Martingale methods for pricing inventory penalties under continuous replenishment and compound renewal demands

Junmin Shi · Michael N. Katehakis ·
Benjamin Melamed

Abstract This paper addresses the problem of inventory penalty pricing under the risk-neutral valuation principle. The underlying production-inventory system has a constant replenishment rate and a compound renewal demand stream (i.e., iid demand interarrival times are independent of iid demand sizes), and is subject to underage and overage penalties. Our pricing approach treats the penalties as a series of perpetual American options, and constructs auxiliary martingale processes in term of the inventory process. We provide a necessary and sufficient martingale condition for general compound renewal demands. Explicit expressions of penalty functions for underage and overage are obtained for the case where demand arrivals follow a Poisson process.

Keywords Inventory penalty functions · Martingale · Overage · Risk-neutral pricing · Underage

1 Introduction

Consider a production-inventory system with a constant replenishment rate. Demands arrive according to a *compound renewal* process, that is, both inter-arrival times and demand sizes are mutually independent random sequences, each being independent and identically

J. Shi (✉)
Dept. of Managerial Sciences, Robinson College of Business, Georgia State University, 35 Broad St.,
Atlanta, GA 30303, USA
e-mail: jmshi@gsu.edu

M.N. Katehakis
Dept. of Management Science and Information Systems, Rutgers Business School – Newark and
New Brunswick, 1 Washington Park, Newark, NJ 07102, USA
e-mail: mnk@andromeda.rutgers.edu

B. Melamed
Dept. of Supply Chain Management and Marketing Sciences, Rutgers Business School – Newark and
New Brunswick, 94 Rockafeller Rd., Piscataway, NJ 08554, USA
e-mail: melamed@business.rutgers.edu

distributed (iid). The system is subject to overage and/or underage penalties, both assessed against the firm (owner of the system). An *underage penalty* is assessed whenever a demand arrival encounters a stockout, whereby all or a portion of the demand cannot be satisfied from on-hand inventory, while an *overage* penalty is assessed whenever the inventory level hits a prescribed *target level* (e.g., the base stock level in a make-to-stock production-inventory system).

We define the *perpetual inventory penalty* as the discounted total penalty over an infinite time horizon. The *standard penalty valuation problem* in production-inventory context is the computation of the expected perpetual inventory penalty. In contrast, the problem we address in this paper is the *inverse penalty valuation problem* of devising proper dynamic *penalty functions* for assessing underage and/or overage penalties, such that the expected perpetual inventory penalty equals a prescribed value; the latter is interpreted as the penalty budget given to an inventory manager at time 0 to defray the cost of future underage and overage penalties. Throughout this paper, we assume that interest is compounded continuously at a given constant rate $r \geq 0$; accordingly, the present value of one unit of cash at time t is e^{-rt}.

Both underage and overage are detrimental to the economics of production-inventory systems. Underage causes tangible and intangible losses that impact both the firm (e.g., lost sales, lost goodwill) and the customer (e.g., reduced service levels with potential business losses incurred by customers). In a similar vein, overage increases holding costs, and empirical evidence shows that even in well-managed businesses, a significant proportion of the inventory constitutes overage at any given time (Relph and Barrar 2003). While it is reasonable to assess overage penalties in terms of inventory holding costs, it is much more challenging to assess the firm's loss of (intangible) goodwill and the costs of business losses sustained by a customer encountering stockouts. Similarly, it is also hard to evaluate the full extent of overage penalty because the supplier may sustain additional losses stemming from the enforcement of supply contracts (e.g., returned items). The latter are often ignored in supply chain studies (Zipkin 2000).

Most inventory optimization models capture overage and underage penalties as important cost components, but rarely address the practical problem of valuating them using real-life metrics. The traditional approach is to assume a prescribed fixed stockout cost that is incurred as a tangible loss by each stockout occurrence (Ernst and Powell 1995). However, intangible losses such as goodwill are not explicitly captured (Walter and Grabner 1975). The literature on the valuation of inventory penalties is relatively sparse. Schwartz (1966, 1970) point out that the effect of goodwill loss is characterized by the fact that disappointed customers react by changing their future purchasing habits. Thus some lingering stockout effects tend to adversely perturb future demand, a phenomenon quite different from the instantaneous imposition of a penalty at its time of occurrence. Early studies of the properties of perturbed demand appear in Schwartz (1966, 1970).

Several statistical studies have recently addressed the evaluation problem of lost sales. For example, in order to estimate the value of lost sales opportunities, Ofenbakh (2008) analyzes historical sales data to assess the underlying lost demand. Other studies meld finance methods and inventory management in a framework that treats inventory as a financial asset, and an inventory policy (or strategy) as a contingent claim. Accordingly, some well-known financial models have been applied to inventory management problems. For example, Stowe and Su (1997) view inventory as an option on future sales and show how some basic inventory problems can be solved using pricing models. The reader is referred to Stowe and Su (1997) for a literature survey of applications of financial theory to inventory decisions.

Our approach to the inverse penalty valuation problem is motivated by the risk-neutral valuation framework for contingent claims, and specifically by the special case of perpetual

American options. *A perpetual American call (put) option* is a legal contract giving the holder of the option the right to buy (sell) a particular stock at a particular price at any time in the future (there is no option expiry date). The principle of *risk-neutral valuation* states that the option is risk-neutral if its expected value, discounted to the present time, equals its present value. It is known that if a contingent claim is risk neutral, then the stochastic process of its discounted version over time forms an exponential martingale process over a modified probability space (Hull 2002); this fact motivates us to apply a martingale approach to the penalty valuation problem and the inverse problem.

In this paper, we shall treat the imposition of an inventory underage or overage penalty as an exercise action of a perpetual American option; accordingly, the inventory penalty values are derived by using the risk-neutral principle. More specifically, the main idea is that inventory penalties are incurred stochastically at stopping times of the inventory process, which are analogous to the exercise times in a series of perpetual American options. To this end, we shall use techniques similar to those employed in financial valuation models to construct auxiliary exponential martingale processes, and derive functional equations for the requisite expected discounted penalties. The aforementioned auxiliary martingale processes will then be further used to derive the Laplace transform of the probability density function (pdf) of the stopping times at which penalties are incurred. To this end, we use techniques similar to those in the insurance model of Gerber and Shiu (1998), and obtain a penalty function by solving an equation with a prescribed present value of the respective penalties over an infinite time horizon. For the case of compound Poisson demands, explicit expressions for the underage penalty and overage penalty are obtained in Theorem 3 and Theorem 5, respectively.

The main contributions of this paper are listed below:

(1) For an underlying production-inventory system with a general compound renewal demand process and constant replenishment, we construct an exponential process and provide a necessary and sufficient condition for this process to be a martingale. This result extends a known result for the special case of compound Poisson demand arrivals, known as *Lundberg's fundamental equation* (Gerber and Shiu 1998).
(2) For the same underlying system with lost sales, we derive explicit expressions for the underage penalty sequence.
(3) For the same underlying system with backorders, we derive explicit expressions for the overage penalty sequence.

Throughout the rest of this paper, we use the following notation. The set of real numbers is denoted by $\mathbb{R}$. For any real number $x \in \mathbb{R}$, $x^+ = \max\{x, 0\}$. The indicator function of a set A is denoted by 1_A. For a random variable X, the functions $f_X(x)$, $F_X(x)$ and $\bar{F}_X(x) = 1 - F_X(x)$ denote respectively its probability density function (pdf), its cumulative distribution function (cdf) and its complementary cdf. The Laplace transform of a real function $f(x)$ is defined by $\tilde{f}(z) = \int_0^\infty e^{-zx} f(x)\, dx$.

The rest of this paper is organized as follows. Section 2 introduces assessment functions and a concise relevant background of the martingale methodology. Section 3 formulates the inventory model with compound-renewal demand arrivals. Sections 4, 5 and 6 treat the special case of compound Poisson demands, with Section 5 addressing underage penalties and Section 6 addressing overage penalties. Finally, Section 7 concludes this paper.

