OPERATIONS RESEARCH
Vol. 56, No. 4, July–August 2008, pp. 1039–1043
ISSN 0030-364X | EISSN 1526-5463 | 08 | 5604 | 1039

DOI 10.1287/opre.1080.0547

TECHNICAL NOTE

Note: A Simple Heuristic for Serial Inventory Systems with Fixed Order Costs

Kevin H. Shang
Fuqua School of Business, Duke University, Durham, North Carolina 27708, khshang@duke.edu

We propose a heuristic for finding base order quantities for stochastic inventory models. The heuristic includes two steps. The first clusters the stages according to cost parameters. The second solves a single-stage problem for each cluster with the original problem data. In a numerical study, we show that the heuristic is near optimal.

Subject classifications: multiecholon; (R, nQ) policies; simple solution.
Area of review: Manufacturing, Service, and Supply Chain Operations.
History: Received August 2004; revisions received July 2005, March 2007; accepted March 2007.

1. Introduction

Consider a continuous-review, serial inventory system with N stages. Material flows from stage N to stage $N-1$, $N-1$ to $N-2$, etc., until stage 1, where stationary compound Poisson demand occurs. Let λ and μ denote the demand rate and the mean of demand size, and $D[t]$ denote the demand over a time interval t. There is a constant lead time L_j for stage j. Define $\tilde{L}_i = \sum_{j=1}^{i} L_j$. Inventory is controlled by echelon-stock (R, nQ) policies. That is, for each stage j, as soon as the echelon inventory-order 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, an order of the smallest integer multiple of base quantity Q_j is placed to raise the inventory position above R_j. We assume that base quantities satisfy integer-ratio constraints, i.e., $Q_j = q_j Q_{j-1}$, $q_j \in \Im^+$, $j = 2, \ldots, N$, where $\Im^+$ denotes the set of positive integers. There is a linear echelon holding cost with rate h_j for stage j. Unsatisfied demand is fully backlogged and incurs a linear penalty cost with rate b. Define $h'_j = \sum_{i=j}^{N} h_i$ as the local holding cost rate for stage j. A fixed order cost k_j occurs for each base quantity Q_j ordered. The objective is to minimize the long-run average systemwide cost.

When demand is deterministic, it is well known that there exists a simple heuristic that guarantees a 94% effective policy, i.e., the cost of this heuristic policy is at most about 6% more costly than any feasible policy. This heuristic includes two steps. The first step is to identify clusters by using cost ratios. (A cluster is a set that includes consecutive stages; see a formal definition below.) The second step solves an economic order quantity (EOQ) problem for each cluster. A heuristic policy is then obtained by converting these EOQ solutions into integer-ratio order quantities. One well-known integer-ratio solution is the so-called "power-of-two" policy (e.g., Maxwell and Muckstadt 1985, Roundy 1985, and Federgruen et al. 1992). We refer the reader to Zipkin (2000) for a detailed discussion of the deterministic model and the power-of-two policy.

In the stochastic demand model, however, finding effective base order quantities is more involved. The difficulty comes from the fact that, unlike the deterministic model, the objective function (the average total cost) is not a sum of separable, convex functions of control variables. To resolve this, Chen and Zheng (1998) construct cost bounds to replace the average total cost in the objective function. These cost bounds are a sum of separable functions of base quantities. Thus, a standard clustering algorithm (e.g., Maxwell and Muckstadt 1985) can be applied to find the optimal solution for these revised problems. They propose heuristics by converting these optimal solutions to power-of-two base quantities.

The objective of this note is to propose a new heuristic that employs the same steps as the heuristic for the deterministic model. As we shall see, this heuristic outperforms the existing ones in a numerical study.

2. The Heuristic

This heuristic includes two steps, *clustering* and *minimization*. In the clustering step, the stages are grouped into disjoint clusters $\{c(1), c(2), \ldots, c(M)\}$ according to cost ratios. (Let $S = \{1, 2, \ldots, N\}$. For any $i, j \in S$ with $i \leqslant j$, the set $\{i, i+1, \ldots, j\}$ is called a cluster.) Specifically, define

$$h[m] = \sum_{i\in c(m)} h_i, \quad h'[m] = \sum_{i\in c(m)} h'_i, \quad \text{and} \quad k[m] = \sum_{i\in c(m)} k_i.$$

These clusters satisfy the following two conditions:

(i) $k[1]/h[1] < \cdots < k[M]/h[M]$,

(ii) for each cluster $c(m) = \{l_1, \ldots, l_2\}$, there does not exist an l with $l_1 \leqslant l < l_2$ so that $k[m^-]/h[m^-] < k[m^+]/h[m^+]$, where $c(m^-) = \{l_1, \ldots, l\}$ and $c(m^+) = \{l+1, \ldots, l_2\}$.

