Optimal (s,S) Inventory Policies for Levy Demand Processes

Robin O. Roundy* and Gennady Samorodnitsky**

School of Operations Research and Industrial Engineering
Cornell University, Ithaca, NY 14853

Abstract

A *Levy jump process* is a continuous-time, real-valued stochastic process which has independent and stationary increments, with no Brownian component. We study some of the fundamental properties of Levy jump processes and develop (s, S) inventory models for them. Of particular interest to us is the gamma-distributed Levy process, in which the demand that occurs in a fixed period of time has a gamma distribution.

We study the relevant properties of these processes, and we develop a quadratically convergent algorithm for finding optimal (s, S) policies. We develop a simpler heuristic policy and derive a bound on its relative cost. For the gamma-distributed Levy process this bound is 7.9% if backordering unfilled demand is at least twice as expensive as holding inventory.

Most easily-computed (s, S) inventory policies assume the inventory position to be uniform and assume that there is no overshoot. Our tests indicate that these assumptions are dangerous when the coefficient of variation of the demand that occurs in the reorder interval is less than one. This is often the case for low-demand parts that experience sporadic or spiky demand. As long as the coefficient of variation of the demand that occurs in one reorder interval is at least one, and the service level is reasonably high, all of the polices we tested work very well. However even in this region it is often the case that the standard Hadley-Whitin cost function fails to have a local minimum.

$*$ Research supported by the NSF grant DDM-921704.
$**$ Research supported by the NSF grant DMS-94-00535 and the NSA grant MDA904-95-H-1036.

Optimal (s,S) Inventory Policies for Levy Demand Processes

Robin O. Roundy and Gennady Samorodnitsky

Section 1: Introduction

We consider (s, S) inventory models in which both time and inventory are modelled as continuous quantities, the lead times are deterministic, and customer service is modelled through cost minimization rather than constraints on service levels. We discuss two important models of this sort below. For a more thorough discussion of the literature see [Zheng 1992]. Also see [Gallego 1994] and [Axsater 1993].

By far the most popular model for (s, S) and (R, Q) policies is originally due to Hadley and Whitin [1963]. This model has been the mainstay of introductory textbooks on inventory theory for over 20 years (see, for example, [Johnson and Montgomery 1974], [Nahmias 1993] and [Vollman, Berry and Whybark 1992]). This model assumes that the inventory position lands on the reorder point rather then jumping over it, so the distinction between (s, S) policies and (R, Q) policies disappears. This model also assumes that the inventory position is uniformly distributed between s and S. The cost function makes use of approximations (see Section 4). This model is conceptually and computationally simple. Any demand distribution can be used, continuous or discrete. The computations required are relatively simple, and the model gives good solutions for most real-world inventory systems.

The main disadvantage of the Hadley-Whitin model is its robustness. When the backorder cost is sufficiently low or the order cost is sufficiently high the cost function can fail to have a local minimum. For example, suppose that the holding cost is 1 dollar per item per day, the mean demand rate is 1 item per day, the order cost is K dollars, the backorder cost is $\hat{p}$ dollars per item, and the demand that occurs during one lead time is uniformly distributed over the interval $[0, 2]$. If $\hat{p}\,(\hat{p} - 2) \leq 2K$ the cost function has no local minimum ((31) in Section 4 is non-decreasing). Thus if $\hat{p} = 10$ and $K \geq 40$, or if $\hat{p} = 40$ and $K \geq 760$, the cost function has no local minimum.

More recently Zheng developed an elegant (R, Q) inventory model [Zheng 1992]. This model differs from that of Hadley and Whitin in that Zheng uses a time-weighted backorder cost of p dollars per item per day rather than $\hat{p}$ dollars per item. Zheng assumes that the cumulative demand process is non-decreasing, has continuous sample paths, and has identically distributed increments. The continuity of the sample paths implies that the inventory position lands on the reorder point rather than jumping over it, and that the inventory position is uniformly distributed between s and S. As in the Hadley-Whitin model, the distinction between (s, S) policies and (R, Q) policies disappears.

Zheng makes no approximations in the cost function, so it is not surprising that Zheng's model is more robust than the Hadley-Whitin model. Zheng's model accomodates any demand distribution. It is more complex than the Hadley-Whitin model, both conceptually and computationally, but it is simple enough to teach in introductory courses and efficient enough to use in large, real-world inventory systems.

A *Levy jump process* is a continuous-time, real-valued stochastic process which has independent and identically distributed (*i.i.d.*) increments, with no Brownian component. According to the Levy Decomposition Theorem [Hida 1970, p. 45], any real-valued Levy process with *i.i.d.* increments can be expressed as $X(t) = \mu \cdot t + \sigma \cdot B(t) + J(t)$, where $\mu \cdot t$ is a deterministic drift, $B(t)$ is a Brownian motion and $J(t)$ is a Levy jump process. Zheng's cumulative demand process is assumed to be non-decreasing and continuous. Brownian motion is not non-decreasing and Levy jump processes are discontinuous, so according to this theorem either Zheng's demand process is deterministic or it fails to have independent increments.

Processes with dependent, identically-distributed increments that satisfy Zheng's assumptions have been studied [Serfozo and Stidham 1978, Browne and Zipkin 89]. For these demand processes Zheng's model produces policies that are optimal within the class of (s, S) policies. However (s, S) policies are no longer an optimal class because past demands are correlated with future demands.

A primary goal of this research is to study some fundamental properties of Levy jump processes and to develop (s, S) inventory models for them. Levy jump processes are a realistic and rich class of stochastic processes in which the inventory position usually jumps over the reorder point rather than landing exactly on it. Of particular interest to us is the gamma-distributed Levy process, in which the demand that occurs in a fixed period of time has a gamma distribution. We study some relevant properties of these processes, and we develop a quadratically convergent algorithm for finding optimal (s, S) policies for them. We develop a simpler policy called the Mass Uniform heuristic, and derive a bound on its relative cost. For the gamma-distributed Levy process this bound is 7.9% if backordering unfilled demand is at least twice as expensive as holding inventory. Our quadratically-convergent algorithm for computing optimal policies can also be used to compute Zheng's (R, Q) policy more efficiently.

Another of our goals in initiating this research was to convince ourselves that Zheng's algorithm is as robust as it appears to be. Our computational tests indicate that it is very robust, much more so than the cost-minimization versions of the Hadley-Whitin model are, but that it can produce poor policies in some realistic scenarios.

Overview

This paper is organized as follows. In Section 2 we develop the key properties of general Levy demand processes and, specifically, of the gamma-distributed Levy process. In Section 3 we develop our inventory model for (s, S) policies for Levy demand processes. We also present a quadratically-convergent algorithm for computing optimal policies, and we discuss service levels. In Section 4 we develop the Mass Uniform heuristic for computing (s, S) policies, we prove that the relative cost of the Mass Uniform heuristic is at most 7.9% if backordering unfilled demand is at least twice as expensive as holding inventory. We develop bounds on the performance of the (s, S) policy derived from Zheng's (R, Q) policy, and we briefly describe the other policies that we have tested. In Section 5 we summarize our computational experiments, and in Section 6 we draw our conclusions. Appendix 1 contains a glossary of notation, and Appendix 2 contains all of the proofs.

Section 2: Levy Demand Processes

We begin this section with two definitions.

Definition 1: A *Levy jump process* is a continuous-time, real-valued, stochastic process $D(t)$ which has the following properties.

$D(0) = 0.$

The sample paths of $D(t)$ are right-continuous functions of t.

For every $0 \leq t_1 < t_2 < \ldots < t_n$, $D(t_1)$, $D(t_2) - D(t_1)$, ..., $D(t_n) - D(t_{n-1})$ are independent.

For every $t, h > 0$, $D(t+h) - D(t)$ and $D(h)$ have the same distribution.

$D(t)$ has no Brownian component.

Definition 2: A *Levy demand process* is a non-decreasing Levy jump process which has positive, finite expected value for all $t > 0$.

We use Levy demand processes to model cumulative demand. Thus $D(t)$ is the total demand that occurs in the time interval $[0, t]$ and $D(t, u)$ is the total demand that occurs in the time interval $(t, u]$. By selecting the units of measure in our inventory models appropriately, we assume without loss of generality that

$$\mathrm{E}[D(t)] = t.$$

Examples of Levy demand processes include the compound Poisson processes. In a Levy demand process, demand for randomly-sized quantities of inventory occurs instantaneously at random points in time, creating "demands" or "jumps" (discontinuities in the cumulative demand process). If the demand quantities are all integer multiples of some number, we say that the jump sizes are arithmetic [Feller 1996, p. 360]. Some of our results require non-arithmetic jump sizes.

Property 1: The jump sizes are non-arithmetic.

Any of three functions can be used to characterize a Levy demand process $D(t)$. The first of these comes from observing the time epochs at which a demand of size greater than x occurs. Since the demand process $D(t)$ has i.i.d. increments, these time epochs form a Poisson process. $\psi^+(x)$ is the rate at which demands of size greater than x occur. Algebraically we have

$$\psi^+(x) = \lim_{t \to 0} \frac{1}{t}[1 - F_{D(t)}(x)] \tag{1}$$

(See Feller, p. 302, Theorem 1. Note that the distributions of $D(t)$ for $t > 0$ form a convolution semi-group as defined in Feller p. 293). Clearly $\psi^+(x)$ is non-increasing. The mean rate per unit time at which demand occurs is $\int_0^\infty x\, d\{-\psi^+(x)\} = \mathrm{E}[D(1)] = 1$. Because $D(t)$ is well-defined and non-decreasing, $\limsup_{x \to 0} x\, \psi^+(x) = 0$. If $\limsup_{x \to \infty} x\, \psi^+(x) > 0$ we have $\mathrm{E}[D(1)] = \infty$, which is contrary to our assumptions. Integration by parts leads to

$$1 = \mathrm{E}[D(1)] = \int_0^\infty x\, d\{-\psi^+(x)\} = \int_0^\infty \psi^+(x)\, dx\ . \tag{2}$$

A Levy demand process can be constructed from any non-negative, non-increasing function $\psi^+(x)$ satisfying (2). Compound Poisson processes are the Levy demand processes in which demands of any size occur at a finite rate, i.e., $\psi^+(0) < \infty$. For a compound Poisson process $\mathrm{P}(D(t) = 0) > 0$ for all $t \geq 0$, and $\psi^+(x)/\psi^+(0)$ is the probability that the size of a given jump is greater than x. On the other hand, if $\psi^+(x) \to \infty$ as $x \to 0$ then any open interval on the time axis contains a countably infinite number of jumps, and $\mathrm{P}(D(t) = 0) = 0$ for all $t > 0$.

The measure $d\{-\psi^+(x)\}$ is called the Levy measure. We define the random variables J and V by

$$F_J(z) = \int_0^z x\, d\{-\psi^+(x)\} \quad \text{and} \quad F_V(z) = \int_0^z \psi^+(x)\, dx\ . \tag{3}$$

The following lemma allows us to interpret J as the demand-weighted jump size.

Lemma 1: For all $T > 0$, $\mathrm{E}\left(\frac{1}{T}\int_0^T 1(D(t) - D(t^-) > z)\, d\{D(t)\}\right) = \overline{F}_J(z).$

All proofs are in Appendix 2. To illustrate Lemma 1, suppose that demands of size 1 arrive at rate $1/3$, and that demands of size 2 arrive at rate $1/3$. Then $\psi^+(x) = 2/3$ for $0 \leq x < 1$, $\psi^+(x) = 1/3$ for $1 \leq x < 2$, and $\psi^+(x) = 0$ otherwise. Half of the jumps are of size 1 and half are of size 2, but J weights the probabilities of the jumps by their size, so $J = 1$ with probability $1/3$, and $J = 2$ with probability $2/3$.

The second function that is used to characterize Levy demand processes is the distribution of $D(t)$. Because $D(t)$ has stationary, independent increments, the Laplace transform $\mathcal{L}_{D(t)}(\gamma)$ of $D(t)$ satisfies

$$\mathcal{L}_{D(t)}(\gamma) = [\mathcal{L}_{D(1)}(\gamma)]^t. \tag{4}$$

Since $D(t)$ is a non-trivial, non-negative random variable,

$$\mathcal{L}_{D(t)}(\gamma) < 1 \text{ for } t > 0. \tag{5}$$

By [Hida 1970][1],

$$\mathcal{L}_{D(t)}(\gamma) = e^{-t\int_0^\infty \gamma e^{-\gamma x}\psi^+(x)\, dx} = e^{-t\int_0^\infty (1 - e^{-\gamma x})\, d\{-\psi^+(x)\}}. \tag{6}$$

The third function that can be used to characterize Levy demand processes is the expected length of time $\theta(x) \equiv \mathrm{E}[D^{-1}(x)]$ required to accumulate x units of demand, starting from a given point in time. $\theta(x)$ is related to the steady-state distribution of the inventory position. Clearly $\theta(0) = 0$. $\theta(0^+) > 0$ if

[1] Theorem 3.3 page 42. Hida's $d\{n(u)\}$ is our $d\{-\psi^+(u)\}$. His $X(t,\omega) + \int_0^\infty \frac{tu}{1+u^2}\, d\{n(u)\}$ is our $D(t)$.

and only if $\psi^+(0) < \infty$, i.e., if and only in $D(t)$ is a compound Poisson process. $\theta(x)$ is left-continuous, and

$$\theta(x)/x \to 1 \text{ as } x \to \infty. \tag{7}$$

One might be tempted to conjecture that $\theta(x) \equiv x$, but this is typically not the case.

Lemma 2: $\theta(x)$ satisfies $\theta(x) = \mathrm{E}[D(D^{-1}(x))] \geq x$ for all $x > 0$. If Property 1 holds then $\theta(x) - x \to \mathrm{E}[V]$ as $x \to \infty$.

In general $\mathrm{E}[V]$ can be either finite or infinite. Lemma 2 holds in either case. Note that $D(D^{-1}(x)) \geq x$ by definition.

Intuitively, the fact that $\theta(x) > x$ can be explained as follows. If we are told that $D(\tau) = d$ then we know that $D(t)$ lands on d rather than jumping over d. This fact increases the expected length of time that $D(t)$ spends in the interval $[d, d+\delta)$, $\delta > 0$.

The overshoot $D(D^{-1}(x)) - x$ is the amount by which a Levy demand process over-shoots a given value x. The expectation of the overshoot is $\theta(x) - x$. Note that many inventory models, including those of Zheng and Hadley and Whitin , assume that there is no overshoot, so $\theta(x) - x$ is equal to zero for all x. Lemma 2 describes the expectation of the overshoot. The following lemma allows us to interpret V as the asymtotic distribution of the overshoot.

Lemma 3: If Property 1 holds then $\mathrm{P}(D(D^{-1}(x)) - x > u) \to \int_u^\infty \psi^+(z)\,dz = \overline{F}_V(u)$ as $x \to \infty$.

Equation (3) and the integration by parts formula imply that $F_V(z) = F_J(z) + z \cdot \psi^+(z) > F_J(z)$ and $\mathrm{E}[J] = 2 \cdot \mathrm{E}[V]$. The means are finite if $\int_0^\infty x\, \psi^+(x)\,dx < \infty$. Thus the asymtotic overshoot V is stochastically smaller than the demand-weighted jump size J, and has a mean that is half of the mean of the demand-weighted jump size.

The Laplace transform of $\theta(x)$ is

$$\begin{aligned}
\mathcal{L}_\theta(\gamma) &\equiv \int_{x=0}^\infty e^{-\gamma\cdot x}\, d\{\theta(x)\} = \gamma \int_{x=0}^\infty \theta(x)\, e^{-\gamma\cdot x}\, dx \\
&= \gamma\, \mathrm{E} \int_{x=0}^\infty D^{-1}(x)\, e^{-\gamma\cdot x}\, dx = \gamma\, \mathrm{E} \int_{x=0}^\infty \left(\int_{t=0}^\infty 1(D(t) \leq x)\, dt \right) e^{-\gamma\cdot x}\, dx \\
&= \int_{t=0}^\infty \mathrm{E}\left(\int_{x=0}^\infty 1(D(t) \leq x)\, \gamma\, e^{-\gamma\cdot x}\, dx \right) dt = \int_{t=0}^\infty \left(\mathrm{E}\, e^{-\gamma D(t)}\right) dt \\
&= \int_{t=0}^\infty \mathcal{L}_{D(t)}(\gamma)\, dt = \int_{t=0}^\infty [\mathcal{L}_{D(1)}(\gamma)]^t\, dt = \frac{-1}{\ln[\mathcal{L}_{D(1)}(\gamma)]}\,.
\end{aligned} \tag{8}$$

The last two equalities follow from (4) and (5).

The Gamma-Distributed Levy Demand Process

We now turn our attention to a specific Levy demand process.

Definition 3: The *gamma-distributed Levy process* is the Levy demand process $D^*(t)$ for which $D^*(1)$ has an exponential distribution with mean one.

Note that $D^*(t)$ has a gamma distribution with shape parameter t and rate parameter 1. This is without loss of generality; in continuous-time inventory models we can choose our units of measure for time and inventory so that the demand that occurs in one day has a mean of 1 and a variance of 1. This process is not a new one (see, for example, [Prabhu 1980] and [Feller 1966, p. 567]). Note that

$$\mathcal{L}_{D^*(t)}(\gamma) = \frac{1}{(1+\gamma)^t}\,. \tag{9}$$

Figure 1: $\psi^+(x)$

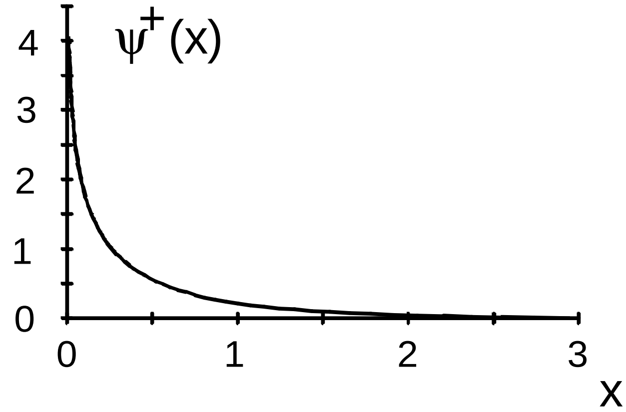
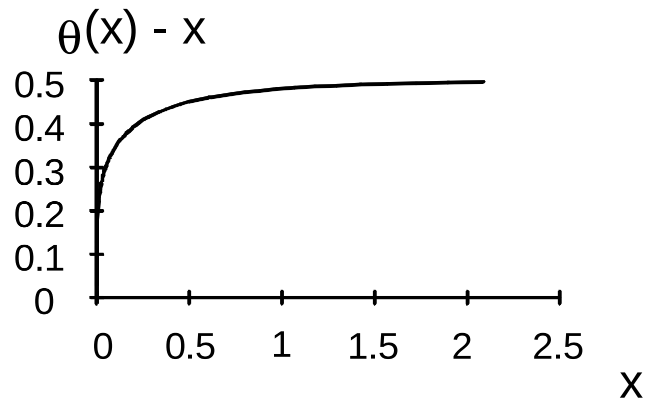


Figure 2: $\theta(x) - x$



Lemma 4: For the gamma-distributed Levy demand process $D^*(t)$,

$$\psi^+(x) = \int_x^\infty \frac{e^{-y}}{y}\, dy \quad \text{and} \tag{10}$$

$$\theta'(x) = \int_{t=0}^\infty \frac{x^{t-1}}{\Gamma(t)} e^{-x}\, dt\,, \tag{11}$$

where $\theta'(x)$ is the derivative of $\theta(x)$, and $\theta(0) = 0$.

