OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0030-364X | EISSN 1526-5463 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

New Policies for Stochastic Inventory Control Models: Theoretical and Computational Results

Gavin Hurley
Goldman Sachs, Peterborough Court, 133 Fleet Street, London EC4A 2BB, England, gavin.hurley@gs.com,

Peter Jackson
School of ORIE, Cornell University, Ithaca, NY 14853., pj16@cornell.edu,

Retsef Levi
Sloan School of Management, MIT, Cambridge, MA, 02139., retsef@mit.edu,

Robin O. Roundy
School of ORIE, Cornell University, Ithaca, NY 14853., robin@orie.cornell.edu,

David B. Shmoys
School of ORIE and Dept. of Computer Science, Cornell University, Ithaca, NY 14853., shmoys@cs.cornell.edu,

Recently Levi, Pál, Roundy and Shmoys introduced a novel, Dual-Balancing policy for the classical single-item, single-location inventory model with backlogged demands and dynamic forecasts of future demands that evolve as time advances. These models are usually computationally intractable due to the enormous size of the state space. The expected cost of the dual-balancing policy is guaranteed to be at most twice the optimal expected cost, but until now, no computational testing of the policy has been done. We propose two extended families of policies, based on cost-balancing techniques and myopic-like policies that generate lower and upper bounds on the optimal base-stock levels. We show that cost-balancing techniques combined together with these lower and upper bounds lead to improved policies. The expected cost of the new policies is also guaranteed to be at most twice the optimal expected cost. Nevertheless, empirically their performance is significantly better. Moreover, all of the new policies can be implemented efficiently in an on-line manner.

We have conducted extensive testing of these policies, with demand forecasts that evolve according to the multiplicative MMFE model. The best of the new generation of policies are very robust. They are consistently better than the classical myopic policy over a broad set of important scenarios, and the improvement can get to up to 30 percent. The computational results demonstrate the effectiveness and computational practicality of the new policies in realistic scenarios.

Subject classifications: Stochastic Inventory Control; Heuristics; Approximation Algorithms.
Area of review: Supply Chain Management.
History: Submitted July 3, 2006, Revised October 23, 2007

1. Introduction

The design of effective inventory control policies for models with stochastic demands and forecast updates that evolve dynamically over time is a fundamental problem in supply chain management. In particular, this has been a very challenging theoretical and practical problem, even for models with a very simple forecast update mechanism. We describe new algorithms that were initially motivated by a theoretical analysis in terms of worst-case performance, and present extensive computational results that demonstrate their superior empirical performance compared to previously known policies.

Most of the existing literature has focused on characterizing the structure of optimal policies. For many of these inventory models, it is well known that there exists an optimal *state-dependent base-stock policy* (Zipkin 2000). In contrast, there has been relatively little progress on how to compute good inventory policies for models with complex demand structures. In particular, finding an optimal base-stock policy is usually computationally intractable. As a result, in most practical situations the default policy has been to use a *Myopic policy*, which computes its decision at the beginning of each period by minimizing the expected cost for the current period, and ignores all future costs. The Myopic policy is attractive since it can be computed efficiently even in complex environments with forecast updates. There are certain settings in which the Myopic policy is even optimal (Veinott 1963, 1965a, Ignall and Veinott 1969, Iida and Zipkin 2001, Lu et al. 2006). However, as was pointed in Levi et al. (2007), it performs very poorly in many important scenarios, such as in settings in which the demand is highly variable.

In recent work, Levi et al. (2007) have introduced *Dual-Balancing policies* for periodic-review, single-item, single-location models with backlogged demands. These policies incorporate several nontraditional ideas. First, they are based on *marginal cost accounting schemes.* Traditional cost accounting schemes associate with a decision in a certain period only those costs that are incurred in that period (or more generally, a lead time ahead); in contrast, a marginal cost accounting scheme associates with each decision *all* costs that are incurred as a result in this and subsequent periods, and are unaffected by any future decision. Secondly, these policies use *cost-balancing* techniques, which balance the following two opposing costs in each period: the conditional expected marginal *holding cost* incurred by maintaining excess inventory due to over-ordering; and the conditional expected *backlogging cost* incurred by not satisfying demand on time due to under-ordering. These policies can be easily implemented and efficiently computed under very general assumptions, including models with dynamic forecast updates. Moreover, it can be shown that Dual-Balancing policies have a *worst-case* performance guarantee of two. That is, the expected cost of the Dual-Balancing policy is guaranteed to be at most twice the optimal expected cost. In several subsequent papers (Levi et al. 2004, 2005b,a), the Dual-Balancing policy and its worst-case analysis have been extended to more general stochastic inventory models.

This paper focuses on the classical uncapacitated *periodic-review stochastic inventory control problem* with nonstationary, correlated and evolving demands; it builds on the results of Levi et al. (2007) and extends them in several important directions. While we do not claim to have done an exhaustive empirical study of all possibilities, we have set up a rigorous set of tests under which we consider two classes of policies. Specifically, we consider *balancing policies* that are based on cost-balancing techniques and *myopic-like* base-stock policies. The reason that we focus attention on these classes of policies is that they can be computed in an on-line manner, that is, the decision in each certain period does not depend on the decisions in future periods. This seems to be an essential property for a policy to be computationally applicable in the presence of dynamic forecast updates. Motivated by our preliminary computational results, we derive new policies based on these two approaches, and we demonstrate them to be superior to previously known approaches across a wide spectrum of demand scenarios in which forecasts evolve over time.

The new algorithmic ideas are based on exploiting computationally tractable upper and lower bounds on the optimal base-stock levels in each period, in order to refine the costs being balanced, as well as to correct the resulting cost-balancing-driven ordering decision. The first idea is that we start by computing the ordering quantity based on the Dual-Balancing policy; however, if the resulting inventory level is lower than the lower bound or greater than the upper bound, we appropriately correct the balancing order quantity to be within the range provided by the upper and lower bounds; we call this the *Interval-Constrained-Balancing policy.* Of course, this idea can also be applied to improve other policies, not just the Dual-Balancing one.

However, we can also use the bounding information in a more subtle way. We can instead balance only the conditional expected marginal holding cost of units ordered beyond the lower bound against the conditional surplus marginal backlogging costs incurred by not ordering up to the upper bound. Observe that the optimal policy orders more units than the respective lower bound. Thus, any holding cost incurred by units ordered up to the lower bound is also incurred by the optimal policy. Similarly, the optimal policy does not order above the respective upper bound, thus, it incurs at least as much backlogging cost as a policy that orders up to that upper bound. This modified balancing procedure ignores certain costs provided it is guaranteed that the optimal policy incurs them as well. These two algorithmic ideas are combined to create the *Surplus-Balancing policies*. The lower and upper bounds that we use are based on myopic-like base-stock polices that can be efficiently computed. All of the new balancing policies can be shown to have a worst-case performance guarantee of two. More importantly, the computational experiments that we conduct indicate that combining cost-balancing techniques and myopic-like base-stock policies leads to significantly more effective policies than using either of the two approaches separately.

Finally, the policies that we consider can be easily extended by introducing parameterized variants; for example, one might not compute an exact balance of the holding and backlogging costs, but instead compute a different set proportion that each should attain. Furthermore, we construct policies that dynamically compute these parameters over time.

As we have already mentioned, the focus of this paper is to investigate the empirical and theoretical performance of different policies in environments with dynamic forecast updates. We chose to perform most of the computational experiments using the *martingale model of forecast evolution* (MMFE) with multiplicative updates as introduced independently by Heath and Jackson (1994) and Graves et al. (1986). This model is very flexible and can capture many different scenarios of evolving forecasts and many other relevant aspects such as auto-correlation and variability. The MMFE maintains a vector of forecasts of future demands in each period. This vector can be viewed as the estimated means of the future demands as seen from the current period. Then we observe a random vector of updates. We generate a new forecast vector for the next period by computing the component-wise product of the initial forecast vector and the vector of updates. We note that the MMFE model has also a variant in which the updates are additive. The additive model is easier mathematically, but is less realistic, since it can lead to negative demand.

The additive variant has been studied by Iida and Zipkin (2001) and by Lu et al. (2006). They have obtained necessary and sufficient conditions for the optimality of the Myopic policy, and proposed several heuristics (see also the work of Dong and Lee (2003)). Several optimization algorithms and heuristics have been proposed for other demand structures such as *exogenous Markov modulated demand* (Song and Zipkin 1993, Chen and Song 2001, Gavireni and Tayur 2001) and *advance demand information* (Özer and Gallego 2001). However, to the best of our knowledge none of these heuristics is implementable in the multiplicative MMFE model. Moreover, even computing good lower bounds on the optimal cost seems to be computationally challenging in the multiplicative MMFE model.

This work is the first extensive computational study of inventory control policies within this important model. The results of our computational experiments provide a strong indication that the typical performance of the new policies is significantly better than the worst-case performance guarantee of two. Moreover, these policies appear to be robust and perform relatively well across a broad set of important scenarios. In particular, they out perform the Myopic policy in almost all scenarios, and the improvement can be as high as 30 percent. The computational results also demonstrate the computational practicality of the new policies in realistic scenarios.

As already mentioned it is prohibitively hard to compute optimal policies or even good lower bounds on the optimal cost for the multiplicative MMFE models. Thus, it is hard to get an accurate empirical estimate of how the new policies perform compared to an optimal policy. To get another indication of the empirical performance of the new policies, we also tested them in a simpler model in which the optimal policy and cost can be computed (see Section 8). The results validate again the robustness of the new balancing policies in a variety of relevant scenarios.

The rest of the paper is organized as follows. In Section 2, we define the general inventory model that is discussed in this paper. In Section 3, we briefly describe the previous work on Dual-Balancing and myopic-like base-stock policies. Then in Section 4, we describe the new policies that we construct and establish several important properties of their performance. In Section 5, we provide the details of the multiplicative MMFE model. Section 6, describes the computational experiments that we conducted in the MMFE model. This is followed by Section 7, in which we present a summary of the computational results and related conclusions. Finally, in Section 8 we discuss the computational comparison of the new policies with an optimal policy in a much simpler demand model that we call the customer retention demand model.

2. Model Definition

In this section, we provide the mathematical formulation of the periodic-review stochastic inventory problem and introduce some of the notation used throughout the paper. As a general convention, we distinguish between a random variable and its realization using capital letters and lower case letters, respectively. Script font is used to denote sets. We consider a finite planning horizon of T periods numbered $t = 1, \ldots, T$ (note that t and T are both deterministic unlike the convention above). The demands over these periods are random variables, denoted by $D_1, \ldots, D_T$.

As part of the model, we assume that at the beginning of each period s, we are given what we call an *information set* that is denoted by f_s. The information set f_s contains all of the information that is available at the beginning of time period s. More specifically, the information set f_s consists of the realized demands $(d_1, \ldots, d_{s-1})$ over the interval $[1, s)$, and possibly some more (external) information denoted by $(w_1, \ldots, w_s)$. The information set f_s in period s is one specific realization in the set of all possible realizations of the random vector $F_s = (D_1, \ldots, D_{s-1}, W_1, \ldots, W_s)$. This set is denoted by $\mathcal{F}_s$. In addition, we assume that in each period s, there is a known conditional joint distribution of the future demands $(D_s, \ldots, D_T)$, denoted by $I_s := I_s(f_s)$, which is determined by f_s (i.e., knowing f_s, we also know $I_s(f_s)$). For ease of notation, D_t will always denote the random demand in period t conditioning on some information set $f_s \in \mathcal{F}_s$ for some $s \leq t$, where it will be clear from the context to which period s we refer. We will use t as the general index for time, and s will always refer to the current period.

The only assumption on the demands is that for each $s = 1, \ldots, T$, and each $f_s \in \mathcal{F}_s$, the conditional expectation $E[D_t | f_s]$ is well defined and finite for each period $t \geq s$. In particular, we allow non-stationarity and correlation between the demands in different periods. We note again that by allowing correlation we let I_s be dependent on the realization of the demands over the periods $1, \ldots, s-1$ and possibly on some other information that becomes available by time s (i.e., I_s is a function of f_s). However, the information set f_s as well as the conditional joint distribution I_s are assumed to be independent of the specific inventory control policy being considered.

All the costs are linear consisting of time-dependent per-unit ordering cost c_t, per-unit holding cost h_t and per-unit backlogging penalty cost p_t. Unsatisfied demand is fully backlogged. Each order placed in period t arrives and becomes available only after a lead time of L periods. We also assume that the cost parameters are *non-speculative*. This is a typical assumption that can be captured through the conditions $c_t \leq c_{t-1} + h_{t+L-1}$ and $c_t \leq c_{t+1} + p_{t+L}$, for each $t = 2, \ldots, T - L$. It is well known (see for example Levi et al. (2007)) that under these conditions, we can assume that, for each $t = 1, \ldots, T$, we have $c_t = 0$, $h_t \geq 0$ and $p_t \geq 0$, without loss of generality. (Note that since the cost parameters are time-dependent we can also incorporate a discount factor $0 < \alpha < 1$ and salvage cost at the end of the planning horizon.) The goal is to find an ordering policy that minimizes the overall expected ordering cost, holding cost and backlogging cost. We consider only policies that are *non-anticipatory*, i.e., at time s, the information that a feasible policy can use consists only of f_s and the current inventory level.

We use $D_{[s,t]}$ to denote the accumulated demand over the interval $[s, t]$, i.e., $D_{[s,t]} := \sum_{j=s}^{t} D_j$. Superscripts P and OPT are used to refer to a given policy P and the optimal policy, respectively. We also use NI_t to denote the *net inventory* at the end of period t, and X_t, Y_t to denote the *inventory position* at the beginning of period t *before* and *after* ordering, respectively. In particular,

$X_t := NI_{t-1} + \sum_{j=t-L}^{t-1} Q_j$ (for $t = 1, \ldots, T$) and $Y_t = X_t + Q_t$, where Q_j denotes the number of units ordered in period j. (We sometimes denote $\sum_{j=t-L}^{t-1} Q_j$ by $Q_{[t-L,t-1]}$.) Note that once we know the policy P and the information set $f_s \in \mathcal{F}_s$, the quantities ni_{s-1}^P, x_s^P and y_s^P are deterministic. (These are the realizations of NI_{s-1}^P, X_s^P and Y_s^P, respectively.)

3. Base-Stock and Dual-Balancing Policies

As already mentioned, our new policies are based both on traditional base-stock policies and on the new algorithmic approach of Dual-Balancing policies introduced by Levi et al. (2007). We next describe the main underlying ideas of these approaches to provide the necessary background for the next section, in which we describe and analyze several new policies that extend the traditional and the balancing ideas in rather significant ways.

It is a well-known fact that, for the model discussed in this paper, there is a *state-dependent base-stock policy* which is optimal (see Zipkin (2000), Levi et al. (2007) for a detailed discussion). A state-dependent base-stock policy can be described by a set of target inventory levels $\{R_t(f_t) : t = 1, \ldots, T,\ f_t \in \mathcal{F}_t\}$, where $R_t(f_t)$ is the target inventory level in period t given that the observed information set is f_t. (It is important to note that $R_t(f_t)$ does not depend on the control policy up to time t.) An optimal base-stock policy can be computed recursively by solving a dynamic program. Unfortunately, in many important scenarios, it is computationally intractable to solve the corresponding dynamic program since its state space explodes. (We refer the reader to Levi et al. (2007) for a detailed discussion.) Due to the apparent difficulty of computing an optimal base-stock policy, researchers have proposed suboptimal policies that can be computed efficiently.

3.1. Myopic Policy

One specific class of suboptimal policies that has attracted a lot of attention is the class of *myopic policies.* In a myopic policy, in each period, we attempt to minimize the expected cost in a single period, a lead time ahead, ignoring the potential effect on the cost in future periods. This gives rise to a *Myopic base-stock policy.* For each period t, given the observed information set f_t, let $R_t^{MY}(f_t)$ be corresponding myopic base-stock level. That is,

$$R_t^{MY}(f_t) = \arg\min_{y \geq 0} E\left[h_{t+L}(y - (D_{[t,t+L]})^+ + p_{t+L}(D_{[t,t+L]} - y)^+ | f_t\right],$$

where $(x)^+ = x$ if $x \geq 0$ and equals 0 otherwise.

