Coordinated admission and inventory controls in a make-to-stock production system

Stratos Ioannidis[1] and Vassilis S. Kouikoglou[2]

Abstract— **We consider a production facility that produces a single product to meet random demand. The system produces until stock reaches a certain level and accepts orders until backlog reaches another level. The problem is to specify the critical stock and backlog levels that minimize the sum of customer rejection, backlog, and inventory carrying costs. The system is modeled as a finite queueing system whose mean cost rate is convex in one control parameter and, under certain conditions, unimodal in the other. A simple algorithm is proposed to find the globally optimum design. Numerical results show that the joint admission/inventory control policy achieves higher profit than some commonly used production control policies.**

***Index Terms*—Production/inventory control, make-to-stock production systems, partly lost sales, global optimization.**

I. Introduction

In production control, questions of the type "when to start or stop producing" and "whether to accept or reject an incoming order" are directly related to net profit, that is, revenue from sales less purchase, production, inventory carrying, and backlog costs. In this paper, we consider *make-to-stock* production systems, which consist of a production facility, a buffer where finished items are stored, and a demand process. A common inventory control policy for make-to-stock systems is one that specifies a target value for the number of finished items. This value is called the *base stock*. When the buffer level reaches the base stock, the production facility is switched off. This policy ensures that the inventory cost is bounded. When the base stock is zero, the corresponding policy is called *zero base stock* or *make-to-order* policy.

Demand during the stockout period is usually either tightly controlled or uncontrolled, that is, customer orders are unconditionally either rejected or accepted (see, e.g., [1]-[3]). A strategy of accepted orders is known as the *complete backordering policy* (CB) whereas rejection of orders corresponds to a *lost sales policy* (LS). When the production rate is less than the demand rate, CB cannot be profitable since the number of outstanding orders would grow without bound. An alternative admission control policy is to always accept customer orders when stock is available and to accept or reject orders in a random manner (e.g. by performing a Bernoulli experiment) during stockouts independent of the current backlog [4]. We then have a *randomized admission control policy* (RAC).

CB and LS are opposite practices. In CB there is no bound on the number of outstanding orders whereas under LS backlog is always zero. An intermediate admission policy is one that rejects orders when backlog reaches a certain limit and accepts them otherwise. The overall control policy of the system is completely specified by two nonnegative integers, a base stock and a base backlog. This is the *partly lost sales policy* (PLS), which generalizes and outperforms CB and LS for systems with exponentially distributed processing and interarrival times ([5], [6]).

In this paper, we provide a rigorous treatment of PLS and examine more general distributions. We prove that the mean cost rate of the system (customer rejection cost plus backlog and inventory carrying costs) is unimodal, though not always convex, in both the base stock and the base backlog. The detailed proof of unimodality is given in [7]. The numerical results show that the proposed policy achieves higher profit than LS, CB, and RAC. This work is a start towards developing and treating rigorously similar strategies for complex production networks.

II. Queueing Model of a Make-to-Stock Production System with Partly Lost Sales

Consider a production facility that produces a single product. Customers arrive at random times and each customer requests one unit of product. The times between successive customer arrivals are independent random variables with mean $1/\lambda$. Processing times are also independent random variables with mean $1/\mu$. Finished items are stored in an output buffer. An arriving order that finds the buffer empty is either backlogged or rejected; otherwise, it is satisfied immediately from the inventory.

The operation of the system is associated with three types of cost:

- P unit rejection cost, which includes the net revenue (selling price less cost of purchasing raw parts and processing) per item and a penalty (if any) per customer rejected,
- h unit holding cost rate, which is the cost per unit time per finished item held in the buffer,
- b unit backlog cost rate, which is the cost per unit time of delay for a pending order.

A simple policy is used to control the number of finished items and pending orders. The facility stops production when

This work was supported in part under grants from the Ministry of Education, Athens, Greece.

The authors are with the Department of Production Engineering and Management,Technical University of Crete, University Campus, GR-73100 Chania, Greece.

