*Probability in the Engineering and Informational Sciences*, **29**, 2015, 507–525.
doi:10.1017/S0269964815000169

# A CONTINUOUS REVIEW MODEL WITH GENERAL SHELF AGE AND DELAY-DEPENDENT INVENTORY COSTS

Awi Federgruen
*Graduate School of Business, Columbia University, 3022 Broadway New York, NY 10025, USA*
*E-mail: af7@columbia.edu*

Min Wang
*LeBow College of Business, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA*
*E-mail: min.wang@drexel.edu*

We analyze a continuous review inventory model with the marginal carrying cost of a unit of inventory given by an increasing function of its shelf age and the marginal delay cost of a backlogged demand unit by an increasing function of its delay duration. We show that, under a minor restriction, an $(r, q)$-policy is optimal when the demand process is a renewal process, and a state dependent $(r, q)$-policy is optimal when the demand is a Markov-modulated renewal process. We also derive various monotonicity properties for the optimal policy parameters $r^*$ and $r^* + q^*$.

## 1. INTRODUCTION

One of the main objectives of any inventory planning model is to analyze the tradeoff between competing risks of overage and underage. This requires an adequate representation of the carrying costs associated with all inventories, as well as the cost and revenue consequences of shortages. Early contributors, for example, the pioneering textbooks by Hadley and Whitin [22] and Naddor [28], discussed possible paradigms to represent the carrying and shortage costs.

One standard paradigm is to assume that carrying costs can be assessed, either continuously or periodically, as a (possibly non-linear) function of the prevailing total inventory, irrespective of its age composition. Similarly, shortage costs are assumed to accrue as a (again, possibly non-linear) function of the total shortfall or backlog, irrespective of the amount of time the backlogged demand units have remained unfilled. We refer to this type of carrying and shortage cost structures as *level-dependent* inventory costs. After the above mentioned early discussions in Hadley and Whitin [22] and Naddor [28], *this* paradigm has been adopted in virtually every inventory model.

There are, however, many settings where carrying costs need to be differentiated on the basis of the inventory's shelf-age composition. First, inventories are often financed by trade credit arrangements, where the supplier allows for a payment deferral of delivered

 0269-9648/15 $25.00

orders, but charges progressively larger interest rates as the payment delay increases. For example, the supplier frequently offers an initial interest-free period (e.g., 30 days) after which interest accumulates. Moreover, interest rates often increase as a function of the item's shelf age. These trade credit schemes have been considered in Gupta and Wang [21] as well as Federgruen and Wang [16]. We refer to the latter for a discussion of how prevalent this practice is. Another setting with shelf age-dependent inventory cost rates arises when the supplier subsidizes part of the inventory cost. For example, in the automobile industry, manufacturers pay the dealer so-called "holdbacks", that is, a given amount for each month a car remains in the dealer's inventory, up to a given time limit (see, e.g., Nagarajan and Rajagopalan [29]). The resulting inventory cost rate for any stocked item is, again, an increasing function of the item's shelf age. Nahmias, Perry, and Stadje [30] consider settings where the sales price of an item decreases as a function of its shelf age; this can be modeled by assuming the sales price is an age-independent constant, but a shelf age-dependent carrying cost is incurred to account for diminution of the item's value.

Even when inventory costs grow as a *linear* function of the loan term or the amount of time the purchased units stay in inventory, *time-varying* purchase prices or interest rates necessitate disaggregating inventory levels according to the time at which the units were purchased, that is, in accordance with the items' shelf age. In the dynamic lot sizing literature, Federgruen and Lee [14], for example, modeled holding costs as proportional to the items' purchasing price, which varies with their purchase period. As a consequence, holding costs depend on the items' shelf age. An an example, assume a firm carries inventories of a commodity which is procured periodically on a commodity exchange. The unit purchase price varies as a function of the exchange index, fluctuating in accordance with a Markov chain. Assume, at the end of a given period, that the firm carries 10 units of stock, 5 of which were purchased at the beginning of the period at price $p_0$, 3 in the prior period at price $p_{-1}$, and the remaining 2 units two periods earlier at price $p_{-2}$; assume the per period interest rate is $\iota$. The period's inventory carrying cost can not be expressed as a function of the total inventory size, but depends critically on its age composition: it is given by $(5p_0 + 3p_{-1} + 2p_{-2})\iota$. However,this setting is easily modeled with shelf age and state-of-the-world dependent inventory cost rates, as in Section 5.

Even more general shelf-age dependencies are assumed in Levi et al. [26] and its generalization, that is, so-called metric holding costs, in Stauffer et al. [39]. Finally, beyond capital costs, inventories often incur physical handling costs (see, e.g.,Richardson [32]); here, too, the marginal rates often vary as a function of the items' shelf age.

Similar to shelf age-dependent holding costs, backlogging costs may also depend on the amount of time by which delivery of a demand unit is delayed. This may reflect the structure of contractually agreed upon penalties for late delivery or, in case of *implicit* backlogging costs, the fact that customers typically become more impatient over time. This type of backlogging costs has been studied by Chen and Zheng [11], Perry and Stadje [31], Rosling [33] and Huh et al. [23].

In this paper, we analyze a continuous review inventory model with general shelf age-dependent carrying and delay-dependent backlogging costs. This means that the marginal carrying cost of a unit of inventory is given by a general increasing function of its shelf age; similarly, the marginal delay cost of a backlogged demand unit is specified by an increasing function of the delay incurred, thus far. The principal features of the model are as follows: demands are generated by a point process. In addition to the above general holding and delay costs, there is a fixed and variable cost incurred for each replenishment order. An order of arbitrary size may be placed at any demand epoch and arrives after a given deterministic leadtime, or a stochastic leadtime generated by a so-called sequential and exogenous leadtime process, defined below.

In our base model, we assume that the times between consecutive demand epochs are i.i.d., that is, the demand process is a renewal process. We show, under a minor policy restriction, that, an $(r, q)$-policy is optimal, that is, an order of fixed size $q$ is placed whenever the inventory position drops to the reorder point $r$. A straightforward formulation of the control problem results in a semi-Markov decision process (SMDP) with a state space of infinite dimension; the state of the system records the vectors of outstanding orders, the times at which these orders were placed as well as the amount of time any backlogging demand units have been waiting, along with the current inventory level. We show, under the above minor policy restriction, that the complexity of the state space notwithstanding, an optimal policy can be found which acts on the inventory position only. Moreover, under the above matching technique, the expected long-run average costs of any such policy can be expressed as a function of the sequence of inventory positions encountered at consecutive decision epochs. This establishes an equivalence with a *one-dimensional* SMDP; even more remarkably, a simple $(r, q)$-policy is optimal for this as well as the original SMDP.

We also derive various monotonicity properties for the optimal policy parameters $r^*$ and $R^* \equiv r^* + q^*$ as a function of various of the model's primitives. The results are then extended to settings where the consecutive inter-demand time distributions fail to be independent, because they depend on an exogenous state of the world that evolves according to a given Markov process. This gives rise to a model with Markov-modulated demands. We show that an $(r, q)$-policy continues to be optimal, except that both the reorder point and the order size depend on the prevailing state of the world.

Our approach is to assess all holding and delay-dependent costs by matching each ordered unit with a unique demand unit, in the spirit of the approach taken by Axsäter [2,3] as well as the single unit decomposition approach in Muharremoglu and Tsitsiklis [27].

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. In Section 3, we describe preliminary results for Markov-modulated renewal Processes. Section 4 addresses our basic model with a renewal demand process, while Section 5 develops the generalization to Markov-modulated renewal demand processes. Conclusions and extensions are in Section 6.

## 2. LITERATURE REVIEW

We start with a review of the few papers that have characterized the structure of an optimal policy in inventory models with shelf age-dependent holding and/or delay-dependent backlogging costs. To our knowledge, the first and only structural result for models with general shelf age-dependent carrying costs was obtained in the above-mentioned, recent paper by Gupta and Wang [21]. This paper considers, at first, a standard periodic review model with i.i.d. demands and *linear*, stationary order costs. Under a broad class of shelf age-dependent inventory carrying cost functions, the authors establish the optimality of a base stock policy. (In terms of the backlogging costs, the authors assume that they are proportional to the amount of time a demand has been backlogged, as in standard inventory models.) Gupta and Wang [21] also considers a special case of the continuous review model considered here: a Poisson demand process with linear order costs and linear backlogging costs. There, the authors *assume* that the optimal policy is of a base stock, or $(r, 1)$-type and develop a method for the evaluation of any given base stock policy. In their concluding section, the authors present, as a future challenge, the identification of the "structure of the optimal policy when the retailer incurs a fixed ordering cost". That question is resolved in our paper, not just under Poisson demands, but under general renewal demand processes.