The Myopic policy is attractive since it yields a base-stock policy that is easy to compute on-line, that is, it does not require information on the control policy in the future periods. Specifically, in each period, we need to solve a relatively simple single-variable convex minimization problem. Because of its simplicity, the Myopic policy is commonly used in practice.

In many cases, the Myopic policy seems to perform well and even be optimal (for details see Veinott (1965b), Ignall and Veinott (1969), Iida and Zipkin (2001), Lu et al. (2006)). However, in many other cases, especially when the demand can drop significantly from period to period, the Myopic policy performs poorly. In particular, Levi et al. (2007) have shown that the Myopic policy can be arbitrarily more expensive than the optimal policy, even if the demands in different periods are independent of each other.

It is a well-known fact (see, for example, Zipkin (2000), Levi et al. (2007)) that the myopic base-stock levels are always higher than the optimal base-stock levels. That is, for each period t and information set $f_t \in \mathcal{F}_t$, we have $R_t^{OPT}(f_t) \leq R_t^{MY}(f_t)$, where $R_t^{OPT}(f_t)$ denotes the corresponding optimal base-stock level. (We note that Lu et al. (2006) have used myopic base-stock levels to develop additional upper and lower bounds on the optimal base-stock levels.)

3.2. Dual-Balancing Policies

Levi et al. (2007) have proposed a new algorithmic approach for uncapacitated single-item, single-location stochastic inventory control models that is very different from the traditional dynamic-programming-based approach. In particular, they have proposed a new class of policies that are called *Dual-Balancing* policies.

Their approach is based on two main ideas. First, they propose a new way to account for the cost in uncapacitated stochastic inventory models, which is called *marginal cost accounting.* The main idea underlying this approach is to account for *all* the expected costs associated with the decision of how many units to order in period t when this decision is made. More specifically, the decision in period t is associated with all the expected costs that, after that decision is made, become independent of any future decision, and are only dependent on future demands. In Levi et al. (2007) it has been shown that in uncapacitated models, these costs are relatively easy to compute already in period t, even though they include costs that are going to be incurred in future periods.

For each feasible policy P, let H_t^P be the holding cost incurred over the interval $[t,T]$ by the Q_t^P units ordered in period t (for $t=1,\dots,T$), and let Π_t^P be the backlogging cost incurred a lead time ahead in period $t+L$ ($t=1-L,\dots,T-L$). That is, $H_t^P=\sum_{j=t+L}^{T} h_j(Q_t^P - (D_{[t,j]}-X_t^P)^+)^+$ and $\Pi_t^P := p_{t+L}(D_{[t,t+L]}-(X_{t+L}^P+Q_t^P))^+$ (where $D_j := d_j$ with probability 1 and $Q_j^P = q_j$ is given as an input for each $j \le 0$). Let the $\mathcal{C}(P)$ be the effective cost of the policy P. Levi et al. (2007) have shown that the effective cost of policy P can be expressed as

$$\mathcal{C}(P) := \sum_{t=1}^{T-L} (H_t^P + \Pi_t^P). \tag{1}$$

The second idea is the use of *cost balancing* techniques. In the Dual-Balancing policy, which is denoted by superscript B, in each period s, conditioned on the the observed information set f_s, the following two opposing costs are balanced

$$l_s^B(q_s^B) = E[H_s^B(q_s^B)|f_s] \tag{2}$$

and

$$\pi_s^B(q_s^B) = E[\Pi_s^B(q_s^B)|f_s]. \tag{3}$$

That is, we order $q_s^B = q_s'$ to make $l_s^B(q_s') = E[H_s^B(q_s')|f_s] = \pi_s^B(q_s') = E[\Pi_s^B(q_s')|f_s]$.

Note that the Dual-Balancing policy can also be computed in an on-line manner, i.e., the ordering decision in period s does not depend on any future decision. Moreover, in most of the common scenarios, there exist efficient procedures for evaluating the functions l_s^B and π_s^B defined above. Since l_s^B is a monotone increasing function of q_s^B and π_s^B is a monotone decreasing function of q_s^B, the balancer q_s' above is relatively easy to compute. As a result the Dual-Balancing policy is easy to implement both conceptually and computationally (see Levi et al. (2007) for a more detailed discussion of the computational aspects).

Levi et al. (2007) have shown that the Dual-Balancing policy has a *worst-case performance guarantee* of 2. That is, for each instance of the problem, the expected cost of the Dual-Balancing policy is guaranteed to be at most twice the expected cost of an optimal policy. However, this is merely a worst-case analysis, and their paper does not explore the typical performance of the policy.

Finally, unlike base-stock policies, the order up to level of the Dual-Balancing policy does depend on the inventory position at the beginning of the period, i.e., it depends on x_s^B.

3.3. Minimizing Policy

Based on their marginal cost accounting approach, Levi et al. (2007) have also described a new base-stock policy that is called a *Minimizing policy* and is denoted by the superscript M.

In each period s, conditioned on the observed information set f_s, we again consider the functions $l_s^M(q_s)$ and $\pi_s^M(q_s)$ defined above. However, instead of ordering to balance these two quantities, the Minimizing policy orders q_s^M to minimize the sum of these two functions. Specifically, $q_s^M = \arg\min_{q_s \geq 0}[l_s^M(q_s) + \pi_s^M(q_t)]$.

Levi et al. (2007) have shown that the Minimizing policy is in fact a base-stock policy, and that the minimizing base-stock levels always provide a lower bound on the corresponding optimal base-stock levels. That is, for each t and $f_t \in \mathcal{F}_t$, $R_t^M(f_t) \leq R_t^{OPT}(f_t)$. It is readily verified that the Minimizing policy can be easily computed in an on-line manner. Thus, the minimizing and the myopic policies provide respective lower and upper bounds on the optimal base-stock levels that can be computed efficiently.

4. New Policies: Description and Performance Analysis

In this section, we present several new policies for the periodic-review stochastic inventory control model, and establish several important and interesting theoretical results about their performance. All of the new policies described in this section are based on intuitive and conceptually simple ideas. Moreover, they can be computed efficiently in an on-line manner; thus, they can be implemented in rather straightforward ways. In Section 7, we shall also present extensive computational results which indicate that these new policies have significantly better typical performance in many important scenarios. The new policies are based on several new ideas. First, we show how to incorporate cost-balancing techniques together with lower and upper bounds on the optimal base-stock levels. Secondly, we use parametrization to enrich and refine the class of policies being used.

4.1. Bounded Cost-Balancing Techniques

We have already seen that, for each period t and information set $f_t \in \mathcal{F}_t$, the optimal base-stock level is bounded between the respective minimizing and myopic base-stock levels. That is, $R_t^M(f_t) \leq R_t^{OPT}(f_t) \leq R_t^{MY}(f_t)$. Next we shall discuss two approaches that use these lower and upper bounds to modify and improve the cost-balancing techniques.

Interval-Constrained Bounding. First we show how to use the bounds of the Myopic and Minimizing policies to construct an *interval-constrained bounding procedure* that can be applied to any feasible policy. Each feasible policy P can be described by specifying its order up-to level, for each possible state (t, f_t, x_t), where again $t = 1, \ldots, T$ is the period, $f_t \in \mathcal{F}_t$ is some observed information set and x_t is the inventory position at the beginning of the period. In each time period in which the inventory position of P after ordering falls outside the respective interval specified by the minimizing and the myopic base-stock levels, the interval-constrained bounding procedure modifies the policy P. Specifically, if for some (t, f_t, x_t) the resulting inventory position after ordering is smaller than the minimizing base-stock level $R_t^M(f_t)$, then the inventory position is augmented up to$R_t^M(f_t)$ (i.e., $y_t = R_t^M(f_t)$) by appropriately increasing the order quantity; if on the other hand the resulting inventory position is higher than the myopic base-stock level $R_t^{MY}(f_t)$, then the inventory position is truncated by decreasing the ordering quantity until $y_t = R_t^{MY}(f_t)$ or $q_t = 0$ (i.e., $y_t = x_t$). In Appendix D, we discuss the effect of the interval-bounding procedure on various policies.

Applying this procedure to the Dual-Balancing policy leads to what we call the *Interval-Constrained-Balancing policy* and denote by superscript ICB. It turns out that the Interval-Constrained-Balancing policy also has a worst-case guarantee of two; moreover, our computational experiments indicate that its typical performance is better than the Dual-Balancing policy. Since the performance of this policy is outperformed by other (new) policies (see below), we discuss its worst-case analysis in Appendix A, and its typical performance in Appendix C.

THEOREM 1. *The Interval-Constrained-Balancing policy has a worst-case guarantee of two.*

Surplus-Balancing. Next we describe a different and more subtle approach to combine cost-balancing techniques with lower and upper bounds on the optimal base-stock levels. We call this approach *surplus-balancing.* This approach gives rise to a general class of policies, and present a general worst-case analysis. We denote this class by superscript SB.

We begin by introducing several conventions and techniques similar to the ones used by Muharremoglu and Tsitsiklis (2001) (see also Levi et al. (2007) for more details). The main idea is that, without loss of generality, we can assume that units of supply are consumed by the demand on a first-ordered-first-consumed basis, and that we can *match* each unit of supply to the specific unit of demand it will be used to satisfy. More rigorously, for each positive k, we identify the $k-th$ *supply unit* as the $k-th$ unit that will be purchased; we also identify the $k-th$ *unit of demand* as the $k-th$ unit that will be demanded. Without loss of generality, we assume that supply units are used on a first-ordered-first-used basis. Thus, the $k-th$ unit of supply is used to satisfy the $k-th$ unit of demand. If the inventory is measured by discrete quantities, then k is an integer. If fractional orders are allowed, then k is a real number, and the corresponding supply and demand units are infinitesimal. Moreover, we can then describe each policy P in terms of the periods in which it orders each supply unit, where all unordered units are "ordered" in period $T+1$. Since the demand is independent from the inventory policy, we can compare any two feasible policies by comparing the respective periods in which each supply unit was ordered. Our exposition of the Surplus-Balancing policies and the worst-case analysis are based on this idea.

Assume that we are given a set of lower bounds $\{L_t(f_t):\ t=1,\ldots,T,\ f_t\in\mathcal{F}_t\}$ and a set of upper bounds $\{U_t(f_t):\ t=1,\ldots,T,\ f_t\in\mathcal{F}_t\}$, such that for each f_t, we have $L_t(f_t)\le R_t^{OPT}(f_t)\le U_t(f_t)$. ($L_t(f_t)$ should not be confused with the lead time L. We will sometimes use $L_t(F_t)$ and $U_t(F_t)$ as random objects depending on the random information set in period t.) For each $t=1,\ldots,T-L$, let $Q_t^L=(L_t(F_t)-X_s^{SB})^+$ be difference between the corresponding lower bound $L_t(F_t)$ and the inventory position of the Surplus-Balancing policy at the beginning of period t, or zero if it exceeds the lower bound. Similarly, let $Q_t^U=(U_t(F_t)-X_s^{SB})^+$ be the difference between the corresponding upper bound $U_t(F_t)$ and the inventory position of the Surplus-Balancing policy, or zero if it exceeds and the upper bound. Note that conditioned on f_t the quantities O_t^L and Q_t^U are known deterministically.

For each period s and information set $f_s\in\mathcal{F}_s$, recall the functions $l_s^{SB}(q_s^{SB})=E[H_s^{SB}(q_s^{SB})|f_s]$, the conditional marginal expected holding costs of the q_s^{SB} units ordered in period s, and $\pi_s^{SB}(q_s^{SB})=E[\Pi_s^{SB}(q_s^{SB})|f_s]$, the conditional expected backlogging cost in period $s+L$ given that we order q_s^{SB} additional units in period s. (See (2) and (3) above.) Instead of balancing l_s^{SB} against π_s^{SB} like the Dual-Balancing policy, we now balance $(l_s^{SB}(q_s^{SB})-l_s^{SB}(q_s^L))^+$ against $(\pi_s^{SB}(q_s^{SB})-\pi_s^{SB}(q_s^U))^+$. The quantity $(l_s^{SB}(q_s^{SB})-l_s^{SB}(q_s^L))^+$ is equal to the conditional expected marginal holding costs incurred by all the units ordered in period s, expect the q_s^L units that were required to raise the inventory position of the Surplus-Balancing policy up to the lower bound $L_s(f_s)$. This implies that in the cost-balancing, we ignore the holding costs associated with these q_s^L units. Intuitively, we ignore these costs because we know these q_s^L units were ordered by OPT in period s or even earlier. (Observe that if $x_s^{SB}\ge L_s(f_s)$, i.e., the inventory position of the Surplus-Balancing policy at the beginning of period s exceeds the lower bound, then $q_s^L=0$ and $l_s^{SB}(q_s^L)=0$.) The quantity $(\pi_s^{SB}(q_s^{SB})-\pi_s^{SB}(q_s^U))^+$ is equal to the conditional expected *additional* backlogging cost incurred by the Surplus-Balancing policy in period $s+L$ due to not ordering up to the upper bound $U_s(f_s)$. The intuitive reason why we consider only this part of the backlogging costs is that the optimal policy's order up to level is lower than $U_s(f_s)$. Thus, if OPT can reach the optimal base-stock level $R_s^{OPT}(f_s)$, the backlogging cost it incurs in period $s+L$ is at least as high as the respective backlogging cost incurred by a policy, which orders up to the upper bound $U_s(f_s)$.

The Surplus-Balancing policy orders q'_s such that $(l_s^{SB}(q'_s) - l_s^{SB}(q_s^L))^+ = (\pi_s^{SB}(q'_s) - \pi_s^{SB}(q_s^U))^+$. Now $(l_s^{SB}(q_s^{SB}) - l_s^{SB}(q_s^L))^+$ is zero for $q_s^{SB} \le q_s^L$ and increasing to infinity as q_s^{SB} grows to infinity, and $(\pi_s^{SB}(q_s^{SB}) - \pi_s^{SB}(q_s^U))^+$ is non-negative, decreasing and equal to 0 for $q_s^{SB} \ge q_s^U$. It follows that q'_s is well defined and that $q_s^L \le q'_s \le q_s^U$. That is, in period s the Surplus-Balancing policy always orders at least up to the corresponding lower bound $L_s(f_s)$, and never exceeds the corresponding upper bound $U_s(f_s)$, while placing a positive order. In the next theorem we show that the Surplus-Balancing policy has a worst-case guarantee of two.

THEOREM 2. *The Surplus-Balancing policy has a worst-case performance guarantee of two. That is, $E[\mathcal{C}(SB)] \le 2E[\mathcal{C}(OPT)]$.*

W e shall prove the worst-case guarantee by comparing the cost of the Surplus-Balancing policy to the cost of an infeasible policy denoted by OPT' that has expected cost lower than OPT. The policy OPT' is a base-stock policy with the same base-stock levels as OPT. However, if for some period s and information set f_s the inventory position of OPT' at the beginning of the period is higher than the corresponding upper bound $U_s(f_s)$ and also higher than the inventory position of the Surplus-Balancing policy x_s^{SB}, it is allowed to scrap inventory with no cost to bring its inventory position down to $\max\{U_s(f_s), x_s^{SB}\}$. (Observe that since the upper bounds are on the optimal base-stock levels and not on the actual inventory position, it is indeed possible that the inventory position of OPT at the beginning of a period is higher than the corresponding upper bound. For example, this can happen if the upper bound $U_s(f_s)$ is smaller than $U_{s-1}(f_{s-1})$.) It is straightforward to verify that the expected cost of OPT' is lower than that of OPT. Note that OPT' never scraps units that were already ordered by the Surplus-Balancing policy in the current period or in previous periods. (The scrapping is bounded from below by $\max\{U_s(f_s), x_s^{SB}\}$.) This follows from the fact that when it scraps inventory it can never go below the inventory position of the Surplus-Balancing policy in that period.