[1]E-mail: dervish@dpem.tuc.gr.

[2]Corresponding author; Tel: +30 28210-37238; fax: -69410; E-mail: kouik@dpem.tuc.gr.

the output buffer contains s items and the system rejects an incoming order if the current number of backorders is c. Thus, s is the base stock and c is the base backlog. The problem then is to find s and c that minimize the mean cost rate of the system, which is the sum of the mean rejection, holding, and backlog cost rates. We derive expressions for these costs by analyzing an equivalent queueing system.

The state of the system is described by an integer n, $0 \le n \le s + c$. When $n \le s$ there are no pending orders and the output buffer contains $s - n$ items. When $n \ge s$ the output buffer is empty and there are $n - s$ orders to be satisfied. The system is modeled as a $G/G/1/m$ queue with capacity $m = s + c$. The arrival and service completion times equal the customer arrival and production completion times respectively. When $n = 0$ the queueing system is empty and its server idle. Correspondingly, in the original system, the output buffer is full, the production facility is stopped, and no pending orders are present. In a dual fashion, when $n = m$ the queueing system is full and all incoming arrivals are lost, whereas, in the original system, the output buffer is empty, $c = m - s$ orders are pending and all incoming orders are rejected.

Let p_n be the stationary probability that the system is in state n, provided a steady state exists. Since exact analytical models for general $G/G/1/m$ queues do not exist, we assume that the stationary probabilities are geometric except for certain boundary states. Specifically, we assume that

$$p_n = \begin{cases} K_m\alpha, \ n = 0 \\ K_m\sigma\beta, \ n = 1 \\ K_m\sigma^n, \ 2 \le n \le m-2 \\ K_m\sigma^{m-1}\gamma, \ n = m-1 \\ K_m\sigma^m\delta, \ n = m \end{cases} \tag{1}$$

for $m \ge 4$. In the above expressions, the stationary probabilities of the internal states n, $2 \le n \le m - 2$, form a geometric progression with parameter σ. The parameters α, β, γ, and δ are associated with the four boundary probabilities p_n, $n = 0, 1, m - 1, m$. We assume that all parameters depend on the statistics of the interarrival and service times but are independent of s, c, and m. Finally, K_m is a constant arising from the normalizing equation $\sum_0^m p_n = 1$. For brevity, we assume that $\sigma \ne 1$. Then

$$K_m = \frac{1}{\alpha + \sigma\beta + \sigma^2 + \ldots + \sigma^{m-3} + \sigma^{m-1}\gamma + \sigma^m\delta} = \frac{1}{A + D\sigma^m}$$

where

$$A = \alpha + \sigma\beta + \frac{\sigma^2}{1-\sigma} \text{ and } D = \delta + \frac{\gamma}{\sigma} + \frac{1}{\sigma(\sigma-1)} \tag{2}$$

Equation (1) may not hold for values of m less than 4, since then there are no internal states. These cases can be analyzed separately using exact or approximate Markovian models with a small number of states which are computationally tractable.

There are a number of models whose stationary probabilities have the form of Eq. (1).

(a) Setting $\alpha = \beta = \gamma = \delta = 1$ and $\sigma = \lambda/\mu$ in Eq. (1) yields the stationary probabilities of the $M/M/1/m$ queue.

(b) For the approximation of $G/G/1/m$ queues proposed in [8], the stationary probabilities are given by

$$p_n = \begin{cases} \dfrac{1-\rho}{1-\rho\sigma^m}, n = 0 \\ \dfrac{\rho(1-\sigma)\sigma^{n-1}}{1-\rho\sigma^m}, n = 1, \ldots, m \end{cases} \tag{3}$$

where $\rho = \lambda/\mu$ and

$$\sigma = \begin{cases} \dfrac{N-\rho}{N}, \rho < 1 \\ \dfrac{N'}{N'-\rho}, \rho > 1 \end{cases}$$

N is an approximation of the mean number of customers in a $G/G/1$ queue for $\lambda < \mu$ and N' is the mean number of customers in the reversed $G/G/1$ queue with arrival rate μ and service rate λ when $\lambda > \mu$. For $\lambda = \mu$ the limiting result of Eq. (3) is derived as $\rho \to 1$. Setting $\alpha = \sigma(1-\rho)/[\rho(1-\sigma)]$ and $\beta = \gamma = \delta = 1$ in Eq. (1) yields Eq. (3).