For the gamma-distributed Levy process, the demand-weighted jump size J has an exponential distribution (see (10) and (3)). The graph of $\psi^+(x)$ in Figure 1 indicates that orders for small quanitites of inventory occur much more often than orders for large quantities, but there is no finite upper bound on the maximum order quantity. By Lemma 3 the same can be said for the overshoot quantities.

A graph of $\theta(x) - x$ appears in Figure 2. From (10) we obtain E$[V] = 1/2$, so by Lemma 2, $\theta(x) - x \to 1/2$ as $x \to \infty$. Suppose that $x > 0.3$. Figure 2 indicates that even though the mean demand rate is equal to one, from a given starting point, the expected amount of time required to accumulate an additional x units of demand is between $x + 0.4$ and $x + 0.5$.

We now list some other properties of the gamma-distributed Levy process $D^*(t)$. The presentation of these properties is designed to facilitate the statements of our lemmas and theorems.

Property 2: $\theta(0^+) = 0$. $\theta(x)$ is non-decreasing and absolutely continuous. $\theta'(x)$ is non-increasing.

Property 3: $\theta(x)$ is continuous, and $\theta'(x) \to 1$ as $x \to \infty$.

Property 4: $\theta(x) > x \cdot \theta'(x)$ for all $x > 0$.

Property 5: $\theta'(x)$ is continuous.

Property 6: For all $x > 0$ there is a y, $0 < y < x$ such that $\theta'(y) > \theta'(x)$.

Lemma 5: Properties 1 through 6 hold for $D^*(t)$.

Note that the first claim of Property 2 implies that $D(t)$ is not a compound Poisson process.

The Mass-Uniform policy, which we develop in section 4, makes use of the following function.

$$\eta(x) \equiv \int_0^x z \cdot [\theta'(z) - 1]\, dz \,. \tag{12}$$

Lemma 6: For the gamma-distributed Levy process, $\eta(x) \to 1/12$ as $x \to \infty$.

In optimizing and evaluating (s, S) inventory policies for the gamma-distributed Levy process, the functions $\theta(x)$, $\theta'(x)$ and $\eta(x)$ are frequently used. They were tabulated using numerical integration, and numerically approximated. The approximations are given in Appendix 3. Let IP be the steady-state distribution of the inventory position. In the next section we will need to compute $\mathrm{E}[f(IP)] = \int_0^Q f(S-x)\frac{\theta'(x)}{\theta(Q)}\, dx$ numerically for several different functions $f(x)$. The functions $f(x)$ of interest are well-behaved, but the probability measure $\frac{\theta'(x)}{\theta(Q)}\, dx$ is ill-behaved in the neighborhood of $x = 0$. In fact, $\theta'(x) \cdot x\,(ln(x))^2 \to 1$ as $x \to 0$. Table 1 gives some values of the probability density $\frac{\theta'(x)}{\theta(Q)}$ for $Q = 1$. A transformation described in Appendix 3 was used to perform these integrations in a numerically stable manner.

Table 1: The density $\theta'(x)/\theta(1)$.

x	$1e-10$	$1e-8$	$1e-6$	$1e-4$	0.01	0.1	1
$\theta'(x)/\theta(1)$	$1.3e+7$	$2.1e+5$	$3.8e+3$	85.9	3.4	1.12	0.70

Section 3: The Inventory Model

We now turn our attention to (s, S) inventory policies. We consider a continuous-review, single-item inventory system with backorders and with a deterministic lead time L. There is a fixed ordering cost $K > 0$ (in dollars), a holding cost h (in dollars per item per day), and a time-weighted backorder cost p (in dollars per item per day). We choose our units of measure so that the demand $D(1)$ that occurs in one day has a mean and a variance of one. Therefore the demand D that occurs in one lead time satisfies

$$\mathrm{E}[D] = L \,.$$

The minimum order quantity is $Q = S - s$. For Levy demand processes orders come in irregular quantities because we usually over-shoot the reorder point. For the gamma-distributed Levy process the actual order quantity is strictly greater than Q with probability one.

Let $NI(t)$ be the net inventory at time t and let $IP(t) \equiv NI(t+L) + D(t,\, t+L]$ be the inventory position at time t. $IP(t)$ and $NI(t)$ are right-continuous. They have steady-state distributions IP and NI satisfying $NI = IP - D$ where D and IP are independent [Zipkin 1986]. Since $\theta(x)$ is equal to the mean time required for x units of demand to accumulate after an order is placed, $\mathrm{P}(IP \geq S - x) = \theta(x)/\theta(Q)$, $0 \leq x \leq Q$. The average order quantity and the mean time between orders are both equal to $\theta(Q)$. The quantity $\theta(Q) - Q$, referred to in Figure 2 and in Table 1, is the mean quantity by which the demand process overshoots the reorder point.

If the inventory position at time t is $IP(t) = x$ then the expected rate at which holding and backorder costs will be incurred at time $t + L$ is

$$G(x) \equiv \mathrm{E}[h(x-D)^+ + p(D-x)^+] = h(x-L) + (h+p)\cdot n_D(x) \tag{13}$$

where $n_D(x) = \mathrm{E}[(D-x)^+]$ is the partial expectation. Note that $G(x)$ is continuous and convex, and that

$$-p \le G'(x) \le h,\ \ G'(x) \to h \text{ as } x \to \infty, \text{ and } G'(x) \to -p \text{ as } x \to -\infty. \tag{14}$$

Let

$$w \equiv \sup\{x : G'(x) < h\}.$$

If D has a gamma distribution then $w = \infty$ and the following properties hold.

Property 7: $G(x)$ is convex, is strictly convex for $0 < x < w$, and $G'(x) < 0$ for $x < 0$.

Property 8: $G'(x)$ is continuous, and $G''(x)$ is continuous except possibly at $x = 0$.

Following [Zheng 1992] and [Zipkin 1986], the expected holding and backorder cost per cycle incurred by an (s, S) policy with $Q = S - s$ is

$$H(S,Q) \equiv \theta(Q)\cdot \mathrm{E}[G(IP)] = \int_0^Q G(S-x)\,\theta'(x)\,dx. \tag{15}$$

The average cost per day incurred by an (s, S) policy is therefore

$$c(S,Q) = \frac{K + H(S,Q)}{\theta(Q)}. \tag{16}$$

Following the notation in (15), we define

$$H'(S,Q) \equiv \int_0^Q G'(S-x)\,\theta'(x)\,dx \tag{17}$$

$$H''(S,Q) \equiv \int_0^Q G''(S-x)\,\theta'(x)\,dx \quad \text{and} \tag{18}$$

$$E(S,Q) \equiv \theta(Q)\cdot G(S-Q) - K - H(S,Q). \tag{19}$$

The first-order optimality conditions are

$$0 = \theta(Q)\cdot \frac{\partial\, c(S,Q)}{\partial\, S} = H'(S,Q) \quad \text{and} \tag{20}$$

$$0 = \frac{\theta(Q)}{\theta'(Q)}\cdot \frac{\partial\, c(S,Q)}{\partial\, Q} = G(S-Q) - c(S,Q) = \frac{E(S,Q)}{\theta(Q)}. \tag{21}$$

Note that (21) states that $G(S-Q)$ is the average cost, which corresponds exactly with Zheng [Zheng 1992]. Zheng's other optimality condition is $G(S-Q) = G(S)$, a simpler, special case of (20).

Lemma 7: Suppose that S is chosen optimally for a given $Q > 0$. Properties 2 and 3 imply that $G(S-Q) \ge G(S) \ge G(S-Q) - h\cdot[\theta(Q) - Q]$. If Properties 4 and 7 also hold then $G(S-Q) > G(S)$.

Recall that $\theta(Q) - Q \to \mathrm{E}[V]$ as $Q \to \infty$. We claim that

$$G(S-Q) - G(S) \to h\cdot \mathrm{E}[V] \text{ as } Q \to \infty.$$

If S and Q are large then (20), Properties 2 and 3, (14), and Lemma 2 indicate that $G(S-Q) = G(S) + \int_0^Q G'(S-x)[\theta'(x) - 1]\,dx \approx G(S) + \int_0^Q h\cdot[\theta'(x)-1]\,dx = G(S) + h\cdot[\theta(Q) - Q\,] \approx G(S) + h\cdot \mathrm{E}[V]$. A rigorous argument can be constructed along similar lines.

We define y_0 by

$$G'(y_0) = 0. \tag{22}$$

$G(y_0)$ is often referred to as the newsvendor cost or the buffer cost, because when the order cost is equal to zero the optimal policy has $Q = 0$ and $S = y_0$, and $G(y_0)$ is the average cost of that policy.

By the convexity of $G(x)$ and by (20), $S \geq y_0 \geq S - Q$.

Lemma 8: Suppose that S is chosen optimally for a given $Q > 0$. If Properties 7 and 8 hold then

$$S > y_0 > S - Q. \tag{23}$$

If Property 2 also holds, the solution to the first-order optimality conditions (20)-(21) is unique.

Algebraically the first-order optimality conditions are most conveniently expressed in terms of S and Q. However cost minimization is most efficiently carried out in terms of $s \equiv S - Q$ and S. Let $C(s, S) \equiv c(S, S - s)$. Then

$$\frac{\partial C(s,S)}{\partial S} = \frac{\partial c(S,Q)}{\partial S} + \frac{\partial c(S,Q)}{\partial Q}\bigg|_{Q=S-s} \quad \text{and} \quad \frac{\partial C(s,S)}{\partial s} = -\frac{\partial c(S,Q)}{\partial Q}\bigg|_{Q=S-s}. \tag{24}$$

By (21) and (24),

$$\begin{aligned}\frac{\partial^2 C(s,S)}{\partial s\, \partial S} &= \left(\frac{\partial}{\partial S} + \frac{\partial}{\partial Q}\right)\left(-\frac{\partial c(S,Q)}{\partial Q}\right)\bigg|_{Q=S-s} \\ &= \frac{\theta'(Q)}{\theta(Q)}\left[\frac{\partial c(S,Q)}{\partial S} + \frac{\partial c(S,Q)}{\partial Q}\right] + \left[c(S,Q) - G(S-Q)\right]\left[\frac{\partial}{\partial Q}\left(\frac{\theta'(Q)}{\theta(Q)}\right)\right]\bigg|_{Q=S-s},\end{aligned} \tag{25}$$

which is equal to zero whenever the first-order conditions are satisfied. This fact suggests that if we alternate Newton steps in s and S, we are likely to get overall quadratic convergence. What we actually do is similar, and it works for the same reason, but the algebra is somewhat simpler. We alternate between Newton steps in S which attempt to find a zero of $H'(S, S-s)$ in (s, S)-space, and Newton steps in s which attempt to find a zero of $E(S, S-s)$ in (s, S)-space. Because the Newton steps are performed in (s, S)-space rather than (S, Q)-space, the derivative of $H'(S, S-s)$ with respect to S and the derivative of $E(S, S-s)$ with respect to s are computed as in (24). The algorithm for systems with discrete demands described in [Federgruen and Zheng 1991] is similar in that it also alternates between improvements in s and improvements in S.

Cost Minimization Algorithm

Step 1: Select an initial policy (s, S).

Step 2: Set $s' \leftarrow s - \frac{E(S, S-s)}{G'(s)\,\theta(S-s)}$

Step 3: Set $S' \leftarrow S - \frac{H'(S, S-s')}{H''(S, S-s') + G'(s')\,\theta'(S-s')}$

Step 4: Set $s \leftarrow s'$ and $S \leftarrow S'$.

Step 5: Either terminate or go to Step 2.

Lemma 9: The Cost Minimization Algorithm is quadratically convergent if Properties 2, 7 and 8 hold, and if $\theta(x)$ is twice continuously differentiable in a neighborhood of the optimal Q.

The proofs of Lemmas 7 through 9 assume that D and $\theta(x)$ have certain properties, but they do not assume that they are related to a common Levy demand process in the sense of Section 2. In particular, if we assume that $\theta(x) \equiv x$ and allow D to be arbitrary, Lemmas 7 through 9 hold. Under these assumptions our (s, S) policies correspond to Zheng's (R, Q) policies. The Cost Minimization Algorithm is therefore quadratically convergent for Zheng's (R, Q) policies, if Properties 7 and 8 hold.

Service Levels

We conclude this section by discussing service levels. In discrete time inventory models the Type 1 Service Level is usually defined as the probability that a time period ends without any backorders. The continuous-time analogue is $\mathrm{P}[NI \geq 0]$, the fraction of the time that the net inventory is non-negative. The Type 2 Service Level is the fill rate, or the fraction of the demand that is met without backorders. In many continuous-time inventory models, including that of [Zheng 1992], these two service measures are equivalent. However they are not equivalent for Levy demand processes. The following is reminiscent of Zheng's fill-rate computation.

Lemma 10: In (s, S) policies for Levy demand processes,

$$\mathrm{P}[NI \geq 0] = \frac{p}{p+h} + \frac{H'(S,Q)}{(h+p)\,\theta(Q)}\,,$$

which is equal to $p/(p+h)$ for an optimal policy.

Lemma 11: For Levy demand processes, the fill rate of an (s, S) policy is

$$1 - \mathrm{E}[(J - (IP - D)^{+})^{+}/J]$$

where J is the demand-weighted jump size (see Lemma 1), and J, IP, and D are independent.

Lemma 11 can be explained intuitively as follows. Suppose that a demand of size $D(t) - D(t^{-})$ occurs at time t. At time t^{-} the on-hand inventory is $[IP(t-L) - D(t-L\,,\,t^{-})]^{+}$. The number of units ordered at time t, but not delivered to the client at time t, is

$$\left\{[D(t) - D(t^{-})] \;-\; [IP(t-L) - D(t-L\,,\,t^{-})]^{+}\right\}^{+}. \tag{26}$$

Note that the random variables $[D(t) - D(t^{-})]$, $IP(t-L)$ and $D(t-L\,,\,t^{-})$ are stochastically independent. $D(t-L\,,\,t^{-}) \sim D$, $IP(t-L)$ has steady-state distribution IP, and the demand-weighted distribution of the jump size $[D(t) - D(t^{-})]$ is J (see Lemma 1). Substituting these distributions into (26) leads to the formula in Lemma 11.

It is interesting to compare these two service measures, $\mathrm{P}(NI > 0)$ and the fill rate. Note that $1(NI \leq 0) \;\leq\; (J - NI^{+})^{+}/J \;\leq\; 1(NI \leq J)$. If β is the fill rate then taking expectations leads to $\mathrm{P}(NI > 0) \geq \beta \geq \mathrm{P}(NI > J)$. If either the lead time L or the order quantity Q is sufficiently large, the standard deviation of $NI = IP - D$ will be large relative to the jump size J, and $\mathrm{P}(NI > J) - \mathrm{P}(NI > 0)$ will be close to zero.

On the other hand these two service measures can be very far apart. Suppose that the order cost K, the order quantity Q, the lead time L and the order-up-to level S are all very small. Then the jump size J will be large relative to NI, and it is possible for $\mathrm{P}(NI < 0)$ to be close to zero and $\mathrm{E}[(J - NI^{+})^{+}/J]$ to be close to one. This would correspond to a "veneer inventory" policy, in only very small orders can be filled from stock. A large fraction of the total demand comes in large orders that must be backordered. The backorder costs are time-weighted and the lead time is assumed to be short relative to the average delay between consecutive large orders. Therefore the average backorder cost incurred is small, and the policy can be economical.

For the gamma-distributed Levy process, the computation of the fill rate requires a two-dimensional numerical integration, assuming that $F_D(x)$ is evaluated using efficient approximations.

Section 4: Policies: Description and Analysis

We consider four different policies for this problem, two policies that are based on the classical Hadley-Whitin inventory model, the (s, S) policy whose parameters are taken from Zheng's (R, Q) policy, and a new policy designed for Levy Jump Processes that we call the Mass Uniform Policy.

The Hadley-Whitin inventory model is currently the mainstay of introductory inventory courses. It differs from our inventory model (and from Zheng's) in that it uses a cost of $\hat{p}$ dollars per item for backorders, rather than p dollars per item per day. The change in units of measure complicates direct comparison of the models. We will compare them by following the common practice of using service level targets to determine the backorder costs.

In discussing the Hadley-Whitin model we use $\lambda = 1$ to represent the mean rate at which demand occurs. The Hadley-Whitin cost function for an (s, S) policy with $Q = S - s$ is

$$K\lambda/Q + h\,(Q/2 + s - \lambda\,L) + (\hat{p}\,\lambda\,n_D(s))/Q \tag{27}$$

(see, for example, [Nahmias 1993] page 258). If the demand process can over-shoot the reorder point, the first term is an approximation. The second and third terms are also approximations. The third term over-states the marginal benefit of an extra unit of safety stock and the second term over-states the marginal cost of an extra unit of safety stock. Although this cost expression never has a global minimum (let $Q > \hat{p}\lambda/h$ and let $s \to -\infty$), it usually has an easily-computed local minimum that corresponds to a very effective policy. However when K (and consequently Q) is sufficiently large, the error in the holding cost term dominates and (27) fails to have a local minimum.

