Electronic Companion

EC.1. Proofs of Technical Lemmas and Theorems

LEMMA 1. *Let $\mathscr{C}(RB)$ be the total cost incurred by the RB policy. Then we have,*

$$E[\mathscr{C}(RB)] \leq 3 \cdot \sum_{t=1}^{T-L} E[Z_t^{RB}]. \tag{EC.1}$$

Proof of Lemma 1. Using the marginal cost accounting in Equation (4) and standard arguments of conditional expectations, we express

$$\begin{aligned} E[\mathscr{C}(RB)] &= \sum_{t=1}^{T-L} E[H_t^{RB}(Q_t^{RB}) + \Pi_t^{RB}(Q_t^{RB}) + K \cdot \mathbb{1}(Q_t^{RB} > 0)] \\ &= \sum_{t=1}^{T-L} E\left[E[H_t^{RB}(Q_t^{RB}) + \Pi_t^{RB}(Q_t^{RB}) + K \cdot \mathbb{1}(Q_t^{RB} > 0) \mid F_t]\right] \\ &= \sum_{t=1}^{T-L} E[2Z_t^{RB} + P_t K] \leq 3 \sum_{t=1}^{T-L} E[Z_t^{RB}]. \end{aligned} \tag{EC.2}$$

The third equality follows directly from (11). To establish the first inequality in (EC.2) above, we shall show that $Z_t \geq P_t K$ almost surely. That is, for each $f_t \in F_t$, $z_t \geq p_t K$. Given any information set f_t, all the quantities x_t, θ_t, ψ_t, ϕ_t and p_t defined above are known deterministically. We split the analysis into two cases:

1. If $\theta_t \geq K$, then $q_t^{RB} = \hat{q}_t$ (the *balancing quantity*) with probability $p_t = 1$ implying $z_t = \theta_t \geq K$. The claim follows.
2. If $\theta_t < K$, then $q_t^{RB} = \tilde{q}_t$ (the *holding-cost-K quantity*) with probability p_t and $q_t^{RB} = 0$ with $1 - p_t$. Thus, by Equations (8) and (9), we have $z_t = p_t K$, and the claim follows.

This conclude the proof of the lemma. □

LEMMA 2. *The overall holding cost and backlogging cost incurred by OPT are denoted by H^{OPT} and Π^{OPT}, respectively. Then we have, with probability 1,*

$$H^{OPT} \geq \sum_t H_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1H} \bigcup \mathscr{T}_{2H}), \qquad \Pi^{OPT} \geq \sum_t \Pi_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1\Pi} \bigcup \mathscr{T}_{2\Pi}). \tag{EC.3}$$

Proof of Lemma 2. The proof is identical to Lemmas 4.2 and 4.3 in Levi et al. (2007). □

LEMMA 3. *The expected holding cost and backlogging cost incurred by OPT plus the expected amount borrowed from the bank account A are at least $\sum_{t=1}^{T-L} E[Z_t^{RB}]$. That is, The following inequality holds*

$$E\left[\left(H^{OPT}+\Pi^{OPT}\right)+A\right] \geq \sum_{t=1}^{T-L} E[Z_t^{RB}]. \tag{EC.4}$$

Proof of Lemma 3. Using linearity of expectation, it suffices to show

$$E\left[H^{OPT}+\Pi^{OPT}\right] \geq \sum_{t=1}^{T-L} E\left[\mathbb{1}(t \in \mathscr{T}_N) \cdot Z_t^{RB}\right]. \tag{EC.5}$$

Using Lemma 2 and standard arguments of condition expectations, we have

$$\begin{aligned} E[H^{OPT}] &\geq E\left[\sum_t H_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1H} \bigcup \mathscr{T}_{2H})\right] \qquad \text{(EC.6)} \\ &= E\left[E\left[\sum_t H_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1H} \bigcup \mathscr{T}_{2H}) \mid F_t\right]\right] \\ &= E\left[\sum_t Z_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1H} \bigcup \mathscr{T}_{2H})\right]. \end{aligned}$$

Similarly, we also have

$$E[\Pi^{OPT}] \geq E\left[\sum_t Z_t^{RB} \cdot \mathbb{1}(t \in \mathscr{T}_{1\Pi} \bigcup \mathscr{T}_{2\Pi})\right]. \tag{EC.7}$$

Equation (EC.5) follows from summing up Equations (EC.6) and (EC.7). $\square$

LEMMA 4. *The following inequality holds*

$$E[A] \leq E\left[\sum_{t=1}^{T-L} K \cdot \mathbb{1}(Q_t^{OPT} > 0)\right]. \tag{EC.8}$$

In other words, the expected borrowing $E[A]$ is less than the total expected fixed ordering cost incurred by OPT.

