MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 9, No. 2, Spring 2007, pp. 185–205
ISSN 1523-4614 | EISSN 1526-5498 | 07 | 0902 | 0185

DOI 10.1287/msom.1060.0136

Inspection and Replenishment Policies for Systems with Inventory Record Inaccuracy

A. Gürhan Kök, Kevin H. Shang
Fuqua School of Business, Duke University, Durham, North Carolina 27708-0120
{gurhan.kok@duke.edu, khshang@duke.edu}

For many companies, inventory record inaccuracy is a major obstacle to achieving operational excellence. In this paper, we consider an inventory system in which inventory records are inaccurate. The manager makes inventory inspection and replenishment decisions at the beginning of each period. There is a cost associated with each inspection. If an inspection is performed, inventory records are aligned with physical inventory. The objective is to develop a joint inspection and replenishment policy that minimizes total costs in a finite horizon. We prove that an inspection adjusted base-stock (IABS) policy is optimal for the single-period problem. In the finite-horizon problem, we show that the IABS policy is near optimal in a numerical study. Under this policy, the manager performs an inspection if the inventory recorded is less than a threshold level, and orders up to a base-stock level that depends on the number of periods since the last inspection. The prevalent approach to deal with inventory inaccuracy in practice is to implement cycle-count programs. Based on the structure of the IABS policy, we propose a new cycle-count policy with state-dependent base-stock levels (CCABS). We show that CCABS is almost as effective as the IABS policy. In addition, we provide guidelines for practitioners to design effective cycle-count programs by conducting sensitivity analyses on the IABS policy. Finally, by comparing the costs associated with these policies and several benchmark systems, we quantify the true value of accurate inventory information, which may be provided by radio-frequency identification (RFID) systems.

Key words: inventory control; inventory inaccuracy; cycle-count policy; RFID
History: Received: February 15, 2006; accepted: September 12, 2006.

1. Introduction

Inventory record inaccuracy is a common problem across industries. Although companies have invested substantial amounts of money to automate and improve their inventory management processes, inventory records and physical inventory are, contrary to popular belief, seldom aligned. There are several causes of discrepancies between inventory records and physical inventory: stock loss or shrinkage, transaction errors (in inbound or outbound processes), and product misplacement.[1] Such misalignment may create significant losses for firms. Consider a large distribution company with an average inventory of $3 billion, which we refer to as Beta to preserve confidentiality. Beta serves the heavy-equipment manufacturing industry and reports severe inventory inaccuracy problems in their distribution centers. Although Beta has been carrying out several programs to improve inventory accuracy, the records were inaccurate by 1.6% of the total inventory value at the end of 2004. Similar problems are faced by the retail industry. According to Raman et al. (2001), records were inaccurate for 65% of the SKUs at a publicly traded retailer and the magnitude of the errors, on average, is 35% of the target stock levels. The direct effect of inventory record inaccuracy is losses resulting from ineffective inventory order decisions. When an out-of-stock item is reported as in stock, an automated replenishment system may not reorder that product, which may result in higher backorder penalties or lost sales. On the other hand, if the records show fewer items than the physical inventory level, more of that product would be ordered, leading to higher inventory costs.

To remedy this problem, most companies implement a cycle-count program. Cycle counting refers to any process that verifies the correctness of inventory

[1] See Piasecki (2003) and DeHoratius and Raman (2004) for an extensive discussion on the causes and the extent of inventory inaccuracy.

quantity data by counting portions of the inventory on an ongoing basis. The major goals of cycle-counting programs are to identify process problems and to correct inventory data. Usually products are classified according to an ABC system, which is based on SKU rankings of a specific measure, such as velocity (frequency of transactions), units sold/consumed, or dollar value of inventory. Then, count cycles are assigned to each group using rough guidelines. The experience of a senior operations manager at EMC[2], an information systems and storage equipment manufacturer, seems to be typical: "We perform cycle-counts for A-items monthly, for B-items every three months, and for C-items every six months. In our facility, we have to run an 8-hour shift of about 100 workers on a Friday to perform a cycle-count for A-items. I am not sure if it is necessary to count all A-items every month. Some of the counted items already have a high level of accuracy" (Pinkerman 2005).

There are several important issues to be discussed in the above inspection practice. First, the trade-off between inspection and inventory-related costs is not carefully considered. Specifically, if inspections are conducted more frequently, accurate inventory information leads to lower inventory-related costs. However, more frequent inspections increase the total inspection cost. Hence, it is critical for inventory managers to choose the right inspection cycle for each product. Another important question almost always neglected in industry hinges upon whether the replenishment policy should be adjusted according to the level of inventory inaccuracy or to the chosen inspection policy. In particular, if the inspection cost is high, it may be better to inspect less frequently but carry more inventory to account for the increased uncertainty stemming from record inaccuracy. Thus, it is important to derive an optimal joint inventory inspection and replenishment policy that takes into account these trade-offs. Second, there is little research on the effectiveness of the commonly used cycle-count approach and how the inspection cycle for each product should be determined. Third, companies may adopt advanced inventory-tracking systems, such as RFID, to minimize inventory inaccuracy. Because adopting such a system requires significant investment, it is important to carefully assess its potential benefit.

In this paper, we investigate these issues in a single-product, single-location periodic inventory system in which inventory records are inaccurate. Inaccuracy is due to a random error that changes the physical inventory level at the end of each period. These errors are not observed by the information system until an inspection is performed. Thus, the discrepancy between the inventory record and the physical inventory equals the total accumulated error since the last inspection. Customer demand in each period is stochastic and unmet demand is backlogged. The manager makes the inventory inspection and replenishment decisions at the beginning of each period. If inventory is inspected, inventory records are aligned with physical inventory, but an inspection cost is incurred. Our objective is to find a joint inventory inspection and replenishment policy that minimizes total inventory-related and inspection costs over a finite horizon.

We first characterize the optimal joint replenishment and inspection policy in a single period. We show that there exists a threshold inventory level for the inspection decision. The manager should perform the inspection if and only if the initial inventory record is less than or equal to the threshold level. As for the inventory replenishment policy, a base-stock policy is optimal. Both the threshold level and the base-stock level depend on the level of inaccuracy in the system. We term this structured optimal policy the "inspection-adjusted base-stock policy," or the IABS policy for short. The optimal policy for the finite-horizon problem may not have a simple form, because the expected future cost function is poorly behaved. It is therefore desirable to construct a simple and effective heuristic policy. In our numerical study, we find that the optimal policy has the structure of the IABS policy in most examples. To reflect this observation, we show that an IABS policy is optimal for a revised dynamic program that provides a close approximation of the exact cost function. We propose a heuristic that employs the IABS policy in each period with parameters obtained from the revised dynamic program. We term this the IABS heuristic and show in a numerical study that it is near optimal: The cost of the IABS heuristic is, on average, within 0.2% of the optimal cost.

Investigating the heuristic solutions in the cases with stationary data, we find that the IABS heuristic tends to conduct inspections in fixed cycles. This result suggests that a cycle-count policy may be effective if the cycle length is similar to that of the IABS heuristic. In practice, however, inventory inaccuracy is often not taken into account in making replenishment decisions between cycle-counts. We model this practice and term it as the standard cycle-count (CC) heuristic. Based on the structure of the IABS heuristic, we propose an improved cycle-count heuristic, in which the base-stock level is adjusted according to the number of periods since last inspection (CCABS). The best cycle length for these cycle-count heuristics can be obtained via simulation or from the IABS policy. In our numerical study, CCABS has an average optimality gap of 0.5% and CC has an optimality gap of 1.8%. The simpler structure of cycle-count policy (i.e., a fixed inspection schedule) can make it an appealing policy for practitioners over the IABS policy. However, if a company chooses to implement a cycle-count program, we recommend that CCABS should be adopted instead of CC. The performance of CCABS slightly deteriorates in cases with nonstationary cost data, higher inspection costs, and larger demand variance. In such cases, IABS heuristic should be used as it achieves near-optimal results in all cases.

Several important insights can be drawn by studying the IABS policy. First, to minimize the total inventory and inspection costs, companies need to carefully adjust base-stock levels across periods: As the number of periods since the last inspection increases, the order-up-to level should be increased to accommodate the added uncertainty caused by inventory inaccuracy. In addition, analytical and numerical sensitivity analyses on solution parameters show that, all else being equal, a product with higher value, larger error variance, smaller inspection cost, or smaller demand variance should be inspected more frequently. If a cycle-count program requires products to be classified into groups with different inspection schedules, these sensitivity analysis results on the IABS heuristic could be useful in designing an ABC classification scheme.

Radio-frequency identification (RFID) technology is being increasingly adopted in supply chain and inventory management applications. One of the benefits of RFID systems is their ability to provide accurate inventory information. Assessing the value of perfect information by comparing a system with perfect information with a system managed under policies that do not take inventory inaccuracy into account may be misleading. The true value of RFID can be assessed only by comparing a perfect information system with a system that mitigates the effects of inaccuracy by means of replenishment and inspection strategies. We compare the optimal costs of a no-error system (i.e., perfect RFID systems are installed) with an inaccurate-record system that is managed near optimally with the IABS heuristic, and an inaccurate-record system that is managed as if the records were accurate. We find that the increase in inventory costs resulting from inaccurate information is significant. Although RFID systems could help eliminate these costs completely, an effective inspection and replenishment policy, such as the IABS policy, can recover a large portion of the benefit that RFID systems promise. In our numerical study, the true value of accurate inventory information provided by RFID systems is about one-third of the total potential losses due to inventory record inaccuracy. This finding is especially important for small- and middle-size companies that may not be able to make large investments in RFID systems.