Through a two-dimensional diagram suggested by Zipkin (2000), the above clusters $\{c(1), c(2), \ldots, c(M)\}$ can be conveniently identified.

In the minimization step, we solve a single-stage problem for each cluster $c(m)$ sequentially, starting with $m=1$. In each problem, the solution of base quantity $Q_{c(m)}$ is restricted to be an integer multiple of $Q_{c(m-1)}$, $m \geqslant 2$. We consider two variants of such single-stage problems.

Variant 1
$Q_{c(m)}$ is the solution of the following problem:

$$\min_Q \quad \left\{ \frac{\lambda\mu k[m] + \sum_{x=1}^{Q} G^{\ell}_{c(m)}(R^{\ell}(Q)+x)}{Q} \right\} \qquad (1)$$
$$\text{s.t.} \quad Q = qQ_{c(m-1)}, \quad q \in \Im^+,\ m > 1,$$

where

$$R^{\ell}(Q) = \arg\min_y \left\{ \sum_{x=1}^{Q} G^{\ell}_{c(m)}(y+x) \right\},$$
$$G^{\ell}_{c(m)}(y) = \sum_{i \in c(m)} G^{\ell}_i(y),$$

and

$$G^{\ell}_i(y) = \mathsf{E}\big[h_i(y - D[\tilde{L}_i]) + (b + h'_i)(y - D[\tilde{L}_i])^-\big]. \qquad (2)$$

Variant 2
Suppose that $c(m)$ contains stages i, $i \in \{v, v+1, \ldots, v+n(m)-1\}$. Alternatively, $Q_{c(m)}$ is the solution of the following problem:

$$\min_Q \quad \left\{ \frac{\lambda\mu k[m] + \sum_{x=1}^{Q} G_{c(m)}(R(Q)+x)}{Q} \right\} \qquad (3)$$
$$\text{s.t.} \quad Q = qQ_{c(m-1)}, \quad q \in \Im^+,\ m > 1,$$

where

$$R(Q) = \arg\min_y \left\{ \sum_{x=1}^{Q} G_{c(m)}(y+x) \right\}$$

and

$$G_{c(m)}(y) = \mathsf{E}\big[h[m](y - D[\tilde{L}_{v+n(m)-1}]) + (n(m)b + h'[m])(y - D[\tilde{L}_{v+n(m)-1}])^-\big].$$

The $Q_{c(m)}$ is solved sequentially, starting with $m = 1$. Then, set $Q^s_i = Q_{c(m)}$ for $i \in c(m)$; these are the heuristic base quantities. The corresponding heuristic reorder points $(R^s_1, \ldots, R^s_N)$ are found by using the algorithm for finding optimal echelon reorder points in Chen (2000).

We briefly describe the derivation of the heuristic. A detailed discussion can be found in Shang (2004). We first decouple the total average cost into each stage by using the induced penalty functions (the penalty cost charged to each upstream stage for holding inadequate stock). The decoupled cost function for stage j is then replaced by $G^{\ell}_j(\cdot)$ defined in (2). Consequently, we solve a revised problem, in which the objective function is a sum of N separable convex functions subject to the relaxed constraints $Q_j \geqslant Q_{j-1}$, $j \geqslant 2$.

Consider any partition of stages for the revised problem. Suppose that the resulting clusters are $\{c(1), c(2), \ldots, c(Z)\}$. An important observation is that the optimal common base quantity $Q_{c(z)}$ for the cluster $c(z)$, $z = 1, \ldots, Z$, is close to the optimal base quantity $\widehat{Q}_z$ solved from a single-stage (R, nQ) system, referred to as the *mapping system*. This mapping system has the following cost parameters: holding cost rate $h[z]$, backorder cost rate $b[z] \stackrel{\text{def}}{=} (n(z)b + h'[z] - h[z])$, and fixed order cost $k[z]$, where $n(z)$ is the number of stages in $c(z)$. The following theorem shows that $\widehat{Q}_z$ increases in $k[z]/h[z]$ with $\omega[z] \stackrel{\text{def}}{=} b[z]/(b[z]+h[z])$ fixed.