(c) As a final example, consider a discrete-time model, presented in [8], of a two-machine transfer line with constant processing times and geometrically distributed times between failures and times-to-repair. In that model, all events of interest (production, failures, repairs) occur at times 1, 2, … In our setting, the second machine represents the production facility and the first machine models a demand pattern that consists of consecutive times during which customers place requests for one item per time unit followed by consecutive times during which no demand occurs. When $n = m$ the first machine is blocked, that is, all incoming orders are rejected but the demand pattern is not affected. Thus, for the first machine of this equivalent system, failures are time-dependent. However, when $n = 0$ the second machine is stopped and it cannot fail. Thus, the failures of second machine are operation-dependent. The stationary probabilities are derived in Section 6.6.1 of [8] and have the form of Eq. (1).

III. MINIMIZATION OF THE MEAN COST RATE

In this section, we use the stationary probabilities from Eq. (1) to express the mean cost rate as a function of the base stock s and the system capacity m, where $m = s + c$. Then, we derive expressions for the optimal base stock for any fixed m. We also show that, under a rather mild condition, the objective function is unimodal in m. Although from numerical results it appears that this function may have a nonconvex, staircase shape, unimodality ensures that the global minimum can be tracked down by careful line search. This gives rise to an efficient algorithm for computing optimal values of s and m.

Since in state $n = 0$ the production facility is stopped, the proportion of time the facility is busy, on the average, is $1 - p_0$. A busy period is defined as the interval between a startup and a stoppage of the production facility. Whenever the level of the output buffer drops from s to $s-1$ the facility is switched on and a new production cycle begins. When a part is completed and the buffer level reaches the base stock, the facility is stopped. From these observations, we see that each busy period of the facility contains an integer number of complete production cycles. These cycles are independent random variables with mean $1/\mu$. Therefore, the mean production rate of the system is $(1-p_0)\mu$. Since, in steady state, this quantity must equal the mean arrival rate of accepted orders, the mean rejection rate is $\lambda-(1-p_0)\mu$ and the mean rejection cost rate is $P(\lambda-\mu+p_0\mu)$. Finally, the mean cost rate of the system is $P(\lambda-\mu)+Pp_0\mu+hH+bB$, where H is the mean number of finished items in the buffer and B is the mean number of pending orders.

Since P, λ, and μ are constant, minimizing the mean cost rate is equivalent to minimizing the objective function

$$J(s, m) = Pp_0\mu + hH + bB$$

for $m = 0, 1, \ldots$ and $s = 0, 1, \ldots, m$. From Eq. (2) we obtain

$$H = \sum_{n=0}^{s-1}(s-n)p_n = K_m[s\alpha + (s-1)\sigma\beta + (s-2)\sigma^2 + \ldots + 1\sigma^{s-1}]$$

$$B = \sum_{n=s+1}^{m}(n-s)p_n$$
$$= K_m[1\sigma^{s+1} + \ldots + (m-s-2)\sigma^{m-2} + (m-s-1)\sigma^{m-1}\gamma + (m-s)\sigma^m\delta]$$

and the objective function is written

$$J(s, m) = K_m\big[P\mu\alpha + h[1\sigma^{s-1} + \ldots + (s-2)\sigma^2 + (s-1)\sigma\beta + s\alpha]$$
$$+ b[1\sigma^{s+1} + \ldots + (m-s-2)\sigma^{m-2} + (m-s-1)\sigma^{m-1}\gamma + (m-s)\sigma^m\delta]\big]$$

Since the problem is two-dimensional, we solve it sequentially. First, we minimize $J(s, m)$ with respect to s for any fixed m and then we track down the optimal value of m.