1.1. Literature Review

The literature on inventory accuracy has not drawn much attention until recently. Iglehart and Morey (1972) may be the first paper to discuss this problem. They consider count frequency and base-stock level to minimize inspection and inventory-holding costs subject to the probability that showing in stock when the item is actually out of stock is less than a prespecified level. There are several recent empirical studies that investigate inventory inaccuracy. Sheppard and Brown (1993) describe a field experiment in order to assess the causes of inventory record errors in a manufacturing setting. Raman and Ton (2004) describe how inventory accuracy is achieved at an electronics distributor by means of process redesign. DeHoratius and Raman (2004) empirically explore the extent of inventory record inaccuracy and find that it increases with sales, the number of stages in the supply chain, and product variety. Analytical models have been developed recently to study inventory inaccuracy problems. Kang and Gershwin (2003) use simulation to demonstrate that even small inventory discrepancies may lead to severe stockouts. They also propose several approaches to mitigate this problem. Bensoussan et al. (2005, 2006) consider inventory models in which inventory information is inaccurate due to information delay and show that optimal ordering policies can be identified through "reference inventory positions." Atali et al. (2005) explicitly model several sources of inaccuracy and present a heuristic replenishment policy under a given cycle-count policy. Camdereli and Swaminathan (2005) discuss optimal stocking decisions and coordinating contracts in a single-period, single-location system with misplaced inventory. DeHoratius et al. (2005) consider inventory management tools that account for record errors using a Bayesian updating of error distribution. None of these recent papers consider the optimization of the inspection decisions. For a review of the literature on RFID-related models and future applications, see Lee and Özer (2005) and Zipkin (2006).

The present model is similar to DeCroix and Mookerjee (1997), who consider a model in which a manager decides at the beginning of each period whether to purchase demand information in order to eliminate variability of demand in the current period, and subsequently decides how much to order. However, there are several distinctions between their model and ours, which make both the analysis and the insights of the two models significantly different. First, in their case, inventory level before ordering is deterministic before making the acquisition decision, but in our model before the inspection it is a random variable. Second, in their model, the current acquisition decision does not affect demand variability in future periods. In our model, the inspection decision affects the future inaccuracy levels, which has the same effect of increasing demand variability.

Our model is also related to machine-repair problems in which the machine deteriorates over time and the production rate depends on the state of the machine. The problem is to determine joint production and repair decisions. The unknown machine state corresponds to the level of inaccuracy in our model, whereas the repair decision corresponds to the inspection decision. There is a rich literature on this issue, e.g., Derman and Lieberman (1967), Ross (1971), Rosenfield (1976), White (1978), and Singh et al. (2004). However, their formulations do not include inventory-related costs. One exception is Venkatesan (1984), who considers a finite-horizon problem with random demands and perfect repair to minimize fixed and variable production, inventory-holding, shortage, and repair costs. He derives conditions on the value functions so that the structured policy in the single-period problem can carry over to the finite-horizon problem. The one-period inventory cost function in Venkatesan (1984) is independent of the state of the machine, whereas it depends on the inaccuracy level in our model.

The literature on production systems with random yield and imperfect quality is also relevant. In random yield models, yield uncertainty is resolved after a delivery is made and before demand is realized. For the imperfect quality problem, Lee and Rosenblatt (1985), Peters et al. (1988), So and Tang (1995), and Chen et al. (2001) develop joint inspection/rework and lot-sizing policies to minimize inspection, sampling, rework, and inventory-related costs. In these models, uncertainty about the number of defective units is resolved when demand is realized. See Yano and Lee (1995) for a comprehensive review.

The rest of the paper is organized as follows: §2 describes the model. Section 3 presents the analysis of the single-period problem and derives the optimal IABS policy. Section 4 considers the finite-horizon problem, proposes a near-optimal heuristic policy, and discusses its linkage to the cycle-count policy. Section 5 presents numerical studies and discusses managerial insights. Section 6 discusses several extensions and limitations of our model, and §7 concludes. All proofs are presented in Appendix B.

2. The Model and Preliminaries

Consider a single-product, single-location inventory system in which inventory records are inaccurate. Let t denote the number of periods to go, $t = T, T-1, \ldots, 1, 0$. Inaccuracy is due to a random error ε_t that changes the physical inventory level at the end of each period t. These errors are not observed by the information system until an inspection is performed. We assume that the errors ε_t across periods are independently and identically distributed and follow a continuous distribution with mean zero[2] and

[2] Note that all sources of errors other than shrinkage can generate errors in a positive or negative direction. DeHoratius and Raman

standard deviation σ_ε. Moreover, the errors are cumulative: Let j denote the number of periods since the last inspection. The level of inventory record inaccuracy at the beginning of a period t, denoted as $\varepsilon_t(j)$, is the sum of the errors from the periods since the last inspection, i.e., $\varepsilon_t(j) = \sum_{i=t+1}^{t+j} \varepsilon_i$. Let D_t denote random customer demand in each period t. D_t are independently and identically distributed across periods and follow a continuous distribution with mean μ and standard deviation σ. Unsatisfied demand is fully backlogged. We assume that replenishment lead time is zero.

The model described above fits manufacturers and distributors well, because they usually maintain long-term relationships and contracts with their customers and backlog the orders if they cannot satisfy customer demand immediately. Other modeling choices can be made for different contexts. In retailing, for example, a model with lost sales and an error distribution with a negative mean (due to shrinkage) may be more appropriate. We discuss these and other possible extensions of the model, and the challenges associated with each, in §6.

2.1. Dynamics

We define the following variables to explain the system dynamics:

x_t = inventory record before ordering at the beginning of period t,

y_t = inventory record after ordering at the beginning of period t,

w_t = physical inventory level before ordering at the beginning of period t,

$= x_t - \varepsilon_t(j)$.

We use a two-dimensional vector (x_t, j) to represent the system state, $1 \le j \le T - t + 1$. The state variables represent the information that is available to the inventory manager.

At the beginning of period t, the manager decides first whether or not to inspect and then how much to order according to the state (x_t, j). If an inspection is conducted, physical inventory w_t is observed and inventory record x_t is aligned with the physical inventory w_t. In this case, the order quantity is $y_t - w_t = y_t - (x_t - \varepsilon_t(j))$ and the ending physical inventory is $y_t - D_t$. The system state at the beginning of the period $t-1$ will be $(y_t - D_t, 1)$. We assume that inspections can be completed immediately and that no counting errors are made.

On the other hand, if an inspection is not conducted, the order quantity is $y_t - x_t$. Because there are j periods of error accumulated in x_t, the physical inventory is a random variable, $w_t = x_t - \varepsilon_t(j)$. The ending physical inventory is $w_t + (y_t - x_t) - D_t = y_t - \varepsilon_t(j) - D_t$. The system state at the beginning of period $t-1$ will be $(y_t - D_t, j+1)$. In other words, the gap between inventory record and physical inventory at the beginning of period $t-1$ is $\varepsilon_{t-1}(j+1)$.

We assume that error information is learned only through inspections. In reality, if a company has a process for updating inventory records when backorders occur, then it is possible that the magnitude of the error can be automatically learned. This extension is discussed in §6.

Throughout this paper E denotes expectation and the subscripts of E list the random variables over which the expectation is taken. Note also that $x^+ = \max(x, 0)$ and $x^- = x - x^+$.

2.2. Costs

The cost parameters are as follows. h_t is the inventory-holding cost per unit per period, b_t is the backorder cost per unit per period, and c_t is the purchase cost per unit. The discount factor is $\gamma \in [0, 1]$.

In addition to the standard inventory costs, there is a cost associated with conducting an inspection. The inspection cost has a fixed component k and a variable component $G(w) = gw^+$ with $g \ge 0$. Although we assume a linear form for G, our results can be easily extended for nondecreasing convex functions of w. The inspection cost given the inventory record is the expectation of the inspection costs over the error terms.

$$K(x) = k + g\mathsf{E}_{\varepsilon(j)}[(x - \varepsilon(j))^+]$$

Apparently, $K(x)$ is nondecreasing and convex in x.

The objective is to characterize the optimal inventory inspection and replenishment policy such that the total cost over a finite number of periods is minimized.

We formulate this problem as a dynamic program. We define the cost functions as follows:

$N_t(x, j)$ = optimal expected discounted cost in periods t through 1 assuming No inspection is performed

(2004) present a frequency diagram of the errors in their study, which supports an error distribution with a zero mean.

in t, the initial inventory record is x, and j periods of error have accumulated.

$I_t(x, j) =$ optimal expected discounted cost in periods t through 1 assuming Inspection is performed in t, the initial inventory record is x, and j periods of error have accumulated.

$V_t(x, j) =$ optimal expected discounted cost in periods t through 1 assuming the initial inventory record is x, and j periods of error have accumulated.

These *expected* cost functions are evaluated based on the physical inventory level at the end of a period. Below we show the detailed formulations of these functions.

Although the model and the analysis can be carried out with nonstationary demand, error, and cost parameters, we present our analysis assuming stationary data. Therefore, we drop the time subscript in D_t, ε_t, h_t, b_t, and c_t.