There are two standard computational approaches to this model - the cost minimization approach and the service-constrained approach. The cost-minimization approach attempts to minimize (27) directly. The service-constrained approach searches for a policy that meets a specified service target, and that is optimal for some $\widehat{p}$. We lack a backorder cost $\widehat{p}$, but we know that for optimal policies, the fraction of time that we have inventory on hand is equal to $p/(p+h)$. According the Hadley-Whitin model this is equivalent to the fill rate, and is equal to $1 - n_D(s)/Q$. So we follow the standard service-constrained approach and obtain a policy such that

$$1 - n_D(s)/Q \;=\; p/(p+h). \tag{28}$$

Setting the derivative of (27) with respect to s equal to zero, we obtain

$$h - \frac{\widehat{p}\,\lambda}{Q}\,\overline{F}_D(s) \;=\; 0. \tag{29}$$

Similarly, the value of Q at a local minimum of (27) must be

$$Q = \sqrt{2(K + \widehat{p}\,n_D(s))\,\lambda/h}\,. \tag{30}$$

Substituting (30) into (27) we obtain an average cost of

$$\sqrt{2h(K + \widehat{p}\,n_D(s))\,\lambda} \;+ h(s - \lambda\,L\,). \tag{31}$$

The derivative of (31) with respect to s converges to $h > 0$ as $|s| \to \infty$. If the derivative ever becomes negative, the supremum of all s for which the derivative is negative is a local minimum of (27). The corresponding policy satisfies (29), and the policy can be obtained through either the cost-minimization approach or the service-constrained approach. However the derivative of (31) may never become negative, in which case the cost minimization approach will fail to produce a policy (in (29), $hQ/\widehat{p}\,\lambda > 1$). The service-constrained approach is more robust. It never fails to produce a policy which meets the target service level, and it produces policies which cannot be obtained using the cost-minimization approach.

We study two policies that are based on the Hadley-Whitin model. Our first policy, called HW-COST, is the standard service-constrained algorithm (see, for example, [Nahmias 1993] pages 263-264). We eliminate $\widehat{p}$ from (29) and (30), and the resulting equation is solved together with (28).

This algorithm fails only when $h \geq p$. The standard cost minimization algorithm, which is equivalent when it works, but which works less often, is as follows. Guess at $\widehat{p}$ and find the local minimum of (27). Then search for a $\widehat{p}$ such that the computed policy satisfies $1 - n_D(s)/Q = p/(p+h)$. It may be that $1 - n_D(s)/Q > p/(p+h)$ for all values of $\widehat{p}$ for which (27) has a local minimum. In that case we say that HW-COST has failed, meaning that to obtain the desired service measure we would need to make $\widehat{p}$ small enough that (27) would not have a local minimum. We compute and evaluate the policy whether this happens or not.

Our second policy, called HW-EOQ, is the common approach of using the EOQ model to select Q and using the service target to select s. Thus $Q = \sqrt{2K\lambda/h}$ and s is selected so that (28) holds (see, for example, [Nahmias 1993] page 262). If $s \geq 0$ then (29) gives an imputed $\widehat{p}$. If $s < 0$ we say that HW-EOQ has failed. In this case, if $Q = \sqrt{2K\lambda/h}$ then the standard cost-minimization algorithm will fail $(hQ/\widehat{p}\,\lambda > 1$ in (29)), and all values of $\hat{p}$ for which (28) has a local minimum yield policies with service levels that are higher than the target.

Our third policy, called ZHENG, is the (s, S) policy whose parameters are taken from Zheng's (R, Q) policy [Zheng 1992]. The policy is computed using the Cost Minimization Algorithm.

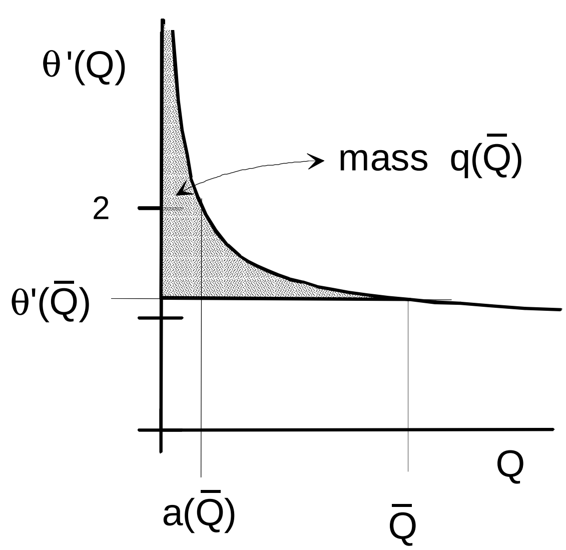
Our fourth policy is called the Mass Uniform policy, or MASS-U. For a given $\overline{Q} > 0$ we approximate $\theta(x)$ with a function $\mu(x)$ defined as follows.

$$\begin{aligned} q(\overline{Q}) &\equiv \int_0^{\overline{Q}} \left(\theta'(x) - \theta'(\overline{Q})\right)\, dx \;=\; \theta(\overline{Q}) - \overline{Q}\cdot\theta'(\overline{Q}), \\ a(\overline{Q}) &\equiv \frac{1}{q(\overline{Q})}\int_0^{\overline{Q}} x\cdot[\theta'(x) - \theta'(\overline{Q})]\, dx \;=\; [\eta(\overline{Q}) - \frac{1}{2}\,\overline{Q}^{\,2}\left(\theta'(\overline{Q}) - 1\right)]/q(\overline{Q}), \\ \mu(x) &\equiv \theta'(\overline{Q})\cdot x + q(\overline{Q})\cdot 1(x \geq a(\overline{Q})). \end{aligned} \tag{32}$$

The dependence of $\mu(x)$ on $\overline{Q}$ is suppressed. The measure $\mu(x)$ takes the area that lies under the curve $\theta'(x)$, $0 \leq x \leq \overline{Q}$ and above $\theta'(\overline{Q})$, and concentrates it into a mass of size $q(\overline{Q})$ located at $x = a(\overline{Q})$ (see Figure 3). $a(\overline{Q})$ is chosen to equalize the first moments of the measures $\theta'(x)\,dx$, $0 \leq x \leq \overline{Q}$ and $d\{\mu(x)\}$, $0 \leq x \leq \overline{Q}$. Since $[\theta'(x) - \theta'(\overline{Q})]/q(\overline{Q})$, $0 \leq x \leq \overline{Q}$ is a probability density, (32) and Properties 2 and 6 imply that

$$0 < q(\overline{Q}) \text{ and } 0 < a(\overline{Q}) \leq \overline{Q}/2 \text{ for all } \overline{Q} > 0. \tag{33}$$

Figure 3: The Mass Uniform Policy



If we use $d\{\mu(x)\}$ in place of $\theta'(x)\,dx$, and if $Q \geq a(\overline{Q})$, then (16) becomes

$$c^{\overline{Q}}(S,Q) \equiv \left[K + q(\overline{Q}) \cdot G(S - a(\overline{Q})) + \theta'(\overline{Q}) \int_0^Q G(S-x)\,dx\right] / \left[q(\overline{Q}) + Q \cdot \theta'(\overline{Q})\right]. \qquad (34)$$

Consider the optimization problem

$(\mathrm{P}^{\overline{Q}})$ min: $c^{\overline{Q}}(S,Q)$

such that: $Q \geq a(\overline{Q})$.

Let $S(\overline{Q})$ and $Q(\overline{Q})$ solve $(\mathrm{P}^{\overline{Q}})$, and suppose that $a(\overline{Q}) < Q(\overline{Q})$. Lemma 9 implies that the Cost Minimization algorithm converges quadratically to $(S(\overline{Q}), Q(\overline{Q}))$. The Cost Minimization Algorithm is easily adapted to solve $(\mathrm{P}^{\overline{Q}})$.

To compute the Mass Uniform policy we select a $\overline{Q}$ and we use the Cost Minimization Algorithm with appropriate modifications to solve $(\mathrm{P}^{\overline{Q}})$. We adjust $\overline{Q}$ to make $\overline{Q} = Q(\overline{Q})$. The approximations in Appendix 3 are helpful. Unlike the computation of optimal policies, no numerical integrations are required. The following Lemma guarantees that the approach works.

Lemma 12: Assume that Properties 2, 3, 5, 6, 7 and 8 hold and that $\overline{Q} > 0$. The optimal solution $(S(\overline{Q}),\, Q(\overline{Q}))$ of $(\mathrm{P}^{\overline{Q}})$ is the unique solution of the first-order optimality conditions for $(\mathrm{P}^{\overline{Q}})$. $S(\overline{Q})$ and $Q(\overline{Q})$ are continuous functions of $\overline{Q}$. If $\overline{Q}$ is sufficiently large then $\overline{Q} > Q(\overline{Q})$, and if $\overline{Q} > 0$ is sufficiently small then $\overline{Q} < Q(\overline{Q})$.

The relative cost of a policy is defined to be $(c' - c^*)/c^*$ where c' is the average cost of the policy policy and c^* is the average cost of an optimal policy.

Lemma 13: The relative cost of Zheng's policy is at most $\min\left\{ \frac{\theta(Q)-Q}{Q+G(y_0)/p} \,,\ \frac{h+p}{p} \ \frac{[\theta(Q)-Q]^2+\eta(Q)}{Q^2} \right\}$, where $Q = S - s$ is the optimal order quantity (not Zheng's order quantity).

For the gamma-distributed Levy process, $\theta(Q) - Q \rightarrow \mathrm{E}[V] = 1/2$ and $\eta(Q) \rightarrow 1/12$ as $Q \rightarrow \infty$, so the relative cost of Zheng's policy is at most $\min\left\{ \frac{0.5}{Q+G(y_0)/p} \,,\ \frac{h+p}{p} \ \frac{1}{3\,Q^2} \right\}$.

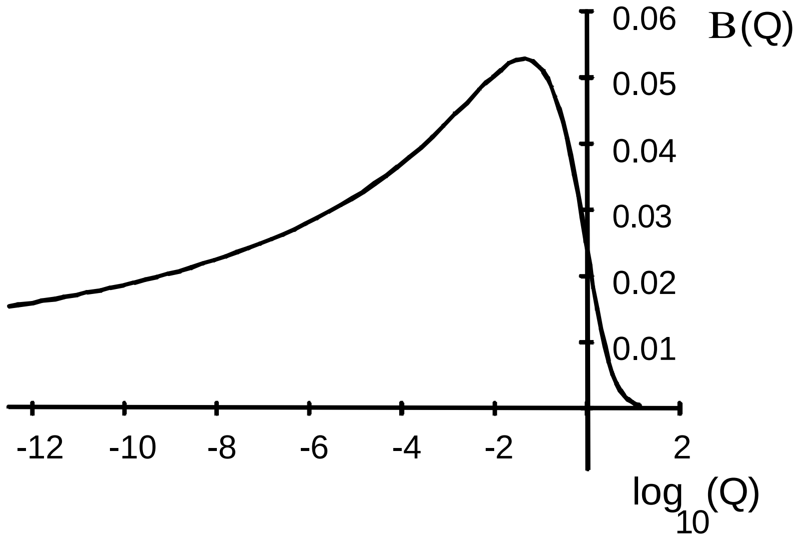
Note that $G(y_0) \rightarrow 0$ as $L \rightarrow 0$, and $G(y_0) \rightarrow \infty$ as $L \rightarrow \infty$. The bound of Lemma 13 implies that if $L > \epsilon > 0$ then Zheng's policy works very well as $Q \rightarrow 0$, as $Q \rightarrow \infty$ and as $L \rightarrow \infty$. However there is no uniform, finite upper bound on the relative cost of Zheng's policy (proof omitted), and the trends illustrated by this bound are similar to the computational results of Section 5. If the lead time is zero or close to zero, and the order cost K is small, Zheng's policy can be far from optimal. This statement probably applies to all existing continuous-time (s, S) inventory models which assume the inventory position to be uniformly distributed.

Lemma 14: The relative cost of the Mass Uniform Policy (S, Q) is at most $\frac{h+p}{p}\, B(Q)$, where

$$B(Q) \equiv \frac{1}{Q \cdot \theta(Q)} \int_0^Q x \cdot (\theta'(x) - \theta'(Q))\,dx \;=\; \frac{q(Q) \cdot a(Q)}{Q \cdot \theta(Q)}.$$

For the gamma-distributed Levy process, the function $B(Q)$ is graphed in Figure 4. Since $B(Q)$ achieves its maximum value of 0.0527 at $Q = 0.0307$, the relative cost of the Mass Uniform policy is at most 0.0791 if $p/h \geq 2$. Since $B(1) = 0.0235$, the relative cost of the Mass Uniform policy is at most 0.0353 if $p/h \geq 2$ and $Q \geq 1$.

Figure 4: The Relative Cost of the Mass Uniform Policy



Section 5: Computational Results

The main purpose of our computational study is to use the gamma-distributed Levy process as a vehicle for testing the robustness of the (s, S) inventory policies described in Section 4. The bound in Lemma 14 is strong enough that MASS-U does not need computational validation, but the performance that can be expected of the other policies is less certain. In addition we want to gain intuition into the nature of optimal policies for the gamma-distributed Levy process and to explore the importance of modelling the overshoot and the non-uniformity of the distribution of the inventory position.

As has been mentioned before, we scale our units of measure for time and inventory so that the demand which occurs in one time period has a mean and a variance of one. In this section we select our unit of measure for money so that the holding cost is $h = 1$. The fact that we re-scaled out units of measure alters the intuitive meaning of the remaining parameters. The backorder cost p is interpreted as a measure of service. Lemma 10 implies that $1/(1+p)$ is the fraction of time that we are out of stock. The lead time L is a measure of the variability of the demand D that occurs in one lead time. D has a squared coefficient of variation of $1/L$. The order cost K is a prime determinant of the minimum order quantity Q. We define the reorder interval to be Q, the time interval corresponding to the minimum order quantity. The squared coefficient of variation of the demand that occurs in a reorder interval is $1/Q$.

Application Range

As a vehicle for interpretating our results we define the "application range", a domain of the parameter space in which most real-world applications of (s, S) inventory systems lie. Since this is primarily based on personal experience it is bound to be somewhat controversial. In our experience lead times can be long or short, so all lead times are included in the application range. Most inventory systems operate with at least moderately high service levels, so for membership in the application range we require $p \geq 3$ (the system is out of stock at most 25% of the time; see Lemma 10).

We define the order costs that lie in the application range indirectly, through the order quantities Q. In our experience most large and moderately large inventory systems have a substantial number of items which experience spikey or sporadic demand. For these items the squared coefficient of variation of the demand that occurs in Q days is often very high. Maintaining inventories for these products tends to be very expensive. In a great many cases distribution systems should be re-designed to eliminate the need to inventory these parts, but this is not always possible. Our application range is intended to include some parts of this type, but not all of them.

We considered the following criteria for an order quantity that falls within the application range. First, the mean order quantity is at most twice the minimum order quantity, i.e., $\theta(Q^Z) \leq 2Q^Z$, where Q^Z comes from ZHENG. (This translates into $Q^Z \geq 0.444$. Taking Q from HW-COST or HW-EOQ gives results that are comparable, but somewhat different.) Second, the demand that occurs in Q^Z days has a squared coefficient of variation of at most 2. (This translates into $Q^Z \geq 0.5$.) Third, the demand that occurs in $\theta(Q^O)$ days has a squared coefficient of variation of at most 2, where Q^O comes from the optimal policy. (This translates into $Q^O \geq 0.134$. Taking Q from MASS-U gives nearly identical results.)

In the context of our computational studies these three criteria turned out to be nearly equivalent. We chose the first one. Thus a problem instance is said to fall within the application range if $p \geq 3$ and if $Q^Z \geq 0.444$, where Q^Z comes from ZHENG.

Results

Figure 5 illustrates how the relative order quantities change as the order cost decreases. For $K \geq 0.25$ it appears that the Q values for ZHENG approximate the optimal value of $\theta(Q)$. This trend has intuitive appeal because the average order quantity is $\theta(Q)$ for these processes, and it is Q in Zheng's paper. However the trend breaks down for smaller order quantities. Qualitatively, HW-COST and HW-EOQ behave similarly to ZHENG, MASS-U behaves like the optimal policy, and the backorder cost and the lead time have little or no impact.

Figure 6 illustrates the fact that the relative cost of ZHENG matches the qualitative behavior of the theoretical bound in Lemma 13. ZHENG, HW-COST and HW-EOQ all have have unbounded relative cost for $L = 0$ and $K \approx 0$. This is true of all policies for which, if we set $L = 0$, $Q/\sqrt{2K\lambda/h}$ fails to converge to 0 as $K \to 0$, including both HW-COST and HW-EOQ. For a fixed $L > 0$ the relative cost of these three policies (and most other (s, S) policies in the literature) converges to 1 as $K \to 0$.

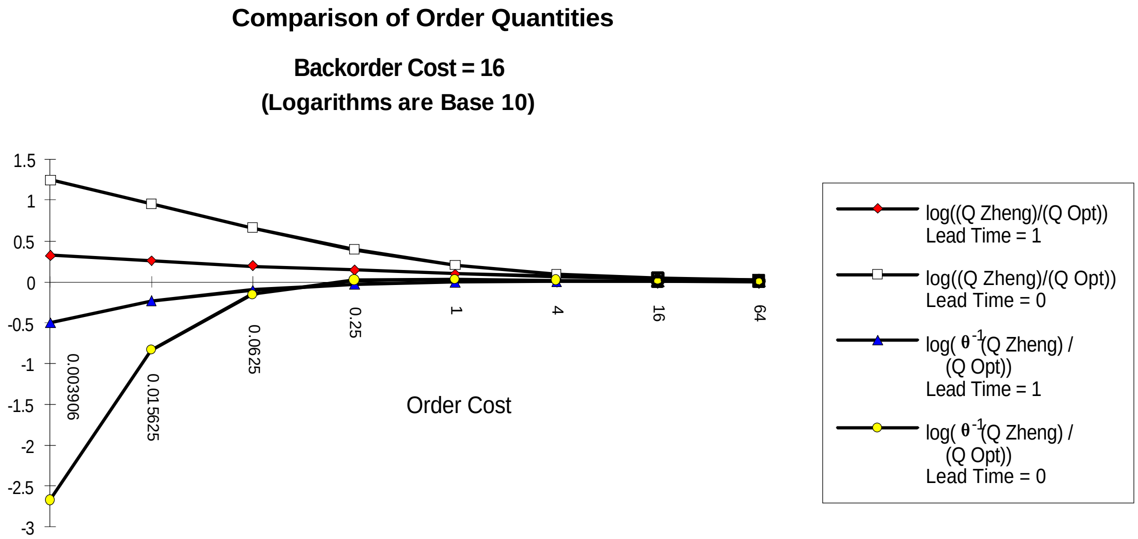
These observations can be explained as follows. Because of the non-uniformity of the distribution of the inventory position, in optimal policies the order-up-to level S is closer to y_0 than it otherwise would be. As K gets small this trend becomes more pronounced. In addition, the expected overshoot grows relative to the minimum order quantity Q, effectively reducing the average order cost incurred per day. Figure 5 confirms this by showing that for ZHENG, Q is too large when the order costs are small. If the lead time is zero or very small these errors can be very costly. On the other hand, if the lead time is positive then the costs of all of these policies converge to the Newsvendor Cost $G(y_0)$ as $K \to 0$, so the relative cost tends to 0.