Proof of Lemma 4. First we define the *reduced information set* f_t^- to be the information up to period t *excluding* the randomized decisions of the RB policy over $[1, t-1]$. In particular, given the entire evolution of demand f_T^-, the sequence of orders placed by OPT is known deterministically. Let $1 \leq t_1 < t_2 < \ldots < t_n \leq T-L$ be the periods in which OPT placed $n = n \mid f_T^-$ orders sequentially. Let $t_0 = 0$ and $t_{n+1} = T-L+1$. We shall show that there are no problematic periods within (t_0, t_1) and that, for each $i = 1, \ldots n$, the expected borrowing within the interval $[t_i, t_{i+1})$ does not exceed K. That is,

$$(t_0, t_1) \bigcap \mathscr{T}_{2M} = \emptyset, \tag{EC.9}$$

$$E\left[\sum_{t \in [t_i, t_{i+1}) \bigcap \mathscr{T}_{2M}} Z_t^{RB} \mid f_T^-\right] \leq K. \tag{EC.10}$$

It is important to note that f_T^- does not include the randomized decisions of the RB policy. Thus, the set $\mathscr{T}_{2M}$ is still random and so is the amount borrowed from the bank. In particular, the expectation in Equation (EC.10) is taken with respect to the randomized decisions of the RB policy. Equations (EC.10) and (EC.9) imply that, for each f_T^-,

$$E\left[\sum_{t \in \mathscr{T}_{2M}} Z_t^{RB} \mid f_T^-\right] \leq K \cdot n \mid f_T^- = K \cdot n, \tag{EC.11}$$

and therefore

$$E[A] \leq K \cdot E[N] = E\left[\sum_{t=1}^{T-L} K \cdot \mathbb{1}(Q_t^{OPT} > 0)\right]. \tag{EC.12}$$

Thus, it suffices to prove Equations (EC.10) and (EC.9). Figure EC.1 gives a graphical interpretation of Equation (EC.10), i.e., we want to show that the fixed ordering cost K incurred by OPT in period t_i will cover the expected amount borrowed from the bank in periods that belong to set $\mathscr{T}_{2M}$ within the interval $[t_i, t_{i+1})$.

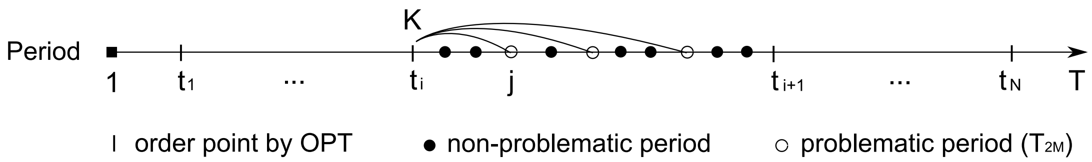

Figure EC.1 Decomposition of the problematic periods in the set $\mathscr{T}_{2M}$ into intervals between ordering points of OPT

Proof of Equation (EC.9). We first show that Equation (EC.9) holds. Recall the definition $\mathscr{T}_{2M} = \left\{t : \Theta_t < K \text{ and } X_t^{RB} < Y_t^{OPT} \leq X_t^{RB} + \tilde{Q}_t^{RB}\right\}$. Since at the beginning of the planning horizon, it is assumed that every feasible policy will have the same initial inventory position, it follows that if period t is in $\mathscr{T}_{2M}$, OPT must have placed an order and overtaken the inventory position of the RB policy. (The two policies face the same sequence of demands.) However, (t_0, t_1) denotes the set of periods in which OPT has not placed any order yet. Thus, the intersection of these two sets is empty.