If no inspection is conducted in period t, then

$$N_t(x, j) = \min_{y\geq x}(c(y-x)+\mathsf{E}_{\varepsilon(j),D}[h(y-D-\varepsilon(j))^+ +b(y-D-\varepsilon(j))^-]+\gamma\mathsf{E}_D[V_{t-1}(y-D, j+1)]).$$

Define

$$L(y, j) = cy + \mathsf{E}_{\varepsilon(j),D}[h(y-D-\varepsilon(j))^+ + b(y-D-\varepsilon(j))^-], \quad (1)$$

$$H_t(y, j) = L(y, j) + \gamma\mathsf{E}_D[V_{t-1}(y-D, j+1)]. \quad (2)$$

Notice that $L(y, j)$ is the single-period cost function and $H(y, j)$ is the cost-to-go function given initial inventory record y and j periods of error in the system. Thus,

$$N_t(x, j) = \min_{y\geq x}\{H_t(y, j)\} - cx. \quad (3)$$

If an inventory inspection is conducted at the beginning of period t, there will be no discrepancy between the records and the physical inventory in period t. Thus,

$$I_t(x, j) = \mathsf{E}_{\varepsilon(j)}\Big[\min_{y\geq w}(c(y-w)+\mathsf{E}_D[h(y-D)^+ + b(y-D)^-] + \gamma\mathsf{E}_D[V_{t-1}(y-D, 1)] + K(x))\Big]$$
$$= \mathsf{E}_{\varepsilon(j)}\Big[\min_{y\geq x-\varepsilon(j)}\{H_t(y, 0)\}\Big] - cx + K(x). \quad (4)$$

The optimal total cost from periods t through 1 given states x and j is given by the following dynamic program (DP):

$$\text{(DP)}\quad V_t(x, j) = \min\{N_t(x, j), I_t(x, j)\}$$
$$= \min\Big\{\min_{y\geq x} H_t(y, j), \mathsf{E}_{\varepsilon(j)}\Big[\min_{y\geq x-\varepsilon(j)} H_t(y, 0)\Big] + K(x)\Big\} - cx.$$

We assume that the boundary conditions are $V_0(x, j) = 0$ for all x and j. This implies that the inventory has no salvage value at the end of the horizon.

First consider the one-period expected cost $L(y, j)$ defined in (1). It is clear that the effect of inventory record inaccuracy is equivalent to having a more variable demand. Define $Z_j = D + \varepsilon(j)$. Let $\zeta_j(\cdot)$ denote the cdf of Z_j. Because $L(y, j)$ is strictly convex in y, the minimizer s_j can be found easily:

$$s_j = \arg\min_y\{L(y, j)\} = \zeta_j^{-1}\left(\frac{b-c}{b+h}\right), \quad j\in\{0, 1, 2, \ldots, T\}.$$

To proceed with our discussion, we need to introduce a concept of univariate variability orders.

Definition 1. If X and Y are random variables with distributions Θ_X and Θ_Y, respectively, we say X is more variable than Y, and write $X \geq_{var} Y$ if and only if

$$\int_a^\infty (1-\Theta_X(x))\,dx \geq \int_a^\infty (1-\Theta_Y(x))\,dx \quad \text{for all } a$$

(Shaked and Shanthikumar 1993).

Lemma 1. *If $Z_j \geq_{var} Z_i$ for $j > i$, then there exists a critical number θ such that when $(b-c)/(b+h) > \theta$, $s_j \geq s_i$ (Song 1994).*

From now on, we assume that $(b-c)/(b+h) > \theta$ and $Z_j \geq_{var} Z_i$ for $j > i$. Note that if D and ε_j are normally distributed, it can be verified that $Z_j \geq_{var} Z_i$ holds for $j > i$ and $\theta = 1/2$.

Here, we would like to draw attention to a plausible policy. Based on the model description and Lemma 1, it is conceivable that an effective policy should conduct an inspection if the cost caused by the accumulated error is higher than the inspection cost and adjust base-stock levels according to the inaccuracy level. This suggests a heuristic policy in which an inspection is conducted in a fixed cycle and the base-stock level depends on the number of periods since the last inspection. Indeed, we shall examine this heuristic policy, referred to as the cycle-count policy with adjusted base-stock levels (the CCABS policy), in detail in §4.3. Although the CCABS policy is appealing because it may be easy to implement, it is necessary to assess its effectiveness. In order to achieve this, we characterize the optimal policy in the following sections. Also, examining the optimal policy can help us develop insights for designing effective cycle-count programs.

3. Analysis of the Single-Period Problem

We consider the problem in the last period. For simplicity, we drop the subscript $t=1$. We first characterize the optimal replenishment policy and derive the resulting cost functions in the cases of inspection and no-inspection. We then characterize the optimal inspection decision by comparing these cost functions.

3.1. Derivation of the Optimal Policy

Consider the case without inspection. Suppose that the error level is $\varepsilon(j)$ at the beginning of the period.

$$N(x,j) = \min_{y\ge x} H(y,j) - cx = H(\max\{x, s_j\}, j) - cx.$$

Note that $H(y,j) = L(y,j)$, because $V_0(y,j) = 0$. Next, consider the case with inspection. In this case, x is aligned with w and

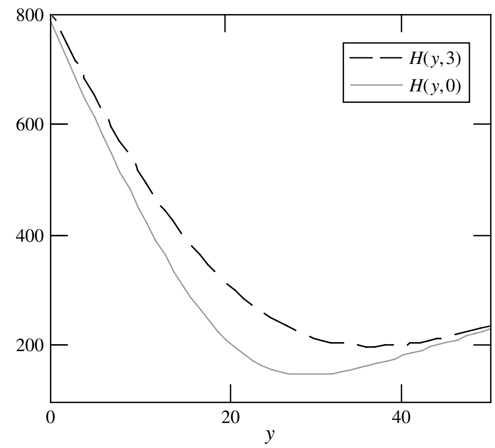
$$I(x,j) = \mathsf{E}_{\varepsilon(j)}\Big[\min_{y\ge x-\varepsilon(j)} H(y,0)\Big] - cx + K(x) = \mathsf{E}_{\varepsilon(j)}[H(\max\{x-\varepsilon(j), s_0\}, 0)] - cx + K(x).$$

To characterize the optimal inspection and replenishment policy, we need to compare the $H(y,j)$ and $H(y,0)$ functions. Figure 1 illustrates these functions for $j=3$.

Proposition 1. (1) *$H(y,j) \ge H(y,i)$, for all y and $j > i$.* (2) *$H(s_j, j) > H(s_i, i)$, for $j > i$.*

Whether an inspection will be made depends on the comparison of $N(x,j)$ and $I(x,j)$. We define the auxiliary functions $N^0(x,j)$ and $I^0(x,j)$, which are simply the backorder and holding costs of the $N(x,j)$ and $I(x,j)$, respectively.

Figure 1 **Single-Period Cost Functions $H(y,0)$ and $H(y,3)$ for $k=5$, $g=0$, $c=4$, $h=1$, $b=39$, $\sigma=8$, $\sigma_\varepsilon=4$, $\mu=20$**



$$N^0(x,j) = H(\max\{x, s_j\}, j),$$
$$I^0(x,j) = \mathsf{E}_{\varepsilon(j)}[H(\max\{x-\varepsilon(j), s_0\}, 0)].$$

The properties stated in the following proposition are helpful in deriving the optimal policy.

Proposition 2.
(1) *$N^0(x,j)$ is constant for $x \le s_j$ and convex increasing in x for $x > s_j$.*
(2) *$I^0(x,j)$ is convex nondecreasing in x.*
(3) $\lim_{x\to\infty}(N^0(x,j) - I^0(x,j)) = 0$.
(4) $\lim_{x\to-\infty}(N^0(x,j) - I^0(x,j)) > 0$.
(5) *$\partial N^0(x,j)/\partial x < \partial I^0(x,j)/\partial x$, for all x.*
(6) $\lim_{x\to-\infty} K(x) = k$.

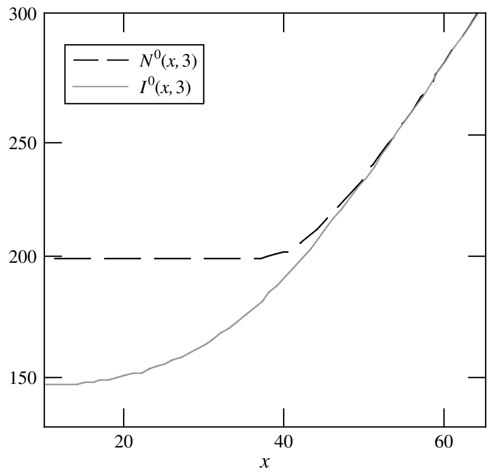
From Proposition 2, it is clear that $N^0(x,j) \ge I^0(x,j)$ for all x and j. Figure 2 illustrates these functions for $j=3$. To see the relationship between $N(x,j)$ and $I(x,j)$, define

$$\bar{k}_j = H(s_j, j) - H(s_0, 0). \tag{5}$$

This quantity is the maximum difference between the cost functions with j errors and no-error. It can be interpreted as the maximum amount we should be willing to pay to switch to the no-error cost function. Consider the three cases below:

Case 1. $k > \bar{k}_j$: From the proof of Proposition 2, Part (4), we have $\lim_{x\to-\infty} I^0(x,j) = H(s_0, 0)$. Hence,

Figure 2 Single-Period Cost Functions After Optimal Replenishment Decision Under No-Inspection, N^0, and Inspection, I^0



$\lim_{x\to-\infty}(N^0(x,j) - I^0(x,j)) > k$ and $N(x,j) \le I(x,j)$ for all x. There is no intersection. Therefore, when $k > \bar{k}_j$, an inventory inspection should not be performed because the fixed cost k is too large to compensate for the benefit of correcting the inventory record. The optimal replenishment policy is a base-stock policy: order up to s_j if $x < s_j$; do not order, otherwise.

Case 2. $0 < k < \bar{k}_j$: From Parts (3)–(5) of Proposition 2, $N(x,j)$ and $I(x,j)$ intersect only once. Let $\bar{x}_j$ be the point where $N(\bar{x}_j, j) = I(\bar{x}_j, j)$. If the inventory record x is less than or equal to $\bar{x}_j$, an inspection should be performed. The optimal replenishment policy is to order up to s_0 if x is less than s_0, and not to order otherwise. On the other hand, if the inventory record x is larger than $\bar{x}_j$, an inspection should not be performed. The optimal replenishment policy is to order up to s_j if x is less than s_j, and not to order otherwise.

Case 3a. $k = 0$ and $K'(x) = 0$ (i.e., $g = 0$): Because $N^0(x,j) \ge I^0(x,j)$ for all x, then $N(x,j) \ge I(x,j)$ and they do not intersect. Thus, an inspection should always be performed, and the optimal replenishment policy is a base-stock policy: Order up to s_0 if $x < s_0$; do not order, otherwise.