Huh et al. [23] address a periodic review model with general delay-dependent backlogging costs, but regular linear holding costs, and with unmet demand fully backlogged. The authors show that, under i.i.d. demands and fixed-plus-linear order costs, as in this paper, an $(s, S)$-policy is optimal. This structural result is obtained under a *restriction* guaranteeing either that no demand is delayed by more than the leadtime plus one periods, or that the incremental backlogging cost rate is constant for delays in excess thereof. (We show that our policy restriction is significantly weaker than the first restriction in Huh et al. [23].) Again, focusing on periodic review models, but allowing full or partial backlogging of stockouts, or lost sales, Federgruen and Wang [18] combine these types of delay-dependent cost functions with general shelf age-dependent holding cost structures. That paper establishes the structure of an optimal policy in a variety of periodic review models. The analysis approaches in Huh et al. [23] and Federgruen and Wang [18] are entirely distinct from the one employed in this paper. For example, the approach in Huh et al. [23] is to show that the part of the end-of-period backlog that is of a given age, may be expressed as a function of the inventory position a leadtime earlier, as well as the intermittent demands. An analogous approach, decomposing the end-of-the-period inventory on the basis of the units' shelf age, fails to work, however, in our continuous review model.

Most continuous review inventory models confine themselves to the case of Poisson or compound Poisson demands where the interarrival times have an exponential distribution. Recent examples include Chao, Xu, and Yang [8], Chao and Zhou [9], Xu and Chao [42] and Shi, Katehakis, and Melamed [37]. Other authors have assumed that the demand process is a Brownian motion; see, for example, Bather [4], Browne and Zipkin [7] and Dai and Yao [13]. Several authors have considered general renewal demand processes in their papers; see, for example, Tijms [40], Sahin [34,35], Federgruen and Schechner [15], Zipkin [44], Rosling [33] and Chen and Simchi-Levi [12]. As mentioned, in Section 5, we consider an even more general demand model where the inter-demand times depend on a state-of-the-world variable, assumed to fluctuate in accordance with an independent Markov process. This model has been studied by several authors, both in periodic and continuous review settings, going back to the seminal papers by Karlin [25] and Iglehart and Karlin [24]. Song and Zipkin [38] discuss how several important considerations, such as product demands that are sensitive to underlying economic conditions, and product demand rates that fluctuate in accordance with the stage in the product life cycle, can be modeled in this general framework. Benjaafar, Cooper, and Mardan [5] and Gayon, Benjaafar, and Véricourt [20] invoke this model, with the state-of-the-world given by the number of advance customer orders at various stages. To our knowledge, all existing inventory models with Markov-modulated demands assume standard, that is, level-dependent inventory carrying and backlogging costs.

## 3. PRELIMINARIES: MARKOV-MODULATED RENEWAL PROCESS

While our base model assumes that demands are generated by a renewal process, we extend our results to settings where the inter-demand times are no longer identically and independently distributed. More specifically, we assume that they depend on a state-of-the-world variable, assumed to fluctuate in accordance with an independent Markov process.

Let $\{W(t), t \geq 0\}$ denote the independent Markov process, according to which the state of the world $W(t) = i$, evolves, with transition rate matrix $Q$. Let $\mathcal{W}$ denote its finite state space and assume that the Markov chain embedded on its state transition epochs has a single recurrent class of states. If the state of the world at a given demand epoch $t$ equals $W(t) = i$, the interval until the next demand epoch is distributed like a random variable $X(i)$, $i \in \mathcal{W}$. For general inter-demand distributions, we assume that the distribution of the inter-demand

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press

time only depends on the state of the world at its initiation. An important special case arises when all inter-demand times are exponentially distributed. This gives rise to a so-called Markov-modulated Poisson process. For the latter, it is usually assumed that the inter-demand distribution is affected by *any* change in the state of the world, that is, *whenever* the state-of-the-world switches to a state $i$, the remaining time until the next demand epoch is distributed like the exponential $X(i)$. Indeed, for Markov-modulated Poisson processes, the alternative assumption of *continuous* dependence of the inter-demand time on the prevailing state of the world may easily be accommodated in our analysis without affecting any of the structural results and without complicating the required conditions.

The state process, embedded on demand epochs, has a transition probability matrix $P \equiv [p_{ii'}]$, where $p_{ii'} = Pr[W(X(i)) = i' | W(0) = i]$. Let $\lambda = \{\lambda_i, i \in \mathcal{W}\}$ denote its steady-state distribution. The steady-state distribution of the embedded Markov chain with probability transition matrix $P$ is, in general, *different* from the time average stationary distribution of the Markov process. This can be seen from the following simple example of a two-state Markov chain: Assume the Markov process alternates between states 1 and 2, that is, $I = \{1, 2\}$. Let the holding time in state 1(2) be exponential with mean 1(2). Under the time average stationary distribution of the Markov process, the latter resides in state 1 with probability 1/3 and in state 2 with probability 2/3. Let $X(1)$ be deterministic with length 0.5 and $X(2)$ be deterministic with length 2.5. The steady-state distribution of the Markov process embedded on demand epochs, has $\lambda_1 > \lambda_2$, contrary to the time-average stationary distribution.

The transition probabilities $\{p_{ii'} : i, i' \in \mathcal{W}\}$ may be computed by first determining the transient distributions $\{q_{ii'}(t) : t \geq 0\}$ as the solution of a system of linear differential equations, the so-called Kolmogoroff's forward differential equations; see, for example, Theorem 4.5.1 in Tijms [41]. See there, for a discussion of numerical methods for the solution of these differential equations. Mixing the transient probability functions $\{q_{ii'}(t), t \geq 0\}$ with the distribution of the $\{X(i), i \in \mathcal{W}\}$ variables enables us to compute the transition probability matrix $P$.

In the analysis below, we often need to know the distribution of $\Xi_j$, the amount of time elapsed between a given demand epoch and the $j$th subsequent demand epoch, both when $j \geq 1$ and when $j < 0$. (When $j < 0$, $\Xi_j$ denotes how much time has elapsed since the $(-j)$th preceding demand epoch.) $\stackrel{d}{=}$ denotes "equality in distribution".

LEMMA 1: *Let* $j \geq 1$.

(a) *The distribution of* $(\Xi_j \mid W(0) = i), i \in \mathcal{W}$, *may be computed recursively as follows: Assume the distributions* $(\Xi_{j-1} \mid W(0) = i')$ *are known.* $(\Xi_j \mid W(0) = i)$ *is the convolution of the random variable* $X(i)$ *and a mixture of the random variables* $\{\Xi_{j-1} \mid W(0) = i'\}$ *with mixing probabilities* $\{p_{ii'} : i' \in \mathcal{W}\}$ *from the matrix* $P$.

(b) $(\Xi_{-j} \mid W(0) = i) \stackrel{d}{=} (\Xi_j \mid W^-(j) \mid W(0) = i)$ *where* $W^-(j)$ *denotes the state-of-the-world* $(-j)$ *renewal epochs before time 0,*

$$\Pr[W^-(j) = i' \mid W(0) = i] = \frac{p^{(j)}_{i'i}\lambda_{i'}}{\sum_{w \in \mathcal{W}} p^{(j)}_{wi}\lambda_w}, \tag{1}$$

*and* $P^{(j)}$ *denotes the* $j$*th power of the transition probability matrix* $P$.

PROOF: Part $(a)$ is immediate, as is the characterization $(\Xi_{-j} \mid W(0) = i) \stackrel{d}{=} (\Xi_j \mid W^-(j) \mid W(0) = i)$ in part $(b)$. It thus suffices to verify the conditional probabilities in Eq. (1).

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press

However, the latter is a direct application of Bayes Law, recognizing that the unconditional steady-state probabilities of the state of the world are given by the vector $\lambda$. ■

Note that, in part $(b)$ of Lemma 1, the conditional distribution $(\Xi_j \mid W^{-}(j))$ can be computed as described in part (a). When the demand process is renewal, the characterization of the intervals $\{\Xi_j\}$ simplifies, majorly, as a direct corollary of Lemma 1:

Corollary 1: *When the demand process is a simple renewal process with all inter-demand times distributed like the random variable $X$, $\Xi_j$ is distributed like the $\mid j \mid$th convolution of $X$.*

## 4. THE BASE MODEL

In our base model, we assume that demands are generated by a renewal process, that is, the interarrival times between demand epochs are i.i.d. random variables $X_1, X_2, \ldots$, distributed like $X$ with mean $\tau$; without loss of generality, time is measured such that $\tau = 1$. A single unit is demanded at each demand epoch. In Section 5, we generalize our results to settings where the times between consecutive demand epochs fail to be identical and independent, because they depend on an underlying state variable which evolves according to an independent Markov process. We assume that stockouts are backlogged.