Proof of Equation (EC.10). Next we show that Equation (EC.10) holds. Recall that f_T^- denotes an entire evolution of the system excluding the randomized decisions of the RB policy. Given the entire evolution of demands f_T^-, construct a decision tree based on the randomized decisions of the RB policy. The root node corresponding to period 1 contains the information set $f_1 = f_1^- \in f_T^-$. The tree is built in layers, each corresponding to a period, where the number of nodes in layer t is 2^{t-1} numbered $l = 1, \ldots, 2^{t-1}$. In particular, a node l in period (layer) t corresponds to some information set $f_t \in \mathscr{F}_t$ which includes the realized reduced information set $f_t^- \subseteq f_T^-$, and the realized randomized decisions up to period $t-1$ of the RB policy. Therefore it is known whether under this state period t belongs to the set $\mathscr{T}_{2M}$ or not.

The edges in the tree represent the different (randomized) decisions that the RB policy may make with their respective probabilities. Each path from the root to a specific node corresponds to a sequence of realized randomized ordering decisions made by the RB policy. For example, consider again some node l in period (layer) t in which the RB policy will order $\tilde{q}_{tl}^{RB}$ units with probability p_{tl} and nothing with probability $1 - p_{tl}$; then the node l in period t (denoted by tl) will have two edges to two children nodes in the next period $t+1$ each containing its distinctive ordering information. Conceptually one can think about the decision tree as a collection of independent coins, each corresponding to a node in the tree. The coin corresponding to node l at layer (period) t has probability of success (ordering) p_{tl}.

Next we partition the nodes in the tree into *problematic* nodes (*pn* nodes), i.e., nodes that correspond to a pair (t, f_t) for which $t \in \mathscr{T}_{2M}$, and *non-problematic* nodes (*nn* nodes). An example of a general decision tree is illustrated in Figure EC.2.

Focus now on a specific time interval $[t_i, t_{i+1})$. Suppose we have constructed the tree from period 1 to T; the number of nodes and paths are clearly finite (possibly exponential). Let the set $\mathscr{G}$ to be the set of all possible outcomes of the randomized decisions in all nodes in layers within the interval $[1, t_i - 1]$ and in all the *nn* nodes within the interval $[1, T]$. In particular, each $g \in \mathscr{G}$ corresponds to a specific set of outcomes in all nodes in layers (periods) within the interval $[1, t_i - 1]$ and in all the *nn* nodes in the tree. Using the terminology of coins proposed before, g corresponds to the

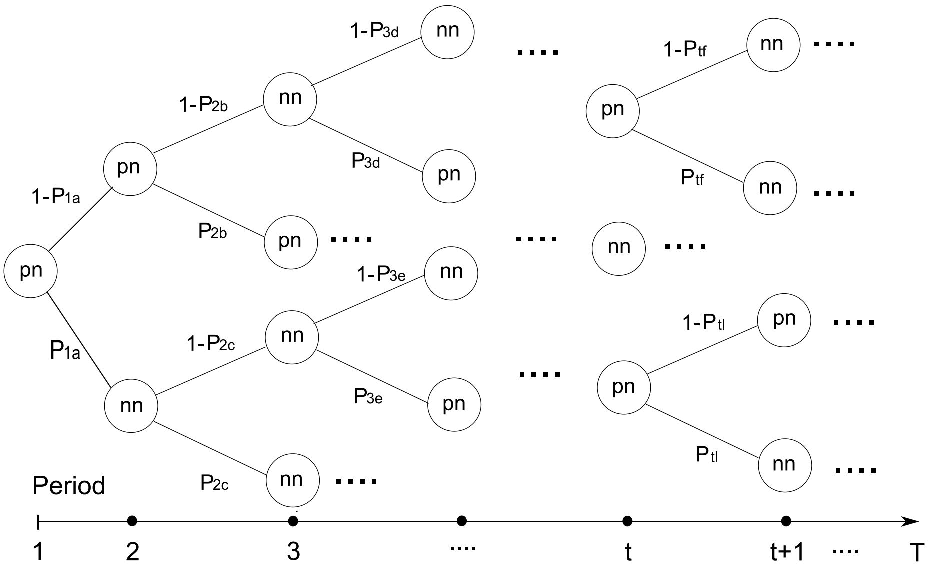

Figure EC.2 An example of a general decision tree

outcome of the respective subset of coins corresponding to all nodes within $[1, t_i - 1]$ and all nn nodes within $[1, T]$.