Case 3b. $k = 0$ and $K'(x) > 0$: $N(x,j)$ and $I(x,j)$ intersect only once. Thus, the optimal inspection and replenishment policies are as in Case 2.

We call this policy the optimal "inspection-adjusted base-stock policy," or the optimal IABS policy for short. From the three cases above, it is clear that $N(x,j)$ and $I(x,j)$ intersect at most once. In cases where they do not intersect, we set $\bar{x}_j$ to $-\infty$ in Case 1, and $+\infty$ in Case 3a. The following theorem formally states the optimal policy.

Theorem 1. *An inspection adjusted base-stock (IABS) policy with parameters $(\bar{x}_j, s_0, s_j)$ is optimal for the single-period problem. If $x < \bar{x}_j$, an inventory inspection is conducted and the optimal base-stock level is s_0. Otherwise, no inspection is conducted and the optimal base-stock level is s_j.*

It may seem counterintuitive that inventory record plays a role in the inspection decision even when inspection costs are fixed (i.e., $K(x) = k$). The inspection decision depends on the interaction between the inventory record and the level of inaccuracy in the system. If the inventory record is high, the manager may choose not to conduct an inspection because it is likely that there is enough inventory to meet the demand. On the other hand, if the number of errors in the system is large or inspection costs are low, then the manager may choose to conduct an inspection even if the inventory record shows a high level of inventory. Under this policy, once the inspection decision is made, the replenishment policy is a base-stock policy that depends on the demand and the number of errors in the system.

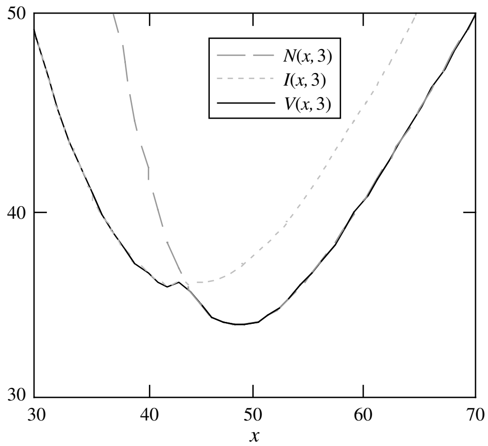
As a result of the optimal policy, $V(x,j)$ is equal to $N(x,j)$ in Case 1, $I(x,j)$ in Case 3a, and $\min\{N(x,j), I(x,j)\}$ in Cases 2 and 3b. In Cases 2 and 3b, $V(x,j)$ is not convex, as demonstrated in Figure 3.

3.2. Properties of Optimal Policies

To see how system parameters affect the inspection decisions as well as the corresponding inventory order decisions, we perform sensitivity analysis on the optimal solution. It is worth noting that $\bar{k}_j$ (defined in Equation (5)) plays an important role for inspection decisions: If $\bar{k}_j < k$, then it is never optimal to perform an inspection (i.e., $\bar{x}_j = -\infty$). On the other hand, when there are j errors accumulated in the system, an inspection is optimal for some values of x if $k < \bar{k}_j$.

Proposition 3. (1) *s_j increases in j and σ_ε,* (2) *$\bar{x}_j$ decreases in k and g,* (3) *$\bar{k}_j$ increases in j,* (4) *when demand and errors follow a Normal distribution, $\bar{k}_j$ decreases in σ, and increases in σ_ε, b, h, or c.*

Figure 3 Single-Period Cost Functions



Notes. *N*: no inspection; *I*: with inspection; *V*: optimal cost.

Part (1) indicates that the optimal base-stock level should be increased—as the number of errors accumulated in the system or the variability of errors increases—to account for increased variability due to inventory inaccuracy. Part (2) states that the interval of x values such that it is optimal to inspect (i.e., $x < \bar{x}_j$) becomes smaller as the inspection costs k or g become larger. Parts (3) and (4) state that $\bar{k}_j$ increases with the number of periods since the last inspection, variance of inventory inaccuracy, purchasing, holding and backorder costs; and decreases with demand variance. It may seem counterintuitive that inspection is less likely for products with high demand variability. In a highly variable environment, the cost functions are relatively flat around the optimal base-stock level, therefore, the benefit of an inspection is relatively less significant. It is not reasonable to incur the inspection cost to resolve the uncertainty due to the inventory records when there is significant uncertainty from other sources.

One could argue that products with higher demand would have higher error variance, which suggests that more frequent inspections are in order. On the other hand, products with higher demand may also have higher demand variance, which suggests less-frequent inspections. It is not clear beforehand which of these effects dominates. We explore this in our numerical study.

To summarize, we conclude that, all else being equal, more valuable products (higher b, c, or h), products with less variable demand, products with high error levels, or products with low inspection costs should be inspected more frequently. The values of these parameters can vary significantly across products. Hence, each of these variables should be taken into account when designing an ABC classification scheme for cycle-count or inventory accuracy programs.

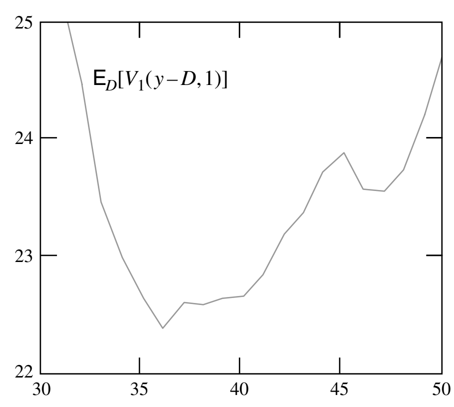
The next proposition shows that monotonic relationships exist between the optimal cost and the level of inventory record inaccuracy. These results are helpful for the analysis of the finite-horizon problem.

Proposition 4. (1) $N(x, j) \geq N(x, i)$, (2) $I(x, j) \geq I(x, i)$, *and* (3) $V(x, j) \geq V(x, i)$ *for all* x *and for* $j > i$.

4. The Finite-Horizon Problem

We first demonstrate that in the finite-horizon case the optimal policy may not have a simple structure. Consider the inventory replenishment decision in $t = 2$. Recall that $H_2(y, j) = L(y, j) + \gamma \mathsf{E}_D[V_1(y - D, j + 1)]$, where y is the inventory record after ordering. However, as seen in the previous section, $V_1(x, j)$ is not necessarily convex, so $\mathsf{E}_D[V_1(y - D, 1)]$ may not be unimodal (see Figure 4 for an example). Thus, the optimal replenishment policy derived from $H_2(y, j)$ may have more than one order-up-to level, each corresponding to a certain interval of x. Hence, the resulting cost functions associated with no-inspection and inspection decisions may intersect more than once. Consequently, the optimal inspection decision cannot

Figure 4 Example of a Not Well-Behaved Future Cost Function, $h = 1$, $c = 2$, $b = 39$, $k = 9$, $g = 0$



Notes. The pdf of demand is $f(x) = (1/16)\beta((x/16), 0.7, 0.3)$, for $0 < x < 16$, where β denotes the pdf of Beta distribution.

be characterized by a single threshold; it has to be specified for each x.

Theoretically, the optimal policy can always be characterized by using backward recursion based on (DP). However, explaining a policy with no simple structure to practitioners could be difficult, hindering the possibility of implementation. More importantly, it may not be possible to derive any managerial insights from such an optimal policy. Therefore, it is desirable to construct a simple and effective heuristic. In preliminary numerical experiments, the optimal policy obtained by solving (DP) had a structure similar to the IABS policy in every period for many commonly assumed demand distributions.[3] This observation motivates us to construct a revised dynamic program, whose cost function is a close approximation of the exact cost function in (DP). In §4.1, we show that the IABS policy is indeed optimal in this revised dynamic program. In §4.2, we implement this IABS policy for the finite-horizon problem as a heuristic. In §4.3, we consider two variants of the commonly used cycle-count policy as alternative heuristics. We shall test the quality of the lower bound and the effectiveness of the heuristics in §5.

4.1. Lower Bound

The optimality of an IABS policy can be established if a base-stock policy is the optimal replenishment policy (after a given inspection decision is carried out) and the inspection and no-inspection cost functions intersect at most once. To obtain this structure, we construct a convex lower bound to the expected cost-to-go function in each iteration of the dynamic program. Thus, the resulting cost functions are convex and base-stock policies are optimal. Further, the revised total cost functions have properties analogous to Parts (1)–(5) of Proposition 2, which ensure the single-intersection property in the finite horizon. Below we describe the details of the procedure.

The construction of lower-bound cost functions starts from $t=2$. Recall that when $t=2$, the optimal cost function without inspection in (DP) is

$$N_2(x, j) = \min_{y\ge x}\{L(y, j) + \gamma \mathsf{E}_D[V_1(y-D, j+1)]\} - cx.$$

[3] Table 1 presents the IABS policy for a few examples.

We construct a convex lower-bound function $B_1(y, j+1)$ for $\mathsf{E}_D[V_1(y-D, j+1)]$. (The construction of these convex lower-bound functions is described in the algorithm in Appendix A.) Replacing $\mathsf{E}_D[V_1(y-D, j+1)]$ with $B_1(y, j+1)$, the resulting cost function $\underline{N}_2(x, j)$ is a convex lower bound to $N_2(x, j)$.