Any order incurs a fixed cost $K$ and a variable per unit cost $c$. It arrives after a leadtime $L$ which is assumed to be constant or, more generally, characterized by a stochastic process $\{L(t) : t \geq 0\}$ with $L(t)$ the leadtime experienced by an order placed at time $t$. We assume the leadtime process is *exogenous*, that is, it is independent of the demand process, as well as *sequential*, that is, $t + L(t) \leq t' + L(t')$ for all $t < t'$, with probability one. Under sequential leadtime processes, orders do not cross. Let $\mathcal{L}$ denote a random variable distributed like the steady-state distribution of the process $\{L(t) : t \geq 0\}$. We refer to Zipkin [45] for an extensive discussion of such processes and their applications.

The shelf age-dependent holding and delay-dependent backlogging cost structure is characterized by two functions $\alpha(\cdot)$ and $\beta(\cdot)$. $\alpha(t)$ denotes the *marginal* inventory cost rate, incurred for an item that has a shelf age $t$, and $\beta(t)$ the marginal backlogging cost rate, when a unit of demand has been waiting for $t$ time units. (The functions $\alpha(\cdot)$ and $\beta(\cdot)$ are defined on the full real line $\mathcal{R}$, with the understanding that $\alpha(s) = \beta(s) = 0$, when $s < 0$.)

Let $H(t) \equiv \int_0^t \alpha(u)\, du$ denote the *total* carrying costs incurred for a unit that remains in stock for $t$ time units and $J(t) \equiv \int_0^{-t} \beta(u)\, du$ the *total* backlogging costs for a demand unit filled with a delay of $-t$ time units. The functions $\alpha(\cdot)$ and $\beta(\cdot)$ are assumed to be *increasing*. The monotonicity of $\alpha(\cdot)$ reflects the fact that the rate at which capital and maintenance costs of inventories are accrued increases as the inventory becomes older. Similarly, the out-of-pocket and goodwill losses associated with delayed deliveries are incurred at an increasingly larger rate, explaining the monotonicity of $\beta(\cdot)$. The monotonicity properties are satisfied in all application settings reviewed in the Introduction. Together with the fact that orders do not cross, monotonicity of $\alpha(\cdot)$ and $\beta(\cdot)$ guarantees that it is optimal to deplete inventories and to clear backlogs on a FIFO base. (On any given sample path, if a pair of supply units $u_1$ and $u_2$ are ordered at time $s_1 \leq s_2$, but $u_1$ is assigned to fill a later demand unit than $u_2$, a cost saving can be achieved by interchanging the assignments.)

Our objective is to minimize long-run average costs, considering the SMDP that arises when embedding the system process on the demand epochs. As with all SMDPs, control of the system is restricted to the sequence of time epochs on which the process is embedded (in this case, the demand epochs); more specifically, an order may be placed at any demand epoch, but not between demand epochs. When the interarrival times of the demands are exponential, that is, when the demand process is Poisson, this restriction is without loss of generality, due to the memorylessness property of the exponential distribution. For general inter-demand time distributions, this restriction is, potentially, with some loss of optimality, in as much as it is standard in all semi-Markov decision problems, in general, and all of the aforementioned inventory models with renewal demands, in particular.

The state space $\Sigma$ of the SMDP is infinite dimensional with the following list of state components:

$$\begin{aligned} I &= \text{inventory in stock;}\\ a_l &= \text{shelf age of the } l\text{th item in stock}, l = 1, \ldots, I;\\ B &= \text{backlog size;}\\ d_l &= \text{delay experienced, thus far, by the } l\text{th most recently backlogged and}\\ &\quad\ \text{still unsatisfied demand unit,}\\ &\qquad l = 1, \ldots, B;\end{aligned}$$

for $k = 1, 2, \ldots,$

$$z_k = \begin{cases} \text{size of the } k\text{th most recent order placed,} & \text{if still in process,}\\ 0, & \text{otherwise;}\end{cases}$$

$$t_k = \begin{cases} \text{time elapsed since the placement of the } k\text{th} & \\ \text{most recent order placed,} & \text{if still in process,}\\ 0, & \text{otherwise.}\end{cases}$$

Let $x = I - B + \sum_{l=1}^{\infty} z_l$ denote the inventory position. Note that since $\alpha(\cdot) \geq 0$ and $\beta(\cdot) \geq 0$, $IB = 0$, that is, it is never beneficial to have a positive inventory along side a positive backlog.

Let $U = \{z \geq 0 : z \text{ is integer}\}$ denote the action set of possible order sizes which may be selected at any state $\sigma \in \Sigma$. As in standard dynamic program textbooks (see, e.g., Bertsekas [6]), we define a policy $\pi$ as a sequence of functions $\pi = (\mu_1, \mu_2, \ldots)$ where each $\mu_t : \Sigma \to U$ maps the state $s$ into an action $\mu_t(\sigma) \in U$. Let $\Pi$ be the set of all policies.

This SMDP with its infinite-dimensional state space is completely intractable. We show that, under *a single minor restriction of the policy space*, this intractable SMDP may be replaced by an SMDP with a one-dimensional state space, referred to as the *transformed SMDP*, which is *equivalent* in the sense that any policy that is optimal in the latter is optimal in the original SMDP as well. The transformed SMDP uses the *inventory position* as the single state variable, similar to the treatment of *standard* inventory models with positive leadtimes, that is, models in which inventories do not need to be differentiated based on the items' shelf age or backlogs based on the delay durations. Moreover, an $(r, q)$-policy is shown to be optimal in the transformed SMDP and, hence, in the original SMDP, in view of the equivalency between the two models.

To introduce the restriction, we first need the following definition.

Definition: *A pair of states $\sigma'$ and $\sigma \in \Sigma$ are defined to be similar if they have the same state components, with the possible exception of the vector $d$ of delay durations.*

Similarity is an equivalency relationship and partitions the state space $\Sigma$ into equivalency classes. Note that similar states have the same inventory position.

*Restriction* equal orders for similar states with negative inventory positions (EOSSNIP) Assume policy $\pi$ prescribes an order of size $z$ on some state $\sigma \in \Sigma$, such that $x(\sigma) + z < 0$. Then policy $\pi$ prescribes the same order size $z$ in all states that are similar to $\sigma$.

This policy restriction may come at some loss of optimality: while, in general, the policy *may* differentiate the order size among similar states, based on the vector of delay durations, such differentiation is precluded when the inventory position after ordering remains negative. We note that a considerably *more* restrictive assumption is imposed by Huh et al. [23] in their treatment of periodic review models with delay-dependent (marginal) backlogging costs. The authors require the inventory position after ordering to be *non-negative*, see their remark after Lemma 3.1, in which case the EOSSNIP imposes *no restriction*, whatsoever. Under Restriction EOSSNIP, negative inventory positions after ordering are permitted; however, as long as a negative inventory position is maintained after ordering, the policy may not discriminate on the basis of the specific vector of delay durations $d$.

Allowing for a negative inventory position after ordering is important: as we demonstrate, an $(r, q)$-policy acting on the inventory position is optimal; however, when the fixed order cost $K$ is sizable relative to the backlogging costs, an optimal reorder point $r^* < 0$ may arise: this means that modest negative inventory positions of less than $r^*$ units are optimally tolerated. Let $\tilde{\Pi}$ denote the class of policies that satisfy EOSSNIP. Finally, consider a further restriction $\Pi^0$ of the policy class, consisting of all policies that prescribe order sizes exclusively based on the current inventory position and, possibly, the index of the current time period. Clearly,

Lemma 2: $\Pi^0 \leq \tilde{\Pi} \leq \Pi$.

We next show that an optimal policy in the class $\tilde{\Pi}$ can be found within the considerably smaller and simpler policy space $\Pi^0$. To this end, we employ a single-unit matching approach, first introduced by Axsäter [2,3] and subsequently used by many others, for example, Muharremoglu and Tsitsiklis [27]. More specifically, observe that every demand unit is filled by a *specific* supply unit and all inventory and delay costs associated with that demand unit are a function of the time elapsed between the arrival of the "assigned" supply unit and the epoch at which the demand unit is requested. More specifically, assume time $t = 0$ corresponds with a demand epoch, after which the system remains empty, that is, without inventory, backlogs or outstanding orders. Since, as demonstrated, ordered units are optimally used on a FIFO basis, the $j$th ordered unit (since time 0) is used to fill the $j$th demand unit (again, since time 0). Without loss of optimality, we may restrict ourselves to policies in $\tilde{\Pi}$ with a long-run average ordering rate, at least equal to the average demand rate. (Otherwise, the policy has infinitely large backlogs in steady state, and infinite long-run average costs, since $\beta^+ \equiv \lim_{t\to\infty} \beta(t) > 0$.) This implies that every demand unit is matched to a supply unit, ordered within a finite amount of time from the demand epoch. This implies that if an order is placed that elevates the inventory position from $x$ to $y > x$ at a given demand epoch, the $(y - x)$ units in the order may be given an index $j = x + 1, x + 2, \ldots, y$, such that the $j$th item is used to satisfy the $j$th $((-j)$th) demand following (preceding) the order epoch if $j > (\leq) 0$.

