda\mu - (\frac{1}{1-\omega})\int_{r(q)}^{r(q)+q} f(y, w)dy.$$

To show this theorem, we need to show $\frac{dq}{d\upsilon}$ is positive. By the implicit function theorem, this means to show $\frac{-\frac{\partial G}{\partial \upsilon}}{\frac{\partial G}{\partial q}}$ is positive. Fix ω and consider $-\frac{\partial G}{\partial \upsilon}$:

$$-\frac{\partial G}{\partial \upsilon} = \lambda\mu > 0.$$

Next,

$$\frac{\partial G}{\partial q} = \frac{1}{1-\omega}\left(-r^{'}(q)f(r(q)+q, \omega) + qf^{'}(r(q)+q, \omega)(r^{'}(q)+1)\right).$$

Since $f^{'}(r(q)+q, \omega) > 0$ and $-1 < r^{'}(q) < 0$, $\frac{\partial G}{\partial q} > 0$. As a result, $\frac{-\frac{\partial G}{\partial \upsilon}}{\frac{\partial G}{\partial q}}$ is positive.

References

[1] Atkins, D. 1990. A survey of lower bounding methodologies for inventory/production models. *Annals Operations Research* **26** 9-28.

[2] Atkins, D. and D. Sun. 1995. 98%-effective lot-sizing for series inventory system with backlogging. *Operations research* **43** 335-345.

[3] Axsäter, S. 1993. Exact and approximate evaluation of batch ordering policies for two-level inventory systems. *Operations research* **41** 777-785.

[4] Axsäter. 2000. Exact analysis of continuous review (r, q) policies in two-echelon inventory systems with compound poisson demand. *Operations research* **48** 686-696.

[5] Axsäter, S., K. Rosling. 1993. Installation vs. echelon stock policies for multi-level inventory control. *Operations Research* **39** 1274-1280.

[6] Badinelli, R. 1992. A model for continuous-review pull policies in serial inventory systems. *Operations Research* **40** 142-156.

[7] Cachon, G. 2001. Exact evaluation of batch-ordering policies in two-echelon supply chains with periodic review. *Operations Research* **49** 79-98.

[8] Chen, F. 1998. Stationary policies in multiechelon inventory systems with deterministic demand and backlogging. *Operations Research* **46** S26-S34.

[9] Chen, F. 2000. Optimal policies for multi-echelon inventory problems with batch ordering. *Operations research* **48** 376-389.

[10] Chen, F., Y. S. Zheng. 1994a. Evaluating echelon stock (r, nq) policies in serial production/inventory systems with stochastic demand. *Management Science* **40** 1262-1275.

[11] Chen, F., Y. S. Zheng. 1994b. Lower bounds for multi-echelon stochastic inventory systems. *Management Science* **40** 1426-1443.

[12] Chen, F., Y. S. Zheng. 1998. Near-optimal echelon-stock (r, nq) policies in multi-stage serial systems. *Operations research* **46** 592-602.

[13] Clark, A., H. Scarf. 1962. Approximate solutions to a simple multi-echelon inventory problem. In *Studies in Applied Probability and Management Science.* K. J. Arrow, S. Karlin, H. Scarf (eds.). Stanford University Press, Stanford, CA.

[14] De Bodt, M., S. Graves. 1985. Continuous review policies for a multi-echelon inventory problem with Stochastic demand. *Management Science* **31** 1286-1295.

[15] Doğru, M. K., G. J. van Houtum, A. G. de Kok. 2005. Newsboy characterizations for the optimal reorder levels of multi-echelon inventory systems with fixed batch sizes. *Working paper*, Technische Universiteit Eindhoven, Eindhoven, Netherlands.

[16] Federgruen, A. M. Queyranne, Y. S. Zheng. 1992. Simple power of two policies are close to optimal in general class of production/distribution networks with general joint setup costs. *Math. Operations Research* **17** 951-963.

[17] Federgruen, A., Y. S. Zheng. 1992. An efficient algorithm for computing an optimal (r, q) policy in continuous review stochastic inventory systems. *Operations research* **40** 808-813.

[18] Gallego, G., P. Zipkin. 1999. Stock positioning and performance estimation in serial production-transportation systems. *Manufacturing & Service Operations Management.* **1** 77-88.

[19] Geoffrion, A. M. 1967. Reducing concave programs with some linear constraints. *SIAM J. Applied Math.* **15** 653-664.

[20] Heyman, D. P., M .J. Sobel. 1984. *Stochastic Models in Operations Research.* Vol. II, McGraw-Hill, New York.

[21] Jackson, L. P., W. Maxwell, J. Muckstadt. 1988. Determining optimal reorder intervals in capacitated production-distribution systems. *Management Science* **34** 938-958.

[22] Maxwell, W., J. Muckstadt. 1985. Establishing consistent and realistic reorder intervals in production-distribution systems. *Operations Research* **33** 1316-1341.

[23] Muckstadt, J., R. Roundy. 1993. Analysis of multistage production systems. *In Handbooks in OR and MS*, Vol. 4, Graves et al. (eds.), Elsevier Science Publishers, North-Holland.

[24] Rosling, K. 1993. 94%-effective lostsizing: user-oriented algorithms and proofs. Working paper, Linköping Institute of Technology, Sweden.

[25] Roundy, R. 1985. 98%-effective integer-ratio lot sizing for one-warehouse multi-retailer systems. *Management Science* **31** 1416-1430.

[26] Roundy, R. 1986. A 98%-effective lot-sizing rule for a multi-product, multi-stage production/inventory system. *Mathematics of Operations Research* **11** 699-727.

[27] Schwarz, L. B. 1973. A simple continuous review deterministic one-warehouse multi-retailer problem. *Management Science* **19** 555-566.

[28] Schwarz, L. B., L. Schrage. 1975. Optimal and system myopic policies for multi-echelon production/inventory assembly systems. *Management Science* **21** 1285-1294.

[29] Serfozo, R., S. Stidham. 1978. Semi-stationary clearing processes. *Stochastic Process. Appl.* **6** 165-178.

[30] Shang, K., J. S. Song. 2003a. Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science* **49** 618-638.

[31] Shang, K., J. S. Song. 2003b. Serial supply chains with economies of scale: bounds and approximations. Working paper.

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

[33] Svoronos, A., P. Zipkin. 1988. Estimating the performance of multi-level inventory systems. *Operations Research* **36** 57-72.

[34] Zheng, Y. S. 1992. On properties of stochastic inventory systems. *Management Science* **38** 87-103.

[35] Zipkin, P. 1986a. Stochastic leadtimes in continuous-time inventory models. *Naval Research Logisitics Quarterly* **33** 763-774.

[36] Zipkin, P. 1986b. Inventory service-level measures: convexity and approximation. *Management Science* **32** 975-981.

[37] Zipkin, P. 2000. *Foundations of Inventory Management.* McGraw-Hill, New York.