For the optimal cost function with inspection $I_2(x, j)$, we replace $\mathsf{E}_D[V_1(y-D, 1)]$ with $B_1(y, 1)$, and the resulting function

$$I'_2(x, j) = \mathsf{E}_{\varepsilon(j)}\Big[\min_{y\ge x-\varepsilon(j)}\{L(y, 0) + \gamma B_1(y, 1)\}\Big] - cx + K(x)$$

is a lower bound to $I_2(x, j)$. Because $L(y, 0) + \gamma B_1(y, 1)$ is convex, by Jensen's inequality,

$$\underline{I}_2(x, j) = \min_{y\ge x}\{L(y, 0) + \gamma B_1(y, 1)\} - cx + K(x)$$

is a lower bound to $I'_2(x, j)$. Thus, $\underline{I}_2(x, j) \le I_2(x, j)$. Consequently, the optimal expected discounted cost

$$\underline{V}_2(x, j) = \min\{\underline{N}_2(x, j), \underline{I}_2(x, j)\}$$

is the lower bound to the exact optimal cost $V_2(x, j)$. Repeat this procedure for $t = 3, 4, \ldots, T$: first construct $B_{t-1}(y, j+1)$ for $\mathsf{E}_D[\underline{V}_{t-1}(y-D, j+1)]$ for all $0 \le j \le T-t+1$. Then, compute $\underline{N}_t(x, j)$, $\underline{I}_t(x, j)$, and then $\underline{V}_t(x, j)$ for all j.

The formal description of the revised dynamic program (RDP) is as follows.

$$\text{(RDP)}\quad \underline{V}_t(x, j) = \min\{\underline{N}_t(x, j), \underline{I}_t(x, j)\}, \quad \text{for } t\ge 2,$$
$$\underline{V}_1(x, j) = \min\{N_1(x, j), I_1(x, j)\},$$

where

$$\underline{N}_t(x, j) = \min_{y\ge x} \underline{H}_t(y, j) - cx, \quad 1\le j\le T-t+1,$$
$$\underline{I}_t(x, j) = \min_{y\ge x} \underline{H}_t(y, 0) - cx + K(x), \quad 1\le j\le T-t+1,$$
$$\underline{H}_t(y, j) = L(y, j) + \gamma B_{t-1}(y, j+1), \quad 0\le j\le T-t+1.$$

The detailed algorithm for constructing $B_{t-1}(y, j+1)$ for $t \ge 2$, and $j \ge 0$ is given in Appendix A. Here, we only summarize the key idea of how these lower bound functions are constructed. $B_{t-1}(y, j+1)$ is a lower convex envelope of $\mathsf{E}_D[\underline{V}_{t-1}(y-D, j+1)]$. It has the same limiting slopes of $\mathsf{E}_D[\underline{V}_{t-1}(y-D, j+1)]$ shown below:

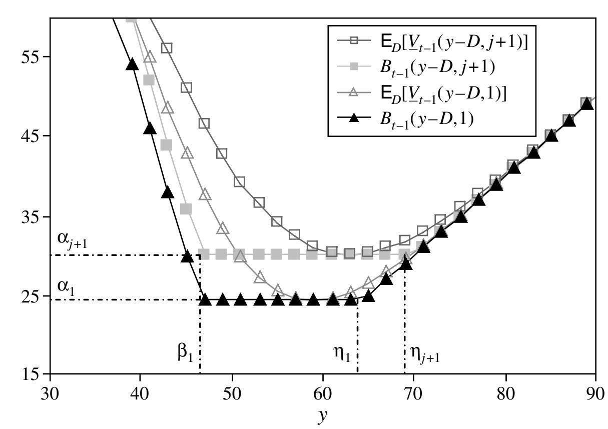
$$\lim_{y\to-\infty}\frac{\partial}{\partial y}\mathsf{E}_D[\underline{V}_{t-1}(y-D,j+1)] = \lim_{x\to-\infty}\frac{\partial}{\partial x}\underline{V}_{t-1}(x,j+1) = -c,$$

$$\lim_{y\to\infty}\frac{\partial}{\partial y}\mathsf{E}_D[\underline{V}_{t-1}(y-D,j+1)] = \lim_{x\to\infty}\frac{\partial}{\partial x}\underline{V}_{t-1}(x,j+1) = \bar{h}_{t-1},$$

where $\bar{h}_t = \sum_{j=1}^{t}\gamma^{j-1}h$. More specifically, $B_{t-1}(y, j+1)$ consists of three linear line segments: the leftmost one has a slope of $-c$, the middle one has a slope of zero, and the rightmost one has a slope of $\bar{h}_{t-1}$. $B_{t-1}(y, j+1)$ is the tightest such bound for $\mathsf{E}_D[\underline{V}_{t-1}(y-D, j+1)]$ for $j > 1$. $B_{t-1}(y, 1)$ may not be the tightest bound to $\mathsf{E}_D[\underline{V}_{t-1}(y-D, 1)]$: The leftmost part may be shifted left to ensure that the flat middle piece does not start before $\min\{\beta_2, \ldots, \beta_{T-t+2}\}$. Figure 5 demonstrates these lower bounds and some of the variables used in the construction for an example. Observe in this example that $B_{t-1}(y, j+1) \geq B_{t-1}(y, 1)$ and $\partial B_{t-1}(y, j+1)/\partial y \leq \partial B_{t-1}(y, 1)/\partial y$ for all y. The last inequality holds because the two functions have the same slope except for the region where $B_{t-1}(y, 1)$ is increasing and $B_{t-1}(y, j+1)$ is flat.

This completes the definition of (RDP). The following can be immediately observed. $\underline{V}_1(x, j) = V_1(x, j)$, $1 \leq j \leq T$. Also, $\underline{V}_t(x, j) \leq V_t(x, j)$ for any t, x and $1 \leq j \leq T-t+1$. Hence, (RDP) provides a lower bound to the (DP), i.e., $\underline{V}_T(x, 1) \leq V_T(x, 1)$.

Figure 5 Expected Future Cost Functions and the Constructed Bounds, $t = 2$ and $j = 5$



Let the minimizers of $\underline{H}_t(y, j)$ be $\underline{s}_{tj}$:

$$\underline{s}_{tj} = \arg\min_y \underline{H}_t(y, j).$$

Define

$$\underline{N}_t^0(x, j) = \min_{y\geq x}\underline{H}_t(y, j) = \underline{H}_t(\max\{x, \underline{s}_{tj}\}, j),$$

$$\underline{I}_t^0(x, j) = \min_{y\geq x}\underline{H}_t(y, 0) = \underline{H}_t(\max\{x, \underline{s}_{t0}\}, 0).$$

Because the cost functions at $t = 1$ are the same in DP and RDP, $\underline{s}_{1j} = s_{1j}$, where s_{1j} is the minimizer of $H_1(y, j)$. The following propositions state several properties that are similar to the single-period case.

Proposition 5. *For any t, (1) $\underline{H}_t(y, j)$ is convex, (2) $\underline{H}_t(y, j) \geq \underline{H}_t(y, 0)$ for all y, and (3) $\underline{s}_{tj} \geq \underline{s}_{t0}$, and (4) $\partial\underline{H}_t(y, 0)/\partial y \geq \partial\underline{H}_t(y, j)/\partial y$ for all $y \geq \underline{s}_{tj}$.*

Now, we show that an IABS policy is optimal for RDP after establishing the single-intersection property in the next proposition.

Proposition 6. *For any $t \geq 2$, and $1 \leq j \leq T-t+1$,*
(1) $\underline{N}_t^0(x, j)$ is constant for $x \leq \underline{s}_{tj}$ and convex increasing for $x > \underline{s}_{tj}$.
(2) $\underline{I}_t^0(x, j)$ is constant for $x \leq \underline{s}_{t0}$ and convex increasing for $x > \underline{s}_{t0}$.
(3) $\lim_{x\to\infty}(\underline{N}_t^0(x, j) - \underline{I}_t^0(x, j)) = 0$.
(4) $\lim_{x\to-\infty}(\underline{N}_t^0(x, j) - \underline{I}_t^0(x, j)) > 0$.
(5) $\partial\underline{N}_t^0(x)/\partial x < \partial\underline{I}_t^0(x)/\partial x$, for $x > \underline{s}_{tj}$.
(6) $\underline{N}_t(x, j)$ and $\underline{I}_t(x, j)$ intersect at most once.

As in the single-period case, we define the intersection point as $\bar{x}_{tj}$ for each j in period t. The following theorem describes the optimal policy for the revised dynamic program. The proof is similar to the analysis that led to Theorem 1.

Theorem 2. *An inspection-adjusted base-stock (IABS) policy with parameters $(\bar{x}_{tj}, \underline{s}_{t0}, \underline{s}_{tj})$ in period t and state j is optimal for (RDP). If $x_t < \bar{x}_{tj}$, an inspection is performed and optimal base-stock level is $\underline{s}_{t0}$. Otherwise, no inspection is performed and optimal base-stock level is $\underline{s}_{tj}$.*

The following proposition describes the impact of system parameters on the optimal policy. Define

$$\bar{k}_{tj} = \underline{H}_t(\underline{s}_{tj}, j) - \underline{H}_t(\underline{s}_{t0}, 0).$$

Proposition 7. *For $t \geq 2$, (1) $\bar{x}_{tj}$ decreases in k and g and increases in j. (2) $\bar{k}_{tj}$ increases in j (3) when demand*

and errors follow a Normal distribution, $\bar{k}_{tj}$ increases with σ_ε.

Part (1) indicates that the intersection point shifts to the left with inspection cost parameters and to the right with the number of errors accumulated. Part (2) describes the effect of system parameters on $\bar{k}_{tj}$. Because $\bar{k}_{tj}$ increases in j without bound, there exists j^* such that $\bar{k}_{tj} > k$ for $j \geq j^*$, meaning that inspections will be not be performed for any value of x if $j < j^*$. Part (3) suggests that inspections should be conducted more frequently for products with higher error variance. Although we are not able to replicate all of the sensitivity results from the single-period case, our numerical study shows that results analogous to Proposition 3 seem to hold for the finite horizon as well.

4.2. The IABS Heuristic

The IABS heuristic uses the solutions from RDP as the IABS policy parameters in each period. If $x < \bar{x}_{tj}$, an inspection is performed and inventory is ordered up to $\underline{s}_{t0}$. Otherwise, no inspection is performed and inventory is ordered up to $\underline{s}_{tj}$. As we shall see in the next section, the lower bounds developed in RDP approximate the original cost functions very well. Therefore, we suspect that the IABS heuristic is an effective solution to the original problem.