In the first step of the analysis, we express the expected cost of the Surplus-Balancing policy using (1) above. Let H_s^{SB} be the marginal holding cost incurred by the Q'_s units ordered by the Surplus-Balancing policy in period s over the entire horizon $[s+L, T]$. Let H_s^L be the holding costs incurred by the $Q_s^L = (L_s(F_s) - X_s^{SB})^+$ units required to raise the inventory position of the Surplus-Balancing policy to $L_s(F_s)$, over the entire horizon. We have already seen that $Q_s^L \le Q'_s$, which implies that $H_s^L \le H_s^{SB}$ since H_s^L captures the holding cost of only some of the units ordered in period s. (If $X_s^{SB} \ge L_s(F_s)$ then $Q_s^L = 0$ and $H_s^L = 0$.) Similarly, let Π_s^{SB} be the backlogging cost incurred by the Surplus-Balancing policy in period $s+L$. In addition, let Π_s^U be the backlogging cost incurred in period $s+L$ by a policy that orders up to $\max\{U_s(F_s), X_s^{SB}\}$. Observe that if $X_s^{SB} \ge U_s(F_s)$ then the Surplus-Balancing policy does not order and $\Pi_s^{SB} = \Pi_s^U$. On the other hand if $X_s^{SB} < U_s(F_s)$, the Surplus-Balancing will order up to at most $U_s(F_s)$, and then $\Pi_s^{SB} \ge \Pi_s^U$. We will call Π_s^U the *minimal backlogging costs* of period s. We can express the expected cost of the Surplus-Balancing policy as

$$\begin{aligned} E[\mathcal{C}(SB)] &= E\left[\sum_{s=1}^{T-L} \left(H_s^{SB} + \Pi_s^{SB}\right)\right] \\ &= E\left[\sum_{s=1}^{T-L} \left(H_s^L + (H_s^{SB} - H_s^L) + \Pi_s^U + (\Pi_s^{SB} - \Pi_s^U)\right)\right]. \end{aligned} \tag{4}$$

For each $s = 1, \ldots, T-L$, let $Z_s = E[H_s^{SB} - H_s^L | F_s] = E[\Pi_s^{SB} - \Pi_s^U | F_s]$. Note that second equality follows from the construction of the Surplus-Balancing policy. Moreover, Z_s is a random variable that is observed at the beginning of period s with the observed information set f_s. Using (4) above, this implies that

$$E[\mathcal{C}(SB)] = \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + \sum_{s=1}^{T-L} E[E[(H_s^{SB} - H_s^L) + (\Pi_s^{SB} - \Pi_s^U)|F_s]] \tag{5}$$
$$= \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + 2\sum_{s=1}^{T-L} E[Z_s].$$

In the second step of the analysis, we show how to amortize the cost of the Surplus-Balancing policy against the cost of OPT'. In particular, we shall show that in expectation at least half of the cost of the Surplus-Balancing policy can be amortized against the cost of OPT'.

Next we partition the periods based on a comparison between the inventory positions of OPT' and the Surplus-Balancing policy. Let $\mathcal{T}_H$ be the set of periods in which the inventory position of OPT' after ordering is no lower than the respective inventory position of the Surplus-Balancing policy. That is $\mathcal{T}_H = \{s:\ Y_s^{SB} \le Y_s^{OPT'}\}$. Let $\mathcal{T}_\Pi$ be the complement set of $\mathcal{T}_H$, i.e., $\mathcal{T}_\Pi = \{t:\ Y_s^{SB} > Y_s^{OPT'}\}$.

In the remainder of the proof we shall show how to amortize the cost of the Surplus-Balancing policy against the cost of OPT'. In particular, we shall show that, in expectation, the cost of OPT' can be used to amortize at least half of the cost of the Surplus-Balancing policy. Specifically, we shall show that $E[\mathcal{C}(OPT')] \ge \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + \sum_{s=1}^{T-L} E[Z_s]$. This and (5) above establish the proof of the theorem.

Let $H^{OPT'}$ be the overall holding costs incurred by OPT'. We claim that these holding costs are higher than the holding costs incurred by units ordered by the Surplus-Balancing policy in periods $s \in \mathcal{T}_H$ and the units ordered in periods $s \in \mathcal{T}_\Pi$ to raise the inventory position of the Surplus-Balancing policy to the corresponding lower bound $L_s(f_s)$. That is, for each complete information set $f_T \in \mathcal{F}_T$ (recall that $f_1 \subset f_2 \subset \cdots \subset f_T$),

$$H^{OPT'} \ge \sum_{s\in\mathcal{T}_\Pi} H_s^L + \sum_{s\in\mathcal{T}_H} H_s^{SB} = \sum_s H_s^L + \sum_{s\in\mathcal{T}_H} (H_s^{SB} - H_s^L). \tag{6}$$

Consider a realization of a complete information set f_T and some period $s \in \mathcal{T}_H$. By definition $y_s^{OPT'} \ge y_s^{SB}$. This implies that the q_s' units ordered by the Surplus-Balancing in period s were ordered by OPT' in period s or even earlier. It follows that the holding cost these units have incurred under OPT' are higher than the respective holding cost they incurred under the Surplus-Balancing policy. Similarly, in each period $s \in \mathcal{T}_\Pi$, we have $y_s^{OPT'} \ge R_s^{OPT}(f_s) \ge L_s(f_s)$. We conclude that the q_s^L units ordered by the Surplus-Balancing policy in period s to raise its inventory position up to $L_s(f_s)$ were ordered by OPT' in period s or even earlier. The proof of (6) is then complete.

Now let $\Pi^{OPT'}$ be the overall backlogging costs incurred by OPT'. We claim that these backlogging costs are higher than the backlogging costs associated with periods $s \in \mathcal{T}_\Pi$ plus the minimal backlogging costs of periods $s \in \mathcal{T}_H$. That is, for each complete information set $f_T \in \mathcal{F}_T$,

$$\Pi^{OPT'} \ge \sum_{s\in\mathcal{T}_H} \Pi_s^U + \sum_{s\in\mathcal{T}_\Pi} \Pi_s^{SB} = \sum_s \Pi_s^U + \sum_{s\in\mathcal{T}_\Pi} (\Pi_s^{SB} - \Pi_s^U). \tag{7}$$

Consider a realization of a complete information set f_T and some period $s \in \mathcal{T}_\Pi$. By definition we know that $y_s^{SB} > y_s^{OPT'}$. Hence, the backlogging cost incurred by OPT' in period $s+L$ is higher than the respective backlogging cost incurred by the Surplus-Balancing policy in that period. For each period $s \in \mathcal{T}_H$, if $y_s^{OPT'} \le U_s(f_s)$ then it is clear that the backlogging cost incurred by OPT' in period $s+L$ are higher than Π_s^U. On the other hand, if $y_s^{OPT'} > U_s(f_s)$, it must be the case that $y_s^{OPT'} = x_s^{SB} = y_s^{SB}$. (The policy OPT' scraps units as long is it inventory position is above $U_s(f_s)$ and x_s^{SB}.) We have already seen that in this case $\Pi_s^U = 0$ and the proof of (7) above follows.

From (6) and (7) it follows that

$$H^{OPT'} + \Pi^{OPT'} \geq \sum_{s=1}^{T-L} (H_s^L + \Pi_s^U) + \sum_{s \in \mathcal{T}_H} (H_s^{SB} - H_s^L) + \sum_{s \in \mathcal{T}_\Pi} (\Pi_s^{SB} - \Pi_s^U). \tag{8}$$

Taking expectations, we see that this implies that

$$\begin{aligned} E[\mathcal{C}(OPT)] &= E[H^{OPT'} + \Pi^{OPT'}] \\ &\geq \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + \sum_{s=1}^{T-L} E\left[1\!\!1(s \in \mathcal{T}_H) \cdot (H_s^{SB} - H_s^L) + 1\!\!1(s \in \mathcal{T}_\Pi) \cdot (\Pi_s^{SB} - \Pi_s^U) \right] \\ &= \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + \sum_{s=1}^{T-L} E\left[E\left[1\!\!1(s \in \mathcal{T}_H) \cdot (H_s^{SB} - H_s^L) + 1\!\!1(s \in \mathcal{T}_\Pi) \cdot (\Pi_s^{SB} - \Pi_s^U) | F_s \right] \right] \\ &= \sum_{s=1}^{T-L} E[H_s^L + \Pi_s^U] + \sum_{s=1}^{T-L} E[Z_s]. \end{aligned} \tag{9}$$

In the second in equality we use a standard conditional expectation argument. The third equality follows from the fact that conditioning on the information set $f_s \in \mathcal{F}_s$, the indicator functions $1\!\!1(s \in \mathcal{T}_H)$ and $1\!\!1(s \in \mathcal{T}_\Pi)$ are known deterministically, and from the definition of Z_s. The proof of the theorem then follows.

Theorem 2 above generalizes the Dual-Balancing policy proposed by Levi et al. (2007). In this case we take the lower bounds $L_s(f_s) = 0$ and upper bounds $U_s(f_s) = \infty$. If one instead uses the base-stock levels of the minimizing policy $\{R_t^M(f_t) : \ t = 1, \ldots, T, \ f_t \in \mathcal{F}_t\}$ as lower bounds with upper bounds $U_s(f_s) = \infty$, we get a Surplus-Balancing policy that has a worst-case performance guarantee of two. By arguments similar to the proof of Theorem 1 it can shown that applying the idea of interval-constrained-bounding described above to the latter Surplus-Balancing policy using the myopic based-stock levels preserve the worst-case guarantee of two. We call this policy the *Truncated Surplus-Balancing Policy* and denote it by TSB. We note that it is possible to use the myopic base-stock levels $\{R_t^{MY}(f_t) : \ t = 1, \ldots, T, \ f_t \in \mathcal{F}_t\}$ as upper bounds in conjunction with the lower bounds of the minimizing policy to get yet another Surplus-Balancing policy with a worst-case guarantee of two. We call this policy the *Pure Surplus-Balancing policy* and denote it by PSB. As we report in Sections 7 and 8 the typical performance of these Surplus-Balancing policies outperform that of the Dual-Balancing policy, the Myopic policy and the Minimizing policy.

4.2. Extended Class of Myopic Policies

For each period t and observed information set f_t, we again define $l_t^P(q_t)$ to be the conditional expected holding costs incurred by the units ordered by policy P over the rest of the horizon $[t+L, T]$. That is, $l_t^P(q_t) = E[H_t^P(q_t)|f_t]$. We have used this function in constructing the minimizing and balancing polices discussed above. Suppose that instead of looking to the end of the horizon, in each period t, we consider the conditional expected holding cost over only the next k periods, for $1 \leq k \leq T - L - t + 1$. That is, we consider the conditional expected holding costs of the units ordered in period t that are incurred over the interval $[t + L, t + L - 1 + k]$.

More generally, the value of k needs not be restricted to be an integer. To count the marginal holding cost over k periods into the future we define H_{tk}^P

$$= \left(\sum_{j=t+L}^{t+L+\lfloor k \rfloor - 1} h_j (Q_t - (D_{[t,j]} - X_t)^+)^+ \right) + (k - \lfloor k \rfloor) \left\{ h_{t+L+\lceil k \rceil} (Q_t - (D_{[t,t+L+\lceil k \rceil]} - X_t)^+)^+ \right\},$$

where the floor function $\lfloor k \rfloor$ is the greatest integer less than or equal to k and the ceiling function $\lceil k \rceil$ is the smallest integer greater than or equal to k. This defines a continuum of random variables parameterized by k. Next define, for each information set f_t, the function $l^P_{tk}(q_t) = E[H^P_{tk}(q_t)|f_t]$. Recall that the minimizing policy computes its ordering quantity, in each period t, by minimizing the conditional expected backlogging costs in period $t+L$, denoted by $\pi^M_t(q_t)$, plus the conditional expected holding costs of the units ordered in period t over the entire horizon denoted by $l^M_t(q_t)$. More generally, we let $M(k)$ denote the following *Minimizing-k policy*: in each period t, it attempts to minimize the conditional expected backlogging costs in period $t+L$ plus the conditional expected holding costs that the units ordered in period t incur over $[t+L, t+L+k-1]$. (We assume that no holding costs are incurred beyond period T, which allows $t+L+k-1$ possibly be larger than T.) That is, the order quantity of the policy $M(k)$ in period t, denoted by $q^{M(k)}_t$ is computed as $q^{M(k)}_t = \arg\min_{q_t \geq 0}[l^{M(k)}_{tk}(q_t) + \pi^{M(k)}_t(q_t)]$. By arguments similar to those used by Levi et al. (2007) regarding the Minimizing policy, one can show that, for each $1 \leq k \leq T-t-L+1$, the policy $M(k)$ is in fact a state-dependent base-stock policy. The base-stock level $R^{M(k)}_t(f_t)$ of the Minimizing-k policy $M(k)$ in period t can be computed as the minimizer of $l^{M(k)}_{tk}(q_t) + \pi_t(q_t)$, assuming that the inventory position at the beginning of the period is 0.

As the next lemma shows, these base-stock levels are decreasing in k (the proof can be found in Appendix A).

Lemma 1. *The base-stock levels of the minimizing-k policies are decreasing in k. That is, for each $k_1 \geq k_2$, we have $R^{M(k_1)}_t \leq R^{M(k_2)}_t$.*

Note that $M(1)$ is the Myopic policy, i.e. $MY = M(1)$. Also, if one chooses k_t dynamically over time, such that $k_t = T-t-L+1$, we get the Minimizing policy M. Thus, for each $t = 1, \ldots, T$ and $f_t \in \mathcal{F}_t$, we have that

$$R^M_t(f_t) = R^{M(T-t-L+1)}_t(f_t) \leq R^{M(T-L)}_t(f_t) \leq \cdots \leq R^{M(2)}_t(f_t) \leq R^{M(1)}_t(f_t) = R^{MY}_t(f_t), \qquad (10)$$

and this induces a parameterized family of myopic-like base-stock policies over the space $[R^M_t(f_t), R^{MY}_t(f_t)]$.

While using a static k (i.e., the same k) in all periods may give a good policy, it is natural to try to think of dynamic methods of choosing k.

Run-Out Time The *run-out time* measures how long a unit stays in the system from the moment it arrives (i.e., becomes available to us) until the moment it is consumed. Assume that the y-th unit was ordered at the beginning of period t. Under the assumption that units are consumed on a first-ordered-first-consumed basis, let $T_t(y)$ be the number of periods from t until the first y units are fully consumed. Then

$$T_t(y) = \sum_{j=t}^{T} \mathbb{1}(y - D_{[t,j]} > 0).$$

Conditioned on the observed information set f_t, we define the *conditional expected post lead-time run-out time* of the y^{th} unit (where again $y > 0$) by $r_t(y) = E\left[(T_t(y) - L)^+ | f_t\right]$. (Note that $r_t(y)$ is always defined with respect to some information set f_t.)

Next we consider two different levels of inventory $0 \leq y_1 \leq y_2$ and denote the difference between their respective expected post-lead-time run-out times by $r_t([y_1, y_2]) = r_t(y_2) - r_t(y_1)$.

We considered and tested several methods for choosing k dynamically. In all of these methods, in each period t, conditioning on the observed information set f_t, we compute k_t, the number of periods we look ahead at time period t, and the resulting base-stock level $R^{M(k_t)}_t(f_t)$. All of these methods compute k_t as a function of the post-lead-time run-out times of different units. Next we describe the three methods:

(i) *Final unit run-out, denoted by M(k-fin).* We consider the run-out time of the final unit being ordered in period t. Specifically, we compute k_t that solves:

$$k_t = r_t(y^{M(k_t)}_t(f_t)),$$

where $y_t^{M(k_t)}$ is the inventory position after ordering, following the $M(k_t)$ policy. This is a circular computation because the last unit ordered in $R_t^{M(k_t)}(f_t)$ depends on the policy $M(k_t)$ in use, which is a function of k_t. So at the start of period t we set $k_t^0 = 1$, (the Myopic policy). If $r_t(y_t^{M(1)}) < 1$, then we follow the Myopic policy, and if $r_t(y_t^{M(T-t)}) = r_t(y_t^M) > T - t$ we follow the minimizing policy. Otherwise, we iteratively compute $k_t^{i+1} = r_t(R_t^{M(k_t^i)}(f_t))$ for increasing iteration indices i. The iterations stop when k_t^i converges, and which time the last of the k_t^i's becomes k_t.

