Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

An Optimal Production and Shutdown Strategy when a Supplier Offers an Incentive Program

Xiuli Chao, Frank Y. Chen,

M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 7, No. 2, Spring 2005, pp. 130–143
ISSN 1523-4614 | EISSN 1526-5498 | 05 | 0702 | 0130

informs®
DOI 10.1287/msom.1040.0066

An Optimal Production and Shutdown Strategy when a Supplier Offers an Incentive Program

Xiuli Chao
Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina 27695-7906, xchao@unity.ncsu.edu

Frank Y. Chen
Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Satin, Hong Kong, yhchen@se.cuhk.edu.hk.

Motivated by the incentive programs that have been offered by energy companies under tight market conditions in the past few years, we consider a production control problem in which time alternates randomly between peak and nonpeak periods. During peak periods, the energy supplier offers the manufacturing firm—the energy user—an incentive program to reduce its energy usage by shutting down its production facility. Participation in the incentive program, however, is totally voluntary, and the user firm is rewarded for each unit of time that it participates in the program. We consider two problems that face the manufacturing firm. The first is whether it is worth shutting down production to participate in the incentive program when it is offered, and in which part (portion) of the peak period the firm should participate. The second problem is how the firm should decide on its production policy in both the peak and nonpeak periods in the presence of such an incentive program. In this paper, we provide simple models to give insight into the nature of these problems. Two cases are studied. In the first case, the peak duration is assumed to be exponentially distributed, and in the second, its length becomes known at the beginning of a peak period. In both cases, the occurrence times of the peak periods are uncertain. We characterize the optimal production and shutdown policy for both the peak and nonpeak periods. We also study the effect of seasonality on the optimal control policies.

Key words: continuous-review inventory systems; incentive program; optimal production strategy; Markov decision processes
History: Received: March 12, 2003; accepted: December 21, 2004. This paper was with the authors 12 months for 3 revisions.

1. Introduction

This work is motivated by the "demand response (DR) programs" that have been offered in the U.S. energy market in the past few years. Under such programs, energy users, such as manufacturers, are offered certain incentives to curtail energy consumption and reduce loads during peak periods in response to system reliability and/or market conditions. Take the electric wholesale markets in the United States as an example. Because of unanticipated increase in demand that is caused by weather or supply disruptions due to natural disasters, it is in the energy supplier's interest during peak periods to induce users to consume less electricity to avoid the occurrence of disastrous events, such as blackouts. Consequently, it is not surprising that, as of December 2002, more than 45 states in the United States had implemented some form of DR program (U.S. Department of Energy 2002). Although these incentive programs emerged under different names, and their implementation may differ across utility firms or governments, they have several features in common. First, because of contract terms, the resource supplier is obliged to maintain a steady supply of the resource to user firms, which implies that participating in any incentive program must be totally voluntary, i.e., users (e.g., manufacturers) can "refuse" to take part in the program in any given peak period. Second, the actual amount of reward that a user receives depends on the length of time that the user participates in the incentive program.

Consider the case of Wisconsin Electric. It has been implementing DR programs—known as energy *buy-back programs*—for several years. One such program, *Dollars for Power* (DFP), is described on the company's webpage as

> a way for Wisconsin Electric to pay participating customers a market-based premium for voluntarily reducing their energy use.... Customers are compensated by Wisconsin Electric at pre-established prices for the portion of the electric load they reduce during periods of program activation.

Similar electricity buy-back programs have been adopted elsewhere (see Wald 2000). For instance, Commonwealth Edison in Illinois and Puget Sound Energy in Washington have offered such programs to their commercial and industrial customers with great success.

Several interesting issues emerge from this situation. The first is whether, from a manufacturing firm's viewpoint, it is worth shutting down a production system to participate in the program at a particular time during a peak period (when the incentive program is activated), and in which part (portion) of the peak period the firm should participate. The second issue is how the firm should dynamically decide on its production policy over time in the presence of such an incentive program. In this paper, we provide simple models to give insight into the nature of these problems.

We address these problems by considering a discrete production system. Products of a single type are manufactured in a fixed batch (or lot) size and stored in inventory, and demand for them arrives according to a compound Poisson process. Stockouts are backlogged. The firm incurs holding costs for the amount of inventory it holds in stock, and shortage costs for the outstanding backorders on the books. Production takes one key resource as the input, such as electricity. The resource is perishable, and cannot be stored for future use, but without it the production system has to be turned off. The firm faces random arrivals of an event that can bring its production system into a special state for a random amount of time, called a *peak period*. Therefore, the time alternates between peak and nonpeak periods. During peak periods, the firm is given an option to participate in an incentive program, which compensates it for not using the resource by means of switching off its production. This means that it will receive a certain amount of financial reward for each unit of time that the system stays idle, whereas continuing production forfeits the reward. The firm has to decide whether to shut down or to produce at any given point in time during both the peak and nonpeak periods. The objective is to find an optimal dynamic production and shutdown policy over time to minimize the total discounted cost.

We consider two cases in this paper. In the first case, we assume that the durations of peak periods are exponentially distributed, and that the firm has no definite information about the time remaining in the current peak period. In the second case, we assume that when a peak period arrives, its duration is announced by the resource supplier, and therefore the firm knows when the current peak period will expire. However, the exact duration is not known in advance by the firm and can vary from period to period. We allow an arbitrary distribution for the peak duration. (Note that in the context of energy buy-back programs, the firm is normally informed of the peak duration at the start of each program activation.) Both settings are formulated as Markov decision processes. We show that the optimal policy is essentially determined by some critical numbers. In the first case of exponential peak-period duration, the production and shutdown strategy is determined by a single threshold level, whereas in the second case the strategy is determined by a sequence of threshold levels depending on the time that remains before the current peak period ends. We also explore how changes in system parameters, in particular the incentive offered by the supplier, affect the optimal strategy and optimal cost.

Our models are related to the literature on inventory and production control, and on control of queues. There is a rich literature on inventory management that characterizes the structure of optimal control policies (see Zipkin 1999). Studies of inventory systems that use Markov decision processes are voluminous, for example, those of Ha (1997), Li (1992), Duenyas and Tsai (2000), and Carr and Duenyas (2000). Production systems have been studied with both discrete and continuous state spaces and from both the queuing and optimal control perspectives. For surveys on the control of queues, the reader is referred to Stidham and Weber (1989) and Kitaev and Rykov (1995), and for a comprehensive treatment based on optimal control theory, to Gershwin (1994) and Bensoussan et al. (1983). One closely related area of study is the control of failure-prone manufacturing systems (see the work of Akella and Kumar 1986, Hu et al. 1994, and the more recent paper by Feng and Yan 2001). However, in a failure-prone manufacturing model, production control is possible only when the machine is in a working state, and when the machine is down, there is no decision to be made. Other papers that are related to ours are Moinzadeh (1997), Zheng (1994), and Feng and Sun (2001). These studies address the inventory control problems with random discount opportunities. In addition to regular ordering costs, the firm faces occasional opportunities that allow it to order at a discount. They assume that any amount of inventory can be replenished all at once, that discounts arise in the fixed cost and/or variable ordering cost, and that the discount opportunities appear according to a Poisson process. However, upon the arrival of a discount opportunity, the firm has to decide immediately whether to order or not, because otherwise the opportunity vanishes instantly. Furthermore, discounts are transaction based and item based. Therefore, our models are significantly different from those of the above-named authors.

This paper extends the literature on the optimal control of make-to-stock inventory systems with finite production capacity. In the queueing context, the earliest work on establishing the optimality of the threshold type of service policy appears to be that of Sobel (1969) (see also Sobel 1982, Stidham and Weber 1989, Kitaev and Rykov 1995 and the references therein). For the optimality of base-stock policies of make-to-stock queues in an inventory context, see Li (1988, 1992). Our results extend the optimality results of make-to-stock inventory systems with a finite capacity to batch Poisson arrivals and batch production processes, and alternating low/high production costs.