Examination of the IABS heuristic solutions leads to some interesting observations. To illustrate, we present the policy parameters from the IABS heuristic for five examples in Table 1. Examining the heuristic parameters confirms our comparative static results. In particular, both the order-up-to levels and the inspection threshold level increase with j in all examples. Comparing the second and third examples with the base case confirms that the inspection threshold level decreases with less variable periodic error and higher fixed inspection cost. Comparing the first periods of the fourth example and the base case, we see that the number of cases where no inspection

Table 1 **Policy Parameters of the IABS Heuristic**

		Example 1			Example 2 $\sigma_\varepsilon = 1$			Example 3 $k = 10$			Example 4 $\sigma = 8$			Example 5 $b/(b+h) = 0.975$		
t	j	$\underline{s}_{t0}$	$\underline{s}_{tj}$	$\bar{x}_{tj}$	$\underline{s}_{t0}$	$\underline{s}_{tj}$	$\bar{x}_{tj}$	$\underline{s}_{t0}$	$\underline{s}_{tj}$	$\bar{x}_{tj}$	$\underline{s}_{t0}$	$\underline{s}_{tj}$	$\bar{x}_{tj}$	$\underline{s}_{t0}$	$\underline{s}_{tj}$	$\bar{x}_{tj}$
6	1	26	27	$-\infty$	26	27	$-\infty$	26	27	$-\infty$	33	34	$-\infty$	28	28	$-\infty$
5	1	26	27	$-\infty$	26	27	$-\infty$	26	27	$-\infty$	33	34	$-\infty$	28	28	$-\infty$
	2	26	28	$-\infty$	26	27	$-\infty$	26	28	$-\infty$	33	34	$-\infty$	28	30	$-\infty$
4	1	26	27	$-\infty$	26	27	$-\infty$	26	27	$-\infty$	33	34	$-\infty$	28	28	$-\infty$
	2	26	28	$-\infty$	26	27	$-\infty$	26	28	$-\infty$	33	34	$-\infty$	28	30	$-\infty$
	3	26	29	27	26	27	$-\infty$	26	29	$-\infty$	33	34	$-\infty$	28	30	28
3	1	26	27	$-\infty$	26	27	$-\infty$	26	27	$-\infty$	33	34	$-\infty$	28	28	$-\infty$
	2	26	28	$-\infty$	26	27	$-\infty$	26	28	$-\infty$	33	34	$-\infty$	28	30	$-\infty$
	3	26	29	$-\infty$	26	27	$-\infty$	26	29	$-\infty$	33	34	$-\infty$	28	30	28
	4	26	29	25	26	27	$-\infty$	26	29	$-\infty$	33	34	27	28	31	30
2	1	26	27	$-\infty$	26	27	$-\infty$	26	27	$-\infty$	33	33	$-\infty$	28	28	$-\infty$
	2	26	28	$-\infty$	26	27	$-\infty$	26	28	$-\infty$	33	33	$-\infty$	28	30	$-\infty$
	3	26	29	$-\infty$	26	27	$-\infty$	26	29	$-\infty$	33	34	31	28	30	$-\infty$
	4	26	29	26	26	27	$-\infty$	26	29	25	33	34	33	28	31	26
	5	26	30	28	26	27	$-\infty$	26	30	27	33	34	34	28	31	28
1	1	23	23	$-\infty$	23	23	$-\infty$	23	23	$-\infty$	25	26	$-\infty$	25	25	$-\infty$
	2	23	23	21	23	23	$-\infty$	23	23	$-\infty$	25	26	$-\infty$	25	25	25
	3	23	24	23	23	24	$-\infty$	23	24	$-\infty$	25	26	$-\infty$	25	26	26
	4	23	24	24	23	24	$-\infty$	23	24	19	25	26	24	25	26	27
	5	23	24	25	23	24	$-\infty$	23	24	22	25	26	25	25	27	28
	6	23	25	25	23	25	$-\infty$	23	25	23	25	26	26	25	27	29

Notes. Example 1 is the base example with parameters $T = 6$, $\mu = 20$, $\sigma = 4$, $\sigma_\varepsilon = 2$, $k = 5$, $g = 0$, and $b/(b+h) = 0.95$ under normal demand and error distribution. Parameters of Examples 2–5 are the same as the base example except for the parameter listed in the column heading.

is performed (i.e., $\bar{x}_{tj} = -\infty$) increases as demand variability increases. This is due to the fact that the importance of data inaccuracy and the value of inspections decrease with an increase in demand variability. The fifth example shows that the number of no-inspection cases decreases with a higher service level (i.e., inspection is more likely).

Most of the time, the implementation of the heuristic policy is simpler than its formal structure suggests. Observe that the IABS heuristic parameters for $t \geq 2$ are almost stationary. The optimal base-stock levels for each j are the same across periods. The inspection threshold levels are also similar across periods. It is also worth noting that the inspection threshold levels are either very close to the base-stock levels or equal to $-\infty$, suggesting "no-inspection." Because demand is nonnegative and the base-stock levels tend to be stationary, it is highly likely that $x_t < \bar{x}_{tj}$ if $\bar{x}_{tj}$ is finite. This implies that the IABS heuristic will recommend inspection for most realizations of x_t if $\bar{x}_{tj}$ is finite. Given that $\bar{x}_{tj}$ is finite when $\bar{k}_{tj} > k$, and $\bar{k}_{tj}$ is increasing in j, the number of periods since last inspection (j) becomes the major factor in determining whether an inspection should be performed. Consider the implementation of the IABS heuristic in Example 1 of Table 1, starting from $t = 6$. Assume a sample path with $x_t \leq 20$ for all t. In Period 6, because $j = 1$, and $\bar{x}_{61} = -\infty$, no inspection is performed. The error will be carried to Period 5, i.e., $j = 2$. Because $\bar{x}_{52} = -\infty$, again no inspection is performed. In Period 4, with $j = 3$, $x_4 < \bar{x}_{43} = 27$, an inspection is performed, and so on. Table 2 summarizes the policy implementation.

In this example, IABS inspects every three periods and adjusts the order-up-to levels according to the number of errors accumulated. This certainly resembles a cycle-count policy with an inspection cycle of three periods. In §5.2, we report a simulation study and find in all stationary examples that the IABS heuristic indeed resembles a cycle-count policy in implementation.

Table 2 Implementation of the IABS Heuristic in Example 1 of Table 1

t	j	Policy implementation in Example 1
6	1	Don't inspect, order up to 27
5	2	Don't inspect, order up to 28
4	3	Inspect, order up to 26
3	1	Don't inspect, order up to 27
2	2	Don't inspect, order up to 28
1	3	Inspect, order up to 23

4.3. Cycle-Count Policy

Cycle-count programs are the most common approach to dealing with inventory inaccuracy in retail, distribution, and manufacturing industries. Briefly, in a cycle-count program, a cycle length m is chosen for a product (mostly for a family of products), and an inspection is performed every m periods. Inventory replenishment follows a base-stock policy based on inventory records. Here, we consider two cycle-count policies that differ in their setting of base-stock levels. The first is a standard cycle-count policy, denoted CC. It employs a base-stock policy with base-stock level s_t, where s_t is the smallest y satisfying

$$\Pr\{D_t \leq y\} \geq b_t/(b_t + h_t).$$

Cycle-count ignores inventory inaccuracy in setting the base-stock levels. This is consistent with how cycle-count policies are implemented in practice. However, we showed in the previous sections that base-stock levels should be adjusted to account for the number of errors accumulated in the system. To reflect this insight, we introduce a cycle-count policy with adjusted base-stock levels, denoted as CCABS. CCABS sets the base-stock levels such that $s_t(j)$ is the smallest y satisfying

$$\Pr\{D_t + \varepsilon_t(j) \leq y\} \geq b_t/(b_t + h_t), \quad 1 \leq j \leq m,$$

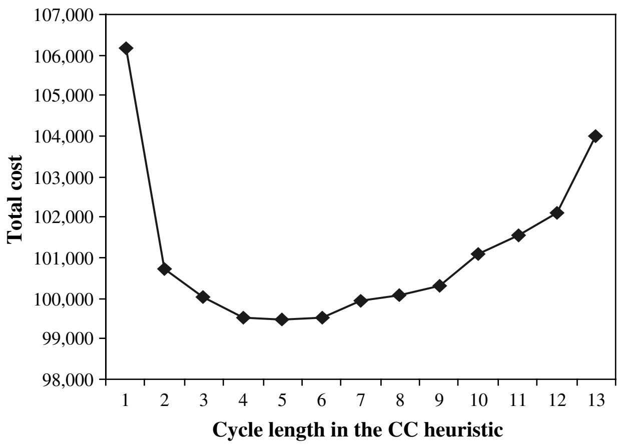
where j is the number of periods since the last cycle count. We illustrate the performance of these cycle-count policies in §5.2.

Figure 6 illustrates how total cost changes with cycle length in the CC heuristic. It is possible to find the best m value for a particular instance of the problem through simulation or computation of the value function. Note that although cycle-count heuristics perform inspections in a fixed cycle, the base-stock levels are adjusted for nonstationarity in the data.

5. Numerical Study

This section is composed of four parts. In §5.1, we investigate the effectiveness of the lower bound cost function and the IABS heuristic. In §5.2, we study

Figure 6 Total Cost from the Cycle-Count Heuristic (CC) in Example 1 of Table 1



how the frequency of inspections of the IABS heuristic changes with respect to system parameters via simulation. In §5.3, we assess the performance of all heuristics and compare them with other benchmarks. In §5.4, we quantify the total loss due to the inventory record inaccuracy problem and assess the true value of RFID systems.