(ii) *Average marginal units run-out, denoted by M(k-mar).* Under this procedure, we look only on the *marginal* average post-lead-time run-out time. That is, we consider the average marginal increase in the run-out time caused by the units ordered in the period. Let x_t be again the inventory position at the beginning of period t, and $q_t^{M(k)} = (R_t^{M(k)}(f_t) - x_t)^+$ be the order quantity in period t if the base-stock policy $M(k)$ is followed. In particular, consider only values of k for which $q_t^{M(k)} > 0$. (If $q_t^{M(1)} = 0$, i.e., the Myopic policy does not order, then order nothing.) Using an iterative procedure similar to (i) above, compute the k_t that solves

$$k_t = \frac{r_t([x_t + 1, y_t^{M(k_t)}])}{q_t^{M(k_t)}}.$$

(iii) *Average total units run-out, denoted by M(k-tot).* Consider the average post-lead-time run-out time of the total inventory position after ordering following an $M(k)$ policy. That is, compute the k_t that solves

$$k_t = \frac{r_t([0, y_t^{M(k_t)}])}{y_t^{M(k_t)}}$$

In Lemma 2 in Appendix A we show that the three methods (i)-(iii) are well defined (i.e., that the procedures described above do converge).

4.3. Parameterized Balancing Policies

In each period t, conditioned on the observed information set f_t, the Dual-Balancing policy described orders q_t' to balance $l_t(q_t') = \pi_t(q_t')$, i.e., to make $E[H^B(q_t')|f_t] = E[\Pi_t^B(q_t')|f_t]$. However, more generally, the order in period t can be chosen to balance the backlogging and holding costs in a different ratio than 1. For each period t and given some information set f_t, let $q_t'(\beta)$ be the order quantity that makes

$$l_t(q_t'(\beta)) = E[H_t^B(q_t'(\beta))|f_t] = \beta\pi_t(q_t'(\beta)) = \beta E[\Pi_t^B(q_t'(\beta))|f_t],$$

where β is some positive number that denotes the desired *balancing ratio*. Clearly, this leads to a rich continuum of balancing policies $B(\beta)$ parameterized by β. Specifically, for $\beta = 1$, we get the original Dual-Balancing policy.

As with the $M(k)$ family of policies, we consider policies based on both fixed balancing ratios of β, and on a dynamic method that chooses different balancing ratios β_t, in each period t. The dynamic method that we consider chooses β_t according to the Myopic policy. Specifically, we set β_t to be the ratio between the conditional expected holding costs and the conditional expected backlogging costs $E[\Pi_t^{MY}|f_t]$ incurred by the Myopic policy in period $t + L$. We denote this policy by $B(\beta\text{-myo})$.

4.4. Summary of Policies

We summarize the policies studied and their short-hand names in Table 1

Table 1 Policies studied

Policy Name	Description
MY or $M(1)$	Myopic Policy
B	Balancing Policy
ICB	Interval-Constrained Balancing Policy
TSB & PSB	Truncated/Pure Surplus-Balancing Policy
$B(\beta)$	Balancing Policy that seeks $q_t'(\beta)$ such that $l_t(q_t'(\beta)) = \beta\pi_t(q_t'(\beta))$
$B(\beta$-myo)	Chooses β_t equal to the ratio of expected holding and backlogging costs of Myopic
M	Minimizing Policy
$M(k)$	Minimizing Policy with holding cost look-ahead of k, for fixed k
$M(k$-fin)	$M(k)$ with $k = k$-fin, the post lead-time run out of the final unit ordered
$M(k$-mar)	$M(k)$ with $k = k$-mar, the average post lead-time run out of marginal units ordered
$M(k$-tot)	$M(k)$ with $k = k$-tot, the average post lead-time run out of current inventory position

5. The Martingale Model of Forecast Evolution

In describing policies for the stochastic inventory problem, we have purposely kept the description of the underlying stochastic process quite general. For computational evaluation, however, we select a specific mechanism. In this section, we describe the model of demand and forecast evolution that we use in the computational study. Specifically, we incorporate forecasting by using the martingale model of forecast evolution (MMFE) as introduced independently by Heath and Jackson (1994) and Graves et al. (1986). The multiplicative variant of the MMFE provides a rich simulation environment for evaluating the performance of diverse policies. We begin by summarizing the multiplicative MMFE model and interpreting its inputs. We then present an approximation method for computing the distribution of the cumulative demand.

For $s \leq t$, let D_{st} be the forecast of the demand in period t, as made at the end of time period s. Thus $D_t = D_{tt}$ is the actual demand in period t. We let $\mathbf{D_{t-1}} = (D_{t-1,t}, D_{t-1,t+1}, \ldots, D_{t-1,T+L})$ be the forecast vector available at the start of period t. (The vector $\mathbf{D_{t-1}}$ corresponds to F_t in Section 2 above. If the current time is t or later we write $\mathbf{d_{t-1}} = (d_{t-1,t}, d_{t-1,t+1}, \ldots, d_{t-1,T+L})$, to indicate that the forecasts are now deterministic.)

The essential assumption of the MMFE is that the forecasts D_{st} evolve as a martingale, i.e., at the end of period s conditioning on $f_{s+1} = \mathbf{d_s}$, we have $d_{st} = E\left[D_t | f_{s+1}\right]$. More generally, $D_{st} = E[D_t|F_{s+1}]$.

The two variants of the MMFE are the *additive* and the *multiplicative* MMFE. In both variants, one of the inputs to the process is an initial forecast vector $\mathbf{d_0} = (d_{01}, d_{02}, \ldots, d_{0,T+L})$. In this paper, we shall use the multiplicative MMFE. Apart from a study done by Heath and Jackson (1994), to the best of our knowledge all other computational studies have used the additive MMFE (see Iida and Zipkin (2001), Lu et al. (2006) for details).

We choose the multiplicative MMFE over the additive version for two reasons. First, for realistic choices of parameters, there is a significant probability that the additive MMFE will give negative demand values; the multiplicative MMFE never does. Secondly, in our experience industry forecasts tend to be updated in a relative sense (as done by the multiplicative MMFE) rather than an absolute sense (as done by the additive MMFE).

Multiplicative MMFE. In the multiplicative MMFE, at the beginning of each period t, we generate an update vector, $\gamma_\mathbf{t} = e^{\epsilon_\mathbf{t}}$, where $\epsilon_\mathbf{t}$ is a multivariate normal random variable with variance-covariance matrix Σ_t and mean $-\frac{\text{diag}(\Sigma_T)}{2}$, where $\text{diag}(\Sigma_t)$ is the vector of diagonal elements of Σ_t. Thus, the random vector $\gamma_\mathbf{t}$ has multivariate lognormal distribution with mean $\mathbf{1} = (1, 1, \ldots, 1)$.(Note that here we deviate from our convention and use lower-case letters to denote random variables.) Writing it component-wise, we have $\gamma_\mathbf{t} = (\gamma_{t,t}, \gamma_{t,t+1}, \gamma_{t,T+L}) = (e^{\epsilon_{t,t}}, e^{\epsilon_{t,t+1}}, \ldots, e^{\epsilon_{t,T+L}})$

Then, at the start of period t we will have the forecast vector, $\mathbf{D_{t-1}}$. For $t=1,2,\ldots,T+L$, the time-t demand will be given by $D_t = \gamma_{t,t} D_{t-1,t}$ and the new forecast vector $\mathbf{D_t}$ will be given by $\mathbf{D_t} = s(\gamma_\mathbf{t}) s(\mathbf{D_{t-1}})$, where $s(\cdot)$ denotes the shift operator: $s((x_1,x_2,x_3,\ldots,x_n)) = (x_2,x_3,\ldots,x_n)$, and where the multiplication is component-wise. The evolution of forecasts and demands is initiated by $\mathbf{D_0} = \mathbf{d_0} = (d_{01}, d_{02}, \ldots,\ d_{0,T+L})$.

The covariance of ϵ_{st} with $\epsilon_{s,t+1}$ is given by $\sigma_{t-s+1,t-s+2}$, an off-diagonal element of Σ. If this element is positive, then before time s, the forecasts D_{st} and $D_{s,t+1}$ will be positively correlated, and the demands D_t and D_{t+1} will be positively correlated. Such correlations exist when, for example, good news causes forecasts for demand in several periods to be revised upwards, or bad news causes forecasts for demand in several periods to be revised downward. A negative correlation arises, for example, when a large forecasted demand is shifted earlier or later in time.

Cumulative Demand Distribution. All of the policies considered, including the Myopic policy (for non-zero leadtimes), need to compute expectations involving the cumulative demand over intervals. In the multiplicative MMFE model the corresponding cumulative distributions consist of the sum of correlated lognormal random variables. There is no closed form expression for the distribution of a sum of lognormal random variables, let alone the sum of correlated lognormal random variables. However, the problem arises in finance (Milevsky and Posner 1998)) and wireless technology (Abu-Dayya and Beaulieu 1994, Beaulieu et al. 1995)). Abu-Dayya and Beaulieu (1994) consider three approximations, and demonstrate that the Wilkinson's method described by Schwartz and Yeh (1982) is the best among the three. As described by Abu-Dayya and Beaulieu (1994), the key ideas in Wilkinson's method are firstly that the sum of lognormal random variables $L = e^{Y_1} + e^{Y_2} + \cdots + e^{Y_n}$ is well approximated by a single lognormal random variable ($L \approx e^Z$, where Z is a normal random variable), and secondly, that we can match the first and second moments of the sum to obtain the appropriate parameters for e^Z. In an appendix to this paper, we present the results of various tests of this approximation conducted in the setting of demand forecasting. Our experiment results indicate that these approximations are very good and the analytical expressions match the simulation.

6. Experimental Design

The space of potential parameter settings for this study is very large. In addition to parameters describing the inventory system, there are many parameters that describe the manner in which forecasts of demand evolve over time. A fully comprehensive study is beyond the scope of this paper. Our goal is to study a broad range of potential application settings, with emphasis on the demand and forecasting processes. The experimental design is oriented around a Base Case and six sets of *scenarios*, each of which expands the Base Case in an interesting dimension. In each set of scenarios we vary specific input parameters. The first three of these scenario sets study *first-order effects*; here, it is the initial forecast $\mathbf{d_0}$ that varies. The final three scenario sets study *second order effects* by varying the variance-covariance matrix Σ in different ways.

We begin this section by discussing the parameters of the Base Case. After that we describe the manner in which the parameters of the Base Case are varied, in each of the six scenario sets.

The Base Case. In all of the experiments, we let our holding and backorder costs per unit per period be stationary and assume values $h_t = 1$ and $p_t = 10$ for all t. As noted in Section 4, we take $c_t = 0$ for all t without loss of generality.

We consider a horizon of length $T = 40$. All experiments are conducted for two different values of the lead-time: $L = 0$ and $L = 4$. Therefore, to facilitate comparison between results, costs are not counted during the first four time periods. Note that when $L = 4$, in the first four time periods the costs incurred are determined by decisions made in the past, and are not influenced by our choice of policy.

The initial demand forecast is flat, with $\mathbf{d_0} = (400, 400, \ldots, 400)$. The horizon over which the user generates forecasts is of length 12, as in monthly forecasts for a year. This implies that we learn nothing about the period-t demand until we are within 12 months of period t, i.e., $D_{t-12,t} = d_{0,t} = 400$ for $t > 12$. Algebraically, recall from Section 5 that the standard multiplicative model updates forecasts using the formula $D_{st} = \gamma_{st} D_{s-1,t}$. The assumption is that for $t > s + 11$ we have $\gamma_{st} = 1$. This implies that at all times s, the first 13 elements of the forecast vector $\boldsymbol{D_s}$ will be different from each other, but the 13-th element and every subsequent element will be equal to 400.

Recall from Section 5 that in the multiplicative MMFE, the period-t update vector is $\gamma_{\mathbf{t}} = e^{\epsilon_{\mathbf{t}}}$, where $\epsilon_{\mathbf{t}}$ is a $T - t + 1$ - dimensional random vector with variance-covariance matrix Σ_t and mean $-\frac{1}{2}\text{diag}(\Sigma_t)$. We obtain the $(T - t + 1) \times (T - t + 1)$ matrix Σ_t from Σ_{t-1}, by dropping the last row and column. The previous paragraph implies that in our experiments, forecast evolution and demand are driven by a 12×12 covariance matrix Σ. We obtain the $T \times T$ matrix Σ_1 from Σ by appending $T - 12$ extra rows and columns to Σ, with 1's on the diagonal and 0's elsewhere. Therefore, for $t > 12$, ϵ_t is a degenerate random variable with mean 0 and variance 0.

In the Base Case, we have *constant learning*, meaning that all of the entries on the diagonal of Σ are equal. (This implies that the variability of the future demand is resolved at a constant pace.) The diagonal elements are selected so that for $t \geq 12$, the coefficient of variation of the demand D_t, seen from the beginning of time period 1, is 0.75. A formula for the coefficient of variation is provided in the description of the Coefficient of Variation Scenarios, below.

The off-diagonal entries of the covariance matrix Σ determine the degree of correlation between the updates that are observed in a given time period, say, time period s. The Base Case assumes that there is some correlation between these updates, modeled by having non-zero, positive values in the first off-diagonal of Σ. Consequently, in the Base Case, if the forecast for the demand in month t goes up in period s (i.e., if $D_{st} > D_{s-1,t}$), then the forecast for demand in month $t + 1$ is likely to increase in period s as well (if $t + 1 \leq s + 11$), but this does not tell us anything about the forecast for demand in month $t + 2$. The values of the non-zero off-diagonal elements are chosen to give a correlation coefficient of 0.5 for each pair of adjacent forecast updates. That is, for each s and each t, $s \leq t \leq s + 10$, the update factors γ_{st} and $\gamma_{s,t+1}$ observed in period s have correlation coefficient 0.5, but γ_{st} and $\gamma_{s,t+2}$ are stochastically independent.

Product Launch Scenarios. In this set of scenarios we study the effect of rising demand, as might be encountered at a product launch. Again, only the initial forecast vector $\mathbf{d_0}$ is varied. For comparison with the base case, we ensure that the mean of the values in $\mathbf{d_0}$ is 400. We consider upward demand trends of +5, +10 and +20 per period. In addition, we consider two examples in which the demand rises in a steeper, non-linear manner, mid-way through the horizon; these are generated using an appropriately scaled normal CDF curve.

End-of-Life Scenarios. Here, we study scenarios associated with products that are in an end of life situation, namely those with decreasing initial forecast vectors. Essentially, these are the reverse of the Product Launch scenarios; we have initial forecast vectors with forecasted demand decreasing by $5, 10$ and 20 per period. We also consider two products whose demands have steeper drop-off curves, generated using the normal complementary CDF curve. In addition, we study a total demand crash, in which the demand is forecast to crash to 0 midway through the time horizon.

Seasonality Scenarios. In the seasonality study, we use the common base-values described above for all parameters except for the initial forecast vector $\mathbf{d_0}$. We conduct experiments with two forms of seasonality, one defined via a sinusoidal function and the other via a step function. In both cases, the maximum value attained is 700 and the minimum is 100. This allows us to compare results with the base case more easily, because the mean of the entries in the initial forecast vector is 400 in all cases.

By the *cycle length*, we mean the number of time periods between two consecutive high-points. We consider cycle lengths with values 2, 4 and 8. For example, for the step-function with period 4, we have $\mathbf{d_0} = (700, 700, 100, 100, 700, 700, 100, 100, \ldots)$.

The above scenario sets test the effect of varying $\mathbf{d_0}$, the initial forecast vector. In the final three scenario sets, we focus instead on varying Σ. In all of these, we take $\mathbf{d_0} = (400, 400, \ldots, 400)$.