The rest of this paper is organized as follows. In the next section we study the problem under the assumption of exponential peak durations, which corresponds to the case in which the firm has no knowledge about how long the peak period will last. In §3 we extend the results to the case with more general peak duration distributions, in which information about the remaining peak time can be utilized to determine the policy. In §4, we further consider the model with seasonality. We conclude the paper with a discussion in §5. Technical proofs are given in the appendix. Throughout the paper, we use the terms *increasing* and *decreasing* in their weak senses, i.e., nondecreasing and nonincreasing, respectively.

2. Exponential Incentive Duration

To facilitate the exposition, we first formulate the model with reference to the buy-back program as mentioned in the introduction. Consider a system that produces one type of product. Demand arrives according to a batch Poisson process with rate λ and batch-size distribution

$$P\{Z=i\}=p_i, \quad i=1,2,\dots.$$

Products are made to stock, and the demand for them is satisfied from the inventory. When the on-hand inventory is insufficient, we assume that unmet demand is backordered. Production is in a batch of K units, which takes an exponentially distributed amount of time with mean μ^{-1}. Clearly, we must have $\mu K > \lambda E[Z]$ to ensure stability. The production cost is incurred at the rate of c per unit of time. Let $X(t)$ denote the inventory level at time t. Thus, $X(t) > 0$ represents the number of units in stock and $-X(t) > 0$ denotes the number of backlogged units at time t. When $X(t) = x$, the holding and shortage cost is incurred at the rate of $h(x)$ per unit of time. We assume $h(x)$ to be convex and $h(x) \to \infty$ when $|x| \to \infty$. Typically, $h(x)=hx^+ + px^-$, where $x^+ = \max\{x,0\}$, $x^- = \max\{-x,0\}$, h denotes the per-unit-time holding cost for each unit of product in stock, and p is the per-unit-time shortage cost for each backlogged unit of product. So far, this is a classical production/inventory control problem.

Departing from the classical production/inventory setting, the system faces the occasional arrival of a special event, each of which lasts for a certain amount of time. During these events, the system is offered an incentive program that rewards it for not producing. Specifically, for each unit of time that the system remains idle, it is rewarded by an amount R. These events were exemplified in the introduction as peak periods of electricity usage, in which the supplier provides incentives for not using the resource by means of shutting down production. For convenience, we retain the terms "peak period" and "nonpeak period" to differentiate the production environments. The activation of the incentive program concurs during peak periods, which means that from the beginning to the end of any peak period, the incentive program is always available.

Peak periods arrive in accordance with a Poisson process with rate β, and each peak period lasts for an exponentially distributed amount of time with parameter γ. There are thus two possible system states, a "peak" state and a "nonpeak" state. The transition rate from nonpeak state to peak state is β, and the transition rate from peak state to nonpeak state is γ (alternatively, the mean length of the nonpeak periods is β^{-1} and the mean duration of the peak periods, or the mean peak duration, is γ^{-1}). The assumption of an exponential peak duration means that the length of a peak period is not only a priori uncertain before the peak arrives, but also a posteriori uncertain, even when the system is already in the peak period. In the next section we will consider a case in which the peak duration is only a priori uncertain.

The set of decision epochs in this model consists of arrival epochs, peak-period expiry epochs, and production completion epochs. A production and shutdown policy specifies at each decision epoch whether the production will be stopped, kept down, resumed, or continued. In particular, at an arrival epoch of a peak period, a decision must be made as to whether to stop production to take advantage of the incentive program. In decision epochs occurring within a peak period, it will be decided whether the system should remain idle if it has been shut down, or whether to stop production if the system is currently producing a batch. In demand arrival (and production completion) epochs during a nonpeak period, it must also be decided whether the system should idle or produce a batch. The objective is to find a policy that minimizes the total discounted cost over an infinite horizon.

We can formulate this optimal production/inventory control problem as a Markov decision process. Define the state to be (x, i), where x denotes the inventory level, $i = 1$ indicates that the system is currently in a peak period, and $i = 0$ indicates that it is in a nonpeak period. Let $V(x, i)$ be the optimal value function when the initial state is (x, i), that is, the minimum discounted total cost over an infinite horizon. Let α denote the discount factor. Because of the exponential assumptions, all of the decision epochs can be generated using a Poisson process with rate $\lambda + \mu + \alpha + \beta + \gamma$. Without loss of generality, we assume

$$\lambda + \mu + \alpha + \beta + \gamma = 1.$$

If this condition is not met, then it can be achieved by time scaling (see, for example, Walrand 1988, chap. 8; Lippman 1975; or Puterman 1994). The value function satisfies the following optimality equations.

$$\begin{aligned} V(x, 0) &= h(x) + \mu \min\{V(x+K, 0) + c, V(x, 0)\} \\ &\quad + \lambda \sum_{j=1}^{\infty} p_j V(x-j, 0) + \beta V(x, 1) \\ &\quad + \gamma V(x, 0), \end{aligned} \tag{1}$$

$$\begin{aligned} V(x, 1) &= h(x) + \mu \min\{V(x+K, 1) + c, V(x, 1) - R\} \\ &\quad + \lambda \sum_{j=1}^{\infty} p_j V(x-j, 1) + \beta V(x, 1) \\ &\quad + \gamma V(x, 0). \end{aligned} \tag{2}$$

Note that when the system is in state $(x, 1)$, the comparison between $V(x+K, 1) + c$ and $V(x, 1) - R$ represents the difference between the cost of production and of nonproduction.

On the right-hand side of Equations (1) and (2), the first block, $h(x)$, represents the expected holding and shortage cost between two potential decision epochs; the block multiplied by μ represents the transition that is generated by a potential batch completion; the block multiplied by λ denotes the transition that is associated with the arrival of demand; and the block multiplied by β represents the transition to the peak state. As there is no transition from peak to nonpeak periods when the state is $(x, 0)$, the block multiplied by γ represents a dummy transition, causing no change of state, but it has to be included because of normalization (Walrand 1988).

Remark 1. The production/inventory control problem studied here has the same economic trade-off of controlling a make-to-stock queue with alternating low-cost and high-cost periods. A low-cost period corresponds to a nonpeak period, and a high-cost period corresponds to a peak period, during which the firm forfeits the guaranteed reward R if it goes into production.

Our main result in this section is the following theorem, which characterizes the structure of the optimal policy.

Theorem 1. *The optimal production and shutdown policy is defined by two critical numbers $A \geq B$, such that during a nonpeak period the optimal production policy is to produce if and only if the inventory level is less than A, and during a peak period, the optimal policy is to produce if and only if the inventory level is less than B. The optimal control parameter A is increasing, and B is decreasing in the rate of incentive R.*

The optimal control policy is thus of the threshold type. The fact that the control parameter B is decreasing in R is intuitively clear, because the higher the incentive offered, the higher the desire to participate in the program, and hence the lower the inventory level that the system is willing to maintain. The reason that A is increasing in R is because, as the latter becomes larger, a higher inventory target level is set during the nonpeak periods so that when a peak period arrives, production can be halted for a longer period of time to take advantage of the incentive program.

As A is increasing in R and B is decreasing in R, a lower bound for A and an upper bound for B can be obtained at $R = 0$. Moreover, when $R = 0$ there will be no difference between the peak and nonpeak periods, and the model is thus reduced to the classical production/inventory control problem with $A = B$. This special case can be represented by an $M^X/M^K/1$ queue with batch arrival and batch services through the transformation $\widetilde{X}(t) = A + K - X(t)$. The optimality of threshold production policy for the special case with $R = 0$, $X \equiv 1$, and $K = 1$ has been established, see, e.g., Li (1988, 1992).