THEOREM 1. *For a single-stage (R, nQ) system with holding cost h, backorder cost b, and fixed order cost k, the optimal base order quantity increases in k/h with $\omega \stackrel{\text{def}}{=} b/(b+h)$ fixed.*

From Theorem 1, if we want to use $k[z]/h[z]$ to determine the order relation of $\widehat{Q}_z$, $\omega[z]$ and the lead times in these Z mapping systems must be equal. In our numerical experience, we find that $\omega[z]$ are similar when b is large and the lead time seems to have little impact on Q_z. For approximation purpose, we simply assume that the condition holds. Because $\widehat{Q}_z$ is close to $Q_{c(z)}$, we may use the cost ratios to identify the clusters for the revised problem.

The heuristic is then derived in a reverse order: we first use the cost parameters h_j and k_j to identify the clusters. In the minimization step, Equations (1) and (3) correspond to the sum of C^{ℓ}_j functions in a cluster and the average total cost of the mapping system, respectively. Finally, for the deterministic model, Shang (2004) finds that the heuristic solution obtained from the bottom-up minimization approach performs, on average, better than the power-of-two heuristic. This motivates us to use the same approach in the minimization step.

3. Numerical Study

We conduct three groups of numerical study.

Group 1. We test the effectiveness of the two variants, using optimal costs as a benchmark (see Chen and Zheng 1998 for an optimization algorithm). The system parameters are $N = 3$, $b = 10, 50$, $h_j = 0.1, 1$, $L_j = 0.5, 2$, $k_j = 10, 100$, for $j = 1, 2, 3$. Demand is Poisson with rate $\lambda = 5$. The total number of instances is 1,024. We also consider a heuristic that uses base quantities obtained from the corresponding deterministic model by applying the clustering and minimization steps mentioned above. (The

Table 1. Performance summary for Variant 1, Variant 2, and the EOQ heuristic in Group 1.

	Variant 1		Variant 2		The EOQ heuristic	
	$b=10$	$b=50$	$b=10$	$b=50$	$b=10$	$b=50$
0%	193	309	137	307	31	5
(0.0%, 0.5%]	262	157	316	153	378	350
(0.5%, 1.0%]	25	26	24	30	57	128
(1.0%, 1.5%]	12	8	11	10	6	15
(1.5%, 2.0%]	4	4	6	4	6	6
(2.0%, 2.5%]	8	0	10	0	20	0
(2.5%, 3.0%]	0	1	0	1	6	4
3.0% above	8	7	8	7	8	4
Average (%)	0.21	0.13	0.24	0.14	0.41	0.40

total cost for nested, stationary policies in the deterministic model with backorders is derived in Chen 1998.) We again use Chen's (2000) algorithm to find the corresponding reorder points. Call this approach the *EOQ heuristic*. Define the percentage error for a heuristic as

$$\text{error\%} = \frac{(\text{heuristic cost} - \text{optimal cost})100\%}{\text{optimal cost}}.$$

Table 1 summarizes the performance of both variants and the EOQ heuristic. The numbers from row 3 to row 10 under each of the heuristic approaches are the distribution of the percentage error for the 512 instances with a specific b value. For example, when $b=50$, there are 157 instances whose percentage errors fall in $(0\%, 0.5\%]$ by using the Variant 1 approach. Remarkably, Variants 1 and 2 generate the optimal solution in 502 and 444 instances, respectively. The average (maximum) percentage errors for Variant 1 and Variant 2 are 0.17% (4.77%) and 0.19% (4.77%). We observe that when $b=10$, each variant generates fewer optimal solutions. On the other hand, although the EOQ heuristic generates the optimal solution in only 36 instances, its average performance is quite effective—the average (maximum) percentage error is 0.40% (5.92%).

Group 2. We use Variant 1, the best heuristic, to compare the heuristic developed by Chen and Zheng (1998). We consider the same instances as those in Tables 1 to 4 of Chen and Zheng (1998)—in particular, the demand is Poisson (compound Poisson) for the 40 instances in Tables 1 and 3 (Tables 2 and 4). Table 2 shows the average percentage error for our heuristic cost and Chen and Zheng's heuristic cost, using the optimal cost as a benchmark. The numbers in parentheses denote the number of optimal solutions found by using the corresponding heuristic. The result shows that our simple heuristic outperforms Chen and Zheng's heuristic.