2 Assessment functions

In this section we provide a brief description of a typical assessment problem. Let $\{Y(t) : t \geq 0\}$ be a real-valued right-continuous stochastic process which models the time evolution of a stock price, inventory level, etc. Let further $\{\tau_i^{(L)} : i \geq 1\}$ be the sequence of successive hitting times of a level L by the process $\{Y(t)\}$, (i.e., $Y(\tau_i^{(L)}-) = L$) at which times certain assessments (e.g., penalties), $W_i^{(L)} = W_i^{(L)}(Y(\tau_i^{(L)}), L)$, are computed as a function of the level L and the state, $Y(\tau_i^{(L)}+) = Y(\tau_i^{(L)})$. The assessment $W_i^{(L)}$ may be a random variable with a distribution that depends on $Y(\tau_i^{(L)})$ or a deterministic function. One typical example is the exercise of a perpetual American option [cf. Gerber and Shiu 1994, 1996], where the owner should exercise the option as soon as the value of the stock price reaches a particular level L. Following our notational conventions, we denote the pdf of $\tau_i^{(L)}$ conditioned on $Y(0)$ by $f_{\tau_i^{(L)}|Y(0)}(t|y)$, and its Laplace transform by $\tilde{f}_{\tau_i^{(L)}|Y(0)}(s|y)$.

Following the standard finance methodology, the risk-neutral value of the assessments conditional on the initial state at time 0 is the total expected discounted value of all the assessments, given by

$$C(L, y) = \mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_i^{(L)}} W_i^{(L)} \,\Big|\, Y(0) = y\right]. \tag{2.1}$$

Of special interest is the expected discounted value of the assessment at the first level hitting time $\tau_1^{(L)}$, namely,

$$C_1(L, y) = \mathbb{E}\left[e^{-r\tau_1^{(L)}} W_1^{(L)} \,\big|\, Y(0) = y\right]. \tag{2.2}$$

Using a renewal argument, the risk-neutral value of Eq. (2.1) can be obtained recursively by conditioning on the value of the state, $Y(\tau_1^{(L)})$, at $\tau_1^{(L)}$.

A classic example of the general standard valuation problem above is the expected discounted deficit at ruin time in an insurance setting (Gerber and Shiu 1998). Specifically, let $Y(t)$ be the surplus of an insurance firm at time t, with $Y(0) = y$ denoting the initial surplus. The firm will encounter bankruptcy at the ruin time $\tau_1^{(0)}$ when its surplus hits zero for the first time.

In the special case of $W_i^{(L)} = w_i(L)$ (deterministic level-dependent assessments), Eq. (2.2) reduces to

$$C_1(L, y) = w_1(L)\mathbb{E}\left[e^{-r\tau_1^{(L)}} \,\big|\, Y(0) = y\right] = w_1(L)\tilde{f}_{\tau_1^{(L)}|Y}(r|y). \tag{2.3}$$

A case in point is the valuation of a perpetual American call option. Here, $Y(t)$ is the underlying stock price at time t, L is the exercise level and K is the exercise (strike) price. Without loss of generality, we assume that the option is out of the money at time 0, that is, $Y(0) < K$, so that exercising the option at time $t = 0$ can be excluded. Then the payoff of the corresponding perpetual American option is given by $w(L) = [L - K]^+ = L - K$.

Recall that the standard valuation problem under the risk-neutral principle is to price the action of "exercising the option" by using Eq. (2.1) with a given payoff function $w(L)$, while the inverse valuation problem is to deduce $w(L)$, given a present valuation, $C(L, y)$. In an inventory budget-planning context, the inverse valuation problem can be interpreted as follows: An inventory manager is given a prescribed perpetual inventory penalty budget to defray the cost of future underage and overage penalties; such a budget can be viewed as a performance metric of the system (e.g., cost component, service-level, etc.). A risk-neutral inventory manager is then faced with the (inverse valuation) problem of deriving

the appropriate penalty values to assess per underage/overage event occurrence under the principle of risk-neutral valuation, that is, subject to the constraint that the total expected discounted value of perpetual inventory penalties equals the prescribed budget.

It follows from Eq. (2.1) that the functional relationship between $C(L, y)$ and $w(L)$ is specified in terms of $\tilde{f}_{\tau_1^{(L)}|Y(0)}(r|y)$, and consequently, this Laplace transform needs to be derived first. To this end, we shall construct in Section 4, an auxiliary martingale sequence $\{M(t) : t \geq 0\}$ to define the requisite sequence of penalty functions. Recall that a martingale process $\{M(t) : t \geq 0\}$ satisfies the regularity condition

$$\mathbb{E}\big[M(t)\big] < \infty, \quad t \geq 0, \tag{2.4}$$

and the martingale property

$$\mathbb{E}\big[M(t) \mid M(z), z \leq s\big] = M(s), \quad s \leq t. \tag{2.5}$$

We shall further make use of the *Optional Stopping Theorem* (also called the *Optional Sampling Theorem*). One version of the theorem is given below [cf. Karr 1993].

Theorem 1 *Let $\{M(t) : t \geq 0\}$ be a martingale process and let τ be a stopping time of $\{M(t) : t \geq 0\}$, such that*

(a) $\mathbb{E}[\tau] < \infty$,
(b) $\{M(t)\}$ *is bounded or uniformly integrable.*

Then,

$$\mathbb{E}\big[M(\tau) \mid M(0)\big] = M(0).$$

Finally, we shall need the following definition

Definition 1 Let $\{X(t), t \geq 0\}$ be a stochastic process and let $\{\tau_i, i \geq 0\}$ be a sequence of increasing hitting times measurable on the sigma-algebra generated by $\{X(t), t \geq 0\}$. Then $\{X(t)\}$ has an *embedded martingale* $\{X(\tau_i) : i \geq 0\}$, if it satisfies for any $0 \leq i \leq j$,

$$\mathbb{E}\big[X(\tau_j) \mid X(t), t \leq \tau_i\big] = X(\tau_i). \tag{2.6}$$

The motivation for defining auxiliary martingale processes as a tool for solving for $\tilde{f}_{\tau_1^{(L)}|Y(0)}(r|y)$ draws on classic martingale methodology as applied in the context of financial derivatives valuation. In particular, the time evolution of an underlying security price process, $\{S(t)\}$, is often modeled as a geometric Brownian motion, that is, $X(t) = \log S(t)$ is a shifted Brownian motion. The reason this model is routinely used in the analysis of financial derivatives (e.g., American options), is its continuity and independent-increment property. In a risk-neutral setting, the discounted-price process of financial derivatives over time is a martingale $\{M(t)\}$ of the form [cf. Luenberger 1998],

$$M(t) = e^{-rt+g(X(t))}, \tag{2.7}$$

where $g(x)$ is a real-valued function, and $e^{g(X(t))}$ is the price of the financial derivatives at time t. By the martingale property, $\mathbb{E}[M(t)|M(0)] = M(0)$. If $\{M(t)\}$ is a bounded process, then

$$M(0) = e^{g(X(0))} = \mathbb{E}\big[e^{-r\tau^{(L)}+g(X(\tau^{(L)}))}\big] = \mathbb{E}\big[e^{-r\tau^{(L)}+g(\log L)}\big] = \tilde{f}_{\tau^{(L)}}(r)e^{g(\log L)}, \tag{2.8}$$

where the second equality is due to the Optional Stopping Theorem applied to the stopping time $\tau^{(L)}$, and the third equality follows since the definition of $\tau^{(L)}$ implies that $S(\tau^{(L)}) = L$. Equation (2.8) readily implies

$$\tilde{f}_{\tau^{(L)}}(r) = e^{g(X(0)) - g(\log L)}. \tag{2.9}$$

Equation (2.9) exhibits the relation between the Laplace transform of $\tau^{(L)}$ and the martingale process of Eq. (2.7), and will be used to solve for the former [cf. Eq. (5.7) in Gerber and Shiu 1998].

The general martingale methodology outlined above will be employed in this paper to assess overage and underage penalties, under the principle of risk-neutral valuation. To this end, the inventory-level process $\{I(t) : t \geq 0\}$ will play a role analogous to $X(t) = \log S(t)$, and the hitting time $\tau^{(L)}$ will map to hitting times of prescribed inventory levels that incur underage and overage penalties. Furthermore, the function $g(x)$ will be chosen as the linear function $g(x) = cx$, where c is a real constant, such that the auxiliary process $\{M(t), t \geq 0\}$, defined by

$$M(t) = e^{-rt + cI(t)}, \tag{2.10}$$

is a martingale. The requisite value of such a constant c is specified in Lemma 1.