The ensuing theorem reveals a qualitative relationship between the optimal value function, or the optimal cost, and some of the system parameters, which is also intuitive.

Theorem 2. *The optimal cost is decreasing in R and μ, and increasing in c.*

However, the value function is not necessarily monotonic with respect to other parameters λ, β, or γ. We show this by way of a counterexample. Consider the following problem, which we will refer to as a "base case." The base case has the parameter values $R = 5$, $c = 10$, $h = 2$, $p = 15$, $\mu = 20$, $\lambda = 3$, $\gamma = 2$, $\alpha = 0.5$, $K = 1$, and $\beta = 5$. We obtain through a numerical study the optimal costs in states $(50, 1)$ and $(50, 0)$: $V(50, 1) = -95$ and $V(50, 0) = -60$. Changing β, the arrival rate of peak periods, to 15 (or 25), but keeping all of the other parameters the same, results in $V(50, 1) = -120$ and $V(50, 0) = -107$ (or $V(50, 1) = -110$ and $V(50, 0) = -106$). This example clearly demonstrates that the value function is not necessarily monotonic with respect to all of the parameters. Similar numerical examples can be constructed for λ and γ.

At first glance, it seems plausible that the optimal cost decreases with β, the rate of transition from peak to nonpeak periods, and increases with γ, the rate of transition from nonpeak to peak periods. Here, we offer an explanation for why this is not true in general. To take advantage of the incentive program in peak periods, which is activated more frequently as β increases, we may lower B and/or raise A to obtain the guaranteed reward R per unit of time. However, the likelihood of having stockout is higher if B is lowered, but a higher A may result in a higher long-term average inventory level. If the dominance of one effect over the other varies as β changes, then the monotonicity will fail. The effect of γ on the optimal cost is similar.

Having characterized the structure of the optimal policy for (1) and (2), we next proceed to analyze the possible ranges of the optimal control parameters. To obtain analytical results, we confine our attention to the special case $h(x) = hx^+ + px^-$. The following result is intuitive.

Theorem 3. *The optimal control parameters A and B are always finite.*

3. Mixed-Erlang Incentive Duration

The previous section assumes that a peak period lasts for an exponentially distributed amount of time, and thus the optimal decision depends only on whether the system is in a peak state or a nonpeak state. In this section, we assume that the duration of the peak period becomes known at the time it begins. This is consistent with the practice of energy buyback programs mentioned in the introduction: The resource supplier announces at the beginning of each peak period when the incentive program will end. As is expected, in this case the optimal policy in the peak state will depend on the amount of time that remains in the current peak period.

The lengths of the peak periods, however, can differ from one to another, and we assume that the duration of the peak period is drawn from a general distribution, say F. More specifically, when the system is in the nonpeak state, the firm is not a priori certain about the exact length of the next peak period, but only knows that it will last for an amount of time that is drawn from a given distribution F. However, as soon as the peak period arrives, its exact duration is known.

As a mixed-Erlang distribution can approximate any nonnegative continuous random variable to any degree of accuracy (see Barbour 1976), we focus on the distribution function

$$F=\sum_{i=1}^{\infty} w_i E(i,\gamma),$$

where $\sum_{i=1}^{\infty} w_i = 1$ and $E(i,\gamma)$ is an Erlang distribution with i phases and each phase is exponentially distributed with a mean of $1/\gamma$.

The state of the system is (x, i), where i takes values $0, 1, \ldots,$ $i=0$ represents the nonpeak state, and $i>0$ represents the number of remaining phases (inclusive of i) before the current peak period ends. The optimal value function then satisfies the following optimality equations, and its optimal policy is described in Theorem 4.

$$\begin{aligned} V(x,0) &= h(x)+\mu\min\{V(x+K,0)+c,V(x,0)\} \\ &\quad +\lambda\sum_{j=1}^{\infty}p_jV(x-j,0)+\beta\sum_{i=1}^{\infty}w_iV(x,i) \\ &\quad +\gamma V(x,0), \qquad (3) \\ V(x,i) &= h(x)+\mu\min\{V(x+K,i)+c,V(x,i)-R\} \\ &\quad +\lambda\sum_{j=1}^{\infty}p_jV(x-j,i)+\beta V(x,i)+\gamma V(x,i-1), \\ &\qquad i=1,2,\ldots. \qquad (4) \end{aligned}$$

Theorem 4. *The optimal production and shutdown policy is defined by a series of critical numbers*: $A \geq B_i$, $i=1,2,\ldots,$ *such that if the system is in state* $(x,0)$, *then the optimal policy is to produce if and only if the inventory level*, x, *is less than* A, *and if the system is in state* (x,i), $i>0$, *then the optimal policy is to produce if and only if the inventory level is less than* B_i. *Moreover*, A *is increasing, and* B_i *is decreasing in the rate of incentive*, R, $i=1,2,\ldots.$

Remark 2. Assume that at a given decision epoch, the system is in state (x, i), where $i>0$. If $x \geq B_i$, then the optimal decision is not to produce, and the system participates in the incentive program. If $x<B_i$, then the system forgoes the incentive program and produces. Note that due to randomness in the finishing time of ongoing production and the arrival of demand, the inventory level can go up or down at the next decision epoch. If the inventory level reaches B_i (or above B_i) before the number of remaining peak phases transits to $i-1$, then the system stops producing. It may also happen that at the next decision epoch the number of remaining phases becomes $i-1$ before the inventory level reaches B_i. In that case, whether to participate in the incentive program depends on the inventory level, x', at the moment of transition: If $x'<B_{i-1}$, then production continues; otherwise, i.e., if $x' \geq B_{i-1}$, then the system is turned off.

Remark 3. If we interpret the result of Theorem 4 in terms of the remaining duration of the peak period, then it is implied that the optimal production policy is defined by a switching curve B_t, where $B_t \leq A$ for all t. Because it is assumed in this section that the supplier announces the length of time for which the incentive program will be available, the firm knows the exact remaining time until the end of the current peak period. If the inventory level is x and there are t units of time left before the current peak period expires, then the system stops production only if $x \geq B_t$. We can call this switching curve the *shutdown threshold curve*.

Remark 4. Intuitively, B_i should increase in i. This is also confirmed by an intensive numerical study. Although unable to prove it formally, we conjecture that such an intuition is generally true.

A special case of mixed-Erlang peak-period distribution is a constant length, that is, all peak periods last for a constant amount of time, say, a units of time. As any deterministic value can be approximated by an Erlang distribution to any degree of accuracy, this can be modeled by the Erlang distribution $E(m,\gamma)$ with $m/\gamma=a$, i.e., the peak period has m phases, and each phase is exponentially distributed with rate γ. The optimal production and shutdown policy is therefore determined by $m+1$ critical numbers, A and B_i, where $A \geq B_i$, $i=1,\ldots,m$, as given in Theorem 4.

4. Seasonal Effects

Even though in reality we cannot know exactly when the next peak period will arrive, it is known that peak periods are more likely to occur in the summer seasons. To take account of the seasonal effect, we therefore assume in this section that time alternates between two seasons, 1 and 0, which can be considered as "summer" and "nonsummer" seasons. For convenience we will call them "hot" and "cool" seasons. As the duration of each season in real life is deterministic, we model them by Erlang distributions (m_i, γ_i), $i = 0, 1$. Therefore, season i goes through m_i phases with each phase being exponentially distributed with parameter γ_i, $i = 0, 1$. There is an arrival process of peak periods in hot seasons that occurs at rate β. However, there is no arrival process of peak periods in cool seasons. For simplicity we only consider the case in which the duration of the peak period is exponentially distributed with parameter ϑ. The case with an arbitrary peak duration can be handled in the same manner as in the last section. The demand and production assumptions, as well as the cost structure, remain the same as in §2.