Coefficient of Variation Scenarios. In this scenario set, we study the effect of varying the magnitude of the variance in the demands and the forecasts. Note that for $t \geq 12$, at the end of time period $t-12$, we have $D_{t,t} = \Gamma_t \; d_{t-12,t}$, where Γ_t is random and has the same distribution as $\Gamma = \Pi_{i=1}^{12} \gamma_i = \exp\left(\sum_{i=1}^{12} \epsilon_i\right)$. The ϵ_i's are independent normal random variables, with mean such that $E[e^{\epsilon_i}] = 1$, and with variance σ_{ii}, the i-th diagonal element of Σ, our forecast update matrix (note that σ_{ii} is a variance, not a standard deviation). Thus, the mean of Γ is one and the variance is $\exp\left(\sum_{i=1}^{12} \sigma_{ii}\right) - 1$. The coefficient of variation of Γ is given by $(\exp\left(\sum_{i=1}^{12} \sigma_{ii}\right) - 1)^{1/2}$, and is equal to 0.75 in the Base Case. In the scenarios where we investigate the effect of variance, we scale the entries of Σ such that the coefficient of variation of this series of twelve updates takes specific values, namely 0.5, 0.7, 1, 2, 4, and 8. This corresponds to different levels of variability in the demands.

Time of Learning Scenarios. Note that the ratio of the sum of the first j diagonal entries of Σ, to the sum of all the diagonal entries, is the fraction of variability in the future demands that is *unresolved* in period $s = t - j$. When all the entries in the diagonal are identical then the variance of each update is the same. This corresponds to what we call *constant learning.* When the values in diag(Σ) are weighted towards the end of the vector, then the unresolved uncertainty is low when j is small (s is close to t). This corresponds to *early learning.* Conversely, when the values in diag(Σ) are weighted towards the beginning of the vector, then this corresponds to *late learning*; that is, most of the uncertainty about the true value of D_t is only resolved in periods s that are close to t. We also consider the setting in which there is more weight in the center of diag(Σ) than at the endpoints. Here, we learn most in the middle of the forecast horizon.

We construct variance-covariance matrices Σ to correspond with these four cases: constant, early, late and mid-horizon learning. In all cases, the values of Σ are scaled to ensure that the coefficient of variation of Γ, and of D_t for $t \geq 12$, remains constant at 0.75.

Correlation Scenarios. In this scenario set we test the effect of different types of correlation between the updates. We vary correlation in two ways. First, we set the number of non-zero off-diagonals of our 12x12 matrix, Σ, to 0 (which corresponds to no correlation), 1, 4 and 8. Secondly, the sign of the off-diagonal elements can be all positive, all negative, or entries alternating between positive and negative. (The base case corresponds to 1 off-diagonal with non-zero elements which are all positive.) As in the base case, the diagonal of Σ corresponds to the constant learning case, and the coefficient of variation of Γ is 0.75.

Table 2 summarizes the scenarios we study. The number of scenarios for each set is given in brackets after the set name; we see that there are 38 in total. We run each of these with $L = 0, 4$ for an overall total of 76 scenario - lead time pairs. For each of the scenarios, we ran $N = 1000$ independent trials for a horizon of length $T = 40$. For the scenarios with a lead time of $L = 4$, our decisions only influenced costs from periods 5 through 40. Therefore, in order to compare costs on an even footing, we consistently computed the total holding and backlogging costs excluding the first 4 periods.

Performance Measures Used. For a fixed policy π, we let $\mathcal{C}_i(\pi)$ denote the cost of the i-th run ($i = 1, \ldots, 1000$), excluding the first 4 periods. Note that since we consider a complex environment and relatively long horizon ($T = 40$), it is not tractable to compute or even evaluate the optimal expected cost. Instead, we use two performance measures of a policy's effectiveness. Both measures are computed relative to the performance of our benchmark, the Myopic policy, MY. The first is the *relative total cost*, given by

$$AT(\pi) = \left(1 - \left(\frac{\sum_{i=1}^{N} \mathcal{C}_i(\pi)}{\sum_{i=1}^{N} \mathcal{C}_i(MY)}\right)\right) * 100\%$$

Table 2 Scenario codes

Topic	Code	Description
Product Launch (5)	$+I$	Increment by I per period, $I \in \{5, 10, 20\}$
	Curve	Increasing scaled normal CDF curve
	S. Curve	Steeper increasing scaled normal CDF curve
End-of-Life (6)	$-I$	Decrement by I per period, $I \in \{5, 10, 20\}$
	Curve	Decreasing scaled normal CDF curve
	S. Curve	Steeper decreasing scaled normal CDF curve
	Crash	Demand crash
Seasonal (7)	Base Case	Initial forecast vector is flat
	Sin(n)	Sinusoidal periodicity with cycle length n, $n \in \{2, 4, 8\}$
	Step(n)	Step-function periodicity with cycle length n
Coeff. of Var. (6)	CV = β	Coefficient of variation equals $\beta = 0.5, 0.7, 1, 2, 4, 8$
Learning Rate (4)	Const	Constant learning
	Late	Late learning
	Early	Early learning
	Mid	Mid-horizon learning
Correlation (10)	None	All off-diagonal elements of Σ are 0
	Pos(n)	First n off-diagonals of Σ have positive entries, $n \in \{1, 4, 8\}$
	Neg(n)	First n off-diagonals of Σ have negative entries
	Mix(n)	First n off-diagonals of Σ have entries alternating positive and negative

whereas the second is the *average relative cost per run*, which is given by

$$AR(\pi) = \left(1 - \left(\frac{1}{N}\sum_{i=1}^{N} \frac{\mathcal{C}_i(\pi)}{\mathcal{C}_i(MY)}\right)\right) * 100\%.$$

Note that both $AT(\pi)$ and $AR(\pi)$ can be positive or negative. If they are positive this implies that they improve upon the myopic policy, and a higher value indicates higher improvement. Conversely, if they are negative, this implies that myopic performs better. (Thus, in the tables given in Section 7 to come, positive numbers indicate better relative performance with respect to the Myopic policy.)

For each run, we also compute a lower bound on the costs to provide an additional reference. Recall that the order-up-to level of the Minimizing policy is always below that of the optimal, whereas that of the Myopic is always above the optimal. Thus, the sum of the holding cost of the Minimizing Policy and the backorder cost of the Myopic Policy gives a lower bound on the cost of the optimal policy. If we imagine that this is the cost of a policy π, we then compute $AT(\pi)$ and $AR(\pi)$ in the manner shown above, but with respect to this lower bound 'policy'. This statistic is denoted LB and is an *upper* bound on the potential relative improvements over the Myopic policy that can be further achieved. For example, in Table 3 in Section 7 we can see that the TSB policy improves upon the Myopic by 0.13%, and the lower bound LB improves by 4.01%. This implies that the TSB policy is within less than 3.88% of optimal. We note that the lower bound that we use is likely to be very loose in many scenarios, especially when the variability in the demand is high. Thus, it might be hard to estimate accurately how far from optimal do the new policies perform. However, to the best of our knowledge there are no other known lower bounds on the optimal cost in the multiplicative MMFE model.

7. Experimental Results

In this section, we present the results of the computational investigation of the average performance of the policies in the multiplicative MMFE model. We demonstrate that three policies in particular, the Truncated Surplus-Balancing policy (denoted TSB), the Balancing-2 policy $B(2)$ which has $\beta = 2$, and the Minimizing policy $M(k\text{-tot})$ which chooses k equal to the average run-out time of all the units present in the system, exhibit superior performance. They achieve an average cost that is up to 30% lower than that of the Myopic policy (our benchmark), they out-perform the Myopic policy in almost every scenario, and they are never much worse than the best performing policy in any scenario (see Table 11). (We note that the Pure Surplus-Balancing Policy was dominated by the Truncated Surplus-Balancing, and this is the reason why we do not report on it in this section.)

The greatest improvements over the Myopic policy occur in contexts where steep demand drops can occur. These contexts include end-of-life scenarios (Tables 4 and 5), seasonality (Table 6), and systems with highly variable demands and forecasts (Table 7). Long lead times make the improvements over Myopic more dramatic. Moreover, comparing to the lower bound, we learn that the average performance of the new policies is significantly better than the worst-case guarantee of two. (This is although the lower bound that we use, is sometimes very loose.)

We note that the standard Balancing and Minimizing policies are greatly improved by the various refinements introduced in this paper. The most universally applicable of these refinements is the Interval-Constrained-Bounding concept presented in Section 4. At the end of this section, we discuss bounding and its effect; otherwise, all of results presented in this section reflect the improvements due to bounding.

Near the end of Section 6, we defined two performance measures, the relative total cost $AT(\pi)$ (which places more weight on randomly generated problem instances in which the total costs are higher), and the average relative cost per run $AR(\pi)$. We prefer $AR(\pi)$ because it weights all problem instances equally. As our accompanying technical report Hurley et al. (2006) indicates, the two measures usually tell similar stories. However $AR(\pi)$ is usually 0-2% higher than $AT(\pi)$. This is because the randomly generated scenarios in which the total costs are highest, are usually ones in which the demand grows unexpectedly, and in these scenarios the Myopic policy performs somewhat better relative to other policies. The strongest exception to the 0-2% rule is the demand crash scenario shown in Table 5 below, for which we report both measures; also see the robustness study below (Table 11). In this scenario, the problem instances with the greatest costs are the ones in which the Myopic policy dramatically over-stocks. These problem instances have a disproportionate impact on $AT(\pi)$, and favor policies that are not myopic.

In our accompanying technical report Hurley et al. (2006), we report the performance of all the policies in Table 1 under each scenario in Table 2, measured by both $AR(\pi)$ and $AT(\pi)$. In this section, we report a subset of those results for the three most successful policies highlighted above, as well as the Balancing policy, B. For each table in this section, there is an accompanying table in Appendix C with a larger set of policies.

This section is organized as follows. First, we present computational results for the scenario sets defined in Section 6 in the following sequence: Product Launch, End of Life, Demand Crash, Seasonality, Coefficient of Variation, Learning, Correlation. We also study the robustness of the different heuristics over the 76 scenarios tested. Finally, we describe a model (the Customer Retention Model) under which we can compute precisely the expected cost of the optimal policy, as well as the Myopic, Minimizing and Surplus Balancing policies. In Appendix D we examine the effect of Interval Constrained-Bounding.

7.1. First Order Effects

Product Launch. In the Product Launch scenarios, the demand is trending upwards strongly. There is little risk of overstocking when this is the case, and hence, the performance of the Myopic policy should be at its peak. The Myopic policy is close to the lower bound when the lead time is short (see Table 3). There is at most a 4.01% improvement possible compared to the LB when

Table 3 Product Launch: $AR(\pi)$, for certain Product Launch scenarios
($AR(\pi)$ is the average percent improvement over Myopic, per run.)

	$L=0$				$L=4$			
Scenario	Flat	+20	Curve	S. Curve	Flat	+20	Curve	S. Curve
B	0.37%	0.34%	0.21%	0.47%	-2.58%	-3.05%	-2.67%	-2.26%
TSB	0.13%	0.11%	0.03%	0.10%	1.91%	1.24%	1.52%	1.79%
$B(2)$	0.46%	0.43%	0.31%	0.53%	1.52%	0.93%	1.25%	1.61%
$M(k\text{-tot})$	0.29%	0.26%	0.22%	0.29%	1.90%	1.32%	1.57%	1.82%
LB	4.01%	3.71%	3.67%	3.79%	25.92%	22.91%	23.18%	23.75%

Table 4 End Of Life: $AR(\pi)$, for certain End Of Life scenarios

	$L=0$				$L=4$			
Scenario	Flat	-20	Curve	S. Curve	Flat	-20	Curve	S. Curve
B	0.37%	0.53%	0.83%	0.66%	-2.58%	-1.65%	0.91%	6.12%
TSB	0.13%	0.16%	0.19%	0.35%	1.91%	3.14%	5.02%	9.93%
$B(2)$	0.46%	0.59%	0.75%	0.87%	1.52%	3.00%	5.25%	10.31%
$M(k\text{-tot})$	0.29%	0.36%	0.40%	0.56%	1.90%	3.15%	5.17%	9.36%
LB	4.01%	4.63%	4.55%	5.19%	25.92%	31.49%	36.14%	44.12%

$L=0$. Each of the new policies improves relative to the Myopic policy even in these scenarios, but the improvement is slight: less than 0.5% for $L=0$.

As the lead time increases, so does the gap between the Myopic policy and the lower bound. Over half of the new policies, especially the recommended policies (TSB, $B(2)$, and $M(k\text{-tot})$), show noticeable improvement (as high as 1.82%) over Myopic in these scenarios. A majority of the policies are markedly worse compared to the Myopic policy with longer lead times. This is a general pattern that is apparent in all scenarios. The most likely explanation for this pattern is that increased lead times magnify errors. For example, in the Base Case (the columns labelled "Flat" in Table 3), all policies except $M(2)$, $B(\beta\text{-myo})$ and the recommended policies, under-order on average when $L=0$, and do so more strongly when $L=4$. These are the policies whose performance deteriorates as L increases from 0 to 4. There are examples in other scenarios where over-ordering becomes more prevalent as the lead time increases.

End of Life Scenarios. The End of Life scenarios are like the Product Launch scenarios, except that the trend is for decreasing demand. In our experiments, the risk of overstocking when using the Myopic policy is low as long as the lead time is short. Table 4 demonstrates that the Myopic policy is close to the lower bound (within 5.5%) for all scenarios with $L=0$.

The results for long lead times ($L=4$) reveal a weakness in the Myopic policy. When the lead time is long and the demand decline is steep, then the new policies perform as much as 10% better than Myopic as can be seen in Table 4.

The Demand Crash Scenario. As could be expected, the new policies perform significantly better than Myopic under the demand crash scenario. In this scenario, improvements over the Myopic policy range between 10% and 20%. This improvement is due to the fact that the policies are much better at avoiding overstocking in the periods after the crash, namely periods 21 through to 40. We measure this by computing the total holding cost incurred by each policy in these periods over all 1000 runs, expressed as a percentage of the same cost incurred by the Myopic policy. We denote this measure of performance by HC. In Table 5 we report the values of $AT(\pi)$ and $AR(\pi)$ and HC for all policies. The fact that the new policies outperform the Myopic policy by up to 80% in periods after the crash is what makes them better overall.

Table 5 Demand Crash: Policy performance in the Demand Crash scenario (Both AT and AR measures are presented)

	$L=0$			$L=4$		
Policy	$AT(\pi)$	$AR(\pi)$	HC	$AT(\pi)$	$AR(\pi)$	HC
B	22.10%	13.12%	75.57%	18.31%	12.29%	64.37%
TSB	19.87%	12.39%	64.21%	20.93%	15.60%	65.32%
$B(2)$	20.19%	12.98%	65.27%	20.20%	16.43%	50.88%
$M(k\text{-tot})$	13.33%	9.39%	40.95%	16.93%	15.37%	41.96%
LB	31.35%	23.94%	-	55.33%	56.03%	-

Table 6 Seasonality: $AR(\pi)$, for certain Seasonality scenarios

	$L=0$				$L=4$			
Policy	Flat	Step(2)	Step(4)	Step(8)	Flat	Step(2)	Step(4)	Step(8)
B	0.37%	5.52%	4.83%	7.20%	-2.58%	-1.97%	-2.52%	2.71%
TSB	0.13%	2.22%	3.39%	4.80%	1.91%	3.01%	3.50%	7.44%
$B(2)$	0.46%	5.89%	5.22%	6.69%	1.52%	2.39%	2.83%	6.99%
$M(k\text{-tot})$	0.29%	4.19%	3.61%	4.02%	1.90%	3.08%	3.79%	7.14%
LB	4.01%	22.93%	20.29%	18.22%	25.92%	30.48%	34.38%	40.82%

Seasonality. Table 6 summarizes the results from the Seasonality scenarios. As the LB row of the table indicates, there is opportunity for improvement over the Myopic policy, particularly as the lead time increases.