3 Inventory model formulation

We consider a general continuous-review production-inventory system. Replenishment occurs at a constant (deterministic) rate $\rho > 0$. The demand stream is a compound renewal process denoted by $\{(A_i, D_i) : i \geq 0\}$, where the demand arrival-time process $\{A_i\}$ and the corresponding demand-size process $\{D_i\}$ are mutually independent. The process $\{A_i : i \geq 0\}$ is a renewal process with arrival rate λ and the convention $A_0 = 0$. Let $\{T_i : i \geq 1\}$ be the corresponding sequence of demand interarrival times, where $T_i = A_i - A_{i-1}$ are iid with common pdf $f_T(t)$ satisfying $f_T(0) = 0$. Let further $\{N_A(t) : t \geq 0\}$ be the counting process of demand arrivals, where $N_A(t) = \sup\{n \geq 0 : A_n \leq t\}$ is the number of demand arrivals in the interval $(0, t]$. The corresponding demand sizes $\{D_i : i \geq 0\}$, where $D_0 = 0$ by convention, are iid random variables with common pdf $f_D(x)$ and cdf $F_D(x)$. Under constant replenishment at rate $\rho > 0$, the inventory process $\{I(t) : t \geq 0\}$ is given by

$$I(t) = I(0) + \rho t - \sum_{i=1}^{N_A(t)} D_i, \tag{3.1}$$

where a negative inventory level means that backordering is in effect.
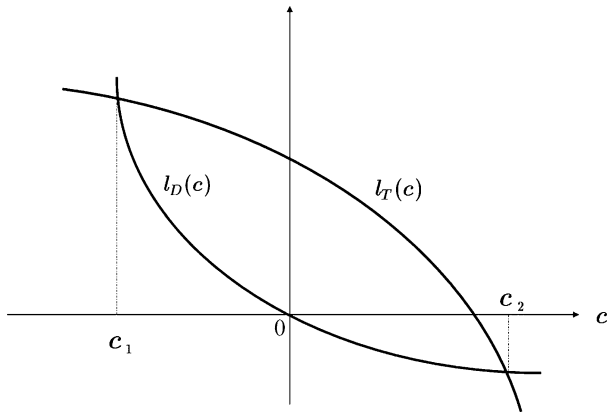
Next, define the auxiliary process

$$M(t) = e^{-rt + cI(t)}, \tag{3.2}$$

where the value of the constant c is chosen so as to ensure that $\{M(A_i)\}$ is an embedded martingale process.

Lemma 1 *The process $\{M(t)\}$ of Eq. (3.2) has an embedded martingale $\{M(A_i)\}$ if and only if c satisfies*

$$\tilde{f}_T(r - c\rho) \tilde{f}_D(c) = 1. \tag{3.3}$$

Fig. 1 The two roots of the
equation $l_T(c) = l_D(c)$

Proof See Appendix.   □

We next investigate the roots of Eq. (3.3). To this end, define

$$l_T(c) = -\log \tilde{f}_T(r - c\rho) \tag{3.4}$$

and

$$l_D(c) = \log \tilde{f}_D(c), \tag{3.5}$$

and note that Eq. (3.3) is equivalent to

$$l_T(c) = l_D(c). \tag{3.6}$$

The following result provides key properties of these functions.

Lemma 2

(a) $l_T(c)$ *is strictly decreasing and strictly concave in c.*
(b) $l_D(c)$ *is strictly decreasing and strictly convex in c.*

Proof See Appendix.   □

The following Proposition establishes a key property for the roots of Eq. (3.3).

Proposition 1 *Equation* (3.3) *has two real roots,* $c_1 \leq 0$ *and* $c_2 \geq 0$.

Proof This follows immediately from Lemma 2 and the fact that $l_D(0) = 0 < l_T(0)$ by
Eqs. (3.4) and (3.5).   □

Figure 1 depicts the functional form of $l_T(c)$ and $l_D(c)$, as well as the relative location of
the roots, c_1 and c_2.

Using the two roots c_1 and c_2, we define the two auxiliary processes

$$M_1(t) = e^{-rt + c_1 I(t)} \tag{3.7}$$

and

$$M_2(t) = e^{-rt + c_2 I(t)}. \tag{3.8}$$

Recall that by Lemma 1, each of these auxiliary processes has an associated embedded martingale at demand arrival times, $\{A_i\}$.

4 Compound Poisson demand arrivals

In this section and the rest of this paper, we consider the case where the demand size is general, but demand arrivals follow a Poisson process with arrival rate $\lambda > 0$, that is, the interarrival times are iid with pdf $f_T(t) = \lambda e^{-\lambda t}$ and Laplace transform $\tilde{f}_T(c) = \frac{\lambda}{\lambda + c}$. Then Eq. (3.3) becomes

$$r - \rho c + \lambda \left[1 - \tilde{f}_D(c) \right] = 0. \tag{4.1}$$

Equation (4.1) is well known in the context of insurance models, where it is referred to as *Lundberg's fundamental equation* (Gerber and Shiu 1998). Accordingly, we shall refer to Eq. (3.3) as the *renewal extension of Lundberg's fundamental equation*.

The following example provides the two roots for the case of Poisson demand arrivals and exponentially distributed demand sizes.

Example 1 Consider the case of exponential demand sizes with parameter $\beta > 0$, that is, $f_D(x) = \beta e^{-\beta x}$. Hence, $\tilde{f}_D(c) = \frac{\beta}{\beta + c}$, and Eq. (4.1) can be rewritten as

$$r - \rho c + \lambda \left[1 - \frac{\beta}{\beta + c} \right] = 0,$$

or equivalently,

$$\rho c^2 + (\rho \beta - r - \lambda)c - \beta r = 0.$$

Finally, this equation has two roots, given by

$$c_1 = \frac{r + \lambda - \rho \beta - \sqrt{(\rho \beta - r - \lambda)^2 - 4\rho \beta r}}{2\rho}, \tag{4.2}$$

$$c_2 = \frac{r + \lambda - \rho \beta + \sqrt{(\rho \beta - r - \lambda)^2 - 4\rho \beta r}}{2\rho}. \tag{4.3}$$

For Poisson demand arrivals, we can prove the following stronger result for the auxiliary processes of Section 3.

Proposition 2 *If the demand arrivals follow a Poisson process and $c_1 \leq 0$ and $c_2 \geq 0$ satisfy Eq. (4.1), then the processes $\{M_1(t)\}$ of Eq. (3.7) and $\{M_2(t)\}$ of Eq. (3.8) are both continuous martingales.*

Proof The proof readily follows from that of Lemma 1 by using the memorylessness property of Poisson processes and by replacing A_i and A_j with deterministic t and s, respectively. □

By Proposition 2, both $\{M_1(t)\}$ and $\{M_2(t)\}$ are martingales if c_1 and c_2 satisfy Eq. (3.3), which in our case becomes Eq. (4.1). We mention that the martingale property does not
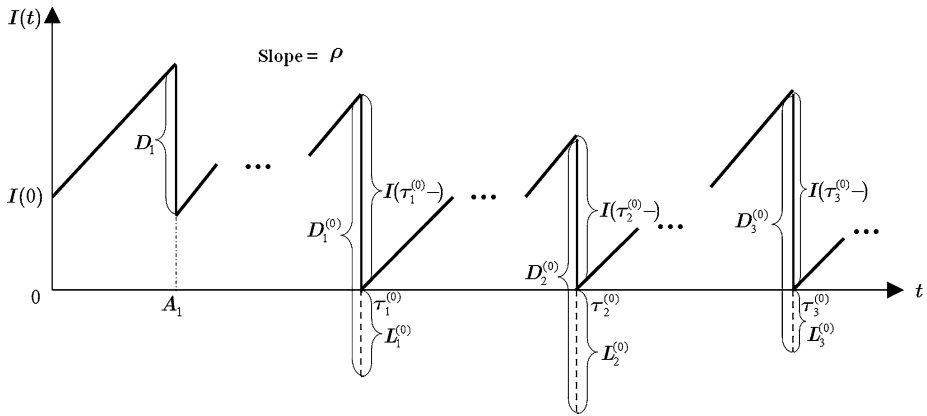
Fig. 2 A sample path of the inventory level subject to lost sales

always hold for processes embedded at stopping times (e.g., underage or overage stopping times). However, the martingale property does hold for $\{M_1(t)\}$ and $\{M_2(t)\}$ when we associate $\{M_1(t)\}$ with underage and $\{M_2(t)\}$ with overage. This is demonstrated in Section 5 and 6 respectively.