The state of the system is (x, i, j), where x is the inventory level; $i = 1, 2, \ldots, m_0$ represents that the system is in phase i of a cool season; $i = m_0 + 1, \ldots, m_0 + m_1$ represents that the system is in phase $i - m_0$ of a hot season; and $j = 0$ or 1 indicates that the system is in a peak or a nonpeak period.

Without loss of generality, we assume that

$$\lambda + \mu + \alpha + \beta + \gamma_0 + \gamma_1 + \vartheta = 1.$$

Let $V(x, i, j)$ denote the minimum total discounted cost when the initial state is (x, i, j). It satisfies the following optimality equations,

$$\begin{aligned} V(x,i,0) = {} & h(x)+\mu\min\{V(x+K,i,0)+c, V(x,i,0)\} \\ & +\lambda\sum_{j=1}^{\infty}p_j V(x-j,i,0)+\beta V(x,i,0) \\ & +\gamma_0 V(x,i+1,0)+\gamma_1 V(x,i,0) \\ & +\vartheta V(x,i,0), \quad i=1,\ldots,m_0, \end{aligned} \tag{5}$$

$$\begin{aligned} V(x,i,0) = {} & h(x)+\mu\min\{V(x+K,i,0)+c, V(x,i,0)\} \\ & +\lambda\sum_{j=1}^{\infty}p_j V(x-j,i,0)+\beta V(x,i,1) \\ & +\gamma_0 V(x,i,0)+\gamma_1 V(x,i+1,0)+\vartheta V(x,i,0), \\ & i=m_0+1,\ldots,m_0+m_1, \end{aligned} \tag{6}$$

$$\begin{aligned} V(x,i,1) = {} & h(x)+\mu\min\{V(x+K,i,1)+c, V(x,i,1)-R\} \\ & +\lambda\sum_{j=1}^{\infty}p_j V(x-j,i,1)+\beta V(x,i,1) \\ & +\gamma_0 V(x,i+1,1)+\gamma_1 V(x,i,1)+\vartheta V(x,i,0), \\ & i=1,\ldots,m_0, \end{aligned} \tag{7}$$

$$\begin{aligned} V(x,i,1) = {} & h(x)+\mu\min\{V(x+K,i,1)+c, V(x,i,1)-R\} \\ & +\lambda\sum_{j=1}^{\infty}p_j V(x-j,i,1)+\beta V(x,i,1) \\ & +\gamma_0 V(x,i,0)+\gamma_1 V(x,i+1,0)+\vartheta V(x,i,0), \\ & i=m_0+1,\ldots,m_0+m_1-1, \end{aligned} \tag{8}$$

$$\begin{aligned} & V(x, m_0+m_1, 1) \\ & \quad = h(x) + \mu\min\{V(x+K, m_0+m_1, 1)+c, \\ & \qquad\qquad V(x, m_0+m_1, 1) - R\} \\ & \qquad + \lambda\sum_{j=1}^{\infty}p_j V(x-j, m_0+m_1, 1) + \beta V(x, m_0+m_1, 1) \\ & \qquad + \gamma_0 V(x, m_0+m_1, 1) + \gamma_1 V(x, 1, 1) \\ & \qquad + \vartheta V(x, m_0+m_1, 0). \end{aligned} \tag{9}$$

We only offer an explanation for optimality Equation (9). The state of the system is $(x, m_0 + m_1, 1)$, which means that the hot season is in its last phase, the firm is in a peak period, and the inventory level is x. The first term on the right-hand side is the holding and shortage cost rate. The second term compares the two options, to produce or not to produce. If the decision is made not to produce, then the system receives the reward R per unit of time as the system is in a peak period, whereas if the decision is made to produce, then a rate c is incurred and the reward R is foregone. The third term on the right-hand side is the result of the batch Poisson arrival process. During a peak period, the arrival process of peak period becomes a dummy event, which is given by the fourth term. The fifth term, which involves γ_0, is a dummy event because the system is in a hot season.

As the hot season is in its last phase, the sixth term gives the transition rate to the first phase of the next cool season. Finally, the last term on the right-hand side is the transition rate from a peak period to a nonpeak period. We note that even though there is no arrival process of a peak period during cool seasons, it is mathematically possible for the system to be in state $(x, i, 1)$ for $i = 1, \ldots, m_0$, as can be seen in Equation (7). This occurs when, during a peak period in a hot season, the hot season ends before the peak period terminates. This is due to our assumption of an exponentially distributed peak-period duration, which can be removed by using more complicated models of peak durations.

Using the same approach as in the previous sections, we can show that the optimal production strategy during a hot (cool) season is determined by the critical numbers $A_i^{(1)}$ and $B_i^{(1)}$ ($A_i^{(0)}$ and $B_i^{(0)}$), $A_i^{(1)} \geq B_i^{(1)}$ ($A_i^{(0)} \geq B_i^{(0)}$), such that when the system is in phase i of a hot (cool) season, if the system is in a peak state, then the optimal production strategy is to produce if and only if the inventory level, x, is less than $B_i^{(1)}$ ($B_i^{(0)}$), and if the system is in a nonpeak state, then the optimal production strategy is to produce if and only if the inventory level, x, is less than $A_i^{(1)}$ ($A_i^{(0)}$). Furthermore, the control parameters $A_i^{(j)}$ are increasing in the rate of incentive R, and $B_i^{(j)}$ are decreasing in R.

The result can be further extended to the case in which a peak period has a nonexponential distribution, in which another dimension will be added in the state space for the number of phases until the end of the peak period. The optimal strategy, then, will be dependent on this additional dimension.

5. Concluding Remarks

Motivated by the incentive programs that have been offered by energy companies under tight market conditions, in this paper we studied a production/inventory control problem in which there is an alternation between peak and nonpeak periods. Using Markov decision processes, we characterized the structure of the optimal production and shutdown policies.

Some extensions are possible. First, in our analysis we assumed that a peak period arrives according to a Poisson process, and thus the time until the next peak period is exponentially distributed. If information is available about the arrival of the next peak period, then we can model the peak-period interarrival time by an arbitrarily distributed random variable using the same technique that we used in §3 for handling nonexponential peak duration times. In this case, the optimal production strategy during a nonpeak period will become dependent on the remaining time until the arrival of the next peak period. Second, the product has been assumed to be made in a fixed batch size of K. It might be more realistic in some applications to assume that the system can produce any quantity up to K. Some of our results can be extended to such a case. For example, consider the model of §2 and let the cost for producing i $(0 \leq i \leq K)$ units be c_i, where $0 = c_0 \leq c_1 \leq \cdots \leq c_K$. The optimal value function for this case is

$$V(x,0) = h(x) + \mu \min_{0\leq i\leq K}(V(x+i,0)+c_i) + \lambda \sum_{j=1}^{\infty} p_j V(x-j,0) + \beta V(x,1) + \gamma V(x,0),$$

$$V(x,1) = h(x) + \mu \min\left\{\min_{1\leq i\leq K}(V(x+i,1)+c_i), V(x,1)-R\right\} + \lambda \sum_{j=1}^{\infty} p_j V(x-j,1) + \beta V(x,1) + \gamma V(x,0).$$

Under certain conditions on the cost parameters, it can be shown that the critical number policy is still optimal. Third, in many applications it may be costly to switch between "on" and "off" because shutdown costs may be incurred. We note that if the shutdown cost is accumulated at a constant rate, it can then be incorporated into our model by modifying the rate of incentive R. It would be interesting to consider a general shutdown cost structure, such as a time-dependent shutdown cost plus a fixed shutdown cost.