Table 2. Performance summary of the heuristics for the instances in Tables 1–4 of Chen and Zheng (1998).

	Table 1	Table 2	Table 3	Table 4
Variant 1	0.28% (8)	0.44% (2)	0.15% (12)	0.20% (3)
Chen and Zheng	1.71% (0)	0.64% (0)	1.76% (1)	1.42% (0)

Group 3. We verify the effectiveness of the Variant 1 heuristic when N increases. We compare our heuristic cost with the best lower-bound cost developed by Chen and Zheng (1998) for $N = 2, 4, 8, 16, 32$, 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

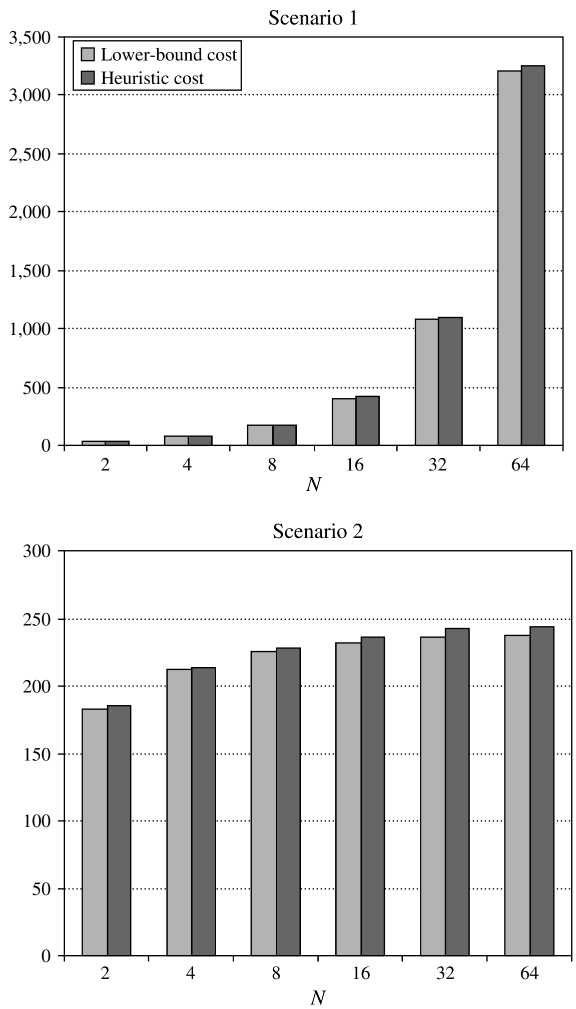
Table 3. The effect of echelon holding cost reduction on the optimal cost.

k_1	k_2	k_3	L_1	L_2	L_3	Optimal cost	Variant 1 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 4. The effect of lead-time reduction on the optimal cost.

k_1	k_2	k_3	L_1	L_2	L_3	Optimal cost	Variant 1 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

Figure 1. The lower-bound cost and the heuristic cost generated from the Variant 1 approach for Group 3 examples.



maximum percentage gaps are 0.54% and 1.48% ($N = 64$), respectively, whereas in the second, they are 1.72% and 2.96% ($N = 64$), respectively. We observe that the percentage error increases in N. We conjecture that this observation resulted from the comparison of our heuristic cost with the lower-bound cost. Overall, the heuristic performs effectively when N is large. Figure 1 shows the heuristic cost with respect to N.

4. Discussion

The above numerical study indicates that the heuristic is near optimal. Thus, we can use it to develop insights for managing the model under consideration.

(1) The heuristic suggests that the base quantity is jointly determined by the fixed order costs and holding costs of the stages in a cluster. Thus, to reduce the base order quantity in a location, it is not sufficient to consider the costs within that location only.

(2) There are 16 instances where percentage errors are larger than 2.5% for Variant 1 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 $b = 10, 50$. Consider the case with $L_1 = 0.5$, $L_2 = 0.5$, $L_3 = 2$, and $b = 50$. In this case, the percentage error is 3.60%. 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 heuristic order quantities are (44, 44, 44), which are exactly the same as the optimal solution. Similar results appear in the other instances. This observation suggests that there may be a way of improving the suggested clustering scheme.