5 Underage penalty functions

In this section, we study production-inventory systems with infinite capacity, subject to the lost-sales policy, that is, the unmet portion of a demand size is lost. The system is subject to an underage (or lost-sale) penalty which is assessed whenever the inventory level hits 0 due to a shortage. In Theorem 3 below we derive explicit expressions for underage penalties in the special case of Poisson demand arrivals and exponential demand sizes. To this end, we shall employ the process $\{M_1(t)\}$, through which we derive some cost properties pertaining to the system.

Let $L(A_i)$ denote the lost-sales size at demand arrival time A_i, given by

$$L(A_i) = \big[D_i - I(A_i-)\big]^+. \tag{5.1}$$

Here the inventory process $\{I(t) : t \geq 0\}$ satisfies the following equation:
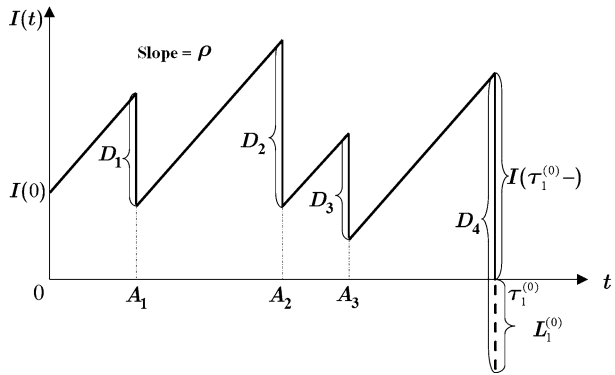
$$I(t) = u + \rho t - \sum_{i=1}^{N_A(t)} \big[D_i - L(A_i)\big], \tag{5.2}$$

where the summation term represents the cumulative amount of demands that have been satisfied. Let the stopping times of shortages (underage times) associated with the inventory process of Eq. (5.2) be defined by

$$\tau_k^{(0)} = \begin{cases} \tau_0^{(0)} = 0, & k = 0, \\ \inf\{A_i : A_i > \tau_{k-1}^{(0)}, L(A_i) > 0\}, & k > 0. \end{cases}$$

Here and in the sequel we use the superscript "0" to denote the hitting of inventory level 0 at lost-sales occurrences. Accordingly, let $D_i^{(0)}$ denote the size of the demand that arrived at the i-th loss occurrence time, $\tau_i^{(0)}$. In a similar vein, let $L_i^{(0)} = L(\tau_i^{(0)})$ denote the lost-sales size at $\tau_i^{(0)}$. Figure 2 illustrates a sample path of the inventory level process.

Fig. 3 A sample path of the inventory level process until the first underage

The following simple lemma will be needed in the sequel.

Lemma 3 *Let the demand size process be iid with a general distribution. Then, for any $x \geq 0$ and $y > 0$,*

$$\mathbb{P}\{L_i^{(0)} > y \mid I(\tau_i^{(0)}-) = x\} = \mathbb{P}\{D_i^{(0)} > x + y \mid I(\tau_i^{(0)}-) = x\}$$
$$= \mathbb{P}\{D_i^{(0)} > x + y \mid D_i^{(0)} > x\}. \qquad (5.3)$$

Proof The first equality follows directly from the identity $D_i^{(0)} = I(\tau_i^{(0)}-) + L_i^{(0)}$.

The second equality readily follows from the equality

$$\mathbb{P}\{D_i^{(0)} > x + y \mid I(\tau_i^{(0)}-) = x\} = \mathbb{P}\{D_i^{(0)} > x + y \mid I(\tau_i^{(0)}-) = x, D_i^{(0)} > I(\tau_i^{(0)}-)\}$$

in view of the fact that $\mathbb{P}\{D_i^{(0)} > I(\tau_i^{(0)}-)\} = 1$. $\qquad \square$

The next theorem establishes an embedded martingale property of $\{M_1(t)\}$.

Theorem 2 *Suppose that demands arrive according to a compound Poisson process. Then the process $\{M_1(t)\}$ in Eq. (3.7) has the martingale property at $\tau_1^{(0)}$.*

Proof By Proposition 2, $\{M_1(t)\}$ is a martingale. Furthermore, on $\{\tau_1^{(0)} < \infty\}$, we have $0 \leq M_1(t) \leq 1$ by Eq. (3.7), since $c_1 \leq 0$ by Proposition 1, implying $-rA_i + c_1 I(A_i) \leq 0$. The proof is completed by applying the Optional Sampling Theorem. $\qquad \square$

Let $w_k^{(0)}(I(\tau_k^{(0)}-), L(\tau_k^{(0)}))$ denote the underage penalty for the k-th shortage occurrence. Our goal is to derive explicit expressions for the underage penalties. To this end, we shall first derive the expected discounted penalty for the first underage in Section 5.1; we then proceed to derive the corresponding subsequent underage sequence over an infinite time horizon in Section 5.2; and finally, we derive the requisite penalty expressions.

5.1 The first underage

In this section, we derive the conditional expected discounted penalty for the first underage. To this end, we revisit the inventory process given by Eq. (5.2). Figure 3 illustrates a sample path of the inventory process.

In view of Theorem 2, we can write

$$M_1(0) = \mathbb{E}\big[M_1\big(\tau_1^{(0)}\big) \,\big|\, I(0) = u\big],$$

which in turn can be rewritten as

$$e^{c_1 u} = \mathbb{E}\big[e^{-r\tau_1^{(0)} + c_1 I(\tau_1^{(0)})} \,\big|\, I(0) = u\big]. \tag{5.4}$$

For the remainder of Section 5.1, we consider the special case where demand arrivals follow a Poisson process and the demand size has an exponential distribution with pdf $f_D(x) = \beta e^{-\beta x}$, $\beta > 0$. Recall that for this case, the two roots of Eq. (3.3) are given by Eqs. (4.2) and (4.3) in Example 1. We then have the following result.

Proposition 3 *Suppose that demand arrivals form a Poisson process and demand sizes are exponentially distributed with rate parameter $\beta > 0$. Then for any $u \geq 0$,*

$$\mathbb{E}\big[e^{-r\tau_1^{(0)}} \,\big|\, I(0) = u\big] = \frac{\beta + c_1}{\beta} e^{c_1 u}. \tag{5.5}$$

Proof See Appendix.   □

In particular, for $u = 0$ in Eq. (5.5), one has

$$\mathbb{E}\big[e^{-r\tau_1^{(0)}} \,\big|\, I(0) = 0\big] = \frac{\beta + c_1}{\beta}. \tag{5.6}$$

More specifically, let the penalty function be a function $w^{(0)}(L(\tau_1^{(0)}))$ of the underage size only, and let D denote the demand size at time $\tau_1^{(0)}$. Then,

$$\begin{aligned}
&\mathbb{E}\big[e^{-r\tau_1^{(0)}} w^{(0)}\big(L\big(\tau_1^{(0)}\big)\big) \,\big|\, I(0) = u, \tau_1^{(0)} = t\big] \\
&= e^{-rt}\mathbb{E}\big[w^{(0)}\big(D - I(t-)\big) \,\big|\, I(0) = u, \tau_1^{(0)} = t, D > I\big(\tau_1^{(0)}-\big)\big] \\
&= e^{-rt}\mathbb{E}\big[w^{(0)}(D)\big].
\end{aligned} \tag{5.7}$$