Theorem 1: *There exists an optimal policy in the class $\tilde{\Pi}$ which belongs to $\Pi^0$, that is, it prescribes orders strictly on the basis of the prevailing inventory position.*

Proof: Consider first any ordering decision which elevates the inventory position to $y \geq x(\sigma)$, with $y \geq 0$, at an arbitrary demand epoch with a given system state $\sigma \in \Sigma$,

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press

and associated inventory position $x(\sigma)$. The expected inventory costs associated with the units ordered are given by:

$$\begin{cases} \sum_{u=x(\sigma)+1}^{y} [\mathbb{E}H(\Xi_u - \mathcal{L}) + \mathbb{E}J(\Xi_u - \mathcal{L})], & \text{if } x(\sigma) \geq 0, \\ \mathbb{E}\{\sum_{j=1}^{-x(\sigma)} J(d_j + \mathcal{L}) + \sum_{u=0}^{y} [\mathbb{E}H(\Xi_u - \mathcal{L}) + \mathbb{E}J(\Xi_u - \mathcal{L})], & \text{if } x(\sigma) < 0. \end{cases} \quad \textbf{(2)}$$

Note that these matching costs are independent of the vector of shelf ages of the items in stock, $(a_1, \ldots, a_I)$, or the vector of outstanding order sizes, $(z_1, z_2, \ldots)$, or the timing of the orders, $(t_1, \ldots, t_I)$, and so are all matching costs incurred at future demand epochs. Therefore, nothing is gained by making the order size decision dependent on this part of the state vector.

When $x(\sigma) \geq 0$, the current and all future inventory costs are also independent of the vector $d$ of backlog durations since, when $x(\sigma) \geq 0$, ordered units are matched with demand units, yet to arise. Such dependence does arise when $x(\sigma) < 0$. However, it follows from Eq. (2) that, for any possible order that results in an inventory $y \geq 0$, the vector $d$ impacts only the constant term $\{J(d_1 + \mathcal{L}), J(d_2 + \mathcal{L}), \ldots, J(d_{-x} + \mathcal{L})\}$ and has no impact on any matching costs incurred at future decision epochs. We conclude that when considering ordering decisions that result in an inventory position $y \geq 0$, nothing is gained by making the decision dependent on *any* part of the state information, other than the current inventory position $x(\sigma)$.

It remains to be shown that the same invariance result prevails when considering order sizes that result in $y < 0$. Such ordering decisions are only relevant if $x(\sigma) \leq y < 0$, in which case the matching costs are given by

$$\sum_{l=-y+1}^{-x} \mathbb{E}J(d_l + \mathcal{L}). \quad \textbf{(3)}$$

Once again, the vector of shelf ages $a$, the vector of outstanding order sizes $z$ and the vector of order times $t$ have *no* impact on the current or any future matching costs; therefore, even when considering possible order sizes resulting in $y < 0$, nothing is gained by making decisions dependent on those parts of the state information, while Restriction EOSSNIP for the policy class $\tilde{\Pi}$ precludes differentiation of the order size decisions based on the vector $d$, either.

We conclude that any order size decision is optimally based on the inventory position $x(\sigma)$, solely, that is, an optimal policy can be found in the class $\Pi^0$. ■

*Optimality of an $(r, q)$-policy*

In view of Theorem 1, we may confine ourselves to the policy class $\Pi^0$. The proof of Theorem 1, in particular expressions Eqs. (2) and (3), show that the expected matching costs incurred when elevating the inventory position $x(\sigma)$ to $y \geq x(\sigma)$, depend on the state vector $\sigma$, only via the vector $d$ of backlog durations, and this of course only when $x(\sigma) < 0$. Each of the duration components $d_j$, $j = 1, \ldots, B$, in this vector $d$ at an arbitrary demand epoch, is simply $\Xi_{-j}$ which, based on Corollary 1 in Section 3, is distributed like the $j$th convolution of $X$. Thus, the *expected* long-run average cost of any policy $\pi \in \Pi^0$ may be determined by replacing in the matching cost expressions Eqs. (2) and (3) the vector $d = (d_1, d_2, \ldots)$ by the random variables $(\Xi_{-1}, \Xi_{-2}, \ldots)$ from which they are drawn. This substitution, also, allows for a uniform representation of the expected matching cost of an ordered unit with index $j$, whether $j \geq 0$, or $j < 0$. Let $A_j$ denote the difference between the arrival time of the demand unit matched to ordered item $j$ and that of item $j$. If $A_j$ is positive, it represents the shelf age of item $j$. If it is negative, the $j$th item in the ordered batch arrives after the associated demand epoch, so that this demand unit experiences a backlog time equal to $-A_j$. Clearly,

$$A_j = \begin{cases} \Xi_j - \mathcal{L} & \text{when } j \geq 0, \\ -\Xi_j - \mathcal{L} & \text{when } j < 0. \end{cases} \quad \textbf{(4)}$$

The total expected inventory and backlogging costs associated with the $j$th ordered unit are thus given by

$$\hat{G}(j) = \mathbb{E}H(A_j) + \mathbb{E}J(A_j), \text{ for any integer } j. \quad \textbf{(5)}$$

*Assumption*: $\hat{G}(j) < \infty$ for any integer $j$.

The Assumption is satisfied, for example, if the functions $\alpha(\cdot)$ and $\beta(\cdot)$ are polynomially bounded and the interarrival time distribution $X$ and the leadtime distribution $\mathcal{L}$ have finite moments.

With the uniform expression $\hat{G}(j)$ being the expected matching costs for the $j$th ordered unit, we have that under any policy $\pi \in \Pi^0$, at any demand epoch, both the expected immediate inventory costs and the transition dynamics only depend on the prevailing inventory position $x(\sigma)$ rather than the full vector $\sigma \in \Sigma$ itself. In other words, for policies in $\Pi^0$, a single-dimensional state description—via the inventory position—suffices, and, by Theorem 1, this class of policies contains an optimal policy within the much broader class $\tilde{\Pi}$. The optimal policy in $\Pi^0$ can thus be found in an SMDP with the one-dimensional state space $\Sigma^0 \equiv \mathcal{Z}$ where $\mathcal{Z}$ denotes the set of integers, action sets $U^0(x) = \{y \geq x : y \in \mathcal{Z}\}$ and one-step expected cost functions:

$$\gamma(x, y) = K\delta(y - x) + c(y - x) + \sum_{j=x+1}^{y} \hat{G}(j), \quad \textbf{(6)}$$

where

$$\delta(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

The first two terms in Eq. (6) represent the fixed and variable procurement costs while the last one denotes the expected carrying and delay costs associated with an order that elevates the inventory position from $x$ to $y \geq x$, according to the above described matching scheme. Finally, when the inventory position is increased to the level $y$, the state at the next decision epoch equals $(y - 1)$ with probability one.