Conditioning on some $g \in \mathscr{G}$ induces a path from the root of the tree (in period 1) up to the earliest pn node, say j, where j corresponds to the period (layer) of that node. Here we abuse the notation ignoring the index of the node within layer j. (Namely, the exact value will be je for some e.) It is straightforward to see that $j \geq t_i$. If j falls outside the interval $[t_i, t_{i+1})$, i.e., $j \geq t_{i+1}$, it follows that there are no pn nodes within the interval $[t_i, t_{i+1})$, and there is no borrowing over the interval. Assume now that j falls within the interval $[t_i, t_{i+1})$ (j can possibly be in period (layer) t_i). We will show that the expected borrowing does not exceed K. That is,

$$E\left[\sum_{s \in [j, t_{i+1}) \bigcup \mathscr{T}_{2M}} Z_s^{RB} \mid f_T^-, g\right] \leq K. \tag{EC.13}$$

The proof of Equation (EC.10) will then follow.

Recall that node j corresponds to some information set $f_j \in \mathscr{F}_j$. It follows that the starting inventory position x_j^{RB} and the corresponding *holding-cost-K* quantity $\tilde{q}_j^{RB}$ are known deterministically. Conditioning on g, the only uncertainty in the evolution of the system depends on the randomized decisions made in pn nodes within $[j, t_{i+1})$. Consider the sub-tree induced by conditioning on g. The

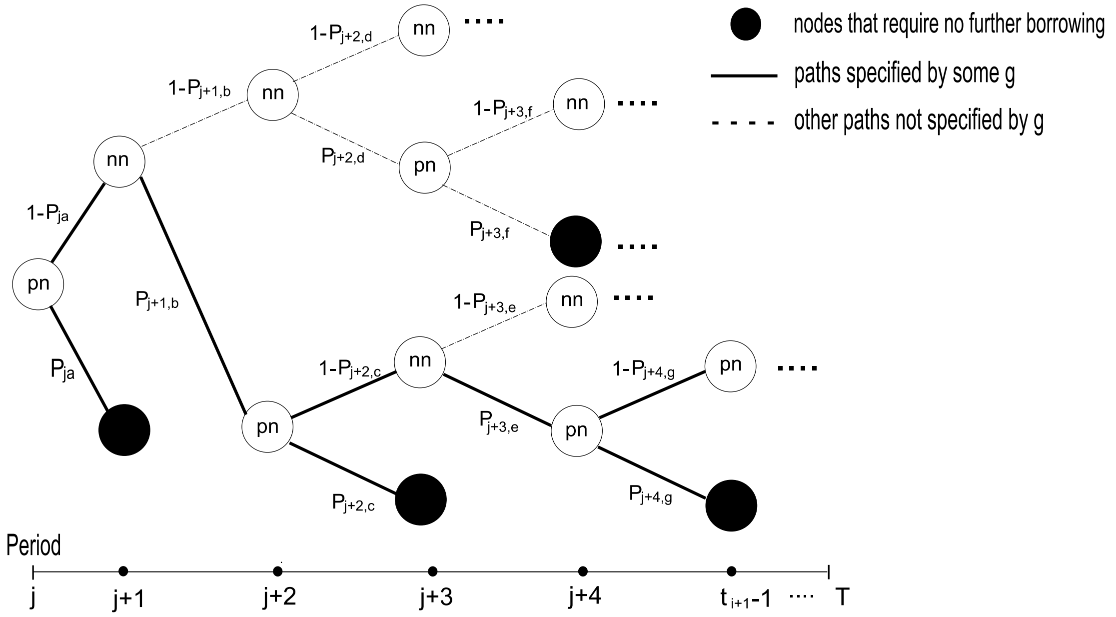

Figure EC.3 An example of a decision subtree: focus on the interval $[t_i, t_{i+1})$ and some $g \in \mathscr{G}$, j is the earliest period in which a problematic node (pn) occurs. According to g, there are two possible outcomes whenever a problematic node (pn) is reached, and there is only one possible outcome whenever a non-problematic node (nn) is reached. If a problematic node (pn) orders, there will not be further borrowing until the next order of OPT in period t_{i+1}.