Our main computational experiment used all combinations of order costs $K = 0.0625$, 0.25, 1, 4, 16, 64, 256, 1024, 4096, 16384, backorder costs of $p = 1$, 2, 4, 8, 16, 32, 64, 128, 256, 512, and lead times $L = 0$, 0.0625, 0.25, 0.5625, 1, 1.5625, 2.25, 3.0625, 4, 5.0625. Since the HW-COST algorithm does not work for $p \leq h = 1$, we substituted $p = 1.2$ for $p = 1$ for the HW-COST policy. In all 1000 parameter sets were tested. Of the 1000 parameter sets, 792 are inside of the application range. 200 parameter sets have backorder costs p that are less than 3. For 9 parameter sets ZHENG produces Q values that are less than 0.444, and both of these criteria apply to one parameter set. Note that Figure 6 includes data sets with smaller order costs than our main experiment, but only for $p = 10$.

Table 2 summarizes the results. MASS-U consistently out-performs the theoretical bound given by Lemma 14, usually by a very substantial margin. It was never more than 3.2% from optimal, and its average relative cost was negligible. On average ZHENG was only 1.3% from optimal, but it was off by as much as 52%, and within the application range it was off by as much as 20.6%. Both HW-COST and HW-EOQ perform well on average, but within the application range they both had maximum relative costs of over 27%. If we had defined the applicaiton range via $Q^Z \geq 1$ rather than $Q^Z \geq 0.444$ the maximum relative costs would have been much smaller.

Our main computational experiement contains 1,000 parameter sets, but the gaps between parameter values are still large enough to make the maximum errors reported in Table 2 unreliable. For example, in the experiment that generated the data for Figure 6 we included the parameter set $K = 0.125$, $L = 0$, $p = 10$. This parameter set falls within the application range ($Q = 0.52$), and it has relative costs of 40% for HW-Cost, 40% for HW-EOQ, and 33% for Zheng. These numbers are substantially larger than the maximum errors reported in Table 2.

Figure 5


Comparison of Order Quantities
Backorder Cost = 16
(Logarithms are Base 10)
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
0.003906
0.015625
0.0625
0.25
1
4
16
64
Order Cost
log((Q Zheng)/(Q Opt)) Lead Time = 1
log((Q Zheng)/(Q Opt)) Lead Time = 0
log(θ-1(Q Zheng) / (Q Opt)) Lead Time = 1
log(θ-1(Q Zheng) / (Q Opt)) Lead Time = 0

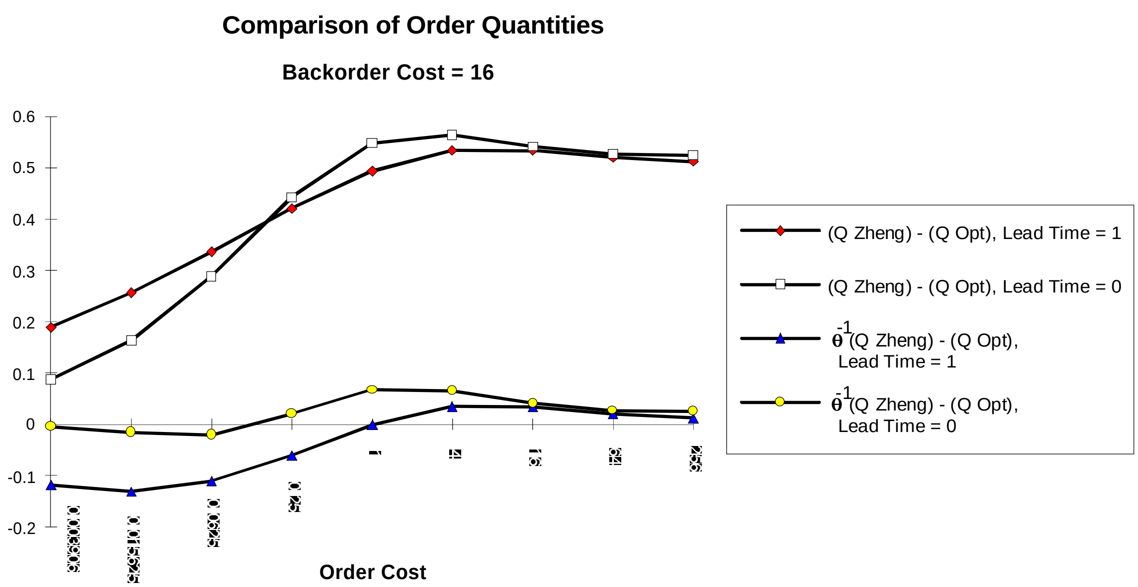
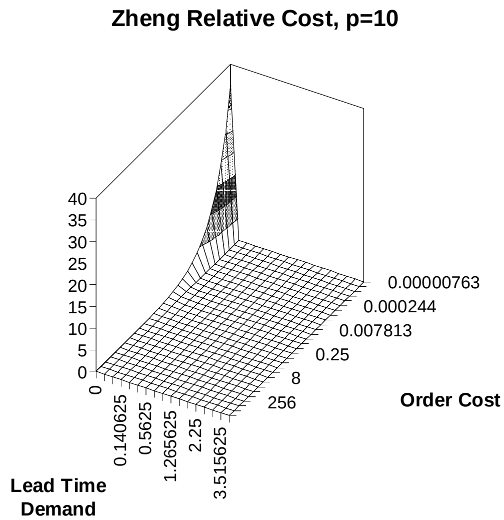
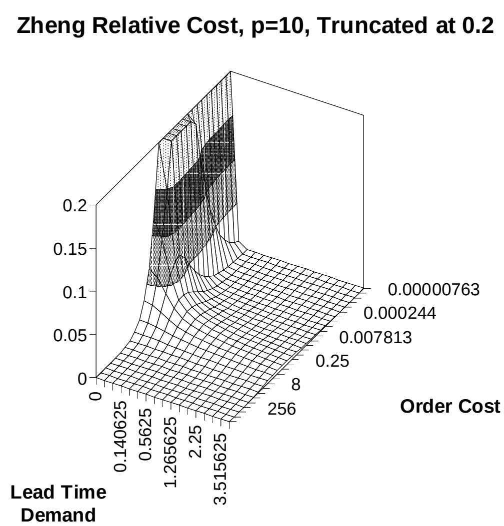

Comparison of Order Quantities
Backorder Cost = 16
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
Order Cost
(Q Zheng) - (Q Opt), Lead Time = 1
(Q Zheng) - (Q Opt), Lead Time = 0
θ-1(Q Zheng) - (Q Opt), Lead Time = 1
θ-1(Q Zheng) - (Q Opt), Lead Time = 0

Figure 6



Zheng Relative Cost, p=10
40
35
30
25
20
15
10
5
0
0.00000763
0.000244
0.007813
0.25
8
256
Order Cost
0
0.140625
0.5625
1.265625
2.25
3.515625
Lead Time Demand



Zheng Relative Cost, p=10, Truncated at 0.2
0.2
0.15
0.1
0.05
0
0.00000763
0.000244
0.007813
0.25
8
256
Order Cost
0
0.140625
0.5625
1.265625
2.25
3.515625
Lead Time Demand

Table 2: Relative Costs of the Policies

Relative Cost = (Policy Cost) / (Opt. Cost) - 1

	HW-Cost	HW-EOQ	ZHENG	MASS-U
All Test Problems				
Mean	0.077	0.051	0.013	9.2E-05
St. Dev.	0.166	0.111	0.057	0.001
Max	1.820	0.929	0.523	0.032
% Heur. Fails	0.452	0.411		
In the Application Range (79% of total)				
Mean	0.017	0.030	0.006	4.9E-06
St. Dev.	0.031	0.061	0.023	4.5E-05
Max	0.275	0.392	0.206	0.001
% Heur. Fails	0.343	0.337		
When Q $\geq$ 1 and p $\geq$ 3				
Max	0.128	0.096	0.070	0.000
When the Heuristic does Not Fail				
Mean	0.025	0.066		
St. Dev.	0.045	0.127		
Max	0.404	0.929		

Even within the application range the cost-minimization versions of HW-COST and HW-EOQ fail for over 33% of the parameter sets, because the cost function does not have a local minimum.

Tables 3, 4 and 5 illustrate the combinations of parameters that cause problems for the different policies. Because MASS-U is uniformly very effective no tables were produced for it. The cost-minimization versions of HW-COST and HW-EOQ fail often, both in and out of the application range, especially for larger order costs K, lower backorder costs p, and lower lead times. The HW-COST policy is more than 20% from optimal only when the backorder cost is outside of the application range ($p < 3$), or when both the order cost and the lead time are small. Relatively speaking, HW-EOQ has more trouble when both the order cost and the lead time are small, and is more robust with small backorder costs. ZHENG is more robust than either of the others, but both inside of the application range and outside of it, Zheng's policy has trouble when both the order cost and the lead time are small.

The errors that occur when both the order cost and the lead time are small were explained when we discussed Figure 6. Both the order quantity and the relative cost of the HW-COST policy grow without bound as p approaches $h = 1$. With $p = 1.2$, HW-COST clearly had problems.

Table 3

The Largest Lead Time for which HW-Cost Fails

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
	1.2	0.25	1	All	All	All	All	All	All	All	All
Back-	2	0.06	0.25	1	1.6	All	All	All	All	All	All
order	4	0.06	0.06	0.25	0.56	1	3.1	All	All	All	All
Costs	8	0	0.06	0.06	0.25	0.56	1	3.1	All	All	All
	16	0	0	0.06	0.06	0.25	0.56	1	3.1	All	All
	32	0	0	0	0.06	0.06	0.25	0.56	1	2.3	All
	64	0	0	0	0	0.06	0.06	0.25	0.56	1	2.3
	128	0	0	0	0	0	0.06	0.06	0.25	0.56	1
	256	0	0	0	0	0	0	0.06	0.06	0.25	0.56
	512	0	0	0	0	0	0	0	0.06	0.06	0.25

Note: When the lead time L is equal to zero, the standard cost-minimization approach leads to a reorder point of 0 and a fill rate of 100%.

Note: In each each row, in the limit as $K \to \infty$, we eventually get ALL.

Lead Times for which HW-Cost has Relative Cost > 20%

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
	1.2	All	All	All	All	All	All	All	All	All	All
Back-	2	*0.25	*0.06	None	None	None	None	None	None	None	None
order	4	0	0	None	None	None	None	None	None	None	None
Costs	8	0	0	None	None	None	None	None	None	None	None
	16	0	0	None	None	None	None	None	None	None	None
	32	0	0	None	None	None	None	None	None	None	None
	64	0	0	None	None	None	None	None	None	None	None
	128	0	0	None	None	None	None	None	None	None	None
	256	0	0	None	None	None	None	None	None	None	None
	512	0	0	None	None	None	None	None	None	None	None
Order	Mean	2.2	2.4	2.9	4.2	6.9	12.6	24.0	46.6	92.0	183
Quan-	Min	0.4	0.7	1.4	2.8	5.7	11.3	22.6	45.3	90.5	181
tities	Max	10.1	10.2	10.3	10.9	12.7	17.5	28.2	50.6	95.7	186

Table 4

The Largest Lead Time for which HW-EOQ Fails

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
Back-order Costs	1	0.06	0.25	0.56	1	2.3	All	All	All	All	All
	2	0.06	0.06	0.25	0.56	1.6	3.1	All	All	All	All
	4	0.06	0.06	0.25	0.56	1	2.3	4	All	All	All
	8	0	0.06	0.06	0.25	0.56	1	2.3	4	All	All
	16	0	0	0.06	0.06	0.25	0.56	1	2.3	All	All
	32	0	0	0	0.06	0.06	0.25	0.56	1	2.3	All
	64	0	0	0	0	0.06	0.06	0.25	0.56	1	2.3
	128	0	0	0	0	0	0.06	0.06	0.25	0.56	1
	256	0	0	0	0	0	0	0.06	0.06	0.25	0.56
	512	0	0	0	0	0	0	0	0.06	0.06	0.25

Note: When the lead time L is equal to zero, the standard cost-minimization approach leads to a reorder point of 0 and a fill rate of 100%.

Lead Times for which HW-EOQ has Relative Cost > 20%

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
Back-order Costs	1	All	$\geq$ 1	$\geq$ 4	None	None	None	None	None	None	None
	2	$\neq$ 0.06	$\geq$ 1.56	None	None	None	None	None	None	None	None
	4	$\neq$ 0.06	$\geq$ 5.06	None	None	None	None	None	None	None	None
	8	$\neq$ 0.06	None	None	None	None	None	None	None	None	None
	16	$\geq$ 0.25, $\leq$ 2.25	None	None	None	None	None	None	None	None	None
	32	0	None	None	None	None	None	None	None	None	None
	64	0	0	None	None	None	None	None	None	None	None
	128	0	0	None	None	None	None	None	None	None	None
	256	0	0	None	None	None	None	None	None	None	None
	512	0	0	None	None	None	None	None	None	None	None
Order Quan-tities	Mean	0.4	0.7	1.4	2.8	5.7	11.3	22.6	45.3	90.5	181
	Min	0.4	0.7	1.4	2.8	5.7	11.3	22.6	45.3	90.5	181
	Max	0.4	0.7	1.4	2.8	5.7	11.3	22.6	45.3	90.5	181

Table 5

Lead Times for which Zheng has Relative Cost > 20%

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
	1	≤ 0	≤ 0	None	None	None	None	None	None	None	None
Back-	2	≤ 0	≤ 0	None	None	None	None	None	None	None	None
order	4	≤ 0.06	None	None	None	None	None	None	None	None	None
Costs	8	≤ 0.06	None	None	None	None	None	None	None	None	None
	16	≤ 0.06	None	None	None	None	None	None	None	None	None
	32	≤ 0.06	None	None	None	None	None	None	None	None	None
	64	≤ 0.06	None	None	None	None	None	None	None	None	None
	128	≤ 0.06	None	None	None	None	None	None	None	None	None
	256	≤ 0.06	None	None	None	None	None	None	None	None	None
	512	≤ 0.06	None	None	None	None	None	None	None	None	None
Order	Mean	0.9	1.4	2.3	3.9	7.0	13.0	25.2	49.7	98.8	197
Quan-	Min	0.4	0.7	1.4	2.8	5.7	11.3	22.7	45.3	90.7	181
tities	Max	1.2	1.9	3.2	5.3	9.0	16.6	32.3	64.2	128	256

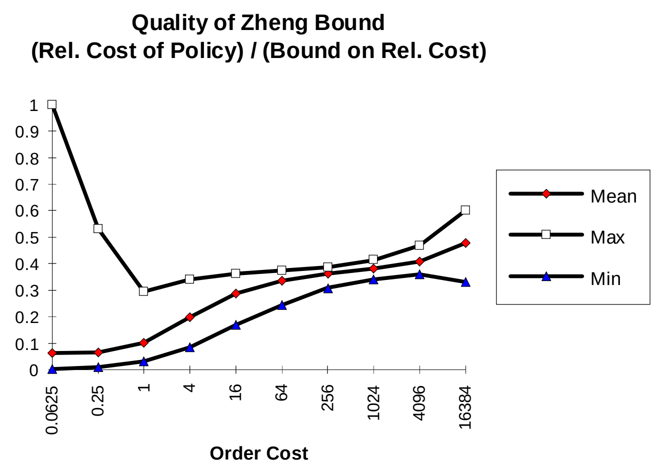
Range of Order Quantities for the Optimal Policy

Lead Times Tested: 0, 0.06, 0.25, 0.56, 1, 1.56, 2.25, 3.06, 4, 5.06

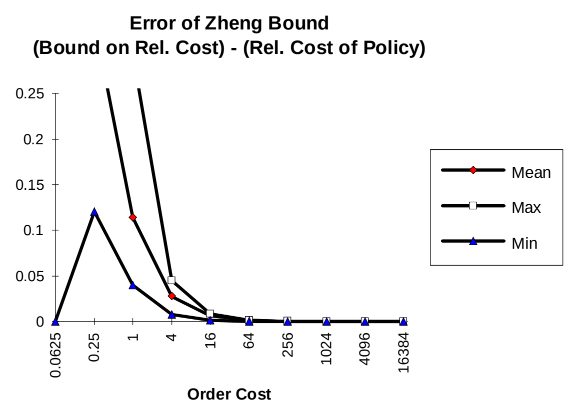
		Order Costs									
		0.06	0.25	1	4	16	64	256	1024	4096	16384
Order	Mean	0.5	1.0	1.8	3.4	6.4	12.5	24.7	49.2	98.3	197
Quan-	Min	0.1	0.3	0.9	2.3	5.1	10.8	22.1	44.8	90.1	181
tities	Max	0.9	1.6	2.8	4.8	8.5	16.1	31.8	63.7	128	256

As Figure 7 illustrates, the ratio of the relative cost of the ZHENG policy to the bound in Lemma 13 is usually much less than one. But in our tests, when the order cost is 16 or more the arithmetic difference between the bound and the relative cost of the policy is at most 1%. Lemma 14 gives a bound on the relative cost of the MASS-U policy. The ratio of this bound to the relative cost of the MASS-U policy is less than 6 in only 2 of the 1000 parameter sets in the test. Usually it is much higher. However the arithmetic difference between the bound and the relative cost of the policy is never more than 5.2%, and is usually much smaller (see Figure 7).