Here, conditioning on the event $\{D > I(\tau_1^{(0)}-)\}$ in the first equality is justified by the fact that a loss occurred at $\tau_1^{(0)}$ by definition. The second equality is a consequence of the extended memorylessness property of the exponential distribution, namely, if X is exponential and Y is any random variable independent of X, then $\mathbb{P}\{X - Y > z | X > Y\} = \mathbb{P}\{X > z\}$ [cf. Ross 1996]; accordingly, we note that each demand size is independent of the history of the inventory process prior to the demand arrival, and therefore D is independent of $I(\tau_1^{(0)}-)$. The expected discounted penalty for the first lost sale can be computed as

$$\begin{aligned}
\mathbb{E}\big[e^{-r\tau_1^{(0)}} w^{(0)}\big(L\big(\tau_1^{(0)}\big)\big) \,\big|\, I(0) = u\big] &= \mathbb{E}_{\tau_1^{(0)}}\big[\mathbb{E}\big[e^{-r\tau_1^{(0)}} w^{(0)}\big(L\big(\tau_1^{(0)}\big)\big) \,\big|\, I(0) = u, \tau_1^{(0)}\big]\big] \\
&= \mathbb{E}\big[w^{(0)}(D)\big]\mathbb{E}\big[e^{-r\tau_1^{(0)}} \,\big|\, I(0) = u\big] \\
&= \mathbb{E}\big[w^{(0)}(D)\big]\frac{\beta + c_1}{\beta} e^{c_1 u}.
\end{aligned} \tag{5.8}$$

Here, the second equality follows from Eq. (5.7) and the last equality follows from Eq. (5.5).

5.2 The subsequent underage sequence

In this section, we continue to investigate the special case where demand arrivals follow a Poisson process and the demand size has an exponential distribution with pdf function $f_D(x) = \beta e^{-\beta x}$, $\beta > 0$. Specifically, we shall study the subsequent underage cost sequence at the hitting times $\{\tau_k^{(0)} : k > 1\}$, and obtain an expression for the conditional expected total discounted underage penalties over an infinite horizon.

For $k > 1$, we have

$$\mathbb{E}\big[e^{-r\tau_k^{(0)}} \mid I(0) = u\big] = \mathbb{E}\big[e^{-r\tau_{k-1}^{(0)}} \mid I(0) = u\big]\mathbb{E}\big[e^{-r[\tau_k^{(0)} - \tau_{k-1}^{(0)}]} \mid I(0) = u\big]$$

$$= \mathbb{E}\big[e^{-r\tau_{k-1}^{(0)}} \mid I(0) = u\big]\mathbb{E}\big[e^{-r\tau_1^{(0)}} \mid I(0) = 0\big]$$

$$= \frac{\beta + c_1}{\beta}\mathbb{E}\big[e^{-r\tau_{k-1}^{(0)}} \mid I(0) = u\big] \tag{5.9}$$

where the last equality holds by Eq. (5.6). Applying Eq. (5.9) recursively to the second term on the rightmost side above with the aid of Eq. (5.5) yields

$$\mathbb{E}\big[e^{-r\tau_k^{(0)}} \mid I(0) = u\big] = \left(\frac{\beta + c_1}{\beta}\right)^k e^{c_1 u}. \tag{5.10}$$

Denoting the prescribed perpetual underage penalty budget by p_U, we have the following result.

Theorem 3 *Suppose that demand arrivals form a Poisson process and demand sizes are exponentially distributed with rate parameter $\beta > 0$. Suppose further that the underage penalty functions $w_k^{(0)}(y) = w^{(0)}(y)$ are functions of the underage size only and independent of k. Then, for $I(0) = u$, one form of $w^{(0)}(y)$ is given by*

$$w^{(0)}(y) = p_U e^{-c_1 u}\big[e^{-c_1 y} - 1\big]. \tag{5.11}$$

Proof By Eqs. (5.8) and (5.10), one has

$$\mathbb{E}\left[\sum_{k=1}^{\infty} e^{-r\tau_k^{(0)}} w_k^{(0)}\big(L\big(\tau_k^{(0)}\big)\big) \mid I(0) = u\right] = -\frac{\beta + c_1}{c_1}e^{c_1 u}\mathbb{E}\big[w^{(0)}(D)\big]. \tag{5.12}$$

By the principle of risk-neutral valuation, the right hand side of Eq. (5.12) equals p_U, which is the amount a risk-neutral manager would budget at time 0 to defray the future expected cost of discounted underage penalties. It follows that

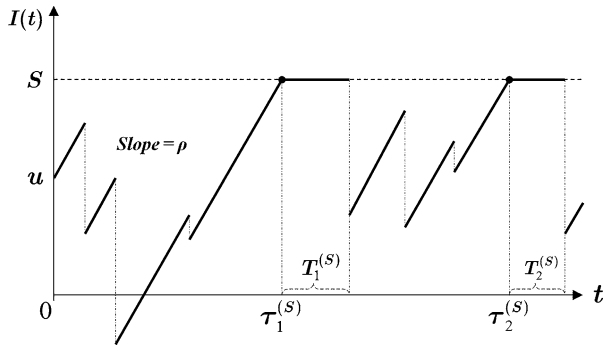$$\mathbb{E}\big[w^{(0)}(D)\big] = p_U \frac{-c_1}{\beta + c_1}e^{-c_1 u}.$$

Noting that exponential demand implies $\mathbb{E}[w^{(0)}(D)] = \beta\tilde{w}^{(0)}(\beta)$, the equation above yields

$$\tilde{w}^{(0)}(\beta) = p_U \frac{-c_1}{(\beta + c_1)\beta}e^{-c_1 u}. \tag{5.13}$$

Finally, taking the inverse Laplace transform of Eq. (5.13) completes the proof.   □

We point out that Eq. (5.13) does not determine $\tilde{w}^{(0)}$ since it holds only at the prescribed β. Consequently, other solutions for $\tilde{w}^{(0)}$ cannot be excluded.

Fig. 4 A sample path of the inventory level process with base stock level S

6 Overage penalty functions

In this section, we study make-to-stock production-inventory systems with a base stock level $S > 0$, that is, replenishment is governed by the base-stock policy as follows: while the inventory level is below S, continuous replenishment restocks the inventory at a constant rate $\rho > 0$; otherwise, replenishment is suspended. All demand shortages are backordered. The system is subject to an overage penalty whenever the inventory level hits level S from below. In Theorem 5 below we derive explicit expressions for overage penalties in the case of Poisson demand arrivals and general demand-size distributions.

To this end, we consider the inventory process $\{I(t) : t \geq 0\}$, which in this case has the representation

$$I(t) = u + \int_0^t \rho 1_{\{I(z)<S\}} dz - \sum_{i=1}^{N_A(t)} D_i. \qquad (6.1)$$

For $u < S$, we define a sequence of stopping times $\{\tau_k^{(S)} : k \geq 1\}$ when the inventory level hits S, namely,

$$\tau_k^{(S)} = \inf\{t > \tau_{k-1}^{(S)} : I(t) = S\}, \qquad (6.2)$$

where $\tau_0^{(S)} = 0$ by convention. Note that $\{I(t)\}$ can only decrease by jumps and only increase continuously, whence with probability 1,

$$I(\tau_k^{(S)}-) = I(\tau_k^{(S)}) = S, \quad k \geq 1. \qquad (6.3)$$

Figure 4 depicts the behavior of a sample path of the inventory process.