non-problematic nodes (nn nodes) in the sub-tree have only one outgoing edge that corresponds to the decision (order/no-order) specified by g to that node. The problematic nodes (pn nodes) have two outgoing edges corresponding to the order/no-order decisions, respectively. (Recall that g does not specify the decisions in these nodes.) Moreover, each pn node $s \in [j, t_{i+1})$ is associated with the probability p_s of ordering. (We again abuse the notation introduced before and omit the index e of the node within the layer/period.) An example of a decision subtree specified by some $g \in \mathscr{G}$ is illustrated in Figure EC.3. Any sequence of randomized outcomes corresponding to the decisions in the pn nodes induces a path of evolution of the system. The resulting cumulative borrowing from the bank account A, corresponding to this path, is equal to K times the sum of probabilities associated with the pn nodes in this path. (For each pn node s in the path, the borrowing is equal to $p_s K = z_s$.)

Next we claim that the sub-tree defined above includes at most one pn node in each layer (period).

This follows from the fact that any path between two pn nodes r, s such that $j \le r < s < t_{i+1}$ in the tree includes only no-ordering edges of pn nodes. To see why the latter is true, observe that if an order is placed by the RB policy in a pn node, the resulting inventory position of the RB policy is higher than OPT. Since both policies face the same sequence of demands, the RB policy will not have higher inventory position than OPT at least until the next order placed by OPT. This excludes the existence of pn nodes in subsequent periods until OPT places another order, i.e., beyond period $t_{i+1} - 1$.

In light of the latter observation, we re-number all the pn nodes in the sub-tree as $1, 2, \ldots, M$ (where 1 corresponds to j, specified before). Moreover, it follows that the probability to arrive at node $m = 1, \ldots, M$ and borrow $p_m K$ is equal to $\prod_{s=1}^{m-1}(1 - p_s)$. (This probability corresponds to no-ordering decisions in all the pn nodes prior to m.) The total expected borrowing is then

$$K \cdot \left\{ p_1^2 + \sum_{m=2}^{M} \left\{ \left(\prod_{s=1}^{m-1} (1 - p_s) \right) p_m \left(\sum_{k=1}^{m} p_k \right) \right\} \right\}. \tag{EC.14}$$

Observe that the probability to borrow exactly $K \cdot \sum_{k=1}^{m} p_k$ is equal to $\left(\prod_{s=1}^{m-1}(1 - p_s) \right) p_m$. Moreover, we have already shown that the expression in (EC.14) is bounded above by K (see Lemma 5). This concludes the proof of the lemma. □

LEMMA 5. *Let $\{p_l\}_{l=1}^{\infty}$ satisfy the condition $0 \le p_l \le 1$ for all l. Then the following inequality holds,*

$$p_1^2 + \sum_{l=2}^{\infty} \left\{ \left(\prod_{s=1}^{l-1} (1 - p_s) \right) p_l \left(\sum_{k=1}^{l} p_k \right) \right\} \le 1. \tag{EC.15}$$

Proof of Lemma 5. We construct an increasing sequence $\{a_m\}$ where

$$a_m = p_1^2 + \sum_{l=2}^{m} \left\{ \left(\prod_{s=1}^{l-1} (1 - p_s) \right) p_l \left(\sum_{k=1}^{l} p_k \right) \right\}. \tag{EC.16}$$

For each m, if we replace p_m by 1, we get

$$\bar{a}_m = p_1^2 + \sum_{l=2}^{m-1} \left\{ \left(\prod_{s=1}^{l-1} (1 - p_s) \right) p_l \left(\sum_{k=1}^{l} p_k \right) \right\} + \left(\prod_{s=1}^{m-1} (1 - p_s) \right) \left(1 + \sum_{k=1}^{m-1} p_k \right), \tag{EC.17}$$

such that $a_m \le \bar{a}_m$. Next we will show by induction that $\bar{a}_m \le 1$ for all m from which the proof of the lemma follows. It is straightforward to verify $\bar{a}_1, \bar{a}_2 \leq 1$. Assume that $\bar{a}_m \leq 1$ for some $m \in \mathbb{Z}^+$, we will show that $\bar{a}_{m+1} \leq 1$.