(3) Scenario 2 in Group 3 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 those related to the corresponding serial system without fixed order costs or with known order quantities (Gallego and Zipkin 1999, Shang and Song 2007), where optimal cost decreases in N.

(4) Tables 3 and 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 lead time L_j from 2 to 0.5. We observe that it is more effective to reduce the echelon holding cost at an upstream stage or reduce the lead time at a downstream stage. This observation is consistent with the findings in Shang and Song (2007).

(5) From Tables 3 and 4, we conclude that reducing fixed order cost at a downstream stage is more effective than at an upstream stage.

(6) 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 a discussion of this property). For example, recall the instance mentioned in (2). Using the heuristic order quantity 33 instead of the optimal order quantity 44 results in only 3.6% of error.

(7) For an assembly system with the assumption that base quantities do not decrease in their total lead times (the sum of the lead times from the current stage to stage 1), Chen (2000) shows that there exists an equivalent serial system. Consequently, our heuristic can be used for such assembly systems.

Appendix. Proof of Theorem 1

We prove this result by the implicit function theorem. For simplicity, we assume that the demand is continuous and the lead time is L. The average total cost for the single-stage problem is

$$C(r, Q) = \frac{k\lambda\mu + \int_r^{r+Q} G(y)\, dy}{Q},$$

where $G(y) = h(y - \mathsf{E}[D[L]]) + (b+h)\int_y^\infty \mathsf{P}(D[L] > \xi)\,d\xi$. Let $r(Q)$ denote the optimal reorder point for given Q. The optimal order quantity Q^* satisfies the following equation:

$$Q^* G(r(Q^*)+Q^*) - k\lambda\mu - \int_{r(Q^*)}^{r(Q^*)+Q^*} G(y)\,dy = 0. \qquad (4)$$

Define $v = k/h$, $\omega = b/h$, and $f(y,\omega) = y - \mathsf{E}[D[L]] + (1+\omega)\int_y^\infty \mathsf{P}(D[L] > \xi)\,d\xi$. Then, Equation (4) can be expressed as $h(Q^* f(r(Q^*)+Q^*) - v\lambda\mu + \int_{r(Q^*)}^{r(Q^*)+Q^*} f(y,w)\,dy) = 0$. Define

$$H(Q,v,\omega) = Q^* f(r(Q^*)+Q^*) - v\lambda\mu + \int_{r(Q^*)}^{r(Q^*)+Q^*} f(y,w)\,dy.$$

$\partial H/\partial v = -\lambda\mu$, and $\partial H/\partial Q^* = Q^*(r'(Q^*)+1) f'(r(Q^*)+Q^*,\omega)$. From Zheng (1992), $-1 < r'(Q^*) < 0$, and $f'(r(Q^*)+Q^*,\omega) > 0$. Thus, $\partial H/\partial Q^* > 0$. Together, we have $\partial Q^*/\partial v > 0$.

Acknowledgments

The author thanks Jan Van Mieghem and the anonymous reviewers for their helpful comments. The author is also grateful to Jing-Sheng Song and Paul Zipkin for useful discussions.

References

Chen, F. 1998. Stationary policies in multiechelon inventory systems with deterministic demand and backlogging. *Oper. Res.* **46** S26–S34.

Chen, F. 2000. Optimal policies for multi-echelon inventory problems with batch ordering. *Oper. Res.* **48** 376–389.

Chen, F., Y. S. Zheng. 1998. Near-optimal echelon-stock (R, nQ) policies in multi-stage serial systems. *Oper. Res.* **46** 592–602.

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

Gallego, G., P. Zipkin. 1999. Stock positioning and performance estimation in serial production-transportation systems. *Manufacturing Service Oper. Management* **1** 77–88.

Maxwell, W., J. Muckstadt. 1985. Establishing consistent and realistic reorder intervals in production-distribution systems. *Oper. Res.* **33** 1316–1341.

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

Shang, K. 2004. A simple heuristic for serial inventory systems with fixed order cost. Unabridged working paper, Fuqua School of Business, Duke University, Durham, NC.

Shang, K., J. S. Song. 2007. Serial supply chains with economies of scale: Bounds and approximations. *Oper. Res.* **55** 843–853.

Zheng, Y. S. 1992. On properties of stochastic inventory systems. *Management Sci.* **38** 87–103.

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