A variation of our model is, instead of considering the reward R, to introduce different production costs c_0 and c_1, respectively, for the nonpeak and peak periods, where $c_0 \leq c_1$. In this case it is more convenient to refer to peak and nonpeak periods as high-cost and low-cost periods (see Remark 1). To illustrate, we consider the model in §2, with exponentially distributed durations of high-cost periods. Under this new cost structure, the optimality equations become

$$\begin{aligned}V(x,0) &= h(x)+\mu\min\{V(x+K,0)+c_0, V(x,0)\}\\&\quad+\lambda\sum_{j=1}^{\infty}p_jV(x-j,0)+\beta V(x,1)+\gamma V(x,0),\\V(x,1) &= h(x)+\mu\min\{V(x+K,1)+c_1, V(x,1)\}\\&\quad+\lambda\sum_{j=1}^{\infty}p_jV(x-j,1)+\beta V(x,1)\\&\quad+\gamma V(x,0).\end{aligned}\tag{10}$$

For this new model, we can prove the following result.

Proposition 1. *The optimal production strategy for the alternating low/high-costs problem with cost rates c_0 and c_1, $c_0 \le c_1$, is determined by two critical numbers $A \ge B$, such that during a low-cost period the optimal production policy is to produce if and only if the inventory level is less than A, and during a high-cost period, the optimal policy is to produce if and only if the inventory level is less than B. The optimal control parameter A is decreasing in cost rate c_0 but increasing in cost rate c_1, while the optimal control parameter B is increasing in c_0 but decreasing in c_1.*

We remark that the structure of the optimal policy here is essentially the same as that given in Theorem 1, which is not surprising in view of Remark 1. However, the problem studied in §2 cannot be considered as a special case of this result because that model cannot be obtained by appropriately choosing c_0 and c_1. One might think that letting $c_0 = c$ and $c_1 = c + R$ in the new model can yield the model of §2. However, after substituting such c_0 and c_1 into the optimality Equation (10), we obtain

$$\begin{aligned}V(x,1) &= h(x)+\mu\min\{V(x+K,1)+c, V(x,1)-R\}\\&\quad+\mu R+\lambda\sum_{j=1}^{\infty}p_jV(x-j,1)\\&\quad+\beta V(x,1)+\gamma V(x,0),\end{aligned}\tag{11}$$

which differs from (2). Clearly, the two models converge only when the rate of incentive R is 0.

Acknowledgments
The research of the first author is partially supported by the NSF under Grants DMI-0196084 and DMI-0200306, by a grant from Progress Energy (formerly CP&L), and by a grant from the National Natural Science Foundation of China under Grant 70228001, and the research of the second author is partially supported by RGC/CUHK 4192/04E. The authors are grateful to the senior editor and two anonymous referees for their detailed comments and constructive suggestions, which have significantly improved the exposition of the paper. Remark 1 was suggested by one of the referees and the senior editor. The authors thank Sandy Stidham for insightful discussions on the control of stochastic systems. Computational assistance from Mr. Sean Zhou is also acknowledged and greatly appreciated.

Appendix
Here, we first define some additional notation that considerably shortens the subsequent analysis. We use the symbols $\uparrow$ and $\downarrow$ as abbreviations for increasing and decreasing, respectively. For any real value function V on the state space, S, define

$T_0V(x,0)=\min\{V(x+K,0)+c, V(x,0)\}$

$T_1V(x,1)=\min\{V(x+K,1)+c, V(x,1)-R\}$

$\Delta V(x,i)=V(x+K,i)-V(x,i)$, $i=0,1,\ldots$

$\Delta_0V(x,0)=T_0V(x+K,0)-T_0V(x,0)$

$\Delta_1V(x,1)=T_1V(x+K,1)-T_1V(x,1)$.

Note that Δ, Δ_0, and Δ_1 represent the additional value of starting with an additional batch (of size K) in the inventory.

We need several lemmas to prove Theorem 1, the first of which is the following result.

Lemma 1. *If* $\Delta V(x,i)\uparrow x\in\{\ldots,-2,-1,0,1,2,\ldots\}$, *then so are* $\Delta_iV(x,i)$, $i=0,1$.

Proof. For brevity, we provide only the proof for Δ_1. The proof for Δ_0 is almost identical.

Note that

$$\begin{aligned}T_1V(x,1) &= c+\min\{V(x+K,1), V(x,1)-R-c\}\\&= c+\min\{V(x+K,1), V(x,1)-b\}.\end{aligned}$$

Thus, to show that $\Delta_1V(x,1)\uparrow x$, it suffices to prove that

$$\Delta_1V(x+1,1)\ge\Delta_1V(x,1).\tag{12}$$

Consider two cases.

Case 1. $T_1V(x+1,1)=V(x+1,1)-R$ (i.e., $\Delta V(x+1,1)\ge -b$). Then, by the hypothesis $\Delta V(x,1)\uparrow x$ we have

$$T_1V(x+K,1)=V(x+K,1)-R$$

and

$$T_1V(x+K+1,1)=V(x+K+1,1)-R.$$

Thus, the left-hand side of (12) equals $\Delta V(x+1,1)$, and the right-hand side is

$$\begin{aligned}&V(x+K,1)-b-\min\{V(x+K,1), V(x,1)-b\}\\&\quad=\max\{-b,\Delta V(x,1)\}.\end{aligned}$$

Therefore, the inequality follows from the assumption $T_1V(x+1,1)=V(x+1,1)-R$ and the hypothesis of the lemma.

Case 2. $T_1V(x+1,1)=V(x+K+1,1)+c$. By the hypothesis of the lemma, it can be seen that $T_1(x,1)=V(x+K,1)+c$. Then, the left-hand side of (12) becomes $\min\{\Delta V(x+K+1,1),-b\}$, and the right-hand side becomes $\min\{\Delta V(x+K,1),-b\}$. Therefore, the inequality follows from the hypothesis of the lemma. □

Remark 5. If $K=1$, then the condition of Lemma 1 is the convexity of $V(x,i)$ for fixed i, and the conclusion is the convexity of $T_1V(x,i)$. This special case has been proved by Lippman and Stidham (1977). For $K>1$, however, counterexamples can easily be given to demonstrate that $T_1V(x,i)$ is not necessarily convex in x, even if $V(x,i)$ is convex. We note that the convexity of function $V(x,i)$ is a stronger requirement than $\Delta V(x,i)$ being increasing, which is what we need in this paper.

Now, consider the n-step version of (1) and (2), which is denoted by $V_n(x,i)$:

$$V_{n+1}(x,0) = h(x)+\mu T_0V_n(x,0)+\lambda\sum_{j=1}^{\infty}p_jV_n(x-j,0) +\beta V_n(x,1)+\gamma V_n(x,0), \tag{13}$$

$$V_{n+1}(x,1) = h(x)+\mu T_1V_n(x,1)+\lambda\sum_{j=1}^{\infty}p_jV_n(x-j,1) +\beta V_n(x,1)+\gamma V_n(x,0), \tag{14}$$

where $V_0(x,i)=0$ for every state (x,i).

Lemma 2. *For a fixed i, the optimal value function $V(x,i)$ has the property that $\Delta V(x,i)\uparrow x$.*

Proof. As $V_n(x,i)\to V(x,i)$ as $n\to\infty$, it suffices to prove that $\Delta V_n(x,i)\uparrow x$ for all $n\ge 0$. We do so by induction on n. As $V_0(x,i)\equiv 0$ for all x and i, V_0 clearly holds the desired property. Suppose that $\Delta V_n(x,i)\uparrow x$. We then proceed to prove ΔV_{n+1}. From (13) and (14), it suffices to show that both $T_0V_n(x,0)$ and $T_1V_n(x,1)\uparrow x$. As the latter requirements have been confirmed in Lemma 1, the proof is completed. □

Two more lemmas are needed to prove Theorem 1.

Lemma 3. *The optimal value function $V(x,i)$ possesses the property that*

$$\Delta V(x,1)+R\ge \Delta V(x,0) \tag{15}$$

for all x.