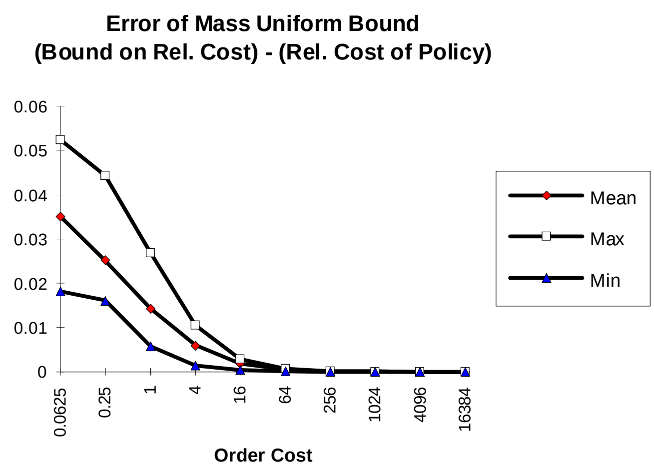
Figure 7



Quality of Zheng Bound
(Rel. Cost of Policy) / (Bound on Rel. Cost)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.0625
0.25
1
4
16
64
256
1024
4096
16384
Order Cost
Mean
Max
Min



Error of Zheng Bound
(Bound on Rel. Cost) - (Rel. Cost of Policy)
0.25
0.2
0.15
0.1
0.05
0
0.0625
0.25
1
4
16
64
256
1024
4096
16384
Order Cost
Mean
Max
Min



Error of Mass Uniform Bound
(Bound on Rel. Cost) - (Rel. Cost of Policy)
0.06
0.05
0.04
0.03
0.02
0.01
0
0.0625
0.25
1
4
16
64
256
1024
4096
16384
Order Cost
Mean
Max
Min

Table 6

Computation Times: Avg. Seconds per Problem Instance				
HW-COST	HW-EOQ	ZHENG	MASS-U	Optimal
.027	.007	.076	.069	36.5
Computations done on a 486-based, 50 megahertz PC				

Table 6 gives the CPU time per problem instance, in seconds. All policies requiring an initial guess were started from the HW-EOQ policy. For Zheng's policy we used the Cost Minimization Algorithm, which is faster than the algorithm that Zheng proposed. All of the policies have very low computation times except for the optimal policy. The optimal policy requires three numerical integrations for each iteration of the Cost Minimization Algorithm.

Recommendations

Heuristic policies are usually measured on the quality of the policies that they produce and on the computational effort that they require. All of the four heuristic policies studied can be computed very efficiently.

The standard Hadley-Whitin cost function often fails to have a local minimum. However the standard algorithms for computing polices with a given fill rate, Zheng's policy, and the Mass Uniform policy are all much more robust. When the service level is reasonably high ($p/(p+h) \geq 2/3$), and $Q \geq 1$, all of these policies perform very well.

The main negative result of these tests is that when $Q < 1$, as it often is with low-demand parts that experience sproadic or spikey demand, only the MASS-U policy can be relied on. (Recall that because of the way we scaled time and inventory, Q should be interpreted as the coefficient of variation of the demand that occurs in one reorder interval.) Other policies that assume the inventory position to be uniformly distributed would almost certainly experience similar problems. If inventory levels are discrete the algorithm of Federgruen and Zheng [1994] should be used. To our knowledge this is the only paper that efficiently computes good (s, S) inventory policies for systems with continuous inventory levels, that does not assume the inventory position to be uniform, and that allows the demand process to over-shoot the reorder point.

Section 6: Conclusions

Levy demand processes are a useful and interesting set of demand processes for inventory models. Numerical approximations for the distribution of D, of θ, and/or both will be required. The gamma-distributed Levy process is particularly attractive, and we have provided the appropriate approximations.

For Levy demand processes the distribution of the inventory position does not need to be uniform, and the demand process is allowed to over-shoot the reorder point. Most easily-computed (s, S) inventory policies require the inventory position to be uniform and assume that there is no overshoot. Our tests indicate that when the coefficient of variation of the demand that occurs in the reorder interval is greater than one, it is important to model the inventory position as non-uniform and to model the overshoot when it occurs. This is often the case for low-demand parts that experience sporadic or spiky demand.

As long as the coefficient of variation of the demand that occurs in one reorder interval is at least one, and the service level is reasonably high, the standard service-constrained Hadley-Whitin (s, S) inventory polices and Zheng's policy work very well. However even in this region it is often the case that the standard Hadley-Whitin cost function fails to have a local minimum.

The Mass Uniform heuristic applies to all Levy demand processes. For the gamma-distributed Levy process it is guaranteed to be within 8% of optimal whenever backorders are at least as expensive as inventory.

For any Levy demand process, the Cost Minimization Algorithm applies to Zheng's (R, Q) inventory model, to the Mass Uniform heuristic and to the computation of optimal policies. The algorithm is quadratically convergent.

Appendix 1: Notation

$a(Q)$ — The location of the point mass associated with the Mass Uniform policy (see (32)).
$B(Q)$ — $\frac{q(Q)\cdot a(Q)}{Q\cdot\theta(Q)}$ (see Lemma 14).
$c(S,Q)$ — The average cost of the (s,S) policy with $S-s=Q$. (see (16)).
$C(s,S)$ — $c(S,S-s)$, the average cost of an (s,S) policy.
$D(t,u)$ — The demand that occurs in the time interval $(t,u]$.
$D(t)$ — The demand that occurs in the time interval $[0,t]$.
D — $D(L)$, the demand that occurs in one lead time.
$E(S,Q)$ — $\theta(Q)\cdot G(S-Q)-K-H(S,Q)$ (see (19)).
$\eta(Q)$ — $\int_0^x z\cdot[\theta'(z)-1]\,dz$ (see (12)).
$F_D(x)$ — $\mathrm{P}(D\le x)$, the cumulative distribution function of the random variable D.
$\bar{F}_D(x)$ — $\mathrm{P}(D>x)$, the complementary cumulative distribution function of the random variable D.
$G(x)$ — $\mathrm{E}[h(x-D)^+ + p(D-x)^+]$, the expected rate at which holding and backorder costs are incurred (see (13)).
$G(y_0)$ — The Newsvendor Cost or the Buffer cost. y_0 minimizes $G(x)$. See (22).
h — The holding cost, in dollars per item per day.
$H(S,Q)$ — $\int_0^Q G(S-x)\,\theta'(x)\,dx$, the expected holding cost incurred per cycle (see (15)).
$IP(t)$ — $NI(t+L)+D(t,t+L)$, the inventory position at time t.
J — The Demand-Weighted Distribution of the Jump Size (see (3) and Lemma 1).
K — The fixed order cost, in dollars.
L — The lead time.
$\mathcal{L}_D(\gamma)$ — $\mathrm{E}[e^{-\gamma D}]$, the Laplace Transform of the random variable D.
$\mathcal{L}_\theta(\gamma)$ — $\int_{x=0}^{\infty} e^{-\gamma\cdot x}\,d\{\theta(x)\}$, the Laplace Transform of the function θ.
λ — The demand rate (assumed to be equal to 1).
$n_D(x)$ — $\mathrm{E}[(D-x)^+]$, the partial expectation of the random variable D at x.
$NI(t)$ — The net inventory at time t.
p — The backorder cost, in dollars per item per day.
Q — The minimum order quantity. $Q=S-s$.
$q(Q)$ — The point mass associated with the Mass Uniform policy (see (32)).
s — The reorder point.
S — The order-up-to level.
$\theta(Q)$ — $\mathrm{E}\,[D^{-1}(Q)]$, the expected time to accumulate Q units of demand.
$\theta'(Q)$ — The derivative of $\theta(Q)$ (see (11)).
t — Used to index time.
V — The Asymtotic distribution of the overshoot (see (3) and Lemma 3).
y_0 — The value that minimizes $G(x)$. See (22).

Appendix 2: Proofs

The following lemma is used in the proof of Lemmas 1 and 11.

Lemma A.1: Let $h(x,y)$ be a non-negative, real-valued, bounded measurable function. Define

$$I(T) = \int_0^T h(D(t) - D(t^-),\ X(t))\, d\{D(t)\}$$

where $(X(t),\ t \geq 0)$ is a measurable process such that for every $t > 0,\ X(t)$ is adapted to the sigma-field $\sigma(D(u),\ u < t)$. Then

$$\mathrm{E}(I(T)) \ = \ \int_{t=0}^T \int_{x=0}^\infty \mathrm{E}[h(x, X(t))]\ d\{-\psi^+(x)\}\ dt = \int_{t=0}^T \mathrm{E}\left[h(J, X(t))\right]\ dt\,, \tag{35}$$

with J independent of $(X(t)\,,\, t \geq 0)$.

Proof: Suppose first that

$$h(x,y) = h_1(x) \cdot h_2(y). \tag{36}$$

with h_1, h_2 non-negative, real-valued, bounded and measurable with $h_1(x) \leq \overline{h}_1$ for all x. We have

$$I(T) = \int_0^T h_1(D(t) - D(t^-))\ \ h_2(X(t))\, d\{D(t)\}.$$

Note that for all $t,\ D(t) - D(t^-)$ and $X(t)$ are independent. Define $\widetilde{D}\,(t)$ by $\widetilde{D}\,(0) = 0$ and

$$d\left\{\ \widetilde{D}\,(t)\right\} = h_1(D(t) - D(t^-))\ \ d\left\{D(t)\right\}.$$

Then $\widetilde{D}\,(t)$ is constant whenever $D(t)$ is constant, $\widetilde{D}\,(t) \leq \overline{h}_1 \cdot D(t)$ for all $t \geq 0$, and $\widetilde{D}\,(t)$ is a Levy demand process. Following (1) let $\psi^+_{\widetilde{D}}(x)$ be the rate of arrivals of jumps of size greater than x, for the process $\widetilde{D}\,(t)$. If $D(t)$ jumps by u at time t then $\widetilde{D}\,(t)$ jumps by $x = h_1(u) \cdot u$ at time t. By (6),

$$\mathcal{L}_{\widetilde{D}(t)}(\gamma) \ = \ e^{-t\int_0^\infty\ (1 - e^{-\gamma x})\, d\{-\psi^+_{\widetilde{D}}(x)\}} \ = \ e^{-t\int_0^\infty\ (1 - e^{-\gamma \cdot h_1(u)\cdot u}\,)\, d\{-\psi^+(u)\}}.$$

Consequently the mean of $\widetilde{D}\,(t)$ is

$$\mathrm{E}[\widetilde{D}\,(t)] = t\int_0^\infty h_1(u) \cdot u\ \ d\{-\psi^+(u)\} \equiv m \cdot t\,.$$

Note that $\hat{D}(t) \equiv \widetilde{D}\,(t) - m \cdot t$ is an $\mathfrak{F}_t$-martingale where $\mathfrak{F}_t$ is the sigma-field $\sigma(D(u), u \leq t)$. Therefore

$$\mathrm{E}\left[\int_0^T h_2(X(t))\ d\left\{\hat{D}(t)\right\}\right] = 0,$$

so

$$\begin{aligned}\mathrm{E}[I(T)] = \mathrm{E}\left[\int_0^T h_2(X(t))\ d\{\ \widehat{D}(t) + mt\}\right] &= 0 + \int_0^T \mathrm{E}[h_2(X(t))]\, m\ dt \\ = \int_{t=0}^T \int_{u=0}^\infty \mathrm{E}[h(u, X(t))] u\ d\{-\,\psi^+(u)\}\ dt,&\end{aligned}$$

proving (35) for the case (36).

Let $\mathfrak{H}$ be the set of bounded measurable functions h for which (35) holds. $\mathfrak{H}$ is a monotone vector space [Sharpe 1988, p. 364]. We have shown that $\mathfrak{H}$ contains all functions of type (36). By the Montone Convergence Theorem [Sharpe 1988, Theorem A0.6], $\mathfrak{H}$ contains all bounded measurable functions. □

Lemma 1: $\mathrm{E}\left(\frac{1}{T}\int_0^T 1(D(t)-D(t^-)>z)\,d\{D(t)\}\right) = \overline{F}_J(z) = \int_z^\infty x\,d\{-\psi^+(x)\}.$

Proof: If we set $h(x,y) = 1(x>z)$ in Lemma A.1 we have

$$\mathrm{E}\left(\frac{1}{T}\int_0^T 1(D(t)-D(t^-)>z)\,d\{D(t)\}\right) = \frac{1}{T}\int_{t=0}^T\int_{x=0}^\infty \mathrm{E}\,[1(J>z)]\,x\,d\{-\psi^+(x)\}\,dt$$

$$= \frac{1}{T}\int_{t=0}^T \overline{F}_J(z)\,dt \;=\; \overline{F}_J(z). \quad \square$$

Lemma 2: $\theta(x)$ satisfies $\theta(x) = \mathrm{E}[D(D^{-1}(x))] \geq x$ for all $x>0$. If Property 1 holds then $\theta(x)-x \to \mathrm{E}[V]$ as $x\to\infty$.

Proof: We define a renewal reward process which renews itself at times Y_n where $Y_0 = 0$ and $Y_{n+1} = D^{-1}(D(Y(n))+x)$. Then $Y_{n+1}-Y_n \sim D^{-1}(x)$, and has mean $\theta(x)$. At renewal epoch n we earn a reward $D(Y_n)-D(Y_{n-1}) \sim D(D^{-1}(x))$. Let $n(t) = \max\{n : Y_n \leq t\}$. By the Renewal Reward Theorem,

$$\mathrm{E}[D(D^{-1}(x))]\,/\,\theta(x) = \lim_{t\to\infty}\left[D(Y_{n(t)})/t\right] = 1.$$

Since $D(D^{-1}(x)) \geq x$, our first assertion follows.

By Lemma 3, $D(D^{-1}(x)) \to V$ as $x\to\infty$. Fatou's Lemma implies that

$$\liminf_{x\to\infty} \mathrm{E}\left[D(D^{-1}(x))\right] \geq \mathrm{E}[V]. \quad (37)$$

If $\mathrm{E}[V] = \infty$ we are done. Therefore we assume that

$$\mathrm{E}[V] = \int_0^\infty x\,d\{-\psi^+(x)\} < \infty. \quad (38)$$

Let $\delta > 0$,

$$D_n^\delta \equiv D(n\delta), \quad n = 0,1,\dots, \text{ and}$$

$$N(x,\delta) \equiv \min\{n \geq 0 : D_n^\delta > x\}.$$

Note that by (6) and (38),

$$\mathrm{E}\left[\left(D_1^\delta\right)^2\right] = \delta^2\int_0^\infty x\,d\{-\psi^+(x)\} + \delta\int_0^\infty x^2 d\{-\psi^+(x)\} = \delta^2 + 2\delta\int_0^\infty x\,\psi^+(x)\,dx < \infty. \quad (39)$$

Let

$$H(x) \equiv \mathrm{E}\left(D_{N(x,\delta)}^\delta - x\right) \text{ and}$$

$$h(x) \equiv \mathrm{E}\left[\left(D_1^\delta - x\right)1\left(D_1^\delta > x\right)\right].$$

Then

$$H(x) = h(x) + \int_{y=0}^x H(y-x)\,d\{F_{D_1^\delta}(y)\}.$$

We claim that $h(x)$ is directly Riemann integrable. Observe that

$$h(x) \le \mathrm{E}\big[D_1^\delta \, 1\big(D_1^\delta > x\big)\big] \equiv g(x).$$

Then $g(x)$ is non-increasing, and

$$\int_0^\infty g(x)\,dx = \mathrm{E}\Big[\big(D_1^\delta\big)^2\Big] < \infty$$

by (39). By Remark 3.10.5, p. 232 of [Resnick 1992], $h(x)$ is directly Riemann integrable. By the Key Renewal Theorem,

$$\begin{aligned}
\limsup_{x\to\infty} \mathrm{E}\big[D(D^{-1}(x))\big] &\le \lim_{x\to\infty} H(x) \\
&= \frac{1}{\mathrm{E}(D(\delta))}\int_0^\infty h(x)\,dx \\
&= \frac{1}{\delta}\int_0^\infty \mathrm{E}\big[\big(D_1^\delta - x\big)1\big(D_1^\delta > x\big)\big]\,dx \\
&= \frac{1}{\delta}\,\mathrm{E}\int_0^{D_1^\delta}\big(D_1^\delta - x\big)\,dx = \frac{1}{2\delta}\,\mathrm{E}\Big(\big(D_1^\delta\big)^2\Big) \\
&= \frac{1}{2\delta}\left[\delta^2 + 2\delta\int_0^\infty x\,\psi^+(x)\,dx\right] = \frac{\delta}{2} + \mathrm{E}[V].
\end{aligned}$$

The last line follows from (39). By (37) and because $\delta > 0$ is arbitrary, the result follows. □

Lemma 3: If Property 1 holds then $\mathrm{P}(D(D^{-1}(x)) - x > u) \to \int_u^\infty \psi^+(z)\,dz = \overline{F}_V(u)$ as $x \to \infty$.

Proof: If $\psi^+(u) = 0$ then $\mathrm{P}[D(D^{-1}(x)) - x > u] = 0 = \overline{F}_V(u)$. Given a $u > 0$ such that $\psi^+(u) > 0$, we view $D(t)$ as a renewal process which renews itself whenever a jump of size greater than u occurs. Suppose that successive renewals occur at times t_1 and t_2, let S represent the total demand that occurs in the interval $(t_1,\, t_2)$, and let U represent the size of the jump that occurs at time t_2. We define $W(x) \equiv \mathrm{P}[D(D^{-1}(x)) - x > u]$ and $w(x) \equiv \mathrm{P}[S < x,\; S + U > x + u]$. This leads to the renewal equation

$$W(x) = w(x) + \int_0^x W(x - z)\; d\{F_{S+U}(z)\}.$$

The Key Renewal Theorem implies that

$$W(x) \to \frac{1}{\mathrm{E}[S+U]}\int_0^\infty w(x)\,dx \quad \text{as } x \to \infty.$$

Since $U > u$,

$$\begin{aligned}
\int_0^\infty w(x)\,dx &= \int_0^\infty \mathrm{P}[S < x,\; S + U > x + u]\,dx \\
&= \mathrm{E}_S\left[\int_S^\infty \mathrm{P}[S + U > x + u]\,dx\right] \\
&= \mathrm{E}_S\left[\int_0^\infty \mathrm{P}[U - u > y]\,dy\right] = \mathrm{E}[U] - u.
\end{aligned}$$

We define the Levy demand process $D^*(t)$ to be the process which has a demand of size $U - u$ whenever $D(t)$ has a demand of size $U > u$. Following (2),

$$\frac{1}{t}\,\mathrm{E}[D^*(t)] = \mathrm{E}[D^*(1)] = \int_0^\infty \psi^+(x + u)\,dx\,. \tag{40}$$

We view $D^*(t)$ as a renewal process that has the same renewals as $D(t)$. The expected amount by which $D(t)$ increases between renewals is $\mathrm{E}[S+U]$, and the expected amount by which $D^*(t)$ increases between renewals is $\mathrm{E}[U]-u$. By (40) and the Renewal Theorem,

$$\int_u^\infty \psi^+(x)\,dx \;=\; \int_0^\infty \psi^+(x+u)\,dx \;=\; \frac{\mathrm{E}[D^*(t)]}{\mathrm{E}[D(t)]} \rightarrow \frac{\mathrm{E}[U]-u}{\mathrm{E}[S+U]} = \lim_{x\to\infty} W(x). \qquad \square$$

Lemma 4: For the gamma-distributed Levy demand process $D^*(t)$,

$$\psi^+(x) \;=\; \int_x^\infty \frac{e^{-y}}{y}\,dy \quad \text{and} \tag{10}$$

$$\theta'(x) \;=\; \int_{t=0}^\infty \frac{x^{t-1}}{\Gamma(t)} e^{-x}\,dt\,. \tag{11}$$

where $\theta'(x)$ is the derivative of $\theta(x)$, and $\theta(x)=0$.