On average, all of the new policies do better relative to the Myopic policy with longer cycle lengths (because the effects of over-stocking last longer, thus hurting the Myopic policy). A closer look reveals that for $L=4$, all of the new policies do much better than Myopic with a cycle length of 8 than they do with shorter cycle lengths. This is because the Myopic policy is less heavily affected by seasonality when the lead time is long enough to include at least one full cycle.

With regard to lead times, the recommended policies ($TSB, B(2)$ and $M(k$-tot)) exhibit a mixed, but fairly stable performance as the lead time grows from $L=0$ to 4. In marked contrast, the policies that are not recommended all suffer as the lead time increases and sometimes perform worse than Myopic. This is consistent with the general pattern discussed in the "Product Launch" scenarios above.

7.2. Second Order Effects

Finally, we consider briefly the remaining three sets of scenarios. The initial forecast vector in these scenarios is flat because the focus is on the investigation of second order effects (that is, the form of the variance-covariance matrix of the updates, Σ).

Coefficient of Variation. In Table 7 we report the average value of $AR(\pi)$ for all policies under the Coefficient of Variation scenarios. They demonstrate clearly that the new policies' performance improvement increases as the coefficient of variation increases. For highly variable forecast change (C.V.=8), the improvement over Myopic can be as high as 30%. This is caused by the fact that for larger coefficients of variation, the demand can fall quickly, resulting in over-stocking by the Myopic policy.

Learning Table 8 gives the average value of $AR(\pi)$, averaged over all policies, for $L=0$, under the various learning scenarios. Since the forecast horizon (12 months) is much longer than the lead time (0-4 months), we would expect that under early learning the Myopic policy would see approximately deterministic demand. For the same reason, late learning should favor the new

Table 7 Coefficient of Variation: $AR(\pi)$, for certain Coefficient of Variation scenarios ($L=0$)

C.V.	0.5	0.7	1	2	4	8
B	0.01%	0.23%	1.59%	9.74%	22.22%	26.84%
TSB	0.00%	0.06%	1.02%	9.65%	22.21%	29.15%
$B(2)$	0.01%	0.29%	1.94%	10.74%	22.26%	29.36%
$M(k\text{-tot})$	0.01%	0.16%	1.48%	10.07%	22.43%	30.17%
LB	0.39%	2.99%	10.27%	35.60%	57.68%	72.40%

Table 8 Time of Learning: Average of $AR(\pi)$ over all policies, for the Time of Learning scenarios ($L=0$)

Scenario	Early	Mid	Const	Late
Average	0.00%	0.04%	0.36%	1.10%

Table 9 Correlation: $AR(\pi)$ for the Correlation scenarios ($L=0$)

Scenario	None	Pos(4)	Neg(4)	Mix(4)	Pos(8)	Neg(8)	Mix(8)
B	1.00%	0.17%	0.93%	2.15%	0.07%	1.69%	2.17%
TSB	0.35%	0.02%	0.34%	0.49%	0.01%	0.54%	0.49%
$B(2)$	1.04%	0.21%	1.07%	2.05%	0.15%	1.61%	2.07%
$M(k\text{-tot})$	0.68%	0.13%	0.78%	1.37%	0.11%	1.07%	1.39%
LB	6.06%	2.73%	7.27%	8.85%	2.72%	7.45%	8.90%

Table 10 Number of scenarios in which each policy performs best (under the $AR(\pi)$ measure)

Policy	B	TSB	$B(0.5)$	$B(2)$	$B(\beta)$	M	$M(2)$	$M(3)$	$M(k\text{-fin})$	$M(k\text{-mar})$	$M(k\text{-tot})$
$L=0$	4	-	1	13	-	-	11	5	2	2	-
$L=4$	-	5	-	8	1	-	1	2	-	-	21

policies. This is born out by Table 8, but the gain is small. Otherwise, the patterns that we have seen in other scenario sets with respect to the performance of different policies and the effect of different lead times, hold here as well.

Correlation Table 9 shows the results for the Correlation scenarios. These show that under the scenarios where updates are positively correlated, the new policies are less of an improvement over the Myopic policy than in the Base Case (where there is no correlation). The improvement over the Base Case is about the same or slightly greater when the off-diagonal elements of Σ are all negative, and is greatest when the signs are mixed. None of the improvements is greater than about 2%. This is due to the fact that with positive correlation the demand is approximately constant. However, in the negative and mixed scenarios the demands are more choppy, and when both large and small demands are present the Myopic policy is more likely to overstock.

7.3. Other Aspects

Robustness It is useful to summarize the performance of the heuristics over the different scenarios. We approached this in two different ways. First, for all of the 37 scenarios in the End of Life and Seasonality categories, and for each of the lead time cases ($L=0$ and $L=4$), we compute the number of times each policy is the best, and report this in Table 10. It is clear that no one policy dominates, so we seek a policy that will be robust in many different settings.

For each of the 76 scenarios, we now compute the percentage by which the average relative cost of each policy exceeds that of the best performing policy for that scenario. Then, for each policy,

Table 11 Robustness statistics: % above the cost of the best heuristic, across the 76 scenarios
Results are for $AR(\pi)$ unless otherwise noted

Policy	Mean	Median	65th of 76	70th of 76	75rd of 76	Highest
B	2.06%	0.79%	4.94%	5.00%	5.95%	6.55%
TSB	0.51%	0.32%	1.04%	1.54%	3.32%	4.03%
TSB, $AT(\pi)$	0.16%	0.10%	0.26%	0.48%	1.54%	1.54%
$B(2)$	0.44%	0.09%	0.65%	0.84%	7.62%	9.65%
$B(2)$, $AT(\pi)$	1.10%	1.11%	1.49%	1.69%	1.72%	1.72%
$M(k\text{-tot})$	0.35%	0.14%	0.81%	1.06%	3.95%	4.44%
$M(k\text{-tot})$, $AT(\pi)$	0.34%	0.05%	0.29%	2.31%	5.07%	5.07%

we compute the mean of these 76 values, along with the median, and the 65th, 70th, 75rd and overall (76th) highest values. This information is reported in Table 11. Note that in this table all the entries are positive and that small numbers indicate a better and more robust performance.

In this study it matters whether we use the average total cost $AT(\pi)$ or our preferred measure, $AR(\pi)$. For the best policies, both measures are reported. Table 11 indicates that the Surplus-Balancing policy TSB is the most robust of the policies studied. Specifically, it never exceeds the cost of the best policy by more than 4.1 percent and, on average, is within 1 percent of the lowest cost. The second robust policy is $M(k\text{-}tot)$. The third policy we recommend, $B(2)$, is often stellar, but it is somewhat less robust than the other policies we recommend, as the last two columns indicate. $B(\beta\text{-myo})$ is another strong performer.

Bounding. We have noted that the optimal order-up-to level is bounded below and above by the Minimizing and Myopic order-up-to levels respectively. We consider briefly the gap between these order-up-to levels, where the gap is defined as $\frac{R^{MY}-R^{M}}{R^{M}}$. Over our 76 scenarios, the average value of this statistic is 4.59%. The quartiles are $\{0\%, 0.72\%, 3.03\%, 8.55\%, 18.72\%\}$. Both the minimum and maximum values come from Learning scenarios; the minimum (0%) from Early Learning with $L = 0$ and the maximum (18.72%) from Late Learning with $L = 4$. This can explain the notable improvement we observed while applying the interval-constrained-bounding procedure to the different policies. This is discussed further in Appendix D.

8. The Customer Retention Model

The worst-case analysis of the Surplus-Balancing policies was established under a very general model of demand and forecast evolution. The computational experiments in Sections 6 and 7 provide evidence about the performance of the different heuristic policies in the multiplicative MMFE model, which is a particular but still very general model. For such general models, it is not practical to compute optimal policies for benchmarking purposes. The state space that must be explored using dynamic programming is too large. In fact even tight lower bounds are very hard to obtain. In this section, we consider a simple but nontrivial model of demand evolution. This model admits practical computation of optimal policies and allows us to quantify more accurately the optimality gap of the Surplus-Balancing policies. We call this demand structure the *customer retention model.*

The details of the model are as follows. Let N_t denote the number of customers in period t who place regular orders for our product. We assume that a customer places a unit-sized order in each period. The demand in period t is, therefore, N_t. New customers are added at the rate of λ per period, according to a Poisson probability distribution, and each existing customer is retained with probability ρ. Given the number of customers in period t, the number of customers in period $t + 1$, is seen to be given by the sum of two independent random variables: a binomial

Table 12 Impact of Demand Rate on Average Optimality Gap of Three Policies

	Demand Rate				
Policy	0.01	0.04	0.07	0.1	Average
MY	105.99%	18.30%	4.10%	1.00%	32.35%
M	1.01%	36.93%	95.97%	92.77%	56.67%
PSB	5.52%	16.59%	30.41%	26.24%	19.69%

random variable with a Poisson random variable. That is, the distribution of N_{t+1} is given by $N_{t+1} \equiv B(N_t, \rho) + Poisson(\lambda)$.

Hence, a single state variable, N_t, is sufficient to represent the information set, f_t, at time t. If we further assume a zero lead time, then the complete state of the system at the beginning of period t, after receipt of the order placed in period $t-1$, is given by (N_t, x_t) , where x_t is the net inventory level. From this, it is straightforward to formulate a dynamic programming recursion to minimize the expected (undiscounted) holding and backlogging costs over a finite time horizon with zero salvage cost. Similarly, it is not difficult to recursively compute the expected cost of following any of the heuristic policies considered in this paper. We omit the details of these recursions for the sake of brevity.

In general, the Myopic policy performs quite well for this particular model of demand evolution. However, it is dramatically sub-optimal in circumstances in which the risk of carrying inventory for a long time is high. For example, if $\lambda = 0.01$, $\rho = 0.1$, $h = 1$, $p = 10$, $T = 100$ and $x_0 = 0$, then the optimal expected cost is 11.1 and the expected Myopic policy cost is 42.4. This represents an optimality gap of 282%. One can see how this comes about. Over a horizon of 100 days, at least one new customer is likely to arrive. The probability of retaining that customer is 10% and so the expected shortage cost next period if no order is placed exceeds the expected one period holding cost of ordering one unit. The Myopic policy will order up to one unit in such a situation. With high probability, however, the customer will not be retained and the unit ordered by the Myopic policy will remain in inventory for a long time, until a new customer appears. The long run holding costs greatly outweigh the expected shortage costs. In such a parameter setting, the optimal policy orders a unit only in rare circumstances. Not surprisingly, the Minimizing policy outperforms the Myopic policy in this instance, achieving an optimality gap of just 1%. Of interest is how well a policy with an analytically provable worst-case bound performs. We focused attention on the Pure Surplus-Balancing policy (denoted by PSB, since in this model it performed better than all the other heuristics. It achieved an optimality gap of 8.6%, which is much better than the 200% theoretical performance guarantee.

Starting from the parameter setting in which the Myopic policy fared poorly, we considered combinations of demand rate and unit backlogging cost up to a ten-fold increase in the demand rate and a five-fold increase in the backlogging cost. Table 12 summarizes the results of increasing the demand rate. Table 13 summarizes the results of increasing the unit backlogging cost. Table 21 in Appendix C presents the complete set of results. In each dimension, the Myopic policy and the Minimizing policy reverse their positions of dominance. What is interesting is the robustness of the Pure Surplus-Balancing policy. It consistently performs much better than its worst case performance guarantee no matter what the parameter environment. Whereas the other two policies experience optimality gaps in excess of 100%, the Pure Surplus-Balancing policy stays within 45% of optimality for all of these parameter combinations.

Appendix A: Section 4 - Proofs

Proof of Theorem 1: The Interval-Constrained-Balancing policy has a worst-case guarantee of 2.

Table 13 Impact of Unit Backorder Cost on Average Optimality Gap of Three Policies

	Unit Backorder Cost					
Policy	10	20	30	40	50	Average
MY	96.47%	31.45%	16.53%	10.44%	6.85%	32.35%
M	4.86%	39.52%	73.80%	99.07%	66.09%	56.67%
PSB	6.97%	17.02%	25.26%	28.90%	20.28%	19.69%

A s we already mentioned, there exists an optimal base-stock policy with the property that, for each t and each $f_t \in \mathcal{F}_t$, the optimal base-stock level is between the corresponding minimizing and myopic base-stock levels, i.e., $R_t^M(f_t) \leq R_t^{OPT}(f_t) \leq R_t^{MY}(f_t)$. Instead of comparing the expected cost of the Interval-Constrained-Balancing policy with the expected cost of OPT, we shall compare it to an infeasible policy with a lower expected cost than OPT, denoted by OPT'. Specifically, OPT' is a base-stock policy with the same base-stock levels as OPT. However, if for some period t and information set f_t, the resulting inventory position of OPT' at the beginning of period t happened to be higher than the the corresponding myopic base-stock level $R_t^{MY}(f_t)$ *and* the inventory position of the Interval-Constrained-Balancing policy x_t^{ICB}, it is allowed scrap enough inventory with no cost to bring its inventory level down to $\max\{R_t^{MY}(f_t), x_t^{ICB}\}$. (If $x_t^{ICB} > x_t^{OPT'} > R_t^{MY}(f_t)$, OPT' keeps its inventory at the same level.) Since $R_t^{OPT}(f_t) \leq R_t^{MY}(f_t)$ the modified inventory level of OPT' is closer to the optimal base-stock level in that period. This implies that OPT' has lower expected cost than OPT. Also observe that OPT' can not scrap units that were already ordered by the Interval-Constrained-Balancing policy in either the current period or previous periods. (The scrapping is bounded from below by $\max\{R_s^{MY}(f_s), x_s^{ICB}\}$.)

For each $s = 1, \ldots, T-L$, let q'_s be the balancing order quantity defined above. That is, $E[H_s^{ICB}(q'_s)|f_s] = E[\Pi_s^{ICB}(q'_s)|f_s]$. Let $\bar{q}_s$ be the actual order placed by the improved Dual-Balancing policy after possibly augmenting or truncating the balancing order quantity as described above. In particular, for each s and f_s, if $q'_s < \bar{q}_s$, we know that $y_s^{ICB} = R_s^M(f_s)$, and if $q'_s > \bar{q}_s$, we know that $y_s^{ICB} = \min\{R_s^{MY}(f_s), x_s^{ICB}\}$.

For each $s = 1, \ldots, T-L$, let Z_s be the following random variables:

$$Z_s = \max\left\{E[H_s^{ICB}(\bar{Q}_s)|F_s], E[\Pi_s^{ICB}(\bar{Q}_s)|F_s]\right\}.$$

The random variable Z_s is realized at the beginning of period s as the information set f_s is observed. It is readily verified that $2Z_s \geq E[H_s^{ICB} + \Pi_s^{ICB}|F_s]$, with probability 1. Thus, using standard arguments of conditional expectation, we get that

$$\begin{aligned} E[\mathcal{C}(B)] &= \sum_{t=1}^{T-L} E[H_t^{ICB} + \Pi_t^{ICB}] = \sum_{t=1}^{T-L} E[E[H_t^{ICB} + \Pi_t^{ICB}|F_t]] \\ &\leq 2\sum_{t=1}^{T-L} E[Z_t]. \end{aligned} \tag{11}$$

To conclude the proof of the theorem, we will show that

$$E[\mathcal{C}(OPT')] \geq \sum_{t=1}^{T-L} E[Z_t],$$

from which the theorem follows.

Partition the periods $1, \ldots, T-L$ into two (random) sets. Let $\mathcal{T}_H$ be all the periods in which either the inventory position of OPT' after ordering was higher than the corresponding inventory position of the Interval-Constrained-Balancing policy, or both inventory positions are equal to the corresponding minimizing base-stock level. That is,

$$\mathcal{T}_H = \{t:\ Y_t^{ICB} < Y_t^{OPT'} \text{ or } Y_t^{ICB} = Y_t^{OPT'} = R_t^M(F_t)\}.$$

Let $\mathcal{T}_\Pi$ be the complement set, i.e.,

$$\mathcal{T}_\Pi = \{t:\ Y_t^{ICB} \geq Y_t^{OPT'} \text{and } Y_t^{ICB} > R_t^M(F_t)\}.$$

Next we shall show how to amortize the cost incurred by the interval-constrained balancing policy against the cost of OPT'. In particular, we shall show that on expectation we can amortize $\sum_{t=1}^{T-L} E[Z_s]$ of cost incurred by the interval-constrained balancing policy against costs incurred by OPT'. This and (11) imply the theorem.