Proof. We only need to prove that

$$\Delta V_n(x,1)+R\ge \Delta V_n(x,0) \tag{16}$$

for all x and $n>0$. The proof proceeds by induction on n. Suppose that (16) has been established for n, and we proceed to $n+1$.

Substituting (13) and (14) into (16) and canceling the common terms, we can see that to prove (16) for $n+1$, it suffices to show that

$$\mu[\Delta_1V_n(x+K,1)+R]+\lambda[\Delta V_n(x-1,1)+R] \ge \mu\Delta_0V_n(x+K,0)+\lambda\Delta V_n(x-1,0).$$

As, by induction,

$$\Delta V_n(x-1,1)+R\ge \Delta V_n(x-1,0)\quad \text{for all } x,$$

it suffices to prove that

$$\Delta_1V_n(x+K,1)+R\ge \Delta_0V_n(x+K,0). \tag{17}$$

We prove (17) by considering two cases.

Case 1. $T_1V_n(x,1)=V_n(x,1)-R$ (i.e., $\Delta V_n(x,1)+c\ge -R$). Then, by the induction hypothesis we have

$$T_1V_n(x+K,1)=V_n(x+K,1)-R.$$

Thus, the left-hand side of (17) equals $\Delta V_n(x,1)+R$, and its right-hand side is

$$\begin{aligned}\Delta_0V_n(x+K,0) &\le V_n(x+K,0)-T_0V_n(x,0)\\ &= \max\{\Delta V_n(x,0),-c\}.\end{aligned} \tag{18}$$

Because of the induction hypothesis

$$\Delta V_n(x,1)+R\ge \Delta V_n(x,0)$$

and the assumption of Case 1, we have

$$\Delta V_n(x,1)+R\ge \max\{\Delta V_n(x,0),-c\}. \tag{19}$$

Thus, (17) follows from (18) and (19).

Case 2. $T_1V_n(x,1)=V_n(x+K,1)+c$. Then, the left-hand side of (17) can be written as $\min\{\Delta V_n(x+K,1)+R,-c\}$. However, by induction hypothesis (16) and the assumption of Case 2, we have $\Delta V_n(x,0)+c<0$. Hence, the right-hand side of (17) becomes

$$T_0V_n(x+K,0)-(V_n(x+K,0)+c)=\min\{\Delta V_n(x+K,0),-c\}.$$

Applying the induction hypothesis $\Delta V_n(x+K,1)+R\ge \Delta V_n(x+K,0)$ yields (17). This completes the proof of Lemma 3. □

Lemma 4. *The optimal value function exhibits the following properties*:

(i) *For fixed $i=0$ or 1, $\Delta V(x,i)\downarrow R$.*
(ii) *For fixed $i=0$ or 1, $\Delta V(x,i)+R\uparrow R$.*

Proof. As usual, we prove (i) and (ii) for the n-step value function $V_n^R(x,i)$ (here, the superscript R is used to make the dependence of V on R explicit). We need to verify four relations:

(i′) $\Delta V_n^R(x,0)\downarrow R$.
(ii′) $\Delta V_n^R(x,1)\downarrow R$.
(iii′) $\Delta V_n^R(x,0)+R\uparrow R$.
(iv′) $\Delta V_n^R(x,1)+R\uparrow R$.

Here, we show that (i′) and (iv′) hold by induction. The verification of (ii′) and (iii′) is similar, and is omitted for brevity.

First, consider (i′). It is clearly satisfied for $n=0$, and we proceed to prove it for $n+1$. With (13) and the induction hypothesis, we only need to prove that $\Delta_0 V_n^R(x,0)\downarrow R$, i.e., for $R'\geq R$, we require

$$\Delta_0 V_n^{R'}(x,0)\leq \Delta_0 V_n^R(x,0). \tag{20}$$

To prove this inequality, we consider two cases.

Case 1. $T_0 V_n^{R'}(x,0)=V_n^{R'}(x,0)$. Then, by the induction hypothesis we have $T_0 V_n^R(x,0)=V_n^R(x,0)$, $T_0 V_n^{R'}(x+K,0)=V_n^{R'}(x+K,0)$, and $T_0 V_n^R(x+K,0)=V_n^R(x+K,0)$. Thus, (20) is reduced to the requirement $\Delta V_n^{R'}(x,0)\leq \Delta V_n^R(x,0)$. This is clearly satisfied because of the induction hypothesis (i′).

Case 2. $T_0 V_n^{R'}(x,0)=V_n^{R'}(x+K,0)+c$. Then, the left-hand side of (20) equals $\min\{\Delta V_n^{R'}(x+K,0),-c\}$, and the right-hand side is

$$\begin{aligned}\Delta_0 V_n^R(x+K,0) &\geq T_0 V_n^R(x+K,0)-\big(V_n^R(x+K,0)+c\big)\\ &= \min\big\{\Delta V_n^R(x+K,0),0,-c\big\}.\end{aligned}$$

Hence, by the induction hypothesis (20) also holds.

We now prove (iv′). By (14) and the induction hypothesis, we require that for $R'\geq R$,

$$\Delta_1 V_n^{R'}(x+K,1)+R'\geq \Delta_1 V_n^R(x+K,1)+R. \tag{21}$$

We can again differentiate between two cases.

Case 1. $T_1 V_n^{R'}(x,1)=V_n^{R'}(x,1)-R'$. By the induction hypothesis, we see that $T_1 V_n^{R'}(x+K,1)=V_n^{R'}(x+K,1)-R'$. Hence, the left-hand side of (21) equals

$$\Delta V_n^{R'}(x,1)+R', \tag{22}$$

and its right-hand side is $\leq V_n^R(x+K,1)-R-T_1 V_n^R(x,1)+R=\max\{\Delta V_n^R(x,1)+R,-c\}$. Thus, (21) holds because of hypothesis (iv′) and the induction hypothesis $T_1 V_n^{R'}(x,1)=V_n^{R'}(x,1)-R'$.

Case 2. $T_1 V_n^{R'}(x,1)=V_n^{R'}(x+K,1)+c$. Then, by the induction hypothesis (iv′) for n, we need $T_1 V_n^R(x+K,1)=V_n^R(x+K,1)+c$. If this is satisfied, then the requirement (21) becomes

$$\min\big\{\Delta V_n^{R'}(x+K,1)+R',-c\big\}\geq \min\big\{\Delta V_n^R(x+K,1)+R,-c\big\}.$$

This is clearly satisfied because of the induction hypothesis. Therefore, the proof is completed. □

We are now ready to prove Theorem 1.

Proof of Theorem 1. Define

$$A=\inf\{x\mid V(x+K,0)\geq V(x,0)-c\},$$

$$B=\inf\{x\mid V(x+K,1)\geq V(x,1)-R-c\},$$

where $\inf\varnothing=\infty$. It follows from Lemma 2 that A and B are well defined. It follows from the optimality equations for the infinite-horizon problem that is represented by (1) and (2) that when the incentive program is not available, it is optimal to produce if and only if $V(x+K,0)<V(x,0)-c$, or $x<A$, and when the incentive program is available, it is optimal to participate if and only if $V(x+K,1)\geq V(x,1)-R-c$, which is equivalent to $x\geq B$. This control policy clearly satisfies the optimality Equations (1) and (2), and thus is optimal.

By Lemma 3 we have

$$\Delta V(x,1)+R+c\geq \Delta V(x,0)+c,$$

which leads to

$$\begin{aligned}&\{x\mid V(x+K,0)\geq V(x,0)-c\}\\ &\qquad\subseteq\{x\mid V(x+K,1)\geq V(x,1)-R-c\},\end{aligned}$$

which implies $B\leq A$.