Proof: For the gamma-distributed Levy demand process $D^*(t)$, (6) and (8) are equivalent to

$$\ln(1+\gamma) \;=\; \int_0^\infty \gamma\, e^{-\gamma x}\psi^+(x)\,dx \quad \text{and} \tag{41}$$

$$\mathcal{L}_\theta(\gamma) \;=\; \frac{1}{\ln(1+\gamma)}. \tag{42}$$

By definition, $\theta(x)=0$. By the theory of Laplace transforms it suffices to prove that $\psi^+(x)$ as defined in (10) satisfies (41), and that $\theta(x)$ as defined by (11) satisfies (42). For (10) we must show that $\ln(1+\gamma)$ is equal to

$$\int_0^\infty \gamma\, e^{-\gamma x}\int_x^\infty \frac{e^{-y}}{y}\,dy\,dx \;=\; \int_0^\infty (1-\,e^{-\gamma y})\,\frac{e^{-y}}{y}\,dy. \tag{43}$$

When $\gamma=0$, (43) and $\ln(1+\gamma)$ are both equal to zero. The derivative of the right-hand side of (43) with respect to γ is $1/(1+\gamma)$, so (10) holds.

From (11) we have

$$\begin{aligned}&\int_{x=0}^\infty e^{-\gamma x}\int_{t=0}^\infty \frac{x^{t-1}}{\Gamma(t)}e^{-x}\,dt\,dx\\ &= \int_{t=0}^\infty \frac{1}{(1+\gamma)^t}\int_{x=0}^\infty \frac{x^{t-1}(1+\gamma)^t}{\Gamma(t)}e^{-(1+\gamma)x}\,dx\,dt\\ &= \int_{t=0}^\infty \frac{1}{(1+\gamma)^t}\,dt \;=\; \frac{1}{\ln(1+\gamma)}\,,\end{aligned}$$

so (42) holds when $\theta(x)$ is given by (11). $\square$

Lemma 5: Properties 2 through 6 hold for $D^*(t)$.

Proof: Clearly Property 1, the first two claims of Property 2, and the first claim of Property 3 hold. Differentiating (11) we see that $\theta''(x)<0$, so Properties 2, 4 and 5 hold. From (7) we obtain $\theta'(x)\rightarrow 1$.

For $0<t\le 1$ there is a positive constant a such that $\Gamma(t)\ge a/t$ ($a=1+e^{-1}$ works). By (11),

$$
\begin{aligned}
\theta'(x) &= \int_{t=0}^{\infty} \frac{x^{t-1}}{\Gamma(t)} e^{-x}\, dt \geq e^{-x} \int_{t=0}^{1} \frac{t \cdot x^{t-1}}{a}\, dt \\
&= \frac{e^{-x}}{a} \left(\frac{t}{ln(x)} - \frac{1}{(ln(x))^2} \right) x^{t-1} \bigg|_{t=0}^{1} \\
&= \frac{e^{-x}}{a} \left(\frac{1}{ln(x)} - \frac{1}{(ln(x))^2} + \frac{1}{x\,(ln(x))^2} \right) \to \infty \;\text{ as }\; x \to \infty\,.
\end{aligned}
$$

□

Lemma 6: For the gamma-distributed Levy process, $\eta(x) \to 1/12$ as $x \to \infty$.

Proof:

$$
\begin{aligned}
\theta''(x) &= \frac{d}{dx} e^{-x} \int_{t=0}^{\infty} \frac{x^{t-1}}{\Gamma(t)}\, dt \\
&= -\, e^{-x} \int_{t=0}^{\infty} \frac{x^{t-1}}{\Gamma(t)}\, dt \,+\, e^{-x} \int_{t=0}^{\infty} \frac{(t-1)\, x^{t-2}}{\Gamma(t)}\, dt \\
&= -\, e^{-x} \int_{t=0}^{1} \frac{(1-t)\, x^{t-2}}{\Gamma(t)}\, dt\,.
\end{aligned}
$$

$$
\begin{aligned}
\eta(x) &\to \int_{0}^{\infty} x\,(\theta'(x) - 1)\, dx \\
&= \frac{x^2}{2}\,(\theta'(x) - 1)\bigg|_{0}^{\infty} - \int_{0}^{\infty} \frac{x^2}{2}\, \theta''(x)\, dx \\
&= \int_{x=0}^{\infty} \frac{x^2}{2}\, e^{-x} \int_{t=0}^{1} \frac{(1-t)\, x^{t-2}}{\Gamma(t)}\; dt\; dx \\
&= \int_{t=0}^{1} \frac{t\,(1-t)}{2} \int_{x=0}^{\infty} \frac{x^{t}}{\Gamma(t+1)}\, e^{-x}\, dx\; dt \\
&= \int_{t=0}^{1} \frac{t\,(1-t)}{2}\; dt \,=\, \frac{1}{12}.
\end{aligned}
$$

□

Lemma 7: Suppose that S is chosen optimally for a given $Q > 0$. Properties 2 and 3 imply that $G(S-Q) \geq G(S) \geq G(S-Q) - \, h \cdot [\theta(Q) - Q\,]$. If Properties 4 and 7 also hold then $G(S-Q) > G(S)$.

Proof: By (14), $0 = H'(S, Q\,) \,=\, \int_0^Q G'(S - x\,)\, \theta'(x)\, dx = G(S) - G(S-Q\,) + \int_0^Q G'(S - x\,)\, [\theta'(x) - 1]\, dx$. If $0 \leq G(S) - G(S-Q\,) = \int_0^Q G'(S-x\,) dx$ then Properties 2 and 3 and the convexity of G imply that $0 \leq \int_0^Q G'(S-x\,)\, [\theta'(x) - 1]\, dx$. Either both of these inequalities are tight or we have a contradiction. If Properties 4 and 7 also hold then $0 < \int_0^Q G'(S - x\,)\, [\theta'(x) - 1]\, dx\,$.

By (14), $0 = G(S) - G(S-Q\,) + \int_0^Q G'(S-x\,)\, [\theta'(x) - 1]\, dx \,\leq G(S) - G(S-Q\,) + h \cdot \int_0^Q [\theta'(x) - 1]\, dx \,= G(S) - G(S-Q\,) + h \cdot [\theta(Q) - Q\,]$. □

Lemma 8: Suppose that S is chosen optimally for a given $Q > 0$. If Properties 7 and 8 hold then

$$
S \,>\, y_0 \,>\, S - Q. \tag{23}
$$

If Property 2 also holds, the solution to the first-order optimality conditions (20)-(21) is unique.

Proof: First we show that $c(S, Q)$ achieves its minimum over the set $(S, Q) \in \Re \times (\Re^{+} \backslash \{0\})$. This follows from the fact that $c(S, Q) \to \infty$ as $\max\{Q,\, 1/Q,\, S\} \to \infty$, and the continuity of $c(S, Q)$.

Let $S(Q)$ be the value of S that satisfies (20) for a given $Q > 0$, i.e., $H'(S(Q), Q) = 0$. By Properties 7 and 8, $S(Q)$ is unique, and (23) holds for $S = S(Q)$. The other first-order optimality condition (21) can be written as $0 = E(S(Q), Q))$. To show that $c(S, Q)$ has a unique minimum it suffices to show that the zero of $E(S(Q), Q)$ is unique. Differentiating (20) with respect to Q and rearranging we obtain

$$\frac{d}{dQ}\,S(Q) \;=\; -\;\frac{G'(S-Q)\,\theta'(Q)}{H''(S(Q),Q)}\,.$$

Therefore

$$\begin{aligned}\frac{d}{dQ}\,E(S(Q),Q) \;&=\; G'(S(Q)-Q)\left(\frac{dS(Q)}{dQ}\;-1\right)\theta(Q)\;+\;G(S(Q)-Q)\,\theta'(Q)\\ &\quad-\;\left\{H'(S(Q),Q)\,\frac{dS(Q)}{dQ}\;+\;G(S(Q)-Q)\,\theta'(Q)\right\}\\ &=\;-\left(\frac{G'(S(Q)-Q)\,\theta'(Q)}{H''(S(Q),Q)}\;+1\right)\theta(Q)\;G'(S(Q)-Q).\end{aligned}$$

Property 7 and (23) imply that $\theta(Q)\;G'(S(Q)-Q)\;<0$ if $Q>0$. If we can show that

$$H''(S(Q),Q) > -\,G'(S(Q)-Q)\,\theta'(Q) > 0 \tag{44}$$

then we will have shown that $\frac{d}{dQ}\,E(S(Q),Q)>0$, and the proof will be complete. By Properties 2 and 7 and by (23),

$$\begin{aligned}H''(S(Q),Q) \geq \int_0^Q G''(S(Q)-x)\,\theta'(Q)\;dx \;&=\; [G'(S(Q))-G'(S(Q)-Q)]\,\theta'(Q)\\ &>\;-\,G'(S(Q)-Q)\,\theta'(Q)>0.\end{aligned}$$

□

Lemma 9: The Cost Minimization Algorithm is quadratically convergent if Properties 2, 7 and 8 hold, and if $\theta(x)$ is twice continuously differentiable in a neighborhood of the optimal Q.

Proof: Differentiating (19) and applying (20), we see that for an optimal policy

$$\frac{\partial}{\partial S}\,E(S,Q)+\frac{\partial}{\partial Q}\,E(S,Q)=0. \tag{45}$$

Let $\overline{\delta}\equiv(\delta_1,\,\delta_2)$ be a row vector, and let $\epsilon\equiv\max(|\delta_1|,\;|\delta_2|)$. Let (S^*,Q^*) be optimal, $s^*=S^*-Q^*$, $S=S^*+\delta_1$ and $Q=Q^*+\delta_2\;=S-s$. By Property 8 there are column vectors $\vec{h}=(h_1,h_2)^{\mathrm{T}}$, $\vec{e}=(e_1,e_2)^{\mathrm{T}}$ such that the functions $H'(S,Q)$ and $E(S,Q)$ can be written as

$$H'(S^*+\delta_1,\;Q^*+\delta_2)\;=\;\overline{\delta}\,\vec{h}+O(\epsilon^2)\quad\text{and}$$

$$E(S^*+\delta_1,\;Q^*+\delta_2)\;=\;\overline{\delta}\;\vec{e}+O(\epsilon^2).$$

By (45), $e_2=-\,e_1$. Note that

$$\begin{aligned}H''(S,Q)+G'(S-Q)\,\theta'(Q)\;&=\frac{\partial}{\partial S}\,H'(S,Q)+\frac{\partial}{\partial Q}\,H'(S,Q)\\ &=\;h_1+h_2+O(\epsilon),\quad\text{and}\end{aligned} \tag{46}$$

$$-\,G'(S-Q)\,\theta(Q)\;=\;\frac{\partial}{\partial Q}\,E(S,Q)=\;(-e_1)+O(\epsilon). \tag{47}$$

By Lemma 8 and Property 7, $e_1\neq 0$. In the proof of Lemma 8 we showed that Properties 2, 7 and 8 imply (44), so $h_1+h_2>0$.

Starting from (S,s), the Cost Minimization Algorithm computes

$$
\begin{aligned}
s' &= s - \frac{E(S,\, S-s)}{G'(s)\,\theta(S-s)} \\
&= s^* + \delta_1 - \delta_2 - \frac{E[S^*+\delta_1,\ Q^*+\delta_2]}{e_1 + O(\epsilon)} \\
&= s^* + \delta_1 - \delta_2 - \frac{\delta_1\, e_1 + \delta_2\,(-e_1) + O(\epsilon^2)}{e_1 + O(\epsilon)} \;=\; s^* + O(\epsilon^2)
\end{aligned}
$$

and

$$
\begin{aligned}
S' &= S^* + \delta_1 - \frac{H'(S, S-s')}{H''(S, S-s') + G'(s')\,\theta'(S-s')} \\
&= S^* + \delta_1 - \frac{H'(S^*+\delta_1\,,\ S^* - s^* + \delta_1 + O(\epsilon^2))}{h_1 + h_2 + O(\epsilon)} \\
&= S^* + \delta_1 - \frac{\delta_1(h_1+h_2) + O(\epsilon^2)}{h_1 + h_2 + O(\epsilon)} \;=\; S^* + O(\epsilon^2).
\end{aligned}
$$

□

Lemma 10: In (s, S) policies for Levy demand processes,

$$
\mathrm{P}[NI \geq 0] = \frac{p}{p+h} + \frac{H'(S,Q)}{(h+p)\,\theta(Q)}\,,
$$

which is equal to $p\,/(p+h)$ for an optimal policy.

Proof: By (13), $G'(x) = h - (h+p)\,\overline{F}_D(x)$. Therefore

$$
\begin{aligned}
H'(S,Q) &\equiv \int_0^Q G'(S-x)\,\theta'(x)\,dx \\
&= h\cdot\theta(Q) - (h+p)\int_0^Q \overline{F}_D(S-x)\,\theta'(x)\,dx \\
&= \theta(Q)\,[h - (h+p)\,\mathrm{P}[D > IP]].
\end{aligned}
$$

Therefore

$$
\mathrm{P}[NI \geq 0] = 1 - \mathrm{P}[D > IP] = \frac{p}{h+p} + \frac{H'(S,Q)}{(h+p)\,\theta(Q)}.
$$

□

Lemma 11: For Levy demand processes, the fill rate of an (s, S) policy is given by

$$
1 - \mathrm{E}[(J - (IP - D)^+)^+/J]
$$

where J is the demand-weighted jump size (see Lemma 1), and J, IP, and D are independent.

Proof: Let $t(x) = D^{-1}(x)$ and let $t^-(x) = (t(x))^-$. The number of items ordered between time 0 and time T that that are backordered is

$$
S(T) \equiv \int_{t=0}^{T} \left[\frac{(D(t) - D(t^-) - [NI(t^-)]^+)^+}{D(t) - D(t^-)} \right] d\{D(t)\}.
$$

Let $h(x,y) = (x - y^+)^+/x$ for $x > 0$, and let $h(x,y) = 0$ for $x = 0$. Applying Lemma A.1 we have

$$
\begin{aligned}
\frac{1}{T}\mathrm{E}(S(T)) &= \frac{1}{T}\int_{t=0}^{T} \mathrm{E}\,[h(J, NI(t^-))]\ dt \\
&\to \mathrm{E}\,[h(J, NI)] \ \text{ as } \ T \to \infty\,,
\end{aligned}
$$

because $h(x, y)$ is bounded and continuous on $\Re \times \Re$, except at $(x, y) \in \{0\} \times \Re$, and because $\mathrm{P}((J, NI) \in \{0\} \times \Re) = 0$. Since $\frac{1}{T} D(T) \to 1$ as $T \to \infty$, the fill rate is

$$\lim_{T \uparrow \infty} \left[1 - \frac{S(T)}{D(T)} \right] = 1 - \mathrm{E}\left[h(J, NI) \right]$$

because of the regenerative nature of $S(T)$. □

Lemma 12: Assume that Properties 2, 3, 5, 6, 7 and 8 hold and that $\overline{Q} > 0$. Then $(\mathrm{P}^{\overline{Q}})$ has an optimal solution $(S(\overline{Q}), Q(\overline{Q}))$ which is the unique solution of the first-order optimality conditions for $(\mathrm{P}^{\overline{Q}})$. $S(\overline{Q})$ and $Q(\overline{Q})$ are continuous functions of $\overline{Q}$. If $\overline{Q}$ is sufficiently large then $\overline{Q} > Q(\overline{Q})$, and if $\overline{Q} > 0$ is sufficiently small then $\overline{Q} < Q(\overline{Q})$.

Proof: Equating zero to the partial derivatives of $c^{\overline{Q}}(S, Q)$ with respect to S and Q, respectively, we obtain

$$q(\overline{Q}) \cdot G'(S - a(\overline{Q})) + \theta'(\overline{Q}) \cdot (G(S) - G(S - Q)) = 0 \text{ and} \tag{48}$$

$$G(S - Q) = c^{\overline{Q}}(S, Q). \tag{49}$$

Claim 1: Let $a(\overline{Q}) \le Q_1 < Q_2$, and let $S(Q, \overline{Q})$ be the value of S that minimizes $c^{\overline{Q}}(S, Q)$ for given values of Q and $\overline{Q}$. Then $S(Q_1, \overline{Q}_1)$ and $S(Q_2, \overline{Q}_2)$ exist, are unique, and satisfy $S(Q_2, \overline{Q}_2) - Q_2 < S(Q_1, \overline{Q}_1) - Q_1 < y_0 < S(Q_1, \overline{Q}_1) < S(Q_2, \overline{Q}_2)$.

Proof of Claim 1: By (33), $Q_1 \ge a(\overline{Q}) > 0$. By Properties 2, 7 and 8 the left-hand side of (48) is continuous in S, is negative if $S \le y_0$, is strictly increasing in S if $S \ge y_0$, and is positive if $S - Q \ge y_0$. Therefore $S(Q_1, \overline{Q}_1)$ and $S(Q_2, \overline{Q}_2)$ exist and are unique, and $S(Q_1, \overline{Q}_1) - Q_1 < y_0 < S(Q_1, \overline{Q}_1)$. $S(Q_2, \overline{Q}_2) - Q_2 < S(Q_1, \overline{Q}_1) - Q_1$ and $S(Q_1, \overline{Q}_1) < S(Q_2, \overline{Q}_2)$ follow from (48) and the strict convexity of $G(x)$. This proves Claim 1.

Let

$$\begin{aligned} \Delta(Q) \equiv{} & (Q \cdot \theta'(\overline{Q}) + q(\overline{Q})) \cdot G(S(Q, \overline{Q}) - Q) - K \\ & - q(\overline{Q}) \cdot G(S(Q, \overline{Q}) - a(\overline{Q})) - \theta'(\overline{Q}) \int_0^Q G(S(Q, \overline{Q}) - x)\, dx. \end{aligned} \tag{50}$$

Then (49) is equivalent to $\Delta(Q) = 0$.