We next proceed to derive the penalty functions for overages. Let $\alpha_k^{(S)} = A_{N_A(\tau_k^{(S)})+1}$ denote the time of the first demand arrival after the occurrence of the k-th overage. Then the corresponding *replenishment idle time*, $T_k^{(S)}$, is given by
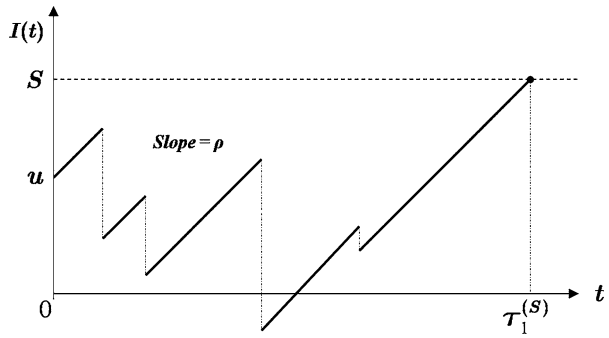
$$T_k^{(S)} = \alpha_k^{(S)} - \tau_k^{(S)}, \quad k \geq 1. \qquad (6.4)$$

By the memorylessness property of the exponential distribution, the random sequence $\{T_k^{(S)}\}$ is iid, and each $T_k^{(S)}$ is exponentially distributed with rate parameter λ, whence

$$\mathbb{E}[e^{-rT_k^{(S)}}] = \frac{\lambda}{\lambda + r}, \quad k \geq 1. \qquad (6.5)$$

Let the k-th overage penalty be denoted by $w_k^{(S)}$. Then the conditional expected total discounted penalty for the overage sequence is

Fig. 5 A sample path of the inventory process until the occurrence of the first overage

$$\mathbb{E}\left[\sum_{k=1}^{\infty} w_k^{(S)} e^{-r\tau_k^{(S)}} \,\bigg|\, I(0) = u\right] = \sum_{k=1}^{\infty} w_k^{(S)} \mathbb{E}\big[e^{-r\tau_k^{(S)}} \,\big|\, I(0) = u\big]. \qquad (6.6)$$

Our goal is to derive explicit expressions for the overage penalties. To this end, we first derive the conditional expected discounted penalty for the first overage in Section 6.1; we then proceed to derive the corresponding subsequent overage sequence over an infinite time horizon in Section 6.2; and finally, we derive the requisite penalty expressions.

6.1 The first overage

In this section, we study the first overage penalty, assuming that the demand arrival process is Poisson. Figure 5 depicts a sample path of the inventory level process.

The next theorem establishes an embedded martingale property of $\{M_2(t)\}$.

Theorem 4 *Suppose that demands arrive according to a compound Poisson process. Then the process $\{M_2(t)\}$ in Eq. (3.8) has the martingale property at $\tau_1^{(S)}$.*

Proof By Proposition 2, $\{M_2(t)\}$ is a martingale. Furthermore, on $\{\tau_1^{(S)} < \infty\}$, we have the inequalities $0 \le M_2(t) \le e^{c_2 u}$ uniformly in t by Eq. (3.8). The proof is completed by applying the Optional Sampling Theorem.   □

From Theorem 4 we have $M_2(0) = \mathbb{E}[M_2(t)|I(0) = u]$. It readily follows that at $t = \tau_1^{(S)}$,

$$M_2(0) = e^{c_2 u} = \mathbb{E}\big[e^{-r\tau_1^{(S)} + c_2 I(\tau_1^{(S)})} \,\big|\, I(0) = u\big] = e^{c_2 S}\mathbb{E}\big[e^{-r\tau_1^{(S)}} \,\big|\, I(0) = u\big], \qquad (6.7)$$

where the second equality follows from Eq. (6.3). It follows that

$$\mathbb{E}\big[e^{-r\tau_1^{(S)}} \,\big|\, I(0) = u\big] = e^{c_2(u-S)}, \qquad (6.8)$$

so the expected conditional discounted penalty for the first overage is

$$\mathbb{E}\big[w_1^{(S)} e^{-r\tau_1^{(S)}} \,\big|\, I(0) = u\big] = w_1^{(S)}\mathbb{E}\big[e^{-r\tau_1^{(S)}} \,\big|\, I(0) = u\big] = w^{(S)} e^{c_2(u-S)}, \qquad (6.9)$$

where the second equality holds by Eq. (6.8).

We mention that Eqs. (6.7)–(6.9) are also valid for $u \le 0$ and $u = S$. Accordingly, we shall assume that the initial inventory level satisfies $u \le S$.

6.2 The subsequent overage sequence

In this section, we study the sequence of subsequent overage penalties, assuming that the demand arrival process is Poisson but the demand size is generally distributed.

For $k > 1$, the hitting times $\{\tau_k^{(S)}\}$ constitute renewal times of the inventory process. Therefore, the time increment $\tau_k^{(S)} - \tau_{k-1}^{(S)}$ is independent of both $\tau_k^{(S)}$ and $I(0)$. Consequently, we can write

$$
\mathbb{E}\big[e^{-r\tau_k^{(S)}} \mid I(0) = u\big] = \mathbb{E}\big[e^{-r\tau_{k-1}^{(S)} - r(\tau_k^{(S)} - \tau_{k-1}^{(S)})} \mid I(0) = u\big]
$$
$$
= \mathbb{E}\big[e^{-r\tau_{k-1}^{(S)}} \mid I(0) = u\big]\mathbb{E}\big[e^{-r(\tau_k^{(S)} - \tau_{k-1}^{(S)})}\big]. \tag{6.10}
$$

Next, write $\tau_k^{(S)} - \tau_{k-1}^{(S)} = T_{k-1}^{(S)} + (\tau_k^{(S)} - \tau_{k-1}^{(S)} - T_{k-1}^{(S)})$, and note that the $(k-1)$-th replenishment idle time, $T_{k-1}^{(S)}$, is independent of the residual time to hit S, namely, $\tau_k^{(S)} - \tau_{k-1}^{(S)} - T_{k-1}^{(S)}$; furthermore, by a renewal argument the latter is distributed as $\tau_1^{(S)}$, given $I(0) = S$. Using these observations, we can write

$$
\mathbb{E}\big[e^{-r(\tau_k^{(S)} - \tau_{k-1}^{(S)})}\big] = \mathbb{E}\big[e^{-rT_{k-1}^{(S)}}\big]\mathbb{E}\big[e^{-r(\tau_k^{(S)} - \tau_{k-1}^{(S)} - T_{k-1}^{(S)})}\big]
$$
$$
= \mathbb{E}\big[e^{-rT_{k-1}^{(S)}}\big]\mathbb{E}\big[e^{-r\tau_1^{(S)}} \mid I(0) = S\big].
$$

Next, observe that on the right-hand side above, the first term is given by Eq. (6.5) and the second term follows from Eq. (6.8), yielding for all $k \geq 1$,

$$
\mathbb{E}\big[e^{-r(\tau_k^{(S)} - \tau_{k-1}^{(S)})}\big] = \frac{\lambda}{\lambda + r} \int_0^\infty f_D(x)e^{c_2(x-S)}dx = \frac{\lambda}{\lambda + r}e^{-c_2 S}\tilde{f}_D(-c_2) = \gamma. \tag{6.11}
$$

Note that $0 < \gamma < 1$ since $0 < \frac{\lambda}{\lambda+r} < 1$, $0 < e^{-c_2 S} < 1$ and $0 < \tilde{f}_D(-c_2) < 1$. Substituting Eq. (6.11) into Eq. (6.10), we obtain the recursive relation

$$
\mathbb{E}\big[e^{-r\tau_k^{(S)}} \mid I(0) = u\big] = \mathbb{E}\big[e^{-r\tau_{k-1}^{(S)}} \mid I(0) = u\big]\gamma, \tag{6.12}
$$

and applying Eq. (6.12) $k-1$ times, we obtain in view of Eq. (6.8)

$$
\mathbb{E}\big[e^{-r\tau_k^{(S)}} \mid I(0) = u\big] = \gamma^{k-1}e^{c_2(u-S)}. \tag{6.13}
$$

Finally, utilizing Eq. (6.13), we have

$$
\mathbb{E}\left[\sum_{k=1}^\infty w_k^{(S)}e^{-r\tau_k^{(S)}} \mid I(0) = u\right] = e^{c_2(u-S)}\sum_{k=1}^\infty w_k^{(S)}\gamma^{k-1}. \tag{6.14}
$$

Denoting the prescribed perpetual overage penalty budget by p_O, we have the following result.

Theorem 5 *If the overage penalty functions $w_k^{(S)} = w^{(S)}$ for level S are constant, independent of k, then the overage penalty is given by*

$$
w^{(S)} = p_O(1 - \gamma)e^{c_2(S-u)}, \tag{6.15}
$$

where $\gamma = \frac{\lambda}{\lambda+r}e^{-c_2 S}\tilde{f}_D(-c_2)$.