$$\begin{aligned}
\bar{a}_{m+1} &= p_1^2 + \sum_{l=2}^{m} \left\{ \left(\prod_{s=1}^{l-1} (1-p_s) \right) p_l \left(\sum_{k=1}^{l} p_k \right) \right\} + \left(\prod_{s=1}^{m} (1-p_s) \right) \left(1 + \sum_{k=1}^{m} p_k \right) \qquad \text{(EC.18)} \\
&= a_{m-1} + \left(\prod_{s=1}^{m-1} (1-p_s) \right) p_m \left(\sum_{k=1}^{m} p_k \right) + \left(\prod_{s=1}^{m} (1-p_s) \right) \left(1 + \sum_{k=1}^{m} p_k \right) \\
&= a_{m-1} + \left(\prod_{s=1}^{m-1} (1-p_s) \right) \left[\left(1 + \sum_{k=1}^{m} p_k \right) (1-p_m) + p_m \sum_{k=1}^{m} p_k \right] \\
&= a_{m-1} + \left(\prod_{s=1}^{m-1} (1-p_s) \right) \left(1 + \sum_{k=1}^{m-1} p_k \right) = \bar{a}_m \leq 1.
\end{aligned}$$

Hence the claim follows by induction. □

EC.2. Performance of the proposed algorithms

The first two columns specify the test instances, namely, fixed ordering cost K, per-unit holding cost h, per-unit backlogging cost p and demand rate vector λ. The third column shows the cost incurred by the optimal policy. The fourth column shows the optimal parameters of parametrized RB policy. The fifth column shows the cost incurred by the parameterized RB policy. The sixth column shows the cost ratio of the parameterized RB policy to the optimal policy. The seventh column shows the cost of unparameterized RB policy (i.e., the original policy without parameter optimization). The eighth columns shows the cost ratio of the unparameterized RB policy to the optimal policy.

(K,h,p)	Demands $(\lambda_0, \lambda_1, \lambda_2)$	Cost of OPT	Optimal $(\beta^*, \gamma^*, \eta^*)$	Cost of param. RB	Cost Ratio	Cost of unparam. RB	Cost Ratio
(0,1,9)	(4,1,4)	46.85	(*,2,*)	49.18	1.0497	58.30	1.2444
(0,1,9)	(4,1,2)	46.39	(*,2,*)	49.30	1.0627	55.24	1.1908
(0,1,9)	(4,1,1)	46.20	(*,2,*)	47.81	1.0348	54.26	1.1745
(0,1,9)	(3,1,2)	41.02	(*,2,*)	41.41	1.0095	49.40	1.2043
(0,1,9)	(2,1,3)	32.88	(*,2,*)	34.42	1.0468	41.51	1.2625
(0,1,9)	(1,1,4)	24.74	(*,2,*)	26.40	1.0671	31.40	1.2692
(5,1,9)	(4,1,1)	102.66	(0.2,2,9)	108.28	1.0547	135.37	1.3186
(5,1,9)	(1,1,4)	86.47	(0.2,2,9)	90.70	1.0489	128.70	1.4884
(5,1,1)	(4,1,1)	71.35	(0.4,1,1)	75.42	1.0570	84.13	1.1791
(100,1,9)	(5,1,0)	427.81	(0.9,*,9)	451.68	1.0558	605.10	1.4144
(100,1,9)	(4,1,1)	424.81	(0.9,*,9)	449.65	1.0585	601.29	1.4154
(100,1,9)	(3,1,2)	421.76	(0.9,*,9)	443.12	1.0506	595.10	1.4110
(100,1,9)	(2,1,3)	418.63	(0.9,*,9)	443.64	1.0597	611.48	1.4607
(100,1,9)	(1,1,4)	415.49	(0.8,*,9)	437.36	1.0526	618.36	1.4883
(100,1,9)	(0,1,5)	412.29	(0.8,*,9)	435.65	1.0567	593.88	1.4404

Table EC.1 Numerical results with lead time $L = 0$ and finite horizon $T = 12$.