Claim 2: Let $a(\overline{Q}) \le Q_1 < Q_2$. Then $\Delta(Q_1) < \Delta(Q_2)$.

Proof of Claim 2: Let $S_1 \equiv S(Q_1, \overline{Q})$ and $S_2 \equiv S(Q_2, \overline{Q})$. Then

$$\begin{aligned} \Delta(Q_2) - \Delta(Q_1) ={} & q(\overline{Q}) \left[G(S_2 - Q_2) - G(S_2 - a(\overline{Q})) - G(S_1 - Q_1) + G(S_1 - a(\overline{Q})) \right] \\ & + \theta'(\overline{Q}) \left[Q_2 \cdot G(S_2 - Q_2) - Q_1 \cdot G(S_1 - Q_1) - \int_0^{Q_2} G(S_2 - x)\, dx + \int_0^{Q_1} G(S_1 - x)\, dx \right] \end{aligned}$$

$$\begin{aligned} ={} & q(\overline{Q})[G(S_2 - Q_2) - G(S_1 - Q_1)] - q(\overline{Q}) \left[G(S_2 - a(\overline{Q})) - G(S_1 - a(\overline{Q})) \right] \\ & + \theta'(\overline{Q}) \cdot Q_1 \cdot [G(S_2 - Q_2) - G(S_1 - Q_1)] + \theta'(\overline{Q}) \cdot (S_1 - Q_1 - S_2 + Q_2) \cdot G(S_2 - Q_2) \\ & + \theta'(\overline{Q}) \cdot (S_2 - S_1) \cdot G(S_2 - Q_2) - \theta'(\overline{Q}) \left[\int_{S_2 - Q_2}^{S_1 - Q_1} G(x)\, dx + \int_{S_1}^{S_2} G(x)\, dx \right] \end{aligned}$$

$$= \left[q(\overline{Q}) + \theta'(\overline{Q}) \cdot Q_1\right] \cdot \left[G(S_2 - Q_2) - G(S_1 - Q_1)\right]$$
$$+ \theta'(\overline{Q}) \int_{S_2 - Q_2}^{S_1 - Q_1} \left[G(S_2 - Q_2) - G(x)\right] dx$$
$$+ \theta'(\overline{Q}) \cdot (S_2 - S_1) \cdot G(S_2 - Q_2) - \theta'(\overline{Q}) \int_{S_1}^{S_2} G(x)\, dx \,-\, q(\overline{Q})\left[G(S_2 - a(\overline{Q})) - G(S_1 - a(\overline{Q}))\right].$$

By (33), $Q_1 \geq a(\overline{Q}) > 0$. Claim 1 and Property 7 imply that the first two terms of this expression are positive. The definition of $S_2 \equiv S(Q_2, \overline{Q})$, Claim 1 and Property 7 imply that

$$\theta'(\overline{Q}) \cdot (S_2 - S_1) \cdot G(S_2 - Q_2) = (S_2 - S_1) \cdot \left[\theta'(\overline{Q}) \cdot G(S_2) + q(\overline{Q}) \cdot G'(S_2 - a(\overline{Q}))\right]$$
$$> \int_{S_1}^{S_2} \left[\theta'(\overline{Q}) \cdot G(x) + q(\overline{Q}) \cdot G'(x - a(\overline{Q}))\right] dx$$
$$= \theta'(\overline{Q}) \int_{S_1}^{S_2} G(x)\, dx \,+\, q(\overline{Q}) \cdot \left[G(S_2 - a(\overline{Q})) - G(S_1 - a(\overline{Q}))\right].$$

Therefore $\Delta(Q_2) - \Delta(Q_1) > 0$, and Claim 2 holds.

Claim 3: For all $\overline{Q} > 0$, ($P^{\overline{Q}}$) has an optimal solution $(S(\overline{Q}),\ Q(\overline{Q}))$ which is the unique solution of the first-order optimality conditions for ($P^{\overline{Q}}$).

Proof of Claim 3: Let $R \equiv \Re \times \{x : x \geq a(\overline{Q})\}$. By (33), $a(\overline{Q}) > 0$. Note that $c^{\overline{Q}}(S, Q)$ is continuous for $(S, Q) \in R$ and that $c^{\overline{Q}}(S, Q)$ converges to infinity as $\max\{Q,\ |S|\} \to \infty$. Thus (34) achieves its minimum in R, and this minimum satisfies the first-order optimality conditions for ($P^{\overline{Q}}$). The first-order optimality conditions are equivalent to

$$S = S(Q, \overline{Q}),\ \ \Delta(Q) \geq 0,\ \ Q \geq a(\overline{Q}),\ \text{and}\ \ 0 = \Delta(Q) \cdot \left(Q - a(\overline{Q})\right).$$

Claims 1 and 2 imply that the solution to these conditions is unique. Thus Claim 3 holds, and $S(\overline{Q})$ and $Q(\overline{Q})$ are well-defined functions of $\overline{Q} > 0$.

Claim 4: $S(\overline{Q})$ and $Q(\overline{Q})$ are continuous functions of $\overline{Q}$.

Proof of Claim 4: By (32) and Properties 3 and 5, $a(\overline{Q})$ and $q(\overline{Q})$ are both continuous. By Properties 3 and 5 and (34), if $(S_m,\ Q_m,\ \overline{Q}_m)$ converges to $(S,\ Q,\ \overline{Q})$ in $\Re \times \{(x, y) : x \geq a(y), y > 0\}$ then $c^{\overline{Q}_m}(S_m, Q_m)$ converges to $c^{\overline{Q}}(S, Q)$. Let $\overline{Q}_n \to \overline{Q} > 0$, let $(S, Q) \equiv (S(\overline{Q}), Q(\overline{Q}))$, and let $(S_n, Q_n) \equiv (S(\overline{Q}_n), Q(\overline{Q}_n))$. Select the subsequence $\{N(j) : j = 1,\ 2,\ \ldots\}$ of $\{n = 1,\ 2,\ \ldots\}$ so that $\liminf_{n\to\infty} c^{\overline{Q}_n}(S_n, Q_n) = \lim_{j\to\infty} c^{\overline{Q}_{N(j)}}(S_{N(j)}, Q_{N(j)})$. For all $\epsilon > 0$, $(S, Q + \epsilon)$ is feasible for ($P^{\overline{Q}_n}$) if n is sufficiently large. Therefore $c^{\overline{Q}_{N(j)}}(S_{N(j)}, Q_{N(j)}) \leq c^{\overline{Q}_{N(j)}}(S, Q + 1) \to c^{\overline{Q}}(S, Q + 1) < \infty$, so the sequence $(S_{N(j)}, Q_{N(j)})$ is bounded. Therefore it has a subsequence $(S_{n(k)},\ Q_{n(k)})$ such that $(S_{n(k)},\ Q_{n(k)},\ \overline{Q}_{n(k)})$ converges to $(S^*,\ Q^*,\ \overline{Q})$ in $\Re^3$. Because $Q_{n(k)} \geq a(\overline{Q}_{n(k)})$, $Q^* \geq a(\overline{Q})$.

Let $\epsilon > 0$.

$$\liminf_{n\to\infty} c^{\overline{Q}_n}(S_n, Q_n) = \lim_{k\to\infty} c^{\overline{Q}_{n(k)}}(S_{n(k)}, Q_{n(k)}) = c^{\overline{Q}}(S^*, Q^*)$$
$$\geq c^{\overline{Q}}(S, Q + \epsilon) + \left[c^{\overline{Q}}(S, Q) - c^{\overline{Q}}(S, Q + \epsilon)\right] = \lim_{n\to\infty} c^{\overline{Q}_n}(S, Q + \epsilon) + \left[c^{\overline{Q}}(S, Q) - c^{\overline{Q}}(S, Q + \epsilon)\right]$$
$$\geq \limsup_{n\to\infty} c^{\overline{Q}_n}(S_n, Q_n) + \left[c^{\overline{Q}}(S, Q) - c^{\overline{Q}}(S, Q + \epsilon)\right].$$

Taking the limit as $\epsilon \to 0$, we see that $c^{\overline{Q}}(S(\overline{Q}), Q(\overline{Q}))$ is continuous in $\overline{Q}$.

Equation (49) implies that $G(S(\overline{Q}) - Q(\overline{Q}))$ is continuous in $\overline{Q}$. By Property 7 and (23), $S(\overline{Q}) - Q(\overline{Q})$ is continuous in $\overline{Q}$. By (48) and Property 5, $q(\overline{Q}) \cdot G'(S(\overline{Q}) - a(\overline{Q})) + \theta'(\overline{Q}) \cdot G(S(\overline{Q}))$ is continuous in $\overline{Q}$. But for $S(\overline{Q}) \geq y_0$, Properties 7 and 8 imply that this expression is continuous and strictly increasing in $S(\overline{Q})$. Thus $S(\overline{Q})$ and $Q(\overline{Q})$ are both continuous in $\overline{Q}$, and Claim 4 holds.

Claim 5: If $\overline{Q}$ is sufficiently large then $\overline{Q} > Q(\overline{Q})$, and if $\overline{Q} > 0$ is sufficiently small then $\overline{Q} < Q(\overline{Q})$.

Proof of Claim 5: Consider the policy

$$Q = 1\ ,\ \ S = y_0 + a(\overline{Q}). \tag{51}$$

Let $p \vee h \equiv \max\{p, h\}$. By (34) and (14), the cost of this policy is

$$\begin{aligned} c^{\overline{Q}}(S,Q) &= \left[K + q(\overline{Q}) \cdot G(S - a(\overline{Q})) +\ \theta'(\overline{Q}) \int_0^Q G(S-x)\,dx\right] / \left[q(\overline{Q}) + Q \cdot \theta'(\overline{Q})\right] \\ &\le \left[K + q(\overline{Q}) \cdot G(y_0) + \theta'(\overline{Q}) \int_0^1 (G(y_0) + (p \vee h) \cdot |x - y_0|)\,dx\right] / \left[q(\overline{Q}) + \theta'(\overline{Q})\right] \\ &\le \left[K + q(\overline{Q}) \cdot G(y_0) + \theta'(\overline{Q})(G(y_0) + (p \vee h)/2)\right] / \left[q(\overline{Q}) + \theta'(\overline{Q})\right] \\ &\le G(y_0) + \left[K + (p \vee h)\,\theta'(\overline{Q})/2\right] / \theta'(\overline{Q}) \\ &= G(y_0) + K/\theta'(\overline{Q}) + \frac{1}{2}(p \vee h)\ \le\ G(y_0) + K + \frac{1}{2}(p \vee h). \end{aligned}$$

The average cost of any policy for which $Q \le \overline{Q}$ is

$$c^{\overline{Q}}(S,Q) > K/\left[q(\overline{Q}) + Q \cdot \theta'(\overline{Q})\right] \ge K/\theta(\overline{Q}).$$

By Property 2 this quantity grows without bound as $\overline{Q} \to 0$. The cost of the policy in (51) is finite, and if $\overline{Q}$ is sufficiently small, this policy is feasible for ($P^{\overline{Q}}$) by (33). Thus $\overline{Q} < Q(\overline{Q})$ for $\overline{Q}$ sufficiently small.

We need to show that $\overline{Q} > Q(\overline{Q})$ if $\overline{Q}$ is sufficiently large. In view of Claim 2, it suffices to show that $\Delta(\overline{Q}) > 0$ if $\overline{Q}$ is sufficiently large. Let $S = S(\overline{Q}, \overline{Q})$. By (50),

$$\begin{aligned} \Delta(\overline{Q}) = q(\overline{Q}) \cdot \left[G(S - \overline{Q}) - G(S - a(\overline{Q}))\right] + \theta'(\overline{Q}) \int_{S-\overline{Q}}^{S} \left[G(S - \overline{Q}) - G(x)\right]^+ dx \\ - \theta'(\overline{Q}) \int_{S-\overline{Q}}^{S} \left[G(x) - G(S - \overline{Q})\right]^+ dx - K. \end{aligned} \tag{52}$$

Assume that $y_0 \le S - a(\overline{Q})$. Claim 1 and (33) imply that $S - \overline{Q} < y_0 \le S - a(\overline{Q}) < S$. By (48), $G(S - \overline{Q}) \ge G(S)$. The convexity of G and (14) imply that the first term of (52) is positive, the third term is zero, and the second term grows without bound as $\overline{Q} \to \infty$. Therefore Claim 5 holds.

Assume that $y_0 > S - a(\overline{Q})$. Claim 1 and (33) imply that $S - \overline{Q} < S - a(\overline{Q}) < y_0 < S$. We claim that the second term of (52) grows without bound as $\overline{Q} \to \infty$. By the convexity of G and (14), this is true if $y_0 - (S - \overline{Q})$ grows without bound as $\overline{Q} \to \infty$. But $y_0 - (S - \overline{Q}) \ge \overline{Q} - a(\overline{Q}) \ge \overline{Q}/2$ by (33). Therefore it suffices to show that

$$A(\overline{Q}) \equiv q(\overline{Q}) \cdot \left[G(S - \overline{Q}) - G(S - a(\overline{Q}))\right] - \theta'(\overline{Q}) \int_{S-\overline{Q}}^{S} \left[G(x) - G(S - \overline{Q})\right]^+ dx \tag{53}$$

is non-negative. By the convexity of G and (48), the first term of (53) is

$$\begin{aligned} q(\overline{Q}) \cdot \left[G(S - \overline{Q}) - G(S - a(\overline{Q}))\right] &\ge -q(\overline{Q}) \cdot \left[(\overline{Q} - a(\overline{Q})) \cdot G'(S - a(\overline{Q}))\right] \\ &= \theta'(\overline{Q}) \cdot (\overline{Q} - a(\overline{Q})) \cdot \left[G(S) - G(S - \overline{Q})\right]. \end{aligned}$$

Equation (48) implies that $G(S) > G(S - \overline{Q})$. Therefore there is a z such that $y_0 < z < S$ and $G(z) = G(S - \overline{Q})$. The second term of (53) is

$$\theta'(\overline{Q})\int_z^S [G(x)-G(S-\overline{Q})]dx \le \theta'(\overline{Q})\cdot\frac{1}{2}\cdot(S-z)\cdot[G(S)-G(S-\overline{Q})]$$
$$< \theta'(\overline{Q})\cdot\frac{1}{2}\cdot a(\overline{Q})\cdot[G(S)-G(S-\overline{Q})].$$

By (33),

$$A(\overline{Q}) > \theta'(\overline{Q})\cdot(\overline{Q}-\frac{3}{2}a(\overline{Q}))\cdot[G(S)-G(S-\overline{Q})] > 0.$$

This completes the proof of Claim 5 and of Lemma 12. □

Lemma 13: The relative cost of Zheng's policy is at most $\min\left\{\frac{\theta(Q)-Q}{Q+G(y_0)/p}\ ,\ \frac{p+h}{p}\ \frac{[\theta(Q)-Q]^2+\eta(Q)}{Q^2}\right\}$, where $Q=S-s$ is the optimal order quantity (not Zheng's order quantity).

Proof: Note that in computing the average cost of an (s,S) policy, Zheng's cost model corresponds to assuming that $\theta(x)\equiv x$. Under this assumption the average cost incurred by an (s,S) policy with $Q=S-s$ is

$$c^z(S,Q) = \frac{1}{Q}\left[K+\int_0^Q G(S-x)\,dx\right].$$

Let

$$H^*(S,Q) \equiv \int_0^Q G(S-x)\,[\theta'(x)-1]\,dx\,.$$

The true average cost incurred by an (s,S) policy can be written as

$$\begin{aligned} c(S,Q) &= \frac{1}{\theta(Q)}\left[K+\int_0^Q G(S-x)\,dx + H^*(S,Q)\right] \\ &= \frac{1}{\theta(Q)}\,[Q\cdot c^z(S,Q)+H^*(S,Q)]. \end{aligned} \tag{54}$$

Let (s_z, S_z) be Zheng's policy and let $Q_z=S_z-s_z$. Let (s_*, S_*) be the optimal policy and let $Q_*=S_*-s_*$. We claim that

$$c(S_*, Q_*) \le c(S_z, Q_z) \le c^z(S_z, Q_z) \le c^z(S_*, Q_*). \tag{55}$$

The optimality of Zheng's policy under his assumptions, and the optimality of the optimal policy, justify the left and right inequalities of (55). The optimality of (s_z, S_z) implies that $G(s_z)=G(S_z)=c^z(S_z,Q_z)$. By (54),

$$c(S_z,Q_z) \le \frac{1}{\theta(Q_z)}\,[Q_z\cdot c^z(S_z,Q_z) + G(s_z)\cdot(\theta(Q_z)-Q_z)] = c^z(S_z,Q_z),$$

which proves (55).

In view of (55) it suffices to prove a bound on

$$\frac{c^z(S_*, Q_*)-c(S_*, Q_*)}{c(S_*, Q_*)}. \tag{56}$$

By (14) and (21),

$$c(S_*, Q_*) = G(s_*) \le G(y_0)+p\cdot(y_0-s_*) \le G(y_0)+p\cdot Q_*\,,$$

$$H^*(S_*\,,\,Q_*) \ge \int_0^{Q_*} [G(S_*) - h\,x] \cdot [\theta'(x) - 1]\,dx \qquad (57)$$
$$= G(S_*) \cdot [\theta(Q_*) - Q_*] - \, h \cdot \eta(Q_*), \quad \text{and}$$

$$H^*(S_*\,,\,Q_*) \ge \int_0^{Q} G(y_0)\,[\theta'(x) - 1]\,dx \,= G(y_0) \cdot [\theta(Q_*) - Q_*].$$

By (54) and (57),

$$\frac{c^z(S_*\,,\,Q_*) - c(S_*\,,\,Q_*)}{c(S_*\,,\,Q_*)} = \frac{\theta(Q_*) \cdot c(S_*\,,\,Q_*) - H^*(S_*\,,\,Q_*)}{Q_* \cdot c(S_*\,,\,Q_*)} - 1$$
$$\le \frac{\theta(Q_*)}{Q_*} - 1 - \frac{G(y_0) \cdot [\theta(Q_*) - Q_*]}{Q_* \cdot [G(y_0) + p \cdot Q_*]}$$
$$= \frac{p \cdot Q_* \cdot [\theta(Q_*) - Q_*]}{Q_* \cdot [G(y_0) + p \cdot Q_*]} = \frac{\theta(Q_*) - Q_*}{G(y_0)/p + Q_*}\,.$$

This is the first of the two upper bounds we need to establish. We claim without proof that for the gamma-dsitributed Levy process, if we set the lead time equal to zero and let the order quantity approach zero, $(S_* - y_0)/Q_* \to 0$ and $\theta(\epsilon \cdot Q)/\theta(Q) \to 1$. Therefore this is an asymtotically tight bound on (56) for small Q, when the lead time is zero.