We now prove our main result, that is, the optimality of an $(r, q)$-policy. However, we first need the following Lemma:

Lemma 3: *The function $\hat{G}(j)$ is convex in $j$.*

Proof: A sequence of random variables $\{E(\theta) : \theta \in \Theta\}$ is stochastically increasing linear (SIL) if $\{E(\theta) : \theta \in \Theta\}$ is stochastically increasing and $\mathbb{E}f(E(\theta))$ is increasing convex in $\theta$ for all increasing convex functions $f(\cdot)$, and is increasing concave in $\theta$ for all increasing concave functions $f(\cdot)$. The family of distributions $\{A_j : j \geq 0\}$ is SIL in $j$; see Example 8.A.16 in Shaked and Shanthikumar [36]. Similarly, we can easily verify that $\{A_j : j \in \mathcal{Z}^-\}$ is SIL in $j$ using Example 8.B.7 and Theorem 8.B.9 there. This implies that $\mathbb{E}H(A_j)$ is a convex function since $H(\cdot)$ is increasing and convex in view of the monotonicity property of $\alpha(\cdot)$, and see Definition 8.A.1(e) in Shaked and Shanthikumar [36]. Similarly, $\mathbb{E}J(A_j) = -\mathbb{E}(-J(A_j))$

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press

is convex in $y$ since $\mathbb{E}(-J(A_j))$ is a concave function of $j$, as $-J(\cdot)$ is increasing and concave in view of the monotonicity property of $\beta(s)$. ■

THEOREM 2:

(a) *An $(r,q)$-policy is optimal among all policies in the class $\tilde{\Pi}$.*

(b) *The long-run average cost under any $(r,q)$ policy is given by*

$$c(r,q) = c + \frac{K + \sum_{j=r+1}^{r+q} \hat{G}(j)}{q}. \tag{7}$$

PROOF: (a) In view of Theorem 1, it suffices to show that an $(r,q)$-policy is optimal within the smaller class of policies $\Pi^0$. Define

$$G(y) \equiv \begin{cases} \displaystyle\sum_{j=1}^{y} \hat{G}(j), & \text{if } y > 0, \\ \displaystyle -\sum_{j=y+1}^{0} \hat{G}(j), & \text{if } y \leq 0, \end{cases}$$

with the convention that $\sum_{j=a}^{b} u(j) = 0$ when $a > b$ for any sequence $\{u(j)\}$. (One can easily verify that, with this definition, $G(y) - G(x) = \sum_{j=x+1}^{y} \hat{G}(j)$, regardless of the signs of $x$ and $y$.)

To show that a particular policy is optimal, it suffices, under certain conditions, to show that the policy prescribes actions which satisfy the minima for a given solution of the model's optimality equation. The long-run average optimality equation in the one-dimensional SMDP associated with this class $\Pi^0$ is given by

$$\begin{aligned} v(x) &= \min_{y \geq x} \{\gamma(x,y) - g + v(y-1)\} \\ &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + G(y) - G(x) - g + v(y-1)\}, x \text{ is integer}, \end{aligned} \tag{8}$$

where $g$ denotes the minimum long-run average cost in the SMDP. Adding $G(x)$ to both sides of the optimality equation and defining $\hat{v}(x) \equiv v(x) + G(x)$, we obtain the following transformed optimality equation in terms of $\{\hat{v}(\cdot), g\}$:

$$\begin{aligned} \hat{v}(x) &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + G(y) - G(y-1) - g + \hat{v}(y-1)\} \\ &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g + \hat{v}(y-1)\}, x \text{ is integer}. \end{aligned} \tag{9}$$

Equation (9) may be interpreted as a variant of the optimality equation in the classical periodic review inventory model, with fixed-plus-linear ordering costs and immediate expected cost $\hat{G}(y)$, whenever the inventory position after ordering equals $y$. By Lemma 3, $\hat{G}(\cdot)$ is convex, that is, $-\hat{G}(\cdot)$ is unimodal. It follows from Zheng [43] that an $(s^*, S^*)$-policy is optimal in this periodic review MDP. Since at every demand epoch a *unit* size demand occurs, this $(s,S)$ policy is of an $(r,q)$-type, say with reorder level $r^* = s^*$ and fixed order size $q^* = S^* - s^*$. Let $g^*$ denote its long-run average costs.

We now show that this $(r^*, q^*)$-policy is also optimal in the SMDP with optimality Eq. (8), that is, the SMDP associated with the policy class $\Pi^0$. Let $\pi = (\mu_1, \mu_2, \ldots) \in \Pi^0$ denote an arbitrary policy. Given a specific starting state $x_1$, let $\{x_t : t > 1\}$ denote the sequence of states adopted under the policy $\pi$. Note that in period $t$, the inventory position is raised from $x_t$ to $y_t \equiv x_{t+1}+1$. Since the $(r^*,q^*)$-policy is optimal (among all Markov strategies) in the transformed periodic review model, we have

$$\begin{aligned}
g^* &\le \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_\pi[K\delta(y_t-x_t)+c(y_t-x_t)+\hat{G}(y_t)] \\
&= \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_\pi[K\delta(y_t-x_t)+c(y_t-x_t)+(G(y_t)-G(y_t-1))] \\
&= \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_\pi[K\delta(y_t-x_t)+c(y_t-x_t)+(G(y_t)-G(x_{t+1}))] \\
&= \liminf_{T\to\infty} \frac{1}{T}\left\{\sum_{t=1}^{T} \mathbb{E}_\pi[K\delta(y_t-x_t)+c(y_t-x_t)+(G(y_t)-G(x_t))] - \mathbb{E}_\pi G(x_{T+1})\right\} \\
&\le \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_\pi[K\delta(y_t-x_t)+c(y_t-x_t)+(G(y_t)-G(x_t))]. \qquad \textbf{(10)}
\end{aligned}$$

This establishes that the long-run average cost under any policy $\pi$ is bounded from below by $g^*$. (To verify the last equality, note that $\lim_{T\to\infty}((G(x_1))/T)=0$. To verify the last inequality, note that $G(\cdot)\ge 0$.)

Similarly, when the $(r^*,q^*)$-policy is selected for the policy $\pi \in \Pi^0$, both inequalities in Eq. (10) become equalities, so that $g^*$ denotes the long-run average cost of the $(r^*,q^*)$ policy in the SMDP with optimality Eq. (8). (When $\pi$ is the $(r^*,q^*)$-policy, the first inequality in Eq. (10) is an equality because $g^*$ denotes the long-run average cost of the $(r^*,q^*)$-policy in the periodic review model. The last inequality in Eq. (10) holds as an equality because $\lim_{T\to\infty}((G(x_{T+1})/T)=0$, as $r^*\le x_{T+1}\le r^*+q^*$.) This proves that the $(r^*,q^*)$-policy is optimal, in our SMDP, among all policies in $\Pi^0$. We conclude that the $(r^*,q^*)$ policy is indeed optimal among *all* policies in the original model.

(b) The proof of part (a) shows that the long-run average cost of any given $(r,q)$-policy is the same as that in the equivalent periodic review model. It is easily verified that the steady-state inventory position after ordering is uniformly distributed on the integers $\{r+1,\dots,r+q\}$. It follows that the long-run average cost $c(r,q)$ is given by Eq. (7). ■

The proof of Theorem 2 by itself does not establish that the optimality Eq. (8) has a solution and that the optimal $(r^*,q^*)$-policy prescribes actions that achieve the minimum in Eq. (8), for every state $x$. Proposition 1, the proof of which is deferred to the Appendix, establishes these results.

Proposition 1: *The optimality Eq. (8) of the SMDP has a solution $\{v^*(\cdot),g^*\}$ and the optimal $(r^*,q^*)$-policy achieves the minimum in Eq. (8) for any starting state $x$.*

Given the cost representation in Eq. (7) and since the function $\hat{G}$ is convex (see Lemma 1), the optimal policy can be computed efficiently with the algorithm in Federgruen and Zheng [19], see also Zipkin [45] for a description of this algorithm. Finally, it is of interest to characterize how various model parameters such as the shape of the marginal inventory cost rate function $\alpha(\cdot)$, that of the marginal backlogging cost rate function $\beta(\cdot)$ and the leadtime or interarrival time distribution impact the optimal policy parameters $r^*$ and $R^*\equiv r^*+q^*$. To investigate the impact of any of these model primitives $\theta$, we write the one-step expected cost function as $\hat{G}(y \mid \theta)$. In the first two examples, $\theta$ is a real-valued function in the space $\Theta$ of all increasing, non-negative functions, which we endow with the partial order implied by point-wise dominance, that is, $\theta_1 \preceq \theta_2$ if and only if $\theta_1(t) \leq \theta_2(t)$ for all $t \geq 0$; in the last two examples, $\theta$ is an element of the space $\Theta$ of all distributions of non-negative random variables, endowed with the $\leq_{st}$ partial order or the $\leq_{cx}$ order. (For any pair of random variables $X$ and $Y$, $X \leq_{cx} Y$ means that $\mathbb{E}f(X) \leq \mathbb{E}f(Y)$ for any convex function $f(\cdot)$; $X \leq_{cx} Y$ implies that $\mathbb{E}(X) = \mathbb{E}(Y)$, while $\mathrm{Var}(X) \leq \mathrm{Var}(Y)$.) We define a function $F(\cdot \mid \theta)$ to have *increasing (decreasing) differences* if $F(y_2 \mid \theta) - F(y_1 \mid \theta)$ is increasing (decreasing) in $\theta \in \Theta$ for all $y_1 < y_2$.