For some fixed m, we seek a value s_m for s such that the following inequalities hold simultaneously:

$$J(s_m, m) \le J(s', m), \text{ for every } s' \text{ such that } 0 \le s' < s_m$$
$$J(s_m, m) < J(s', m), \text{ for every } s' \text{ such that } s' > s_m$$

Due to the special form of the boundary probabilities involving parameters α, β, γ, and δ, the expression of $J(s_m, m)$ for $s_m = 0, 1, m-1$, or m differs from that for $2 \le s_m \le m-2$. The following theorem provides closed-form expressions for s_m.

Theorem 1. For any fixed m, $m \ge 4$, $J(s, m)$ assumes its minimum value at the point s_m, which is given by

$$s_m = \begin{cases} 0, & \text{if } \frac{b}{h+b} < K_m\alpha \\ 1, & \text{if } K_m\alpha \le \frac{b}{h+b} < K_m(\alpha+\sigma\beta) \\ \left\lfloor \dfrac{\ln\left[\frac{(bD\sigma^m - hA)(\sigma-1)}{h+b}\right]}{\ln\sigma} \right\rfloor, & \text{if } K_m(\alpha+\sigma\beta) \le \frac{b}{h+b} < K_m\left(A - \frac{\sigma^{m-1}}{1-\sigma}\right) \\ m-1, & \text{if } K_m\left(A - \frac{\sigma^{m-1}}{1-\sigma}\right) \le \frac{b}{h+b} < K_m\left[A + \sigma^{m-1}\left(\gamma - \frac{1}{1-\sigma}\right)\right] \\ m, & \text{if } K_m\left[A + \sigma^{m-1}\left(\gamma - \frac{1}{1-\sigma}\right)\right] \le \frac{b}{h+b} \end{cases} \quad (4)$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x.

The proof of the above theorem is given in [7].

Having found the optimal value s_m for each m we minimize $J(s_m, m)$ with respect to m, where $m \in \{0, 1, \ldots\}$. Clearly, when $m = 0$ the system does not operate and, therefore, $s_0 = 0$ and $p_0 = 1$. In that case, the objective function becomes $J(0, 0) = P\mu$ and the mean cost rate of the system equals $P(\lambda-\mu) + J(0, 0) = P\lambda$, that is, the cost rate of rejecting all customers. For $m = 1, 2$, or 3 the optimal value s_m and the corresponding value of the objective function can be found by exhaustive search since the number of candidate (s, m) pairs is small ($s \le m$). It then remains to minimize $J(s_m, m)$ on the set $M = \{4, 5, \ldots\}$. If the minimum is less than $P\mu$ then the system is profitable, otherwise it pays more not to operate it.

First, we divide the set M into five subsets given by

$$M_0 = \{m \in M: s_m = 0\}$$
$$M_1 = \{m \in M: s_m = 1\}$$
$$M_2 = \{m \in M: 2 \le s_m \le m-2\}$$
$$M_3 = \{m \in M: s_m = m-1\}$$
$$M_4 = \{m \in M: s_m = m\}$$

The problem then becomes one of finding

$$\min_{m \in M,\, s \le m} J(s, m) = \min_{i=0,\ldots,4}\left[\min_{m \in M_i} J(s_m, m)\right]$$

Second, we show that $J(s_m, m)$ is unimodal on each set M_i. To prove unimodality, we assume that condition

$$A(1-\sigma) \ge 0 \text{ and } D(\sigma-1) \ge 0 \quad (5)$$

holds, where A and D are given by equations (2). Although the above condition may not hold for every $G/G/1/m$ system, we have verified its validity for the three models examined in Section 2.

The following lemma establishes certain properties of the sets M_i and monotonicity of the optimal base stock s_m.