Proof Substituting $w_k^{(S)} = w^{(S)}$ into Eq. (6.14), one has

$$\mathbb{E}\left[\sum_{k=1}^{\infty} w_k^{(S)} e^{-r\tau_k^{(S)}} \,\Big|\, I(0) = u\right] = e^{c_2(u-S)} w^{(S)} \sum_{k=1}^{\infty} \gamma^{k-1} = \frac{e^{c_2(u-S)} w^{(S)}}{1-\gamma}. \qquad (6.16)$$

By the principle of risk-neutral valuation, Eq. (6.16) equals p_O, which is the amount a risk-neutral manager would budget at time 0 to defray the future cost of discounted overage penalties. Therefore,

$$p_O = \frac{e^{c_2(u-S)} w^{(S)}}{1-\gamma}. \qquad (6.17)$$

The result readily follows via simple algebra. □

Theorem 5 shows that in the special case, $I(0) = u = S$, the overage penalty becomes $w^{(S)} = p_O(1 - \frac{\lambda}{\lambda+r} e^{-c_2 S} \tilde{f}_D(-c_2))$.

7 Concluding remarks

The main goal of this investigation was to formulate the inventory penalty function problem and derive explicit expressions for overage and underage penalties under the principle of risk-neutral valuation for production-inventory systems with constant replenishment and compound renewal demand stream, with special emphasis on Poisson demand arrivals and exponential demand sizes.

The problem we addressed in this paper was the *inverse valuation problem*, namely, devising proper dynamic *penalty functions* for assessing overage and underage penalties, given a prescribed penalty budget. Our solution approach treated inventory overage and underage penalties as a series of perpetual American options. To price such options, we first constructed auxiliary exponential martingale processes in terms of the underlying inventory process. Each auxiliary martingale process was then used to establish some intermediate results which were further used to derive the expected discounted perpetual inventory penalties, conditioned on the initial inventory level. Under the risk-neutral valuation principle, the obtained conditional expected total discounted perpetual inventory penalty equals a prescribed value of the penalty budget that a risk-neutral inventory manager would set aside at time zero to defray the cost of future penalties. Accordingly, we derived the requisite penalty functions for both overage and underage penalties, subject to a prescribed penalty budget.

For general compound renewal demands, we obtained the requisite necessary and sufficient martingale condition. We proceeded to prove that this equation has two roots: one negative and one positive. We then used the negative root to construct a martingale process, associated with underage penalties, while the positive root was used to construct a martingale process, associated with overage penalties. Finally, we obtained explicit expressions for underage and overage penalty functions for some special cases of compound Poisson demand arrivals and exponential demand sizes.

Appendix

Proof of Lemma 1

Substituting Eq. (3.1) into Eq. (2.10) and noting that $I(t)$ and $M(t)$ determine each other for each t, we rewrite the left hand side of Eq. (2.6) for any $0 \le i \le j$ as

$$\mathbb{E}\big[M(A_j) \mid M(t), t \le A_i\big] = \mathbb{E}\big[e^{-rA_j+cI(A_j)} \mid M(t), t \le A_i\big]$$

$$= \mathbb{E}\big[e^{-rA_i+cI(A_i)}e^{-r[A_j-A_i]+c[I(A_j)-I(A_i)]} \mid I(t), t \le A_i\big]$$

$$= M(A_i)\mathbb{E}\big[e^{-r[A_j-A_i]+c(\rho[A_j-A_i]-c\sum_{k=i+1}^{j} D_k)} \mid I(A_i)\big]$$

$$= M(A_i)\mathbb{E}\big[e^{-(r-c\rho)(A_j-A_i)-c\sum_{k=i+1}^{j} D_k} \mid I(A_i)\big]$$

$$= M(A_i)\mathbb{E}\big[e^{-(r-c\rho)\sum_{k=i+1}^{j} T_k}\big]\mathbb{E}\big[e^{-c\sum_{k=i+1}^{j} D_k}\big]. \qquad (8.1)$$

For the embedded martingale condition to hold, it is necessary and sufficient that the right-most side of Eq. (8.1) be equal to $M(A_i)$, which is equivalent to the condition

$$1 = \mathbb{E}\big[e^{-(r-c\rho)\sum_{k=i+1}^{j} T_k}\big]\mathbb{E}\big[e^{-c\sum_{k=i+1}^{j} D_k}\big] = \tilde{f}_{\sum_{k=i+1}^{j} T_k}(r-c\rho)\,\tilde{f}_{\sum_{k=i+1}^{j} D_k}(c)$$

$$= \tilde{f}_T^{j-i}(r-c\rho)\tilde{f}_D^{j-i}(c) = \big[\tilde{f}_T(r-c\rho)\tilde{f}_D(c)\big]^{j-i}.$$

Equation (3.3) now readily follows, since $\tilde{f}_T(r-c\rho)\tilde{f}_D(c)$ is non-negative. □

Proof of Lemma 2
To prove part (a), we show that the derivatives of $l_T(c)$ satisfy

$$l'_T(c) = -\frac{\frac{\partial}{\partial c}\tilde{f}_T(r-c\rho)}{\tilde{f}_T(r-c\rho)} < 0, \qquad (8.2)$$

$$l''_T(c) = -\frac{\tilde{f}_T(r-c\rho)\frac{\partial^2}{\partial c^2}\tilde{f}_T(r-c\rho) - [\frac{\partial}{\partial c}\tilde{f}_T(r-c\rho)]^2}{[\tilde{f}_T(r-c\rho)]^2} < 0. \qquad (8.3)$$

Equation (8.2) follows from the facts

$$\tilde{f}_T(r-c\rho) = \mathbb{E}\big[e^{-(r-c\rho)T_i}\big] > 0, \qquad (8.4)$$

$$\frac{\partial}{\partial c}\tilde{f}_T(r-c\rho) = \rho\int_0^\infty te^{-(r-c\rho)t}f_T(t)dt = \rho\mathbb{E}\big[T_ie^{-(r-c\rho)T_i}\big] > 0, \qquad (8.5)$$

where Eq. (8.4) is an immediate consequence of the Laplace transform, and Eq. (8.5) follows by the Leibniz integral rule.

To prove (8.3), note that the denominator is positive, so it remains to show that the numerator is positive too. Differentiating Eq. (8.5) with the aid of the Leibniz integral rule yields

$$\frac{\partial^2}{\partial c^2}\tilde{f}_T(r-c\rho) = \rho^2\int_0^\infty t^2e^{-(r-c\rho)t}f_T(t)dt = \rho^2\mathbb{E}\big[T_i^2e^{-(r-c\rho)T_i}\big]. \qquad (8.6)$$

Substituting (8.4), (8.5) and (8.6) into the numerator of Eq. (8.3), the latter becomes

$$\mathbb{E}\big[e^{-(r-c\rho)T_i}\big]\rho^2\mathbb{E}\big[T_i^2e^{-(r-c\rho)T_i}\big] - \rho^2\mathbb{E}^2\big[T_ie^{-(r-c\rho)T_i}\big]. \qquad (8.7)$$

Finally, applying the Cauchy-Schwarz inequality (Karr 1993) to the product of expectations in the first term of Eq. (8.7) results in the inequality,

$$\mathbb{E}\big[e^{-(r-c\rho)T_i}\big]\mathbb{E}\big[T_i^2e^{-(r-c\rho)T_i}\big] > \mathbb{E}^2\big[T_ie^{-(r-c\rho)T_i}\big], \qquad (8.8)$$

where the strict inequality follows from the fact that the random variables on the left-hand side above are not proportional to each other. Equation (8.8) establishes that the numerator of Eq. (8.3) is positive, thereby completing the proof of part (a).

To prove part (b), we apply an argument similar to that for part (a). Accordingly, we show that the derivatives of $l_D(c)$ satisfy

$$l'_D(c) = \frac{\tilde{f}'_D(c)}{\tilde{f}_D(c)} < 0, \tag{8.9}$$

$$l''_D(c) = \frac{\tilde{f}''_D(c)\tilde{f}_D(c) - \tilde{f}'_D(c)\tilde{f}'_D(c)}{[\tilde{f}_D(c)]^2} > 0. \tag{8.10}$$

Equation (8.9) follows from the facts

$$\tilde{f}_D(c) = \mathbb{E}\big[e^{-cD_i}\big] = \int_0^\infty e^{-cx} f_D(x)dx > 0, \tag{8.11}$$

$$\tilde{f}'_D(c) = \int_0^\infty (-x)e^{-cx} f_D(x)dx = \mathbb{E}\big[-D_i e^{-cD_i}\big] < 0, \tag{8.12}$$

where Eq. (8.11) is an immediate consequence of the Laplace transform, and Eq. (8.12) follows by the Leibniz integral rule.