Let $H^{OPT'}$ be the overall holding costs incurred by OPT'. We claim that these holding costs are higher than the holding costs incurred under the Interval-Constrained-Balancing policy by units it orders in periods $t \in \mathcal{T}_H$, i.e., $H^{OPT'} \geq \sum_{t\in\mathcal{T}_H} H_t^{ICB}$, with probability 1. To see why this claims holds, recall the definition of $\mathcal{T}_H$ above. It follows that, for each $t \in \mathcal{T}_H$, we have $Y_t^{ICB} \leq Y_t^{OPT'}$, which implies that the units ordered by the Interval-Constrained-Balancing policy in period t were ordered by OPT' in period t or even earlier. Note that even when OPT' scraps unit from inventory, it can never go below the inventory position of the Interval-Constrained-balancing policy. This implies that the units ordered by the Interval-Constrained-balancing policy in period t, can not be scrapped by OPT' in period t or later. Thus, it is clear that the holding costs these units incur under OPT' are at least as high as the holding costs they incur under the improved Dual-Balancing policy. The claim then follows.

Similarly, let $\Pi^{OPT'}$ be the overall backlogging cost incurred by OPT'. We claim that these costs are higher than the backlogging costs incurred under the Interval-Constrained-Balancing policy that are associated with periods $t \in \mathcal{T}_\Pi$, i.e., $\Pi^{OPT'} \geq \sum_{t\in\mathcal{T}_\Pi} \Pi_t^{ICB}$, with probability 1. By similar arguments we know that, for each $t \in \mathcal{T}_\Pi$, we have $Y_t^{ICB} \geq Y_t^{OPT'}$, which implies that OPT' will incur backlogging costs higher than the improved Dual-Balancing policy in period $t+L$. The claim then follows.

From the above two claims it follows that

$$\begin{aligned} E[\mathcal{C}(OPT')] &\geq \sum_{t=1}^{T-L} E\left[\mathbb{1}(t\in T_H)\cdot H_t^{ICB} + \mathbb{1}(t\in T_\Pi)\cdot \Pi_t^{ICB}\right] \\ &= \sum_{t=1}^{T-L} E\left[E[\mathbb{1}(t\in T_H)\cdot H_t^{ICB}|F_t] + E[\mathbb{1}(t\in T_\Pi)\cdot \Pi_t^{ICB}|F_t]\right]. \end{aligned} \tag{12}$$

To complete the proof of the theorem it is sufficient to show that, for each $s = 1, \ldots, T-L$, the inequalities

$$\mathbb{1}(s\in T_H)\cdot E[H_s^{ICB}|F_s] \geq \mathbb{1}(s\in T_H)\cdot Z_s$$

and

$$\mathbb{1}(s\in T_\Pi)\cdot E[\Pi_s^{ICB}|F_s] \geq \mathbb{1}(s\in T_\Pi)\cdot Z_s$$

hold with probability 1.

Consider the first inequality and some observed information set f_s. There is nothing to prove unless the indicator $\mathbb{1}(s\in T_H)$ is equal 1. However, if $\bar{q}_s < q'_s$ this implies that $y_s^{ICB} = \max\{x_s^{ICB}, R_s^{MY}(f_s)\} \geq y_s^{OPT'}$, hence, $s \in \mathcal{T}_\Pi$. Thus, $s \in \mathcal{T}_H$ implies that $\bar{q}_s \geq q'_s$, which implies that $z_s = E[H_s^{ICB}|f_s]$.

The second inequality follows by similar arguments. In particular, for each information set f_s such that $s \in \mathcal{T}_\Pi$, we know that $\bar{q} \le q'_s$. (If $\bar{q} > q'_s$ then $y_s^{ICB} = R_s^M(f_s)$ and $s \in \mathcal{T}_H$.) However, this implies that $z_s = E[\Pi_s^{ICB}|f_s]$. It follows from (12) above that indeed $E[\mathcal{C}(OPT')] \ge \sum_{t=1}^{T-L} E[Z_t]$, from which the theorem follows.

We note that if we replace $R_s^M(f_s)$ and $R_s^{MY}(f_s)$ with any sequence of lower and upper bounds on the respective optimal base-stock levels $R_s^{OPT}(f_s)$, then the proof still holds.

Proof of Lemma 1: The base-stock levels of the minimizing-k policies are decreasing in k. That is, for each $k_1 \ge k_2$, we have $R_t^{M(k_1)} \le R_t^{M(k_2)}$.

D efine $m_t^k(q_t) = l_t^{M(k)}(q_t) + \pi_t(q_t)$ under the assumption that the inventory position at the beginning of the period is 0. Note that this is a convex function. The minimizing-k policy $M(k)$ chooses q_t^k to minimize this function. We assume that it always chooses the smallest minimizer, i.e. $q_t^{M(k)} = \min\{\arg\min_{q_t \ge 0}[m_t^k(q_t)]\}$. We also note that $q_t^{M(k)} = R_t^{M(k)}$.

Note that $\pi_t(q_t)$ is common for all values of k. Thus, from (10) above it follows that right-hand-side derivative of $m_t^k(q_t)$, denoted by $(m_t^k(q_t))'$, is increasing in k (for a fix q_t). Specifically,

$$\left(m_t^{k_1}(q_t)\right)' \ge \left(m_t^{k_2}(q_t)\right)',$$

for each q_t. Thus, we conclude that $q_t^{M(k_1)} \le q_t^{M(k_2)}$.

Methods to Dynamically Choose k

Next we show that the procedures $(i)-(iii)$ for dynamically choosing k are well defined, i.e., then converge. We give the proof for procedure (i); the proofs for procedures $(ii)-(iii)$ are similar.

Lemma 2. *If $r_t(y_t^{M(1)}) \ge 1$ and $r_t(y_t^{M(T-t)}) = r_t(y_t^M) \le T-t$, then the equation*

$$k = r_t(y_t^{M(k)})$$

has a solution. That is, procedure (i) above is well-defined.

F irst recall that if $r_t(y_t^{M(1)}) < 1$ we follow the Myopic policy. Now assume that $r_t(y_t^{M(1)}) \ge 1$. We would like to compute a zero of the function $f(k) = r_t(y_t^{M(k)}) - k$. We have already seen (from (10) above) that $y_t^{M(k)}$ is decreasing in k; hence $r_t(y_t^{M(k)})$ is also decreasing in k and so is $f(k)$.

By the assumption, for $k=1$, we have $r_t(y_t^{M(1)}) \ge 1$ and for $k = T-t+1$, we have $r_t(y_t^{M(T-t)}) = r_t(y_t^M) \le T-t$. If $r_t(y_t^{M(1)}) = 1$ or $r_t(y_t^{M(T-t+1)}) = r_t(y_t^M) = T-t$, there is nothing to prove. Otherwise $f(1) > 0$ and $f(T-t) < 0$. In addition, $f(k)$ is continuous and hence there is a k_t such that $f(k_t) = 0$. Moreover, k_t can be computed by bi-section search.

The proofs for the other two cases are similar. However, we need the additional assumption that $E[D_t] \ge 1$ for each t. This ensures that $r_t([0,y])$ increases at a faster rate than y and hence that the corresponding $f(k)$ function is decreasing in k.

Appendix B: Tests of LogNormal Sum Approximation

To test the validity of Wilkinson's approximation we considered the demand model with $\mathbf{D_0} = (400, 400, \ldots, 400)$ and the variance-covariance matrix, Σ, as described in the seasonal demand part of Section 6. We first sampled 10,000 realizations of $(D_1, D_2, \ldots, D_8)$, computed $D_{[1,8]}$ and sorted these. This gives the empirical quartiles of the distribution of $D_{[1,8]}$. We then generated these same quartiles from the approximation scheme.

In Figure 1 we give a scatter plot (a p-p plot) of the quartiles of the approximating and empirical distribution. The approximating distribution corresponds to readings on the X-axis, the empirical distribution to those on the Y-axis. If the approximating distribution were perfectly accurate, this plot would be a straight line.

In fact, there is a slight trend above the 45-degree line for higher value points. This indicates that the approximating distribution underestimates the probability of very high demands. However this behavior only occurs for the final 50 or so points (of 10,000) and so is not very significant.

Our standard implementation of the policies we tested uses the approximating distribution. However, as a further test of that approximation's validity, we also implemented the policies using a Monte Carlo scheme to compute the distribution of cumulative demand, and then ran both of these implementations on 1,000 demand paths. There was a difference of only 0.042% in expected cost between the Monte Carlo and Wilkinson Approximation methods.

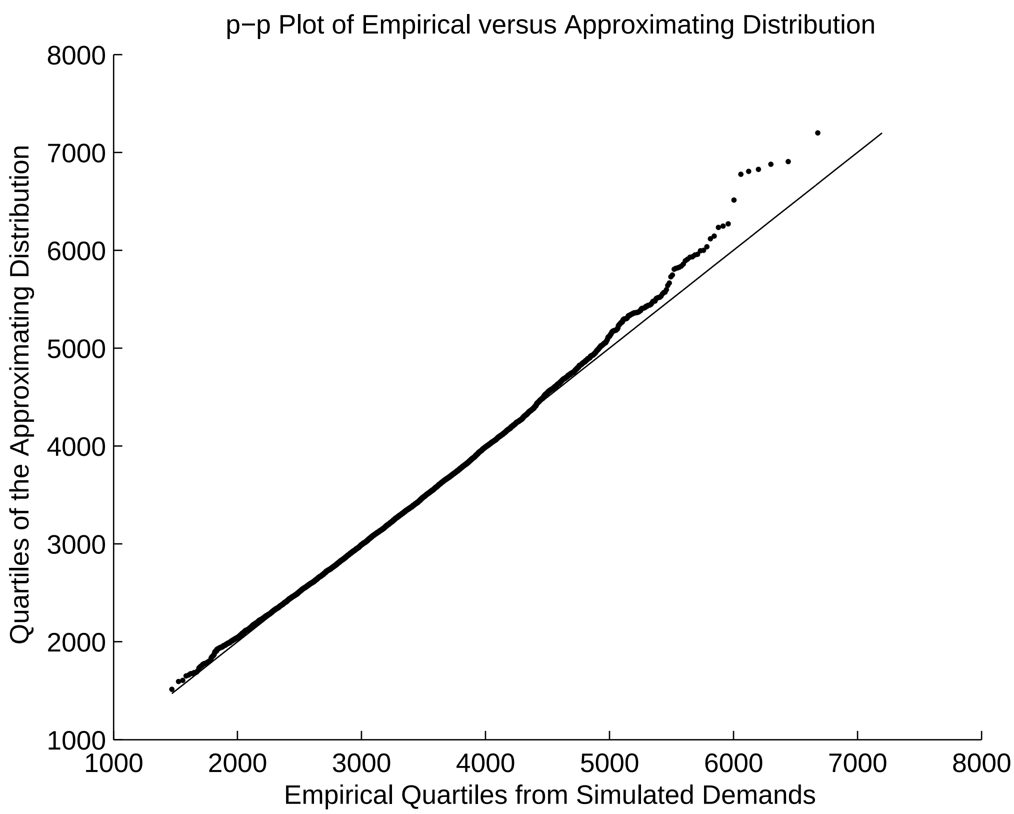

Figure 1 p-p plot of Empirical versus Approximating Distribution of Cumulative Demand.

Appendix C: Section 7 - Full Tables

The tables in this section correspond to those in Section 7 but with more policies included.

Table 14 Product Launch: $AR(\pi)$, for certain Product Launch scenarios
($AR(\pi)$ is the average percent improvement over Myopic, per run.)

	$L=0$				$L=4$			
Scenario	Flat	+20	Curve	S. Curve	Flat	+20	Curve	S. Curve
B	0.37%	0.34%	0.21%	0.47%	-2.58%	-3.05%	-2.67%	-2.26%
SB	0.13%	0.11%	0.03%	0.10%	1.91%	1.24%	1.52%	1.79%
$B(0.5)$	0.36%	0.34%	0.21%	0.47%	-4.15%	-4.15%	-3.75%	-3.44%
$B(2)$	0.46%	0.43%	0.31%	0.53%	1.52%	0.93%	1.25%	1.61%
$B(\beta\text{-myo})$	0.39%	0.37%	0.26%	0.41%	1.47%	0.98%	1.13%	1.32%
M	0.36%	0.34%	0.21%	0.47%	-4.32%	-4.24%	-3.86%	-3.54%
$M(2)$	0.46%	0.42%	0.33%	0.51%	0.82%	0.04%	0.36%	0.75%
$M(3)$	0.44%	0.41%	0.26%	0.51%	-1.71%	-2.39%	-2.02%	-1.57%
$M(k\text{-fin})$	0.44%	0.40%	0.29%	0.51%	-1.89%	-2.36%	-2.00%	-1.60%
$M(k\text{-mar})$	0.29%	0.26%	0.26%	0.29%	-0.51%	-1.03%	-0.70%	-0.33%
$M(k\text{-tot})$	0.29%	0.26%	0.22%	0.29%	1.90%	1.32%	1.57%	1.82%
LB	4.01%	3.71%	3.67%	3.79%	25.92%	22.91%	23.18%	23.75%

Table 15 End Of Life: $AR(\pi)$, for certain End Of Life scenarios

	$L=0$				$L=4$			
Scenario	Flat	-20	Curve	S. Curve	Flat	-20	Curve	S. Curve
B	0.37%	0.53%	0.83%	0.66%	-2.58%	-1.65%	0.91%	6.12%
SB	0.13%	0.16%	0.19%	0.35%	1.91%	3.14%	5.02%	9.93%
$B(0.5)$	0.36%	0.52%	0.82%	0.58%	-4.15%	-4.24%	-2.59%	1.26%
$B(2)$	0.46%	0.59%	0.75%	0.87%	1.52%	3.00%	5.25%	10.31%
$B(\beta\text{-myo})$	0.39%	0.48%	0.53%	0.68%	1.47%	2.50%	4.47%	8.59%
M	0.36%	0.52%	0.82%	0.58%	-4.32%	-4.58%	-3.07%	0.36%
$M(2)$	0.46%	0.62%	0.76%	0.90%	0.82%	2.63%	4.99%	8.53%
$M(3)$	0.44%	0.62%	0.81%	0.91%	-1.71%	0.05%	3.10%	8.65%
$M(k\text{-fin})$	0.44%	0.60%	0.83%	0.83%	-1.89%	-0.93%	1.57%	7.28%
$M(k\text{-mar})$	0.29%	0.37%	0.48%	0.60%	-0.51%	0.70%	3.17%	8.85%
$M(k\text{-tot})$	0.29%	0.36%	0.40%	0.56%	1.90%	3.15%	5.17%	9.36%
LB	4.01%	4.63%	4.55%	5.19%	25.92%	31.49%	36.14%	44.12%

Appendix D: The Effect of Bounding

We consider here the effect of the Interval-Constrained-Bounding improvement scheme that we discussed in Section 4. Note that this applies only to the Balancing policies, as the Minimizing policies fall within the bounds by definition. Also note that the Surplus-Balancing Policy SB might exceed the upper bound, but it cannot order less than the lower bound. Table 22 lists the average improvement in $AR(\pi)$ generated by bounding for these policies, as well as the minimum and maximum improvement. Note that the improvement is always positive, as we would expect from Theorem 1.

The two policies B and $B(0.5)$ are dramatically improved by bounding. The reason for this is that more often than not they fall outside of the known limits on the optimal order-up-to levels, provided by the Minimizing and Myopic policies. To demonstrate this, Table 22 also includes the percentage of all order-up-to levels that fall either below the Minimizing level or above the Myopic level. The policies B and $B(0.5)$ fall outside this range 82.91% and 95.15% of the time, respectively.