PROPOSITION 2:

(a) *When the incremental inventory cost rate function $\alpha(\cdot)$ is replaced by a new function $\hat{\alpha}(\cdot)$ that is point-wise larger, that is, $\alpha(t) \leq \hat{\alpha}(t)$ for all $t \geq 0$, the optimal values $r^*$ and $R^*$ decrease.*
(b) *When the incremental backlogging cost rate function $\beta(\cdot)$ is replaced by a new function $\hat{\beta}(\cdot)$ that is point-wise larger, that is, $\beta(t) \leq \hat{\beta}(t)$ for all $t \geq 0$, the optimal values $r^*$ and $R^*$ increase.*
(c) *When the leadtime distribution $\mathcal{L}^1$ is replaced by $\mathcal{L}^2 \geq_{st} \mathcal{L}^1$, the optimal values $r^*$ and $R^*$ increase.*
(d) *Assume the functions $\alpha(\cdot)$ and $\beta(\cdot)$ are convex. When the interarrival time distribution $X^1$ is replaced by $X^2 \geq_{cx} X^1$, the optimal values $r^*$ and $R^*$ decrease.*

Thus, while both $r^*$ and $R^*$ are guaranteed to increase when the leadtime distribution increases in the general $\leq_{st}$ ordering sense, we obtain the opposite monotonicity properties with respect to the inter-demand time distribution, only when the distribution increases in the more restrictive $\leq_{cv}$ ordering. (The latter implies in particular that $\mathbb{E}(X^1) = \mathbb{E}(X^2)$.) One reason behind that restriction is that, if $\mathbb{E}(X^1) \neq \mathbb{E}(X^2)$, the average demand rate changes and its impact on the optimal policy parameters is ambiguous, a phenomenon already noted in the standard inventory model, see Federgruen and Wang [17]. Proposition 2 covers all of the model primitives with the exception of the fixed cost $K$. For the latter, Zheng [43] already showed that an increase of $K$ results in an increase of $R^*$ and a decrease of $r^*$. (The *variable* order costs are *independent* of the choice of the policy parameters; in other words, $r^*$ and $R^*$ are independent of $c$.)

## 5. MARKOV-MODULATED RENEWAL PROCESSES

We now return to the general demand model described in Section 3, where demands are generated by a Markov-modulated renewal process. In this case, the state of the system in the original SMDP, includes all of the components listed in Section 4, as well as the state of the "world" $W(t) = w \in \mathcal{W}$, which evolves according to the continuous-time Markov chain with transition rate matrix $Q$. Restriction EOSSNIP confines the policy space to $\tilde{\Pi}$. A straightforward extension of Theorem 1 establishes that an optimal policy within $\tilde{\Pi}$ may be found in the much smaller policy space $\Pi^0$, now defined as all policies which prescribe an order size, based exclusively on this prevailing inventory position, the index of the current time period, as well as the state of the world $w$.

Moreover, analogous to the derivation in Section 3, we show that for policies $\pi \in \Pi^0$, a two-dimensional state description – via the inventory position $x(\sigma)$ and the state of the world $W$ – suffices. The optimal policy in $\Pi^0$ can thus be found in an SMDP with the two-dimensional state space $S^0 = Z \times \mathcal{W}$ and action set $U^0(x, w) = \{y \geq x : y \in \mathcal{Z}\}$. The one-step expected cost when at an arbitrary demand epoch with a prevailing inventory position $x(\sigma)$ and state of the world $w$, an order is placed to elevate the inventory position to $y \geq x(\sigma)$ units is given at follows:

$$\gamma(x, y, w) = K\delta(y - x) + c(y - x) + \sum_{j=x+1}^{y} \hat{G}_w(j), \tag{11}$$

where $\hat{G}_w(j)$ denotes the expected matching costs for the ordered unit with index $j$, which is matched with the $j$th (-$j$th) demand following (preceding) the current demand epoch, when the current state of the world is $w \in I$. Analogous to the function $\hat{G}(\cdot)$ in the basic model, we have $\hat{G}_w(j) = \mathbb{E}[H(A_j(w)) \mid W(0) = w] + \mathbb{E}[J(A_j(w)) \mid W(0) = w]$, and $A_j(w)$ again refers to the (positive or negative) *age* of the supply unit with index $j$, at the time it satisfies its matched demand unit. More specifically, for $j \geq 0$, let

$$A_j(w) = [\Xi_j \mid W(0) = w] - \mathcal{L}, \tag{12}$$

where the expression within squared brackets denotes the sum of $j$ consecutive inter-demand times, given that $W(0)$, the state of the world at the current order epoch, is $w$. When $j < 0$,

$$A_j(w) = -[\Xi_j \mid W(0) = w] - \mathcal{L}. \tag{13}$$

See Section 3 for characterizations of the distributions of the random variables $(\Xi_j \mid W(0) = w)$ and $(\Xi_j \mid W(0) = w)$ as well as an efficient procedures to compute their distributions.

It follows, again, from the proof of Lemma 3 that the functions $\hat{G}_w(\cdot)$ are convex for all $w \in \mathcal{W}$. By the proof of Theorem 2, our model is therefore equivalent to a periodic review inventory model with state-dependent holding and backlogging cost functions $\hat{G}_w(\cdot)$. It follows from Chen and Song [10], Section 4, that a state-dependent $(s, S)$-policy is optimal. Since the demand process is a point process, this implies that a state-dependent $(r, q)$-policy is optimal. We conclude:

THEOREM 3: *In the model with a Markov-modulated demand process, a state-dependent $(r, q)$-policy is optimal among all policies in $\tilde{\Pi}$.*

## 6. CONCLUSIONS AND EXTENSIONS

We have discussed a continuous-time inventory model in which demands are generated by a counting process. The distinguishing feature of the model is that the marginal inventory carrying cost of a unit of inventory depends on its shelf age, in accordance with a general, increasing function $\alpha(\cdot)$. Similarly, the marginal delay cost incurred for a backlogged unit, depends on the delay experienced, thus far, again in accordance with a general increasing function $\beta(\cdot)$.

When the counting process is a renewal or a Markov-modulated renewal process, we have shown, under standard fixed-plus-linear order costs, that an $(r, q)$-policy is optimal under the long-run average cost criterion. (When the demand process is a Markov-modulated process, both the reorder point $r^*$ and the order quantity $q^*$ are state-dependent.) We have shown that a standard efficient algorithm can be used to compute the optimal parameter values and we have derived monotonicity properties for the dependence of the optimal policy parameter values, on the various model primitives.

One natural extension of our results involves non-unit demand sizes at the demand epochs. The simplest and most frequently used such process is a so-called *compound renewal* process. Here, the demand sizes at consecutive demand epochs are generated by an i.i.d. sequence of variables $\{D_i\}$, which is also independent of the sequence of demand epochs. Let $N(\cdot)$ denote the renewal process associated with this sequence of i.i.d. random variables, that is, $N(j) = \min\{k \geq 1 : \sum_{i=1}^{k} D_i \geq j\}$. The above-described unit matching approach continues to be optimal, where the expected cost of the $j$th unit continues to be given by $\hat{G}(y) = \mathbb{E}H(A_j) + \mathbb{E}J(A_j)$. The only difference is that under compound renewal demands, the age variable is now given by:

$$A_j = \begin{cases} \displaystyle\sum_{i=1}^{N(j)} X_i - \mathcal{L}, & \text{when } j > 0, \\ \displaystyle -\sum_{i=1}^{N(-j)} X_i - \mathcal{L}, & \text{when } j \leq 0. \end{cases}$$

Thus, assuming again that orders can be placed at demand epochs (only), the model can again be formulated as an SMDP with the system's inventory position as its *one-dimensional* state variable, as opposed to having to disaggregate the *inventory level* according to the items' shelf age or the backlogged units' delays, along with maintaining separate information about the complete vector of outstanding orders. The *one-dimensionality* of the state space allows one to compute the order policy with standard solution methods for SMDPs; see for example, Aviv and Federgruen [1]. One might conjecture that an $(s, S)$-policy would now be optimal, the natural extension of $(r, q)$-policies under demands of arbitrary rather than unit size. However, the above proof technique can not be extended because the function $\hat{G}(\cdot)$ is no longer convex, that is, Lemma 3 may fail to hold. We refer to Federgruen and Wang [18] for an alternative approach to address this model, under certain restrictions.

*Acknowledgement*

The authors are thankful for the helpful comments by the editor.