Lemma 1. If condition (5) holds, then
(a) the sets M_i are disjoint and convex; specifically

$$M_i = \{m_{i-1}, m_{i-1}+1, \ldots m_i - 1\}$$

where m_i are the extreme solutions of the inequalities in Eq. (4) and are given by

$$m_i = \begin{cases} \min\left\{\left\lfloor \dfrac{\ln A_i}{\ln\sigma} \right\rfloor, m_{i-1}\right\}, & \text{if } A_i > 0 \\ \infty, & \text{if } A_i \le 0 \end{cases}, \quad i = 0, \ldots, 3$$

$$m_{-1} = 4,\ m_4 = \infty$$

and

$$A_0 = \frac{(h+b)\alpha - bA}{bD} \quad A_1 = \frac{(h+b)(\alpha+\sigma\beta) - bA}{bD}$$

$$A_2 = \frac{hA}{bD - \frac{h+b}{\sigma(\sigma-1)}} \quad A_3 = \frac{hA}{bD - \frac{h+b}{\sigma(\sigma-1)} - \frac{h+b}{\sigma}\gamma}$$

if $m_i \le m_{i-1}$, then we have $M_i = \varnothing$;

(b) if m and $m+1$ belong to M_2, then $s_m \le s_{m+1} \le s_m + 1$.

The proof of the above is given in [7]. Part (b) of Lemma 1 offers an alternative approach to finding the optimal base stock for each m. From the definition of M_i, for $m \in M_i$, $i = 0, 1, 3$, or 4, the optimal base stock is $s_m = 0, 1, m-1$, or m, respectively. For the set M_2, if s_{m-1} is the optimal base stock when the capacity of the system is $m-1$, then s_m is obtained by comparing the values of the objective function at $s = s_{m-1}$ and $s = s_{m-1} + 1$. This reduces the search effort for s_m.

In [7] we use Lemma 1 to prove the next result, which ensures that an appropriate local minimizer m of each subproblem is also a global minimizer in the corresponding subset M_i.

Theorem 2. If condition (5) holds, then the function $J(s_m, m)$ is unimodal on each subset M_i, $i = 0, \ldots, 4$. Specifically, suppose a point $m \in M_i$ exists such that

$$J(s_m, m) \le J(s_k, k) \text{ for every } k < m,\ k \in M_i,$$
$$\text{and } J(s_m, m) \le J(s_{m+1}, m+1).$$

Then,

$$J(s_{m+1}, m+1) \le J(s_k, k)$$

for every $k > m+1$, $k \in M_i$.

This theorem has an important algorithmic consequence. The optimal solution (s^*, m^*) to the original problem can be tracked down as follows.

Step 1. Compute the extreme points of all sets $M_i = \{m_{i-1}, \ldots, m_i - 1\}$. Set $i = 0$, $J^* = \infty$ and go to Step 4.

Step 2. If $J_i^* < J^*$, then update the globally optimal values: $J^* = J_i^*$, $s^* = s_i^*$, and $m^* = m_i^*$.

Step 3. Set $i = i + 1$.

Step 4. If $M_i = \varnothing$ go to Step 3; otherwise set $m = m_{i-1}$.

Step 5. Compute s_m from Eq. (4) and $J(s_m, m)$. If $m > m_{i-1}$ and $J(s_m, m) \ge J_i^*$, then abort set M_i and go to Step 2; $m - 1$ is the optimal base stock for M_i. If $m = m_{i-1}$ or $J(s_m, m) < J_i^*$, then initialize or update, respectively, the globally optimal parameters of set M_i: $J_i^* = J(s_m, m)$, $m_i^* = m$, and $s_i^* = s_m$.

Step 6. Set $m = m + 1$. If $m \le m_i - 1$ go to Step 5; otherwise go to Step 2.

IV. Numerical Results

We compare the proposed partly lost sales policy (PLS) with the following policies: complete backordering (CB), lost sales (LS), and randomized admission control (RAC).