By Jensen's Inequality,

$$G(x) \equiv \mathrm{E}[h(x - D)^+ + p(D - x)^+] \ge h(x - \mathrm{E}(D))^+ + p(\mathrm{E}(D) - x)^+ \equiv G_0(x). \qquad (58)$$

Given a value of Q we define $s(Q)$ by $G(s(Q)) = G(Q + s(Q))$, and we define $s_0(Q)$ by $G_0(s_0(Q)) = G_0(Q + s_0(Q))$. By Lemma 7, $G(s_*) \ge G(S_*)$, so $G(s_*) \ge G(s(Q_*))$. Therefore

$$c(S_*\,,\,Q_*) = G(s_*) \ge G(s(Q_*)) \ge G_0(s_0(Q_*)) = \frac{p \cdot h}{p + h}\,Q_*\,. \qquad (59)$$

By (57), Lemma 7 and (59),

$$\frac{c^z(S_*\,,\,Q_*) - c(S_*\,,\,Q_*)}{c(S_*\,,\,Q_*)} = \frac{\theta(Q_*) \cdot c(S_*\,,\,Q_*) - H^*(S_*\,,\,Q_*)}{Q_* \cdot c(S_*\,,\,Q_*)} - 1$$
$$\le \frac{\theta(Q_*)}{Q_*} - 1 - \frac{1}{Q_* \cdot c(S_*\,,\,Q_*)}\,(G(S_*) \cdot [\theta(Q_*) - Q_*] - \, h \cdot \eta(Q))$$
$$\le \frac{\theta(Q_*) - Q_*}{Q_*} - \frac{1}{Q_* \cdot c(S_*\,,\,Q_*)}\,\big(c(S_*\,,\,Q_*) \cdot [\theta(Q_*) - Q_*] - h \cdot [\theta(Q_*) - Q_*]^2 - \, h \cdot \eta(Q)\big)$$
$$= h\,\frac{[\theta(Q_*) - Q_*]^2 + \eta(Q)}{Q_* \cdot c(S_*\,,\,Q_*)} \le \frac{p + h}{p}\;\frac{[\theta(Q_*) - Q_*]^2 + \eta(Q)}{Q_*^2}.$$

This is our second bound, and concludes the proof. □

Lemma 14: The relative cost of the Mass Uniform Policy is at most $\frac{h+p}{p}\,B(Q)$ where

$$B(Q) \equiv \frac{1}{Q \cdot \theta(Q)} \int_0^{Q} x \cdot (\theta'(x) - \theta'(Q))\,dx \,= \frac{q(Q) \cdot a(Q)}{Q \cdot \theta(Q)},$$

where $Q = S - s$ is the mass-uniform order quantity (not the optimal order quantity).

Proof: Since $\overline{Q}$ is adjusted to make $Q = \overline{Q} > a(\overline{Q})$, we can view the Mass Uniform policy as the policy computed by the original Cost Minimzation Algorithm, using the measure $d\{\mu(x)\}$ in place of $\theta'(x)\,dx$. We suppress the dependence of $\mu(x)$ on $\overline{Q}$. The cost cost function to be minimized is

$$c^{\mu}(S,Q) \;=\frac{1}{\theta(Q)}\left[K+q(Q)\cdot G(S-a(Q))+\theta'(Q)\int_0^Q G(S-x)\;dx\right].$$

Let

$$H^{\mu}(S,Q)\equiv\int_0^Q G(S-x)\;[\theta'(x)-\theta'(Q)]\;dx\,.$$

The true average cost incurred by an (s,S) policy with $Q=S-s$ can be written as

$$\begin{aligned} c(S,Q) \;&=\frac{1}{\theta(Q)}\left[K+\theta'(Q)\int_0^Q G(S-x)\;dx\;+\;H^{\mu}(S,Q)\right] \\ &=\;\frac{1}{\theta(Q)}\,[\,H^{\mu}(S,Q)-q(Q)\cdot G(S-a(Q))]+c^{\mu}(S,Q). \end{aligned} \tag{60}$$

Let $(s_{\mu}\,,S_{\mu})$ be the Mass Uniform policy and let $Q_{\mu}=S_{\mu}-s_{\mu}$. Let $(s_*\,,S_*)$ be the optimal policy and let $Q_*=S_*-s_*$. We claim that

$$c(S_{\mu}\,,\,Q_{\mu})\ge c(S_*\,,\,Q_*)\ge c^{\mu}(S_*\,,\,Q_*)\ge c^{\mu}(S_{\mu}\,,\,Q_{\mu}). \tag{61}$$

The optimality of the Mass Uniform policy for the cost function $c^{\mu}(S,\;Q)$, and the optimality of the optimal policy, justify the left and right inequalities of (61). Note that $[\theta'(x)-\theta'(Q)]/q(Q),\;0\le x\le Q$ is a probability density with first moment $a(Q)$. By Jensen's inequality

$$H^{\mu}(S,Q)\;\ge\;q(Q)\cdot G(S-a(Q)).$$

By (60), (61) holds.

We derive two more inequalities. First, by the convexity of G and (32),

$$\begin{aligned} H^{\mu}(S_{\mu},Q_{\mu}) &=\int_0^{Q_{\mu}} G(S_{\mu}-x)\;[\theta'(x)-\theta'(Q_{\mu})]\;dx \\ &\le\int_0^{Q_{\mu}}[G(S_{\mu})\;+\;x\cdot(G(s_{\mu})-G(S_{\mu}))/Q_{\mu}]\cdot[\theta'(x)-\theta'(Q_{\mu})]\;dx \\ &=\;q(Q_{\mu})\cdot[G(S_{\mu})+a(Q_{\mu})(G(s_{\mu})-G(S_{\mu}))/Q_{\mu}]. \end{aligned} \tag{62}$$

By (58),

$$\begin{aligned} \frac{h\,Q_{\mu}-G(S_{\mu})}{G(s_{\mu})} &\le\frac{h\,Q_{\mu}-G_0(S_{\mu})}{G_0(s_{\mu})} \\ &=\frac{h\,Q_{\mu}-h(S_{\mu}-\mathrm{E}(D))}{p\,(\mathrm{E}(D)-s_{\mu})}\;=\;\frac{h}{p}\,. \end{aligned} \tag{63}$$

In view of (61) it suffices to prove a bound on $[c(S_{\mu}\,,\,Q_{\mu})-c^{\mu}(S_{\mu}\,,\,Q_{\mu})\,]/c^{\mu}(S_{\mu}\,,\,Q_{\mu})$. By (21), $c^{\mu}(S_{\mu}\,,\,Q_{\mu})=G(s_{\mu})$. By (60), (62), (14) and (63),

$$
\begin{aligned}
\frac{c(S_\mu\,,\,Q_\mu)-c^\mu(S_\mu\,,\,Q_\mu)}{c^\mu(S_\mu\,,\,Q_\mu)} &= \frac{H^\mu(S_\mu,Q_\mu)-q(Q_\mu)\cdot G(S_\mu-a(Q_\mu))}{\theta(Q_\mu)\cdot G(s_\mu)} \\
&\le \frac{q(Q_\mu)\cdot[G(S_\mu)-\,G(S_\mu-a(Q_\mu))\,+\,a(Q_\mu)(G(s_\mu)-G(S_\mu))/Q_\mu]}{\theta(Q_\mu)\cdot G(s_\mu)} \\
&\le \frac{q(Q_\mu)\cdot[h\cdot a(Q_\mu)\,+\,a(Q_\mu)(G(s_\mu)-G(S_\mu))/Q_\mu]}{\theta(Q_\mu)\cdot G(s_\mu)} \\
&= \frac{q(Q_\mu)\cdot a(Q_\mu)}{Q_\mu\cdot\theta(Q_\mu)}+\frac{q(Q_\mu)\cdot a(Q_\mu)}{\theta(Q_\mu)}\,\frac{h-G(S_\mu)/Q_\mu}{G(s_\mu)} \\
&\le \frac{q(Q_\mu)\cdot a(Q_\mu)}{Q_\mu\cdot\theta(Q_\mu)}\cdot\left[\frac{p+h}{p}\right]=B(Q)\cdot\left[\frac{p+h}{p}\right]. \qquad \square
\end{aligned}
$$

Appendix 3: Numerical Approximations of $\theta(x)$ and $\theta'(x)$.

In this appendix we describe our approach to the numerical integrations that are needed to compute an optimal policy. We also give the polynomial approximations for $\theta'(x)$, $\theta(x)$ and $\eta(x)$ for the gamma-distributed Levy process, which we developed and used.

Numerical Integrations

We need to evaluate the integrals $\int_0^Q G(x)\,\theta'(x)\,dx$, $\int_0^Q G'(x)\,\theta'(x)\,dx$ and $\int_0^Q G''(x)\,\theta'(x)\,dx$. Let $D^*(t) = D^*(0,t)$ be the demand that occurs in t days. $D^*(t)$ has a gamma distribution with rate parameter 1 and shape parameter t. Let L be the lead time. The functions $f(x)$ that are of interest can be expressed in terms of the density $f_{D(L)}(x)$, the complementary distribution function $\overline{F}_{D(L)}(x) = P(D(L) > x)$, and the partial expectation $n_{D(L)}(x) = \mathrm{E}[(D(L)-x)^+]$. We define the partial second moment by $n^2_{D(t)}(x) = \mathrm{E}\{\frac{1}{2}[(D(L)-x)^+]^2\}$. We use the following identities, which apply to gamma demand distributions: $n_{D^*(L)}(x) = L\,\overline{F}_{D^*(L+1)}(x) - x\,\overline{F}_{D^*(L)}(x)$, and $n^2_{D^*(L)}(x) = [(L^2+L)/2]\cdot\overline{F}_{D^*(L+2)}(x) - L\,x\,\overline{F}_{D^*(L+1)}(x) + (x^2/2)\,\overline{F}_{D^*(L)}(x)$.

We evaluate $\int_0^Q G(x)\,\theta'(x)\,dx$ as $\int_0^Q G(x)\,dx + \int_0^Q \mathrm{G}(x)\,[\theta'(x)-1]\,dx$. The first of these two integrals is evaluated in closed form using the identity $\frac{d}{dx}n^2_D(x) = -n_D(x)$. This identity also simplifies the computations required for Zheng's policy. Integrals of the form $\int_0^x f(x)\,[\theta'(x)-1]\,dx$ are written as $\int_0^{h^{-1}(x)} f(h(y))\,[\theta'(h(y))-1]\,h'(y)\,dy$ where

$$h(y) = e^{(1-1/y)} \text{ if } y \le 1, \quad \text{and} \quad h(y) = 1 - \ln(2-y) \text{ if } 1 \le y \le 2. \tag{64}$$

This is numerically advantageous because both $[\theta'(h(y))-1]\,h'(y)$, $0 \le y \le 2$ and the interval of integration are bounded. The integrals of $G'(x)$ and $G''(x)$ are managed similarly, using the identities $\frac{d}{dx}n_D(x) = -\overline{F}_D(x)$ and $\frac{d}{dx}\overline{F}_D(x) = -f_D(x)$.

Polynomial Approximation for $\theta'(x)$

Accuracy: $\pm 0.1\%$

Algorithm

$y = (1-e^{-3h^{-1}(x)})/3$ (see (64))

$\nu = c[x,0] + c[x,1]y + c[x,2]y^2 + c[x,3]y^3$

$\theta'(x) = 1 + \nu/h'(y)$ (see (64))

Data

x Range	$c[x,0]$	$c[x,1]$	$c[x,2]$	$c[x,3]$
$0 < x \le 3.25\times10^{-6}$	1	3.138513	8.15216	0
$3.25\times10^{-6} < x \le 0.00027$	1.130977	-3.496022	119.1526	-615.058
$0.00027 < x \le 0.0015$	0.653346	9.77922	-0.2834	-272.15
$0.0015 < x \le 0.0073$	0.619633	10.5112	-4.024	-273.3
$0.0073 < x \le 0.065$	-1.44249	58.05855	-370.532	670.977
$0.065 < x \le 0.237$	1.142996	16.91995	-152.0991	283.902
$0.237 < x \le 0.79$	5.650726	-41.11935	97.60782	-75.0501
$0.79 < x \le 1.0$	3.531086	-20.14113	28.72104	0
$1.0 < x \le 1.13$	0.2771337	-0.7399454	-0.097316	0
$1.13 < x \le 1.26$	0.348694	-0.95677	-0.114	0
$1.26 < x \le 1.5$	0.4634541	-1.346797	0	0
$1.5 < x \le 2.0$	-8.801133	55.67633	-87.7363	0
$2.0 < x \le 2.7$	1.383476	-4.13484	0	0
$2.7 < x \le 7.0$	1.1125	0	-10	0
$7.0 < x$	0	0	0	0

Polynomial Approximation for $\theta(x)$

Comment: Note that by Lemma 2, $[\theta(Q) - Q] \to \mathrm{E}[V] = 1/2$ as $Q \to \infty$.

Accuracy: $\pm 0.1\%$

Algorithm

$y = \ln(x)$

If $x \leq 0.00125$ then $\nu = [c[x,1] - c[x,3](-y)^{c[x,0]}] \,/\, [c[x,2] - c[x,4](-y)^{c[x,0]}]$

If $0.00125 < x \leq 7.0$ then $\nu = c[x,0] + c[x,1]y + c[x,2]y^2 + c[x,3]y^3 + c[x,4]y^4$

If $7.0 < x$ then $\nu = [x + exp(-2)]/x$

$\theta(x) = x - 1/\ln[\nu \cdot x\,/(1 + e^2\,x)]$

Data

x Range	$c[x,0]$	$c[x,1]$	$c[x,2]$	$c[x,3]$	$c[x,4]$
$0 < x \leq 2 \times 10^{-7}$	-1.2839	6.6085	3.76234	-30.11	-30.11
$2 \times 10^{-7} < x \leq 0.00125$	0.5292	-5.963	-2.934	-3.157	-1.78
$0.00125 < x \leq 0.029$	1.45469	0.523461	0.180076	0.0213841	0.00088453
$0.029 < x \leq 0.15$	1.04877	0.013856	-0.058843	-0.0283627	-0.0030053
$0.15 < x \leq 0.62$	1.050675	0.024767	-0.038544	-0.014739	0
$0.62 < x \leq 3.0$	1.04993	0.02395	-0.03346	-0.003828	0.0069
$3.0 < x \leq 7.0$	1.07574	-0.031276	-0.002662	0.002	0

Polynomial Approximation for $\eta(x) \equiv \int_0^x z \cdot (\theta'(z) - 1)\, dz$

Accuracy: $\pm 0.25\%$

Algorithm

$y = \ln(x)$

If $x \leq 0.0025$ then $\nu = \left[1 - c[x,2] \cdot (c[x,4] - y)^{c[x,0]}\right] / \left[c[x,1] - c[x,3] \cdot (c[x,4] - y)^{c[x,0]}\right]$

If $0.0025 < x \leq 1.4$ then $\nu = c[x,0] + c[x,1]y + c[x,2]y^2 + c[x,3]y^3 + c[x,4]y^4$

If $x \leq 1.4$ then $\eta(x) = \nu\,(1 - e^{-x}) \,/\, \{11.93707[1 - (8x+1)e^{-8x}] + (ln(\min(x,0))^2\}$

If $1.4 < x$ then $\eta(x) = c[x,4] \cdot \left\{1 - (c[x,0]\,x + c[x,1])e^{-c[x,2]x}/(x + c[x,3])^2\right\}$

Data

x Range	$c[x,0]$	$c[x,1]$	$c[x,2]$	$c[x,3]$	$c[x,4]$
$0 < x \leq 0.0025$	0.772	0.9062	0.2774	0.2774	1.85
$0.0025 < x \leq 0.035$	-0.0103	-0.16226	0.00445	0.001533	
$0.035 < x \leq 0.105$	1.7647	1.338	0.42813	0.0415	
$0.105 < x \leq 0.36$	0.6827	-0.9955	-1.3737	-0.5603	-0.0741536
$0.36 < x \leq 1.4$	0.9877	0.057	-0.0774	0.063	
$1.4 < x$	-2.02	17.454	0.6573	4.558	1/12

References

Asxater, S. "Using the Deterministic EOQ Formula in Stochastic Inventory Control". Working paper, Lulea University of Technology, S-951 87 Lulea, Sweden, 1993.

Browne, S. and P. Zipkin. "Inventory Models with Continuous, Stochastic Demands", *The Annals of Applied Probability*, 1 (1991), 419-435.

Federgruen, A. and Y. Zheng. "Finding Optimal (s, S) Policies is About as Simple as Evaluating a Single Policy", *Operations Research*, 39 (1991), 654-665.

Feller, W. *An Introduction to Probability and its Applications, Volume II*. Wiley, New York, New York, 1966.

Gallego, G. "Distribution Free Bounds and Heuristics for (Q, r) Policies", Working paper, Department of Industrial Engineering and Operations Research, Columbia University, New York, NY, 1994. Forthcoming in *Management Science*.

Hadley, G.J. and T.M. Whitin. Analysis of Inventory Systems. Prentice Hall, Englewood Cliffs, NJ, 1963.

Hida, T. *Stationary Stochastic Processes*. Princeton University Press, Princeton, New Jersey, 1970.

Johnson, L.A. and D.C. Montgomery. *Operations Research in Production Planning, Scheduling, and Inventory Control*. John Wiley and Sons, New York, New York, 1974.

Nahmias, S. *Production and Operations Analysis*, Second Edition. Irwin, Homewood Illinois, 60430, 1993.

Prabhu, N.U. *Stochastic Storage Processes*. Springer-Verlag, New York, New York, 1980.

Resnick, S. *Adventures in Stochastic Processes.* Birkhauser, Boston, Massachusetts, 1992.

Serfozo, R. and S. Stidham. "Semi-Stationary Clearing Processes", *Stochastic Processes and their Applications*, 6 (1978), 165-178.

Sharpe, M. *General Theory of Markov Porcesses*. Academic Press, Boston Massachusetts, 1988.

Vollman, T.E., W.L. Berry and D.C. Whybark. *Manufacturing Planning and Control Systems*, Third Edition. Irwin, Homewood Illinois, 60430, 1992.

Zheng, Y. "On Properties of Stochastic Inventory Systems", *Management Science* 38 (1992), 87-103.

Zipkin, P. "Stochastic Lead Times in Continuous-Time Inventory Models", *Naval Research Logistics Quarterly* 33 (1986), 763-774.