To prove Eq. (8.10), note that the denominator is positive, so it remains to show that the numerator is positive too. Differentiating Eq. (8.12) with the aid of the Leibniz integral rule yields,

$$\tilde{f}''_D(c) = \int_0^\infty x^2 e^{-cx} f_D(x)\,dx = \mathbb{E}\big[D_i^2 e^{-cD_i}\big] > 0. \tag{8.13}$$

Substituting Eqs. (8.11), (8.12) and (8.13) into the numerator of Eq. (8.10), the latter becomes

$$\mathbb{E}\big[D_i^2 e^{-cD_i}\big]\mathbb{E}\big[e^{-cD_i}\big] - \mathbb{E}^2\big[D_i e^{-cD_i}\big]. \tag{8.14}$$

Finally, applying the Cauchy-Schwarz inequality to the first term of Eq. (8.14) results in the inequality,

$$\mathbb{E}\big[D_i^2 e^{-cD_i}\big]\mathbb{E}\big[e^{-cD_i}\big] > \mathbb{E}^2\big[D_i e^{-cD_i}\big], \tag{8.15}$$

where the strict inequality follows from the fact that the random variables on the left-hand side above are not proportional to each other. Equation (8.15) establishes that the numerator of Eq. (8.10) is positive, thereby completing the proof of part (b). □

Proof of Proposition 3
Let D denote a generic exponential demand size with pdf $f_D(x) = \beta e^{-\beta x}, \beta > 0$. By Lemma 3 we can write

$$f_{L_i^{(0)}|I_i^{(0)}-}(y|x) = f_{D_i^{(0)}|I_i^{(0)}-}(x + y|x) = f_D(y), \quad x \geq 0, \ y > 0, \tag{8.16}$$

where the second equality follows from the memorylessness of the exponential distribution.

Define the functions

$$g(x, y|u) = \int_0^\infty e^{-rt} f_{I(\tau_1^{(0)}-), L(\tau_1^{(0)}), \tau_1^{(0)}|I(0)}(x, y, t|u)\,dt, \tag{8.17}$$

and

$$h(x|u) = \int_0^\infty e^{-rt} f_{I(\tau_1^{(0)}-), \tau_1^{(0)}|I(0)}(x, t|u)\,dt. \tag{8.18}$$

Next, we rewrite Eq. (8.17) as

$$g(x, y|u) = \int_0^\infty e^{-rt} f_{L(\tau_1^{(0)})|I(\tau_1^{(0)}-), \tau_1^{(0)}, I(0)}(y|x, t, u) \times f_{I(\tau_1^{(0)}-), \tau_1^{(0)}|I(0)}(x, t|u) \, dt$$

$$= \int_0^\infty e^{-rt} f_D(y) \times f_{I(\tau_1^{(0)}-), \tau_1^{(0)}|I(0)}(x, t|u) \, dt$$

$$= f_D(y) \times \int_0^\infty e^{-rt} f_{I(\tau_1^{(0)}-), \tau_1^{(0)}|I(0)}(x, t|u) \, dt, \tag{8.19}$$

where the first equality holds by the chain rule of conditional probabilities,

$$f_{I(\tau_1^{(0)}-), L(\tau_1^{(0)}), \tau_1^{(0)}|I(0)} = f_{L(\tau_1^{(0)})|I(\tau_1^{(0)}-), \tau_1^{(0)}, I(0)} \times f_{I(\tau_1^{(0)}-), \tau_1^{(0)}|I(0)},$$

and the second equality follows from Eq. (8.16) combined with the fact that $f_{L(\tau_1^{(0)})|I(\tau_1^{(0)}-), \tau_1^{(0)}, I(0)} = f_{L(\tau_1^{(0)})|I(\tau_1^{(0)}-)}$. In view of Eqs. (8.18) and (8.19), we now have

$$g(x, y|u) = h(x|u) f_D(y). \tag{8.20}$$

Furthermore, integrating both sides of Eq. (8.18) with respect to x from zero to infinity, we get

$$\mathbb{E}\left[e^{-r\tau_1^{(0)}} \mid I(0) = u\right] = \int_0^\infty h(x|u) \, dx. \tag{8.21}$$

Next, we write

$$\mathbb{E}\left[e^{-r\tau_1^{(0)} + c_1 I(\tau_1^{(0)})} \mid I(0) = u\right]$$

$$= \mathbb{E}\left[e^{-r\tau_1^{(0)} - c_1 L(\tau_1^{(0)})} \mid I(0) = u\right]$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty e^{-rt - c_1 y} f_{I(\tau_1^{(0)}-), L(\tau_1^{(0)}), \tau_1^{(0)}|I(0)}(x, y, t|u) \, dx \, dy \, dt$$

$$= \int_0^\infty \int_0^\infty e^{-c_1 y} g(x, y|u) \, dx \, dy = \int_0^\infty \int_0^\infty e^{-c_1 y} h(x|u) f_D(y) \, dx \, dy$$

$$= \int_0^\infty e^{-c_1 y} f_D(y) \, dy \, \mathbb{E}\left[e^{-r\tau_1^{(0)}} \mid I(0) = u\right] = \frac{\beta}{\beta + c_1} \mathbb{E}\left[e^{-r\tau_1^{(0)}} \mid I(0) = u\right]. \tag{8.22}$$

Here, the first equality follows from the fact that $L(\tau_1^{(0)}) = -I(\tau_1^{(0)})$, the second and third equality hold by definition, the fourth equality holds by Eq. (8.20), the fifth equality follows from (8.21), and the sixth equality is a consequence of exponentially distributed demand sizes. Finally, the above equation equals $e^{c_1 u}$ by Eq. (5.4), which completes the proof. $\qquad\square$

References

Ernst, R., & Powell, S. G. (1995). Optimal inventory policies under service-sensitive demand. *European Journal of Operational Research*, 87(2), 316–327.

Hull, J. C. (2002). *Options, futures, and other derivative securities* (5th ed.). New York: Prentice Hall.

Gerber, H. U., & Shiu, E. S. W. (1994). Martingale approach to pricing perpetual American options. *ASTIN Bulletin*, 24(2), 195–220.

Gerber, H. U., & Shiu, E. S. W. (1996). Martingale approach to pricing perpetual American options on two stocks. *Mathematical Finance*, 6(3), 303–322.

Gerber, H. U., & Shiu, E. S. W. (1998). On the time value of ruin. *North American Actuarial Journal*, 2, 48–78.

Karr, A. F. (1993). *Probability*. Berlin: Springer.

Luenberger, D. G. (1998). *Investment science*. London: Oxford University Press.

Ofenbakh, I. (2008). Estimation of lost sales revenue opportunity due to out of stock inventory. In INFORMS, 29 October, 2008, NY, USA.

Relph, G., & Barrar, P. (2003). Overage inventory—how does it occur and why is it important? *International Journal of Production Economics*, *81–82*, 163–171.

Ross, S. M. (1996). *Stochastic processes* (2nd ed.). New York: Wiley.

Schwartz, B. L. (1966). A new approach to Stockout penalties. *Management Science*, *12*, 538–544.

Schwartz, B. L. (1970). Optimal inventory policy in perturbed demand models. *Management Science*, *16*, B509–B518.

Stowe, J. D., & Su, T. (1997). A contingent-claims approach to the stocking decision. *Financial Management*, *26*(4), 42–58.

Walter, C. K., & Grabner, J. R. (1975). Stockout cost models: empirical tests in a retail situation. *The Journal of Marketing*, *39*(3), 56–60.

Zipkin, P. (2000). *Foundations of inventory management* (pp. 25–26). Boston: McGraw Hill.