The system parameters in this study are: $T = 24$, $c = 2$, $h = 1$, $b \in \{5, 9, 39\}$, $\mu = 20$, $\sigma \in \{2, 8\}$, $\sigma_\varepsilon \in \{\sigma/4, \sigma/2\}$, $k \in \{5, 40\}$, and $g = 0$ (i.e., $K(x) = k$). The holding and backorder cost rates are chosen to reflect 84%, 90%, and 97.5% service levels, respectively. These 24 experiments are repeated under three scenarios, resulting in a total of 72 experiments. In the first scenario, denoted by S, all data is stationary. In the second scenario, denoted by NS1, purchase and inventory costs are nonstationary such that $c_t = c(1.5 - t/T)$ and $h_t = h(1.5 - t/T)$. In the third scenario, denoted by NS2, demand parameters are nonstationary such that $\mu_t = \mu(0.5 + t/T)$, $\sigma_t = \sigma(0.5 + t/T)$, and $\sigma_{\varepsilon t} = \sigma_{\varepsilon t}(0.5 + t/T)$. Clearly, there are many other plausible scenarios for the nonstationary case and it is impossible to consider all. NS1 and NS2 together span many parameters and two reasonable trends. In all cases, the error follows a Normal distribution and demand follows a truncated Normal distribution. These parameters will be used for all tests in this section.

5.1. The Effectiveness of IABS Heuristic

We verify the effectiveness of the lower-bound cost by comparing the optimal expected discounted cost $V_T(x, 1)$ with the lower-bound cost $\underline{V}_T(x, 1)$, the solutions from the (DP) and (RDP), respectively. For each experiment, we consider the set $X = \{-2\mu, -2\mu + 1, \ldots, 4\mu\}$ and find the average and maximum percentage cost error within this range. Specifically,

Table 3 Average Performance of the Lower-Bound Cost and the IABS Heuristics

		Lower bound		IABS heuristic	
Scenario	T	Average (%)	Maximum (%)	Average (%)	Maximum(%)
S	6	−1.9	−8.2	0.2	0.3
	12	−0.9	−4.1	0.2	0.2
	24	−0.4	−2.3	0.1	0.2
NS1	6	−2.2	−8.6	0.2	0.3
	12	−1.0	−4.0	0.2	0.2
	24	−0.4	−1.7	0.1	0.2
NS2	6	−8.8	−28.8	0.3	0.5
	12	−5.4	−9.5	0.3	0.3
	24	−3.1	−5.0	0.2	0.2

$$\operatorname*{Max}_{x\in X} (\underline{V}_T(x, 1) - V_T(x, 1))/V_T(x, 1); \qquad (6)$$

$$\operatorname*{Average}_{x\in X} (\underline{V}_T(x, 1) - V_T(x, 1))/V_T(x, 1). \qquad (7)$$

Note that the lower-bound cost is smaller than the optimal cost for any value of x. Thus, the differences are negative. We then compute the average of the maximum percentage gaps (and the average percentage gaps) over 72 experiments for each T. The results are reported in Table 3 under the heading Lower bound. In general, the lower bound provides a close approximation to the exact cost function.

Next, we examine the effectiveness of the IABS heuristic policy. We calculate the total expected discounted cost function for the corresponding IABS policy. We denote the total cost as $V_T^*(x, 1)$. We compute the respective maximum and average percentage error of the IABS heuristic by substituting $\underline{V}_T(x, 1)$ in (6) and (7) with $V_T^*(x, 1)$. The results are reported in Table 3 under the heading IABS. As seen in the table, the cost from the IABS heuristic is within 0.2% of the optimal cost on average. The performance of the IABS heuristic is similar in stationary and nonstationary scenarios. This suggests that the IABS heuristic is near optimal in these numerical tests.

5.2. Sensitivity Analysis on Inspection Cycles

We study the implementation of IABS heuristic via simulation with 2,000 replications for each of the

Table 4 Average and Variance of the Cycle Length from the Implementation of the IABS Heuristics

		Cycle length of IABS		Cycle length of DP	
Parameter	Value	Average	Variance	Average	Variance
k	5	5.6	0.4	5.7	0.3
	40	16.5	0.1	16.5	0.2
σ_ε	$\sigma/4$	13.6	0.2	13.7	0.2
	$\sigma/2$	8.5	0.3	8.5	0.3
$\sigma^\dagger$	2	14.8	0.2	14.8	0.2
	8	7.3	0.3	7.3	0.3
$\sigma^\ddagger$	2	5.5	0.8	5.3	0.6
	8	8.8	0.8	8.9	0.2
b	5	12.2	0.2	12.1	0.2
	9	11.3	0.4	11.4	0.2
	39	9.7	0.3	9.8	0.3
Scenario	S	12.0	0.1	11.9	0.1
	NS1	11.7	0.4	11.9	0.4
	NS2	9.6	0.3	9.4	0.2

†The averages of sigma are from the complete set of numerical experiments with $\sigma_\varepsilon = \sigma/4$ or $\sigma/2$.

‡The averages of σ are from an extra set of experiments with of $\sigma_\varepsilon = 1$, $k = 5$, $b = 19$ only.

72 instances. We report statistics on inspection intervals of the IABS heuristic and the optimal policy in Table 4. The inspection cycles resulting from these policies are similar in all cases. From this observation and the examination of the optimal decisions of these policies for several examples, we conclude that the IABS policy not only results in costs that are close to the optimal cost, but also its inspection and replenishment decisions are similar to those of the optimal solution.

Table 4 shows that the average of the cycle length standard deviation is almost zero in the stationary case, meaning that the cycle length of each experiment is fairly consistent across periods and between simulation replications. This suggests that both the optimal solution and the IABS heuristic performs inspections in fixed cycles. However, the variance of the cycle lengths increases in the nonstationary cases. The average cycle lengths reported in Table 4 confirm our analytical findings in Propositions 3 and 7: Inspections are less frequent as k increases or as σ_ε and service level decrease. The effect of σ alone can be seen in the part of Table 4 marked with ‡: When σ_ε is kept constant, the average cycle length increases (inspection frequency decreases) with σ as Proposition 3 suggests. Products with higher demand may have both higher demand and error variances. When σ_ε is proportional to σ, the effect of an increase in σ is a combination of the effects of σ and σ_ε. As seen in the part of Table 4 marked with †, the effect of σ_ε dominates and more frequent inspections are warranted. Most importantly, the analytical findings based on the IABS policy hold for the optimal solution as well, validating our effort to find a heuristic with a simple structure. We would like to emphasize that these sensitivity results can be instrumental in designing an inspection schedule or an ABC classification scheme for a cycle-count program.

5.3. Effectiveness of Cycle-Count Policies and Comparison with Other Benchmarks

We compare the performance of the following heuristics and benchmark policies.

NE: No-error system,

DP: The optimal solution,

IABS: IABS heuristic,

CCABS: Cycle-count heuristic, base-stock levels adjusted for errors,

CC: Cycle-count heuristic, base-stock levels not adjusted for errors,

NI: No inspection, base-stock levels adjusted for errors,

IG: Ignore errors: No inspection, base-stock levels not adjusted for errors,

AI: Always inspect.

The no-error system (NE) provides the lowest possible cost that can be achieved through preventive strategies, such as the implementation of RFID systems. Note also that a base-stock policy is optimal in the NE case. All other policies are applied to the system with inventory record inaccuracy. DP, IABS, CCABS, and CC are described in detail in the previous sections. NI never performs an inspection, but keeps track of the accumulated error and adjusts the base-stock level accordingly. IG ignores inaccuracy issues and simply implements the NE solution. We posit IG as the worst-case scenario. AI performs an inspection in every period. Let V^{policy} denote the cost function of a policy or system. Clearly, we have $V^{AI} = V^{NE} + Tk$.

We implement the CCABS heuristic with three parameters for each experiment. "m from IABS" denotes the average cycle length in the IABS implementation for a particular experiment, rounded to

Table 5 Average Cost of the Proposed Heuristics and Benchmark Policies Compared to the No-Error System

	Scenario			
Policy—NE (%)	S	NS1	NS2	Average
DP	6.2	6.5	6.2	6.3
IABS	6.3	6.7	6.4	6.5
CCABS(best *m*)	6.5	7.1	6.6	6.8
CCABS(*m* from IABS)	6.7	7.3	6.8	7.0
CCABS(worst *m*)	47.9	54.9	44.1	49.3
CC(best *m*)	7.9	8.9	7.5	8.1
NI	11.0	12.2	13.3	12.2
IG	18.4	20.9	25.4	21.5
AI	77.3	88.6	69.2	78.5

the nearest integer. "Best m" ("worst m") denotes the value of m that yields the lowest (highest) cost for each experiment under CCABS policy. For the cycle-count heuristic, we only report the "best m" case.

Table 5 reports the simulation results for all policies. NE is taken as the benchmark and the values reported for each policy are the average percentage cost difference with the NE system, i.e., $100\% \times (V^{policy}/V^{NE} - 1)$.

We make the following observations from the table. First, IABS achieves costs that are within 0.2% of the optimal solution. Second, CCABS(best m) performs remarkably well and achieves only 0.5% higher than the optimal solution on average. CC(best m) does not perform as well: Its average cost is 1.6% higher than the IABS solution. Comparing the results of CCABS with best m and worst m, we see that not choosing the right cycle length can be very costly. On the other hand, CCABS(m from IABS) achieves results very close to the CCABS(best m). This suggests that using the average number of periods between inspections from the IABS heuristic may be a good method for choosing the proper cycle length for the cycle-count policy. Although cycle-count heuristics perform very well in these examples, it must be noted that their performance deteriorates in certain cases. Table 6 details the performance of the cycle-count heuristics and the IABS heuristic with respect to changes in each parameter. We see that the performance of the standard cycle-count heuristic deteriorates significantly as k, σ_ε increase, and as the service level decreases. CCABS seems more robust, its performance deteriorates slightly with higher k, σ, and under nonstationary cost data. The reason that CCABS still performs

Table 6 Average Percentage Cost of the Cycle-Count Heuristics Compared to the IABS Heuristic

		Policy—DP (%)		
Parameter	Value	IABS	CCABS(best *m*)	CC(best *m*)
k	5	0.2	0.3	0.5
	40	0.1	0.6	3.1
σ_ε	$\sigma/4$	0.2	0.4	0.8
	$\sigma/2$	0.2	0.5	2.8
σ	2	0.2	0.3	2.1
	8	0.1	0.6	1.5
b	5	0.1	0.5	0.8
	9	0.2	0.5	1.3
	39	0.2	0.4	3.3
Scenario	S	0.1	0.3	1.7
	NS1	0.1	0.6	2.4
	NS2	0.2	0.4	1.3
$g^\dagger$	0	0.2	0.5	1.5
	0.2	0.4	0.6	1.6

†The averages for $g > 0$ are from a new set of 72 experiments.

well in the NS2 case may be due to the cycle lengths staying relatively stable due to the effects of increased σ and σ_ε balancing each other out.