*References*

1. Aviv, Y. & Federgruen, A. (1999). The value-iteration method for countable state Markov decision processes, *Operations Research Letters* 24: 223–234.
2. Axsäter, S. (1990). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research* 38: 64–69.
3. Axsäter, S. (1993). Exact and approximate evaluation of batch-ordering policies for two-level inventory systems. *Operations Research* 41: 777–785.
4. Bather, J.A. (1966). A continuous time inventory model. *Journal of Applied Probability* 3: 538–549.
5. Benjaafar, S., Cooper, W.L., & Mardan, S. (2011). Production-inventory systems with imperfect advance demand information and updating. *Naval Research Logistics* 58(2): 88–106.
6. Bertsekas, D. (2005). *Dynamic Programming and Optimal Control*, Vol 1. Nashua, NH: Athena Scientific.
7. Browne, S. & Zipkin, P. (1991). Inventory models with continuous stochastic demands. *Annals of Applied Probability* 1: 419–435.
8. Chao, X., Xu, Y., & Yang, B. (2012). Optimal policy for a production-inventory system with setup cost and average cost criterion. *Probability in the Engineering and Informational Sciences* 26: 457–481.
9. Chao, X. & Zhou, S.X. (2006). Joint inventory-and-pricing strategy for a stochastic continuous-review system. *IIE Transactions* 38: 401–408.
10. Chen, F. & Song, J.-S. (2001). Optima policies for multechelon inventory with Markov-modulated demand. *Operations Research* 49: 226–234.

11. Chen, F., & Zheng, Y.-S. (1993). Inventory models with general backorder costs. *European Journal of Operations Research* 65(2): 175–186.
12. Chen, X. & Simchi-Levi, D. (2006). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: the continuous review model. *Operations Research Letters* 34: 323–332.
13. Dai, J.G. & Yao, D. (2012). Brownian inventory models with convex holding cost (part 1): Average-optimal controls. *Stochastic Systems* 3: 442–499.
14. Federgruen, A. & Lee, C.-Y. (1990). The dynamic lot size model with quantity discount. *Naval Research Logistics* 37: 707–713.
15. Federgruen, A. & Schechner, Z. (1983). Cost formulas for continuous review inventory models with fixed delivery lags. *Operations Research* 31(5): 957–965.
16. Federgruen, A. & Wang, M. (2012). *Inventory subsidy versus supplier trade credit.* Working paper, Columbia University.
17. Federgruen, A. & Wang, M. (2013). Monotonicity properties of stochastic inventory systems. *Annals of Operations Research* 208: 155–186.
18. Federgruen, A. & Wang, M. (2015). *Inventory model with shelf age and delay dependent inventory costs. Operations Research* 63(3): 701–715.
19. Federgruen, A. & Zheng, Y.S. (1992). An efficient algorithm for computing an optimal (r, Q) policy in continuous review stochastic inventory systems. *Operations Research* 40(4): 808–813.
20. Gayon, J., Benjaafar, S., & de Véricourt, F. (2009). Using imperfect advance demand information in production-inventory systems with multiple customer classes. *Manufacturing & Service Operations Management* 11(1): 128–143.
21. Gupta, D. & Wang, L. (2009). A stochastic inventory model with trade credit. *Manufacturing Service Operations Management* 11(1): 4–18.
22. Hadley, G. & Whitin, T. (1963). *Analysis of inventory systems*. Englewood Cliffs, NJ: Prentice-Hall.
23. Huh, W.T., Janakiraman, G., Muharremoglu, A., & Sheopuri, A. (2011). Optimal policies for inventory systems with a generalized cost model. *Operations Research* 59(4): 1040–1047.
24. Iglehart, D. & Karlin, S. (1962). Optimal policy for dynamic inventory process with nonstationary stochastic demands. In *Studies in Applied Probability and Management Science* (K. Arrow, S. Karlin, H. Scarf, Eds). Chapter 8, Stanford, CA: Stanford University Press.
25. Karlin, S. (1960). Dynamic inventory policy with varying stochastic demands. *Management Science* 6: 231–258.
26. Levi, R., Roundy, R., Shmoys, D., & Sviridenko, M. (2008). A constant approximation algorithm for the one-warehouse multi-retailer problem. *Management Science* 54: 763–776.
27. Muharremoglu, A. & Tsitsiklis, J. (2008). A single-unit decomposition approach to multi-echelon inventory systems. *Operations Research* **56**(5): 1089–1103.
28. Naddor, E. (1966). *Inventory Systems.* New York: Wiley.
29. Nagarajan, M. & Rajagopalan, S. (2008). Contracting under vendor managed inventory systems using holding cost subsidies. *Production and Operations Management* 17(2): 200–210.
30. Nahmias, S., Perry, D., & Stadje, W. (2004). Actual valuation of perishable inventory systems. *Probability in the Engineering and Information Sciences* 18: 219–232.
31. Perry, D. & Stadje, W. (1999). Perishable inventory systems with impatient demands. *Mathematical Methods of Operations Research* 50: 77–90.
32. Richardson, H. (1995). Control your costs then cut them. *Transportation and Distribution* 36(12): 94–96.
33. Rosling, K. (2002). Inventory cost rate functions with nonlinear shortage costs. *Operations Research* 50(6): 1007–1017.
34. Sahin, I. (1979). On the stationary analysis of continuous review (s, S) inventory systems with constant lead times.*Operations Research* 27: 717–730.
35. Sahin, I. (1983). On the continuous review (s; S) inventory model under compound renewal demand and random lead times. *Journal of Applied Probability* 20: 213–219.
36. Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic Orders.* New York: Springer.
37. Shi, J., Katehakis M.N., Melamed, B., & Xia, Y. (2014). Production-Inventory systems with lost sales and compound Poisson demands. *Operations Research* 6(5): 1048–1063.
38. Song, J.-S. & Zipkin, P. (1993). Inventory control in a fluctuating demand environment. *Operations Research* 41: 351–370.
39. Stauffer, G., Massonnet, G., Rapine, C., & Gayon, J.-P. (2011). A simple and fast 2-approximation for deterministic lot-sizing in one warehouse multi-retailer systems. Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 67–79.

40. Tijms, H.C. (1972). *Analysis of (s, S): Inventory Models.* Mathematical Centre Tracts 40, Amsterdam: Mathematical Centre.
41. Tijms, H.C. (2003). *A First Course in Stochastic Models.* West Sussex, England: Wiley.
42. Xu, Y. & Chao, X. (2009). Dynamic pricing and inventory control for a production system with average profit criterion. *Probability in the Engineering and Informational Sciences* 23: 489–513.
43. Zheng, Y.S. (1991). A simple proof for optimality of (s, S) policies in infinite-horizon inventory systems. *Journal of Applied Probability* 28(4): 802–810.
44. Zipkin, P. (1986). Stochastic leadtimes in continuous time inventory models. *Naval Research Logistics* 33: 763–774.
45. Zipkin, P. (2000). *Foundations of Inventory Management.* New York: McGraw Hill-Irwin.

## APPENDIX

Proof of Proposition 1: Zheng [43] established that the existence of a solution $\overline{v}^*(\cdot)$ to a relaxed version of the optimality Eq. (9) of the transformed periodic review MDP, where the constraints $y \geq x$ are relaxed:

$$\overline{v}(x) = \min_y \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g + \overline{v}(y-1)\}. \tag{A.1}$$

The solution $\overline{v}^*(\cdot)$ is, in fact, bounded. A solution $w^*$ to the optimality Eq. (9) can be constructed, recursively, from $\overline{v}^*(\cdot)$ as follows:

$$w^*(x) \equiv \begin{cases} \overline{v}^*(x), & \text{if } x \leq R^* \equiv r^* + q^*, \\ \hat{G}(x) - g^* + w^*(x-1), & \text{if } x > R^*. \end{cases} \tag{A.2}$$

We now show that the function $w^*(\cdot)$ satisfies the optimality Eq. (9). Note, first, from Federgruen and Zheng [19, Lemma 2] that:

$$g^*\tau \leq \hat{G}(R^*+1) < \hat{G}(R^*+2) < \hat{G}(R^*+3) < \cdots. \tag{A.3}$$

Equation (A.3) implies, by induction, that $w^*(y)$ is increasing for $y \geq R^*$. For $x \leq R^*$ it is optimal, in the relaxed problem, to select $x \leq y \leq R^*$. Thus, for all $x \leq R^*$,

$$\begin{aligned} w^*(x) = \overline{v}(x) &= \min_y \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + \overline{v}^*(y-1)\} \\ &= \min_{x \leq y \leq R^*} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + \overline{v}^*(y-1)\} \\ &= \min_{x \leq y \leq R^*} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + w^*(y-1)\} \\ &= \min_{x \leq y} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + w^*(y-1)\}, \end{aligned}$$

where the last equality follows from $w^*(R^*) \leq w^*(R^*+1) \leq w^*(R^*+2) \leq \ldots$ and the second to last equality follows from $y-1 \leq R^*-1$, so that $\overline{v}^*(y-1) = w^*(y-1)$. It remains to be shown that $w^*(\cdot)$ satisfies the optimality Eq. (9) for $x > R^*$. For all $x > R^*$ and $y \geq x$:

$$w^*(x) = \hat{G}(x) - g^* + w^*(x-1) \leq K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + w^*(y-1).$$