(K,h,p)	Demands $(\lambda_0, \lambda_1, \lambda_2)$	Cost of OPT	Optimal $(\beta^*, \gamma^*, \eta^*)$	Cost of param. RB	Cost Ratio	Cost of unparam. RB	Cost Ratio
(0,1,9)	(4,1,4)	93.81	(*,2,*)	98.32	1.0481	120.14	1.2807
(0,1,9)	(4,1,2)	88.27	(*,2,*)	94.25	1.0677	108.24	1.2262
(0,1,9)	(4,1,1)	85.48	(*,2,*)	90.21	1.0553	93.97	1.0993
(0,1,9)	(3,1,2)	80.04	(*,2,*)	89.73	1.1211	90.40	1.1294
(0,1,9)	(2,1,3)	73.98	(*,1.5,*)	84.42	1.1411	90.99	1.2625
(0,1,9)	(1,1,4)	70.96	(*,1.5,*)	81.40	1.1471	87.60	1.2345
(5,1,9)	(4,1,1)	137.66	(0.2,2,9)	153.97	1.1185	161.10	1.1703
(5,1,9)	(1,1,4)	121.47	(0.2,2,9)	140.26	1.1525	148.47	1.2223
(5,1,1)	(4,1,1)	78.18	(0.4,1,1)	90.42	1.1566	97.47	1.2467
(100,1,9)	(5,1,0)	434.30	(0.9,*,9)	479.03	1.1030	614.17	1.4142
(100,1,9)	(4,1,1)	431.87	(0.9,*,9)	466.33	1.0798	611.96	1.4170
(100,1,9)	(3,1,2)	429.41	(0.9,*,9)	453.24	1.0555	551.00	1.2832
(100,1,9)	(2,1,3)	426.86	(0.9,*,9)	451.17	1.0570	644.13	1.5090
(100,1,9)	(1,1,4)	424.25	(0.9,*,9)	466.43	1.0994	623.56	1.4698
(100,1,9)	(0,1,5)	421.56	(0.9,*,9)	461.65	1.0951	595.40	1.4124

Table EC.2 Numerical results with lead time $L = 2$ and finite horizon $T = 12$.

(K,h,p)	Demands $(\lambda_0, \lambda_1, \lambda_2)$	Cost of OPT	Optimal $(\beta^*, \gamma^*, \eta^*)$	Cost of param. RB	Cost Ratio	Cost of unparam. RB	Cost Ratio
(0,1,9)	(4,1,4)	57.71	(*,2,*)	58.23	1.0090	61.92	1.0730
(0,1,9)	(4,1,2)	57.71	(*,2,*)	58.36	1.0113	60.94	1.0560
(0,1,9)	(4,1,1)	57.71	(*,2,*)	58.30	1.0102	60.38	1.0463
(0,1,9)	(3,1,2)	50.19	(*,2,*)	51.49	1.0259	53.62	1.0683
(0,1,9)	(2,1,3)	41.27	(*,2,*)	41.96	1.0167	43.63	1.0572
(0,1,9)	(1,1,4)	30.55	(*,2,*)	30.88	1.0108	31.66	1.0363
(5,1,9)	(4,1,1)	128.17	(0.2,2,9)	133.91	1.0448	166.10	1.2959
(5,1,9)	(1,1,4)	101.70	(0.2,2,9)	107.34	1.0555	148.85	1.4636
(5,1,1)	(4,1,1)	86.07	(0.4,1,1)	90.51	1.0516	104.24	1.2111
(100,1,9)	(5,1,0)	535.14	(1.1,*,9)	566.23	1.0581	663.61	1.2401
(100,1,9)	(4,1,1)	533.51	(1.1,*,9)	570.65	1.0696	659.29	1.2358
(100,1,9)	(3,1,2)	529.77	(1.1,*,9)	566.09	1.0686	682.76	1.2888
(100,1,9)	(2,1,3)	523.94	(1.1,*,9)	555.57	1.0604	729.15	1.3917
(100,1,9)	(1,1,4)	520.03	(1.0,*,9)	550.36	1.0583	744.45	1.4316
(100,1,9)	(0,1,5)	516.05	(1.0,*,9)	550.65	1.0670	711.22	1.3782

Table EC.3 Numerical results with lead time $L = 0$ and finite horizon $T = 15$.