Table 16 Demand Crash: Policy performance in the Demand Crash scenario (Both AT and AR measures are presented)

	$L=0$			$L=4$		
Policy	$AT(\pi)$	$AR(\pi)$	HC	$AT(\pi)$	$AR(\pi)$	HC
B	22.10%	13.12%	75.57%	18.31%	12.29%	64.37%
SB	19.87%	12.39%	64.21%	20.93%	15.60%	65.32%
$B(0.5)$	22.66%	12.66%	80.23%	14.93%	5.81%	74.85%
$B(2)$	20.19%	12.98%	65.27%	20.20%	16.43%	50.88%
$B(\beta\text{-myo})$	19.08%	12.31%	61.14%	15.82%	13.90%	46.74%
M	22.66%	12.65%	80.33%	14.13%	3.92%	81.51%
$M(2)$	10.65%	8.36%	32.37%	11.18%	11.39%	26.20%
$M(3)$	15.31%	11.19%	47.80%	15.02%	14.10%	40.10%
$M(k\text{-fin})$	21.44%	13.25%	71.14%	19.74%	13.59%	69.90%
$M(k\text{-mar})$	14.53%	10.23%	44.92%	20.30%	15.15%	64.57%
$M(k\text{-tot})$	13.33%	9.39%	40.95%	16.93%	15.37%	41.96%
LB	31.35%	23.94%	-	55.33%	56.03%	-

Table 17 Seasonality: $AR(\pi)$, for certain Seasonality scenarios

	$L=0$				$L=4$			
Policy	Flat	Step(2)	Step(4)	Step(8)	Flat	Step(2)	Step(4)	Step(8)
B	0.37%	5.52%	4.83%	7.20%	-2.58%	-1.97%	-2.52%	2.71%
SB	0.13%	2.22%	3.39%	4.80%	1.91%	3.01%	3.50%	7.44%
$B(0.5)$	0.36%	5.49%	4.82%	7.18%	-4.15%	-3.67%	-5.16%	-0.96%
$B(2)$	0.46%	5.89%	5.22%	6.69%	1.52%	2.39%	2.83%	6.99%
$B(\beta\text{-myo})$	0.39%	5.20%	4.81%	5.89%	1.47%	2.21%	2.57%	5.12%
M	0.36%	5.50%	4.82%	7.18%	-4.32%	-3.85%	-5.36%	-1.18%
$M(2)$	0.46%	6.01%	5.15%	5.93%	0.82%	2.20%	3.19%	7.20%
$M(4)$	0.41%	5.65%	4.93%	7.33%	-3.19%	-2.24%	-2.83%	3.26%
$M(6)$	0.38%	5.54%	4.86%	7.24%	-4.14%	-3.55%	-4.64%	-0.32%
$M(k\text{-fin})$	0.44%	5.76%	5.14%	7.22%	-1.89%	-1.05%	-2.03%	2.91%
$M(k\text{-mar})$	0.29%	4.38%	3.98%	4.61%	-0.51%	0.88%	0.41%	5.05%
$M(k\text{-tot})$	0.29%	4.19%	3.61%	4.02%	1.90%	3.08%	3.79%	7.14%
LB	4.01%	22.93%	20.29%	18.22%	25.92%	30.48%	34.38%	40.82%

We demonstrate this graphically in Figure D, which plots the evolution of order-up-to level of the Myopic, Minimizing, and three Balancing policies in a single run, for one of the Coefficient of Variation scenarios. To highlight the difference, we subtract from each order-up-to level, the order-up-to level of the Minimizing Policy (which corresponds to the heavy, horizontal line at 0). The other heavy line corresponds to the Myopic policy. It is interesting to note that $B(2)$ (the dashed line) closely tracks the optimal range, while $B(0.5)$ (dashed and dotted line) is always below it. The Surplus-Balancing Policy (SB, the lighter solid line) is often above the range - Table 22 indicates that this occurs 65.24% of the time, on average.

Acknowledgments

Table 18 Coefficient of Variation: $AR(\pi)$, for certain Coefficient of Variation scenarios ($L = 0$)

C.V.	0.5	0.7	1	2	4	8
B	0.01%	0.23%	1.59%	9.74%	22.22%	26.84%
SB	0.00%	0.06%	1.02%	9.65%	22.21%	29.15%
$B(0.5)$	0.01%	0.23%	1.58%	8.63%	19.21%	20.46%
$B(2)$	0.01%	0.29%	1.94%	10.74%	22.26%	29.36%
$B(\beta\text{-myo})$	0.01%	0.24%	1.78%	10.77%	22.87%	29.01%
M	0.01%	0.23%	1.58%	8.62%	18.98%	18.81%
$M(2)$	0.01%	0.28%	1.92%	10.09%	19.32%	24.54%
$M(3)$	0.01%	0.27%	1.86%	11.03%	22.96%	29.81%
$M(k\text{-fin})$	0.01%	0.27%	1.87%	10.42%	22.74%	27.73%
$M(k\text{-mar})$	0.01%	0.15%	1.68%	10.45%	22.98%	30.23%
$M(k\text{-tot})$	0.01%	0.16%	1.48%	10.07%	22.43%	30.17%
LB	0.39%	2.99%	10.27%	35.60%	57.68%	72.40%

Table 19 Time of Learning: Average of $AR(\pi)$ over all policies, for the Time of Learning scenarios ($L = 0$)

Scenario	Early	Mid	Const	Late
Average	0.00%	0.04%	0.36%	1.10%

Table 20 Correlation: $AR(\pi)$ for the Correlation scenarios ($L = 0$)

Scenario	None	Pos(4)	Neg(4)	Mix(4)	Pos(8)	Neg(8)	Mix(8)
B	1.00%	0.17%	0.93%	2.15%	0.07%	1.69%	2.17%
SB	0.35%	0.02%	0.34%	0.49%	0.01%	0.54%	0.49%
$B(0.5)$	1.00%	0.17%	0.94%	2.15%	0.07%	1.69%	2.17%
$B(2)$	1.04%	0.21%	1.07%	2.05%	0.15%	1.61%	2.07%
$B(\beta\text{-myo})$	0.89%	0.15%	0.94%	1.69%	0.12%	1.33%	1.68%
M	1.00%	0.17%	0.93%	2.15%	0.07%	1.69%	2.17%
$M(2)$	0.99%	0.21%	1.00%	2.10%	0.16%	1.58%	2.15%
$M(3)$	1.02%	0.19%	0.97%	2.16%	0.12%	1.64%	2.20%
$M(k\text{-fin})$	1.02%	0.20%	0.98%	2.15%	0.13%	1.68%	2.18%
$M(k\text{-mar})$	0.71%	0.17%	0.83%	1.44%	0.14%	1.14%	1.46%
$M(k\text{-tot})$	0.68%	0.13%	0.78%	1.37%	0.11%	1.07%	1.39%
LB	6.06%	2.73%	7.27%	8.85%	2.72%	7.45%	8.90%

We thank the Associate Editor and the anonymous referee for many constructive comments that improve the content and the exposition of the paper.
This research was partially conducted while the first and third authors were PhD students in the ORIE department at Cornell University.
The research of the thris author was supported partially by NSF grants CCR-9912422, CCR-0430682 and DMS-0732175
The research of the fourth author was supported partially by NSF grants DMI-0075627 & DMI-0500263, and the Querétaro Campus of the Instituto Tecnológico y de Estudios Superiores de Monterrey. The research of the fifth author was supported partially by NSF grants CCR-9912422, CCR-0430682, DMI-0500263, DMS-0732196 and CCF-0635121.

Table 21 Impact of Demand Rate and Unit Backorder Cost on Optimality Gap of Three Policies

		Unit Backorder cost					
Policy	Demand Rate	10	20	30	40	50	Average
MY		281.96%	119.90%	65.96%	39.07%	23.05%	105.99%
M	0.01	0.98%	0.97%	0.98%	1.02%	1.08%	1.01%
SB		8.60%	6.85%	5.33%	4.00%	2.81%	5.52%
MY		85.69%	5.78%	0.02%	0.02%	0.02%	18.30%
M	0.04	0.95%	1.46%	29.64%	61.72%	90.87%	36.93%
SB		6.18%	1.06%	16.30%	26.35%	33.08%	16.59%
MY		18.25%	0.00%	0.05%	0.05%	2.14%	4.10%
M	0.07	1.05%	51.31%	106.83%	149.25%	171.41%	95.97%
SB		2.34%	23.83%	37.21%	43.64%	45.02%	30.41%
MY		0.00%	0.11%	0.09%	2.62%	2.19%	1.00%
M	0.1	16.47%	104.34%	157.76%	184.27%	1.00%	92.77%
SB		10.78%	36.36%	42.20%	41.64%	0.21%	26.24%
MY		96.47%	31.45%	16.53%	10.44%	6.85%	32.35%
M	Average	4.86%	39.52%	73.80%	99.07%	66.09%	56.67%
SB		6.97%	17.02%	25.26%	28.90%	20.28%	19.69%

Table 22 Improvement in $AR(\pi)$ due to bounding

Policy	Mean (Min, Max)	[< Min, > Myo]
B	3.43% (0.04%, 11.54%)	[70.21%, 10.76%]
SB	1.96% (0.03%, 5.00%)	[0.00%, 62.60%]
$B(0.5)$	17.51% (0.39% 33.21%)	[86.02% 7.05%]
$B(2)$	0.27% (0.00%, 5.26%)	[38.75%, 19.12%]
$B(\beta\text{-myo})$	1.95% (0.00%, 5.06%)	[0.00%, 40.07%]

References

Abu-Dayya, A.A., N.C. Beaulieu. 1994. Outage probability in the presence of correlated lognormal interferers. *IEEE Transactions on Vehicular Technology* **43**(1) 164–173.

Beaulieu, N.C., A.A. Abu-Dayya, P.J. McLane. 1995. Estimating the distribution of a sum of independent lognormal random variables. *IEEE Transactions on Communications* **43**(12) 2869–2873.

Chen, F., J.S. Song. 2001. Optimal policies for multi-echelon inventory problems with Markov-modulated demand. *Operations Research* **49** 226–234.

Dong, L., H. L. Lee. 2003. Optimal policies and approximations for a serial multiechelon inventory system with time-correlated demand. *Operations Research* **51** 969–980.

Gavireni, S., S. Tayur. 2001. An efficient procedure for non-stationary inventory control. *IIE Transactions* **33** 83–89.

Graves, S. C., H. Meal, S. Dasu, Y. Qin. 1986. Two-stage production planning in a dynamic environment. S. Axsater, C. Schneeweiss, E. Silver, eds., *Multi-Stage Production Planning and Control*. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, 9–43.

Heath, D. C., P. L. Jackson. 1994. Modeling the evolution of demand forecasts with application to safety stock analysis in production distribution-systems. *IIE Transactions* **26**(3) 17–30.

Hurley, Gavin, Peter Jackson, Retsef Levi, Robin Roundy, David Shmoys. 2006. A theoretical and computational study of new algorithms for the stochastic inventory control problem. Tech. rep., School of Operations Research and Industrial Engineering, Cornell University.

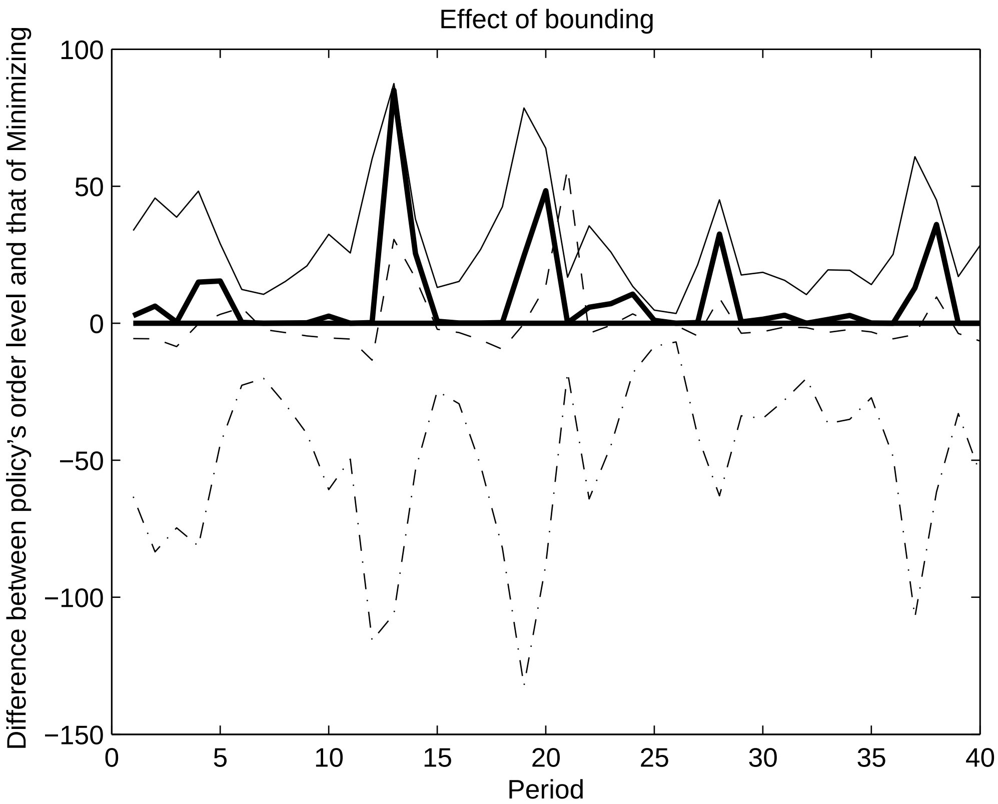

Figure 2 Difference between the order-up-to levels of the following policies and Minimizing, before bounding: 1) Myopic (thick solid line), 2) $B(0.5)$ (dashed and dotted line), 3) $B(2)$ (dashed line) and 4) SB (solid line).

Ignall, E., A. F. Veinott. 1969. Optimality of myopic inventory policies for several substitute products. *Managment Science* **15** 284–304.

Iida, T., P. Zipkin. 2001. Approximate solutions of a dynamic forecast-inventory model. Working paper.

Levi, R., G. Janakiraman, M. Nagarajan. 2005a. A 2-approximation algorithm for stochastic inventory control models with lost sales. To appear in Mathematics of Operations Research.

Levi, R., Martin Pál, R. O. Roundy, D. B. Shmoys. 2007. Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research* **32** 284–302.

Levi, R., R. O. Roundy, D. B. Shmoys, V. A. Truong. 2004. Approximation algorithms for capacitated stochastic inventory control models. Submitted.

Levi, R., R. O. Roundy, V. A. Truong. 2005b. Provably near-optimal balancing policies for multi-echelon stochastic inventory control models. Working paper.

Lu, X., J. S. Song, A. C. Regan. 2006. Inventory planning with forecast updates: approximate solutions and cost error bounds. *Operations Research* **54** 1079–1097.

Milevsky, M. A., S. E. Posner. 1998. Asian options, the sum of lognormals and the reciprocal gamma distribution. *Journal of Financial and Quantitative Analysis* **33**(3) 409–422.

Muharremoglu, A., J. N. Tsitsiklis. 2001. A single-unit decomposition approach to multi-echelon inventory systems. Working paper.

Özer, Ö., G. Gallego. 2001. Integrating replenishment decisions with advance demand information. *Managment Science* **47** 1344–1360.

Schwartz, S. C., Y.S. Yeh. 1982. On the distribution function and moments of power sums with lognormal components. *Bell Syst. Tech. J.* **61** 1441–1462.

Song, J., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Operations Research* **41** 351–370.

Veinott, A. 1965a. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Management Science* **12** 206–222.

Veinott, A. F. 1965b. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Management Science* **12** 206–222.

Veinott, A. H. 1963. Optimal policies with non-stationary demands. H. E. Scarf, D. M. Gilford, M. W. Shelly, eds., *Multistage inventory models and techniques*. Stanford University Press, 85–115.

Zipkin, P. H. 2000. *Foundations of inventory management*. The McGraw-Hill Companies, Inc.