(The equality follows from Eq. (A.2); the inequality follows from the fact that the first two terms to its right are non-negative, while $\hat{G}(y) \geq \hat{G}(x)$ and $w^*(y-1) \geq w^*(x-1)$ for any $y \geq x > R^*$.) Thus, for $x \geq R^*$:

$$w^*(x) = \min_{x \leq y} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g^* + w^*(y-1)\},$$

verifying that $w^*(\cdot)$ satisfies the optimality Eq. (9), also for $x > R^*$. Moreover, the policy $(r^*, q^*)$ achieves all minima in this optimality equation. Furthermore, it follows from the proof of Theorem 2

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press

that the function $v^*(x) \equiv w^*(x) - G(x)$ satisfies the optimality Eq. (8) and the policy $(r^*, q^*)$ achieves all minima in this optimality equation as well. ■

PROOF OF PROPOSITION 2: (a) By Theorem 2 from Federgruen and Wang [17], it suffices to show that $\hat{G}(y \mid \alpha(\cdot))$ has increasing differences in $(y, \alpha(\cdot))$, or that $\mathbb{E}H(A_y \mid \alpha(\cdot))$ has increasing differences in $(y, \alpha(\cdot))$. Since

$$\begin{aligned} \mathbb{E}H(A_y \mid \alpha(\cdot)) &= \int_0^\infty \int_0^t \alpha(s)\, ds\, dF_{A_y}(t) = \int_0^\infty \int_s^\infty \alpha(s)\, dF_{A_y}(t)\, ds \\ &= \int_0^\infty \alpha(s)(1 - F_{A_y}(s))\, ds. \end{aligned} \tag{A.4}$$

We get

$$\begin{aligned} \mathbb{E}H(A_{y+1} \mid \alpha(\cdot)) - \mathbb{E}H(A_y \mid \alpha(\cdot)) &= \int_0^\infty \alpha(s)(F_{A_y}(s) - F_{A_{y+1}}(s))\, ds \\ &\le \int_0^\infty \hat{\alpha}(s)(F_{A_y}(s) - F_{A_{y+1}}(s))\, ds \\ &= \mathbb{E}H(A_{y+1} \mid \hat{\alpha}(\cdot)) - \mathbb{E}H(A_y \mid \hat{\alpha}(\cdot)), \end{aligned}$$

where the inequality follows from the point-wise dominance $\alpha(\cdot) \le \hat{\alpha}(\cdot)$ and the fact that the SIL property of $A_y$ in $y$, see the proof of Lemma 3, implies that $A_y \le_{st} A_{y+1}$.

(b) Analogous to (a), it suffices to show that $\mathbb{E}J(A_y \mid \beta(\cdot))$ has decreasing differences in $(y, \beta(\cdot))$. Since

$$\begin{aligned} \mathbb{E}J(A_y \mid \beta(\cdot)) &= \mathbb{E} \int_{-\mathcal{L}}^0 \int_0^{-t} \beta(s)\, ds\, dF_{A_y}(t) = \mathbb{E} \int_0^{\mathcal{L}} \int_{-\mathcal{L}}^{-s} \beta(s)\, dF_{A_y}(t)\, ds \\ &= \mathbb{E} \int_0^{\mathcal{L}} \beta(s) F_{A_y}(-s)\, ds, \end{aligned}$$

similar to part (a), we have

$$\begin{aligned} \mathbb{E}J(A_{y+1} \mid \beta(\cdot)) - \mathbb{E}J(A_y \mid \beta(\cdot)) &= \mathbb{E} \int_0^{\mathcal{L}} \beta(s)(F_{A_{y+1}}(-s) - F_{A_y}(-s))\, ds \\ &\ge \int_0^\infty \hat{\beta}(s)(F_{A_{y+1}}(s) - F_{A_y}(s))\, ds \\ &= \mathbb{E}J(A_{y+1} \mid \hat{\beta}(\cdot)) - \mathbb{E}J(A_y \mid \hat{\beta}(\cdot)). \end{aligned}$$

(c) Once again, it suffices to show that $\hat{G}(y \mid \mathcal{L}) = \mathbb{E}H(A_y \mid \mathcal{L}) + \mathbb{E}J(A_y \mid \mathcal{L})$ has decreasing differences in $(y, \mathcal{L})$. We prove this for the first term $\mathbb{E}H(A_y \mid \mathcal{L})$; the proof for the second term $\mathbb{E}J(A_y \mid \mathcal{L})$ is analogous. Note that

$$\begin{aligned} \mathbb{E}H(A_{y+1} \mid \mathcal{L}) - \mathbb{E}H(A_y \mid \mathcal{L}) &= \mathbb{E}\left( \int_{\sum_{i=1}^{y} X_i - \mathcal{L}}^{\sum_{i=1}^{y+1} X_i - \mathcal{L}} \alpha(s)\ ds \right) \\ &= \mathbb{E}_{\{X_i\}} \left\{ \mathbb{E}_{\mathcal{L}} \int_{\sum_{i=1}^{y} X_i - \mathcal{L}}^{\sum_{i=1}^{y+1} X_i - \mathcal{L}} \alpha(s)\ ds \right\}. \end{aligned} \tag{A.5}$$

Also, for any given realization of the renewal process of $\{X_i\}$, the function $\int_{\sum_{i=1}^{y} X_i - L}^{\sum_{i=1}^{y+1} X_i - L} \alpha(s)\ ds$ is decreasing in $L$, since the function $\alpha(s)$ is increasing. This implies that

$$\mathbb{E}_{\mathcal{L}^2} \int_{\sum_{i=1}^{y} X_i - \mathcal{L}^2}^{\sum_{i=1}^{y+1} X_i - \mathcal{L}^2} \alpha(s)\, ds \le \mathbb{E}_{\mathcal{L}^1} \int_{\sum_{i=1}^{y} X_i - \mathcal{L}^1}^{\sum_{i=1}^{y+1} X_i - \mathcal{L}^1} \alpha(s)\, ds,$$

whenever $\mathcal{L}^2 \ge_{st} \mathcal{L}^1$, thus implying that $\mathbb{E}H(A_{y+1} \mid \mathcal{L}^2) - \mathbb{E}H(A_y \mid \mathcal{L}^2) \le \mathbb{E}H(A_{y+1} \mid \mathcal{L}^1) - \mathbb{E}H(A_y \mid \mathcal{L}^1)$.

(d) Similar to part (c), it suffices to show that $\hat{G}(y \mid X) = \mathbb{E}H(A_y \mid X) + \mathbb{E}J(A_y \mid X)$ has increasing differences in $(y, X)$ where $X$ is the interarrival time distribution. As we did in part (c), we only prove this for the first term of $\hat{G}(y \mid X)$, since the proof for the second term is analogous. Define $\phi(r,t) \equiv \int_r^{r+t} \alpha(s)\, ds$. Note that $\phi(r,t)$ is a convex function of $r(t > 0)$ for any given $t > 0(r)$ since $\partial^2\phi(r,t)/\partial r^2 = \alpha'(r+t) - \alpha'(r) \ge 0$ and $\partial^2\phi(r,t)/\partial t^2 = \alpha'(r+t) \ge 0$. The first inequality follows from the convexity of $\alpha(\cdot)$, and the second inequality follows from the monotonicity of $\alpha(\cdot)$. Hence, it follows from Eq. (A.5) that

$$\begin{aligned}
\mathbb{E}H(A_{y+1} \mid X^1) - \mathbb{E}H(A_y \mid X^1) &= \mathbb{E}\phi\left(\sum_{i=1}^{y} X_i^1 - \mathcal{L}, X_{y+1}^1\right) \le \mathbb{E}\phi\left(\sum_{i=1}^{y} X_i^2 - \mathcal{L}, X_{y+1}^1\right) \\
&\le \mathbb{E}\phi\left(\sum_{i=1}^{y} X_i^2 - \mathcal{L}, X_{y+1}^1\right) \\
&= \mathbb{E}H(A_{y+1} \mid X^2) - \mathbb{E}H(A_y \mid X^2),
\end{aligned}$$

where the inequalities follow from $\sum_{i=1}^{y} X_i^1 \le_{cv} \sum_{i=1}^{y} X_i^2$ and $X_{y+1}^1 \le_{cv} X_{y+1}^2$.

■

https://doi.org/10.1017/S0269964815000169 Published online by Cambridge University Press