Let $(s, c)_{\mathrm{PLC}}$ be a partly lost sales policy with a base stock s and a base backlog c. It turns out that CB is the same as $(s, \infty)_{\mathrm{PLC}}$ and LS is $(s, 0)_{\mathrm{PLC}}$. Another commonly used control policy is the zero base stock policy, denoted $(0, c)_{\mathrm{PLC}}$. The cost rate under such policy equals the cost rate incurred by the reverse system operating under a lost sales policy, $(c, 0)_{\mathrm{PLC}}$, and so it suffices to study the performance of LS.

We consider two production systems, an $M/M/1/m$ queue and a $G/G/1/m$ queue. The standard parameters for both systems are $\lambda = 9.5$, $\mu = 10$, $h = b = 0.5$, and $P = 10$.

In the second system, the coefficients of variation of the interarrival times and the service times are both equal to 0.5. Using these values and an approximate method proposed by [8] we can compute the parameter σ and, by Eq. (3), the stationary probabilities for this system.

This method cannot be applied directly to systems operating under RAC. Thus, RAC was tested only for the $M/M/1/m$ case, for which the mean cost rate has a closed form, by solving a simple Markov model. This model depends on s and the rejection probability during the stockout period. Optimal values for these parameters are computed by exhaustive search.

We investigate the effects of varying $\rho = \lambda/\mu$, b, and h on the mean production cost rate. Figures 1 through 3 and Table I show the results for the $M/M/1/m$ system and Figs. 4 through 6 show the results for the $G/G/1/m$ system.

TABLE I
Control Parameters and Performance of Production Control Strategies for Various Values of ρ ($M/M/1/m$ system)

	ρ	0.875	0.925	0.975	1.025	1.075	1.125
PLS	*s*	5	7	11	16	23	30
	c	30	22	16	11	8	5
	Cost	2.56	3.87	5.85	8.44	11.61	15.38
RAC	*s*	5	8	13	18	24	31
	rejection pr.	0	0.01	0.08	0.16	0.24	0.34
	Cost	2.59	4.43	6.93	9.45	12.38	15.93
LS	*s*	13	15	17	21	26	32
	Cost	6.56	7.58	8.92	10.72	13.17	16.42
CB	*s*	5	8	27	-	-	-
	Cost	2.59	4.44	13.69	∞	∞	∞

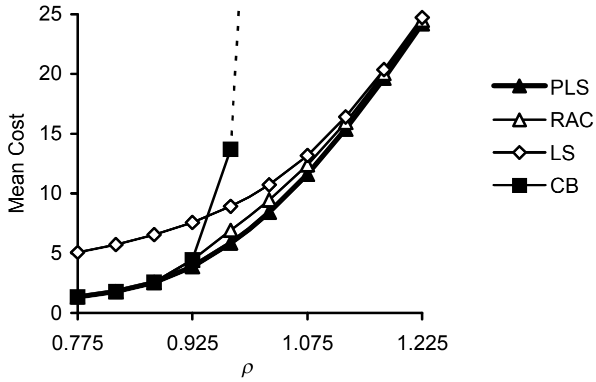

Fig. 1. Mean cost rate for an $M/M/1/m$ system versus ρ.

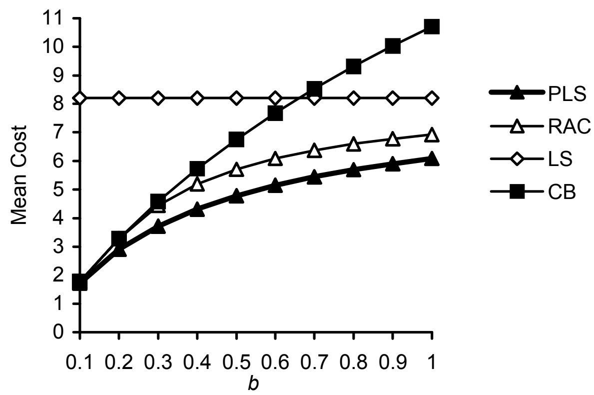

Fig. 2. Mean cost rate for an $M/M/1/m$ system versus b.