To emphasize the dependency on the incentive R, we write control policies A and B as $A(R)$ and $B(R)$, and the value function as $V^R(x,i)$, i.e.,

$$A(R)=\inf\{x\mid V^R(x+K,0)\geq V^R(x,0)-c\},$$

$$B(R)=\inf\{x\mid V^R(x+K,1)\geq V^R(x,1)-R-c\}.$$

If $R'\geq R$, then by Lemma 4 we have

$$\Delta V^{R'}(x,0)+c\leq \Delta V^R(x,0)+c$$

and

$$\Delta V^{R'}(x,0)+R'+c\geq \Delta V^R(x,0)+R+c.$$

Hence,

$$\big\{x\colon T_0 V^{R'}(x,0)=V^{R'}(x,0)\big\}\subseteq\big\{x\colon T_0 V^R(x,0)=V^R(x,0)\big\}$$

and

$$\begin{aligned}&\big\{x\colon T_0 V^{R'}(x,0)=V^{R'}(x,0)-R'\big\}\\ &\qquad\supseteq\big\{x\colon T_0 V^R(x,0)=V^R(x+K,0)-R\big\}.\end{aligned}$$

Therefore, it follows from the definition of $A(R)$ and $B(R)$ that $A(R)\leq A(R')$ and $B(R)\geq B(R')$. This completes the proof of Theorem 1. □

Proof of Theorem 2. The proof that V is increasing in c and decreasing in R can be performed by first proving that it is true to V_n by induction on n, and then letting $n\to\infty$. To prove that V decreases with μ, we can first show that the policy of Theorem 1 is not only optimal in the current setting, but is also optimal under the service capacity of "intensity control" (see Chen and Yao 1990). Consider two rates μ_1 and μ_2 with $\mu_1\leq\mu_2$. One notes that the optimal policy for the problem with the service rate μ_1 is also a feasible policy for the problem with service rate μ_2 (see Carr and Duenyas 2000). As the optimal value function for the system with rate μ_2 is at least as good as the value function that corresponds to this feasible policy, the result thus follows. □

Proof of Theorem 3. It follows from Lemma 2 that $\Delta V(x,i)$ is increasing in x, and, as a result, converges as x goes to infinity. We first consider the case in which $\Delta V(x,i)$ converges to a finite number for $i=0,1$. Let $\lim_{x\to\infty}\Delta V(x,i)=\Delta V_i$, $i=0,1$. If $A=\infty$, i.e., if $T_0V(x,0)=V(x+K,0)+c$ (or $\Delta V(x,0)\le -c$) for all x, then for $x>0$, we have

$$V(x,0) = hx+\mu(V(x+K,0)+c)+\lambda\sum_{j=1}^{\infty}p_jV(x-j,0) +\beta V(x,1)+\gamma V(x,0).$$

Hence,

$$\Delta V(x,0) = hK+\mu\Delta V(x,0)+\lambda\sum_{j=1}^{\infty}p_j\Delta V(x-j,0) +\beta\Delta V(x,1)+\gamma\Delta V(x,0).$$

Letting $x\to\infty$ yields

$$\Delta V_0=Kh+\mu\Delta V_0+\lambda\Delta V_0+\beta\Delta V_1+\gamma\Delta V_0. \tag{23}$$

Because $\alpha+\lambda+\mu+\beta+\gamma=1$, Equation (23) becomes

$$(\alpha+\beta)\Delta V_0=Kh+\beta\Delta V_1. \tag{24}$$

Similarly, if $A<\infty$, then for all $x>\max\{A,0\}$ we have

$$V(x,0) = hx+\mu V(x,0)+\lambda\sum_{j=1}^{\infty}p_jV(x-j,0) +\beta V(x,1)+\gamma V(x,0)$$

and

$$\Delta V(x,0) = Kh+\mu\Delta V(x,0)+\lambda\sum_{j=1}^{\infty}p_j\Delta V(x-j,0) +\beta\Delta V(x,1)+\gamma\Delta V(x,0).$$

Letting $x\to\infty$, we also obtain (24). Using the same argument for Equation (2), we can show that the following equation is always satisfied:

$$(\alpha+\gamma)\Delta V_1=Kh+\gamma\Delta V_0. \tag{25}$$

Solving Equations (24) and (25) yields

$$\Delta V_0=\Delta V_1=Kh/\alpha.$$

As $A=\infty$ if and only if $\Delta V(x,0)\le -c$ for all x, and $B=\infty$ if and only if $\Delta V(x,1)\le -c-R$ for all x, then $\Delta V_0\le -c$ and $\Delta V_1\le -c-R$. This is a contradiction, which proves that $A<\infty$ and $B<\infty$.

We then consider the case in which $\Delta V(x,i)$ converges to infinity for either $i=0$ or $i=1$. It is easily seen from (23) that $\Delta V(x,1-i)$ also converges to infinity. Again, as $A=\infty$ if and only if $\Delta V(x,0)\le -c$ for all x, and $B=\infty$ if and only if $\Delta V(x,1)\le -c-R$ for all x, contradictions arise. This proves that both A and B must be finite. □

Several lemmas are needed to prove Theorem 4. Again, we first need to consider the problem with n steps. The length of each step is exponentially distributed, with a mean of 1. The n-step value function satisfies the following optimality equations.

$$V_{n+1}(x,0)=h(x)+\mu T_0V_n(x,0)+\lambda\sum_{j=1}^{\infty}p_jV_n(x-j,0) +\beta\sum_{i=1}^{\infty}w_iV_n(x,i)+\gamma V_n(x,0), \tag{26}$$

$$V_{n+1}(x,i)=h(x)+\mu T_iV_n(x,i)+\lambda\sum_{j=1}^{\infty}p_jV_n(x-j,i) +\beta V_n(x,i)+\gamma V_n(x,i-1),\quad i=1,2,\ldots, \tag{27}$$

where $T_iv(x,i)=\min\{V(x+K,i)+c,V(x,i)-R\}$. Note that $\Delta_iV(x,i)$ can be extended similarly.

The proof of the following lemma is identical to that of Lemma 2, and hence is omitted.

Lemma 5. *The value function $V(x,i)$ satisfies the requirement that $\Delta V(x,i)$ is increasing in x for given $i=0,1,\ldots$.*

Lemma 6. *The value function $V(x,i)$ satisfies the following condition.*

$$\Delta V(x,i)+R\ge\Delta V(x,j),\quad i,j=0,1,\ldots \tag{28}$$

for all x.

Proof. As in the proofs of Lemma 3 and Lemma 4, it suffices to prove (28) for the n-step value function $V_n(x,i)$. That is,

$$\Delta V_n(x,i)+R\ge\Delta V_n(x,j),\quad i,j=0,1,\ldots \tag{29}$$

for all $n\ge 0$. Because of the different expressions for $V_n(x,i)$ for $i=0$ and $i\ne 0$, we will prove the following three inequalities separately.

$$\Delta V_n(x,i)+R\ge\Delta V_n(x,0),\quad i\ne 0, \tag{30}$$

$$\Delta V_n(x,i)+R\ge\Delta V_n(x,j),\quad i,j\ne 0, \tag{31}$$

$$\Delta V_n(x,0)+R\ge\Delta V_n(x,j),\quad j\ne 0. \tag{32}$$

As before, the proof is carried out by induction on n. We assume that it is true for n, and proceed to prove it for $n+1$. As the proof of this lemma is quite lengthy, we shall only provide the proof for (31). The proofs for the other two cases can be found in Chao and Chen (2002).

To prove (31), we need to show that, for $j,i>0$,

$$\Delta V_{n+1}(x,i)+R\ge\Delta V_{n+1}(x,j). \tag{33}$$

In view of (27), the following requirement is sufficient for (33) to hold:

$$T_iV_n(x+K,i)+R\ge T_jV_n(x+K,j),\quad i,j>0. \tag{34}$$

We again differentiate two cases.

Case 1. $T_i V_n(x,i) = V_n(x,i) - R$ (or $\Delta V_n(x,i) \ge -R-c$). Then, by the induction hypothesis, we have $T_i V_n(x+K,i) = V_n(x+K,i) - R$. Thus, the left-hand side of (34) equals $\Delta V_n(x,i)+R$, and the right-hand side is

$$\begin{aligned} T_j V_n(x+K,j) &\le V_n(x+K,j) - R - T_j V_n(x,j) \\ &= \max\{\Delta V_n(x,j), -R-c\}. \end{aligned}$$

By the induction hypothesis, it can be seen that $\Delta V_n(x,i) + R \ge \Delta V_n(x,j)$. With the assumption of Case 1, we have $\Delta V_n(x,i) + R \ge -c \ge -c-R$, and thus (34) is satisfied.

Case 2. $T_i V_n(x,i) = V_n(x+K,i)+c$ (or $\Delta V_n(x,i) < -R-c$). Then, by induction hypothesis (31) we must have $\Delta V_n(x,j) \le -c$, and the left-hand side of (34) thus equals

$$\begin{aligned} &T_i V_n(x+K,i) - (V_n(x+K,i)+c) + R \\ &\quad = \min\{\Delta V_n(x+K,i)+R, -c\}. \end{aligned}$$

To compare this with the right-hand side of (34), we consider two subcases.

Subcase 1. $T_j V_n(x,j) = V_n(x+K,j)+c$ (or $\Delta V_n(x,j) < -R-c$). In this subcase, the right-hand side of (34) equals

$$\begin{aligned} &T_j V_n(x+K,j) - (V_n(x+K,j)+c) \\ &\quad = \min[\Delta V_n(x+K,j), -R-c]. \end{aligned}$$

The induction hypothesis deduces that $\Delta V_n(x+K,i) + R \ge \Delta V_n(x+K,j)$. Combining this with the relationship $-c > -R-c$, we see that (34) holds.

Subcase 2. $T_j V_n(x,j) = V_n(x,j) - R$. In this subcase we must have, by the induction hypothesis, $T_j V_n(x+K,j) = V_n(x+K,j) - R$. Then, the right-hand side of (34) can be rewritten as

$$\Delta_j V_n(x+K,j) = \Delta V_n(x,j).$$

Note that $-c \ge \Delta V_n(x,j)$ under the assumption of Case 2, and by the induction hypothesis, the following holds:

$$\Delta V_n(x+K,i) + R \ge \Delta V_n(x+K,j) \ge \Delta V_n(x,j),$$

where the inequalities follow from the induction hypothesis. Hence, it must be true that

$$\min\{\Delta V_n(x+K,i)+R, -c\} \ge \Delta V_n(x,j),$$

and thus (34) also holds. □

The proof of the next lemma closely follows that of Lemma 4, and is thus omitted here.

Lemma 7. *The value function $V(x,i)$ has the following properties*:

(i) $\Delta V(x,i) \downarrow R$, *for* $i = 0, 1, \ldots,$

(ii) $\Delta V(x,i) + R \uparrow R$, *for* $i = 0, 1, \ldots.$

We are now ready to prove Theorem 4.

Proof of Theorem 4. Define

$$A = \inf\{x \mid V(x+K,0) \ge V(x,0) - c\},$$

$$B_i = \inf\{x \mid V(x+K,i) \ge V(x,i) - R - c\}, \quad i = 1, \ldots,$$

where $\inf \varnothing = \infty$. It follows from Lemma 6 that $A \ge B_i$, $i = 1, \ldots,$ and it follows from Lemma 7 that A is increasing in the incentive R and B_k is decreasing in R for $i = 1, 2, \ldots.$ The proof is similar to that of Theorem 1, and the details are skipped for brevity. □

References

Akella, R., P. R. Kumar. 1986. Optimal control of production rate in a failure-prone manufacturing system. *IEEE Trans. Automatic Control* **31** 116–126.

Barbour, A. D. 1976. Networks of queues and the method of stages. *Adv. Appl. Probab.* **8** 584–591.

Bensoussan, A., M. Crouhy, J-M. Proth. 1983. *Mathematical Theory of Production Planning*. North-Holland, Amsterdam, The Netherlands.

Carr, S., I. Duenyas. 2000. Optimal admission control and sequencing in a make-to-stock/make-to-order production system. *Oper. Res.* **8**(5) 709–720.

Chao, X., F. Y. Chen. 2002. Optimal production and shutdown strategy when supplier offers incentive program. Technical report, Department of Industrial Engineering, North Carolina State University, Raleigh, NC.

Chen, H., D. D. Yao. 1990. Optimal intensity control of a queueing system with state-dependent capacity limits. *IEEE Trans. Automatic Control* **35** 459–464.

Duenyas, I., C. Y. Tsai. 2000. Optimal control of a manufacturing system with random product yield and downward substitutability. *IIE Trans.* **32** 785–795.

Feng, Y., J. Sun. 2001. Computing the optimal replenishment policy for inventory systems with random discount opportunities. *Oper. Res.* **49**(5) 790–795.

Feng, Y., H. Yan. 2001. Optimal production control in a discrete manufacturing system with unreliable machines and random demands. *IEEE Trans. Automatic Control* **45** 100–114.

Gershwin, S. B. 1994. *Manufacturing Systems Engineering*. Prentice Hall, Englewood Cliffs, NJ.

Ha, A. Y. 1997. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Sci.* **43** 1093–1103.

Hu, J. Q., P. Vakili, G. X. Yu. 1994. Optimality of hedging point policies in the production control of failure-prone manufacturing systems. *IEEE Trans. Automatic Control* **39** 1875–1880.

Kitaev, M. V., V. V. Rykov. 1995. *Controlled Queueing Systems*. CRC Press, Boca Raton, FL.

Li, L. 1988. A stochastic theory of a firm. *Math. Oper. Res.* **13** 447–466.

Li, L. 1992. The role of inventory in delivery-time competition. *Management Sci.* **34** 182–197.

Lippman, S. A. 1975. Applying a new device in the optimization of exponential queuing systems. *Oper. Res.* **23** 687–710.

Lippman, S. A., S. Stidham. 1977. Individual versus social optimization in exponential congestion systems. *Oper. Res.* **25** 233–247.

Moinzadeh, K. 1997. Replenishment and stocking policies for inventory systems with random deal offerings. *Management Sci.* **43** 334–342.

Puterman, M. L. 1994. *Markov Decision Processes*. John Wiley & Sons, New York.

Serfozo, R. 1979. An equivalence between continuous and discrete time Markov decision processes. *Oper. Res.* **27** 616–620.

Sobel, M. J. 1969. Optimal average-cost policy for a queue with start-up and shut-down costs. *Oper. Res.* **17** 145–162.

Sobel, M. J. 1982. The optimality of full service policies. *Oper. Res.* **30** 636–649.

Stidham, S., R. R. Weber. 1989. A survey of Markov decision models for control of networks of queues. *Queueing Systems* **13** 291–314.

U.S. Department of Energy. 2002. http://pnnl-utilityrestructuring. pnl.gov / energy-management / energymanagement.htm, or http://pnnl-utilityrestructuring. pnl.gov/Electric/electric.htm.

Wald, M. L. 2000. Utilities try new ways to vary energy pricing. *New York Times* (July 17).

Walrand, J. 1988. *Introduction to Queueing Networks*. Prentice Hall, Englewood Cliffs, NJ.

Zheng, Y.-S. 1994. Optimal-control policy for stochastic inventory systems with Markovian discount opportunities. *Oper. Res.* **42**(4) 721–738.

Zipkin, P. H. 1999. Fundamentals of inventory management. Irwin/ McGraw-Hill, New York.