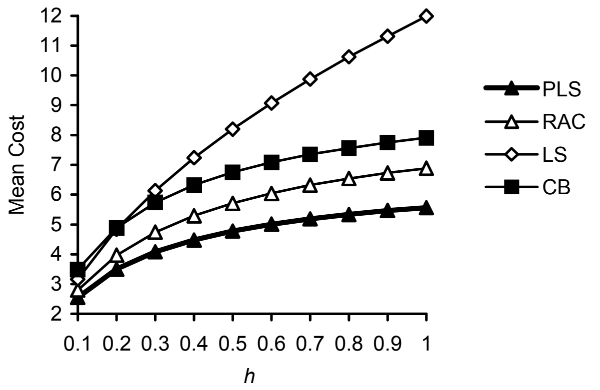

Fig. 3. Mean cost rate for an $M/M/1/m$ system versus h.

The numerical results suggest that PLS incurs the smallest cost rate. For LS and CB, this result agrees with intuition since LS and CB are special cases of PLS. When ρ is close to 1, the costs of LS and CB are higher than that of PLS by 50%. When ρ is less than 1 (this includes the standard case $\rho = 0.95$), LS degrades significantly with increasing h and CB degrades with increasing b. For $\rho \to 0$, PLS achieves the same cost rate as CB. In this case, the mean backlog is very small and any effort to limit it would not achieve a substantial cost reduction. When ρ is much greater than one the cost rate achieved by PLS is approximately equal to the cost rate achieved by LS. In this case, customers arrive frequently and, therefore, the rejection cost rate is negligible compared to the backlog cost rate. Hence, the optimal base backlog for PLS tends to zero. In general, as the contribution of the backlog cost to the overall cost increases, the cost rate of LS converges to that of PLS and when b tends to zero PLS and CB perform alike.

Finally, from Figs. 1-3, we see that PLS and RAC have the same cost rates when $b \to 0$, $\rho \to 0$, or $\rho >> 1$. In the remaining cases PLS provides a cost reduction of 10-20% over RAC.

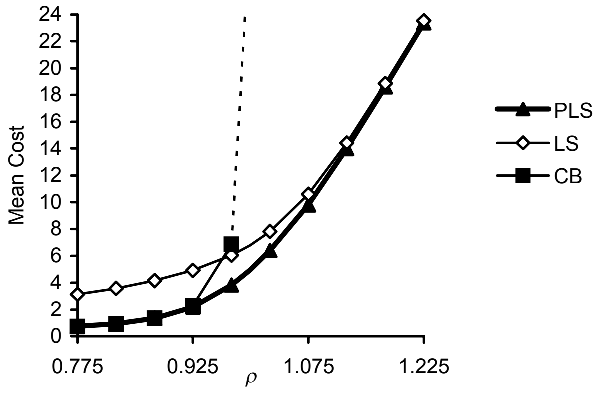

Fig. 4. Mean cost rate for a $G/G/1/m$ system versus ρ.

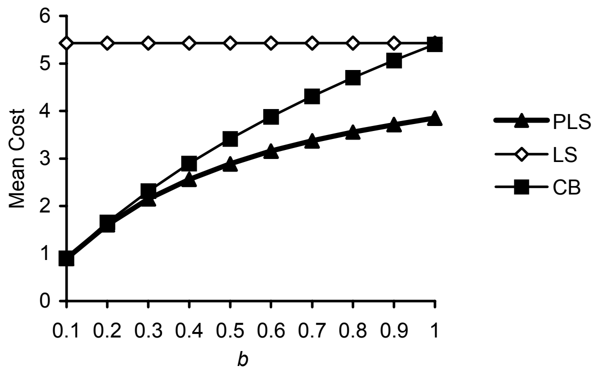

Fig. 5. Mean cost rate for a $G/G/1/m$ system versus b.

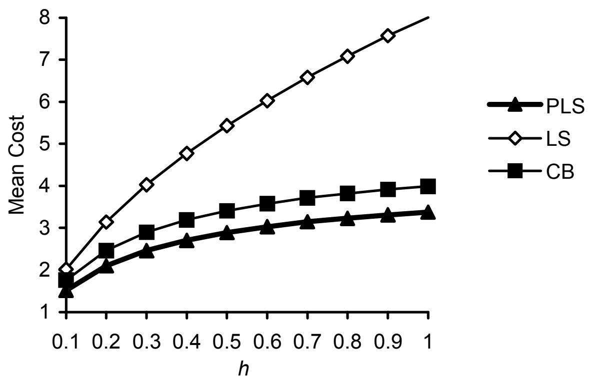

Fig. 6. Mean cost rate for a $G/G/1/m$ system versus h.

V. Concluding Remarks

The idea of coordinating admission and inventory decisions could be applied to production systems with setup times, several demand classes, and complex production networks. For systems that incur a setup cost or setup delay during a startup of the production facility, a practical control policy would be to combine an (s, S) inventory policy ([9], [10]) and the proposed customer admission strategy. In an (s, S) inventory policy, the production facility starts when the level of the output buffer drops to s and stops when it reaches S (the base stock policy is an $(s-1, s)$ inventory policy). Combining this policy and PLS requires three control parameters s, S, and c, and more elaborate tools to establish second-order properties of the objective function. Ha ([11], [12]) has studied the problem of stock rationing under CB or LS when there are several demand classes with different profit or backlog cost structures. Each demand class is associated with a critical stock level above which we accept an incoming order under LS or satisfy a pending order from that class under CB. These two cases could be combined or somehow extended to a PLS policy. Future research will focus on such problems. Current research is underway on proving unimodality properties of PLS in flexible manufacturing systems, in which the production control policy is simple and the stationary probabilities have closed-form expressions.

References

[1] G. Hadley, and T.M. Whitin, *Analysis of Inventory Systems*, Prentice Hall, Englewood Cliffs, NJ, 1963.

[2] S.A. Smith, "Optimal inventories for an ($S-1$, S) system with no backorders", *Management Science*, vol. 239, pp. 522-528, 1977.

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

[4] K. Moinzadeh, "Operating characteristics of the ($S-1$, S) inventory system with partial backorders and constant resupply times", *Management Science*, vol. 35, pp. 472-477, 1989.

[5] R.A. Caldentey, "Analyzing the make-to stock queue in the supply chain and e-business settings", Ph.D. dissertation, Sloan School of Management, MIT, Massachusetts, 2001.

[6] V.S. Kouikoglou, and Y.A. Phillis, "Design of product specifications and control policies in a single-stage production system", *IIE Transactions*, vol. 34, pp. 590-600, 2002.

[7] S. Ioannidis, and V. S. Kouikoglou, "Coordinated admission and inventory controls in a make-to-stock production system", working paper.

[8] J.A. Buzacott, and J.G. Shanthikumar, *Stochastic Models of Manufacturing Systems*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[9] D.L. Iglehart, "Optimality of (s, S) policies in the infinite horizon inventory problem", *Management Science*, vol. 9, pp. 259-267, 1963.

[10] S.P. Sethi, and F. Cheng, "Optimality of (s, S) policies in inventory models with markovian demand", *Operations Research*, vol. 45, pp. 931-939, 1997.

[11] A.Y. Ha, "Inventory rationing in a make-to-stock production system with several demand classes and lost sales", *Management Science*, vol. 43, pp. 1093-1103, 1997.

[12] A.Y. Ha, "Stock-rationing policy for a make-to-stock production system with two priority classes and backordering", *Naval Research Logistics*, vol. 44, pp. 457-472, 1997.