Finally, to assess the effectiveness of IABS and cycle-count heuristics in the case of increasing inspection costs, we tested the original 72 experiments, this time with $g = 0.2$. Comparing the $g > 0$ case with the $g = 0$ case in Table 6, we see that the performances of IABS, CCABS, and CC do not seem to be affected significantly.

Given the relative simplicity of cycle-count programs, the performance of CCABS makes it an appealing policy. If a company implements a cycle-count program, we recommend that CCABS is adopted instead of the standard cycle-count policy. The best CCABS cycle length for each product can be chosen via simulation or by using the average cycle length of the IABS policy. If a cycle-count program requires products to be classified into groups with different inspection schedules, the sensitivity analysis results on the IABS heuristic could be useful in designing an ABC classification scheme. CCABS performance slightly deteriorates when inspections are expensive, demand is more variable, and when there is nonstationarity in the data. In those cases, using the IABS policy is the best option, because IABS provides near-optimal solutions for all cases in the numerical study.

5.4. Value of RFID Systems

As mentioned before, one of the benefits of RFID systems is their ability to eliminate or significantly reduce inventory record inaccuracy. Although industry reports indicate that RFID systems have not yet been perfected, we assume here that RFID systems are able to achieve perfect accuracy. The difference between IG and NE is the measure of the potential benefit of eliminating inaccuracy problems completely, which is on average 21.5% in our numerical study. However, this difference does not represent the true value of RFID systems, as IG is not the best one can do without RFID. Rather, it is the loss due to inventory inaccuracy and policies that do not take inaccuracy into account. A simple policy such as the NI policy that takes inaccuracy into account when choosing replenishment levels reduces the loss due to record inaccuracy to, on average, 12.2%. A more relevant comparison is between NE and a system that has inaccurate records but is managed near optimally. Therefore, we compare NE with IABS.[4] Near-optimal inspection and replenishment policies such as the IABS heuristic secure more than two-thirds of this potential benefit by reducing costs due to record inaccuracy to, on average, within 6.6% of NE. This suggests that the "true" value of accurate inventory information enabled by RFID systems in this numerical study is 6.6% of the inventory-related costs and one-third of the increased costs due to inaccuracy. We conclude that corrective strategies can be quite effective in dealing with inventory record inaccuracy problems, even if a company cannot invest in advanced technologies or other preventive strategies.

6. Discussion

Below we discuss extensions and limitations of the basic model. Some of the extensions point towards additional interesting issues, which are beyond the scope of the current paper.

6.1. Exception Reports

In reality, it is possible that the magnitude of the error can be automatically learned when backorders occur, e.g., the material handler finds an item out of stock but the inventory record shows it to be in stock. This may be viewed as a free inspection. Another example would occur when business software generates an exception report when the shipping of an item with a positive inventory record has been delayed for a long time. (It must be noted that most companies do not have processes in place for updating inventory records in such cases.) Our model assumes that this information is not used to update inventory records. However, the dynamic program can be modified to incorporate this scenario by replacing the original $H_t(y,j)$ function with the following:

$$\begin{aligned} H_t(y,j) = {} & L(y,j)+\mathsf{P}(y\ge D+\varepsilon(j)) \\ & \cdot\gamma \mathsf{E}_{D,\varepsilon(j)}[V_{t-1}(y-D,j+1)\mid y\ge D+\varepsilon(j)] \\ & +\mathsf{P}(y< D+\varepsilon(j)) \\ & \cdot\gamma \mathsf{E}_{D,\varepsilon(j)}[V_{t-1}(y-\varepsilon(j)-D,1)\mid y< D+\varepsilon(j)]. \end{aligned}$$

The expectations over $\varepsilon(j)$ are conditional expectations. The first term is the expected inventory costs from the current period, the second term is the expected future total cost if no backorder occurs, and the last term is the expected future total cost if backorders occur. Note that in the last term, if backorders occur, the manager aligns the inventory record with physical inventory. The analysis and the results in this paper can be shown to be valid in this setting. Interestingly, the manager now has an incentive to order less than the system without exception reports in order to induce backorders and get a free inspection.

Note that the above formulation is not a full learning model, because it does not allow for learning from the nonoccurrence of exception reports. Simply knowing that no backorders have occurred in the last j time periods imposes a downward bias on the distribution of $\varepsilon(j)$ so that its mean will no longer be zero. Also, the distribution of the error in the system depends on the sequence of demand observations since the last inspection. A correct formulation for this model requires a dynamic program with a state space that carries the history of demand and exception reports since the last inspection. This leads to a state dimensionality problem.

[4] Comparing IABS (which has the option to perform inspections) with NI and NE may not be a fair comparison if the fixed inspection cost were too small. This is not the case in our numerical study as the average cost of the AI policy is 80% higher than NE.

6.2. Demand-Dependent Errors

Our model can be extended to the setting in which error depends on demand realization as long as the error mean remains zero. For example, the dynamic program can be modified by replacing $\mathsf{E}_D[V_{t-1}(y - D, j+1)]$ with $\mathsf{E}_D[V_{t-1}(y - D, j+\alpha(D))]$ where $\alpha(D)$ determines the number of errors added in that period. Our formulation and the analytical results hold. As a simple example, suppose that $\alpha(D) = \{0, \text{ if } D \le \mu;\ 1, \text{ otherwise}\}$, then we can redefine

$$H_t(y, j) = L(y, j) + \mathsf{P}(D \le \mu)\mathsf{E}_{D,\varepsilon}[V_{t-1}(y-D, j) \mid D \le \mu] + \mathsf{P}(D > \mu)\mathsf{E}_{D,\varepsilon}[V_{t-1}(y-D, j+1) \mid D > \mu],$$

and all our results are valid for this system as well. If the error distribution in each period is correlated with demand, then j alone would not be able to carry sufficient information about the error distribution. Again, the dynamic program must be reformulated with a state space including the demand observations since the last inspection.

6.3. Imperfect Inspection

In some cases, the inspector may miscount inventory or enter incorrect information when updating data. Specifically, suppose that the error that remains in the system after conducting an inspection is $\alpha\varepsilon(j)$, where $0 \le \alpha \le 1$. Our model can be extended to accommodate this scenario by replacing $V_{t-1}(y - D, 1)$ with $V_{t-1}(y - D, 1 + \alpha j)$ in (DP). The optimality of the IABS policy for the revised dynamic program can be verified in this model.

However, the partial inspection may change our belief about the distribution of the error remaining in the system, which will result in a more complex problem than above. Also, the partial-inspection model leads to other statistical issues, such as determining the sample size and making statistical inferences about the distribution of remaining error.

6.4. General Error Distribution

We assumed an error distribution with zero mean in our model. This may be a restrictive assumption in some situations. If a transaction error of one type is repeated frequently (e.g., entering the SKU number of one product when another is sold), the error mean for one product would be positive, whereas the error mean for the other would be negative. Our results can be generalized for the single-period problem by considering a general error distribution. In particular, the IABS policy is optimal and our results on the cost functions and the sensitivity of the solutions hold with the exception of Proposition 1, Part (1), and Proposition 3, Part (1). (The proofs are available upon request from the authors.) In the finite-horizon model with general error distributions, the property of the single-intersection point for inspection decision breaks down, because Proposition 6, Part (3), does not hold.

6.5. Lost Sales

Our analysis does not apply to the lost-sales case. With lost-sales assumption, the physical inventory level must be nonnegative, and therefore the errors are no longer independent of the inventory level, which breaks down the current structure of the dynamic program. Nonetheless, for a product with a high service level, the backlog model may be a reasonable approximation. Modeling errors that depend on sales or inventory level encounters the same problem with the lost sales case.

6.6. Positive Lead Time

In the case of positive lead times, inspection decisions during a lead-time period will affect the error accumulated in the period, which, in turn, affects total demand variance during the lead time. Thus, the inventory replenishment decision at the current period depends on the future inspection decisions, and our basic model cannot accommodate this situation.

7. Concluding Remarks

In this study, we consider a single-stage inventory system with inventory record inaccuracy. In the single-period problem, we show that the inspection-adjusted base-stock (IABS) policy is optimal. In the finite-horizon problem, we show that the IABS policy is optimal for a revised dynamic program. We propose the IABS heuristic, whose policy parameters are generated from the revised dynamic program. In a numerical study, we find that it is near optimal and it tends to conduct inspections in a fixed cycle. This observation motivates us to propose an improved cycle-count policy (CCABS), in which inventory inaccuracy is considered when making inventory replenishment decisions. With stationary data, the CCABS heuristic can be as effective as the IABS heuristic.

Some important conclusions can be drawn from our study. First, conducting sensitivity analysis on the IABS heuristic parameters reveals that, all else being equal, products with higher value, higher error variance, lower inspection cost, or lower demand variability should be inspected more frequently. Thus, these variables should be taken into account when companies design ABC classification schemes in the cycle-count programs. Second, the standard cycle-count heuristic, which is prevalent in practice, results in significantly higher cost because it does not consider inventory inaccuracy when making replenishment decisions. The CCABS heuristic should be appealing to practitioners because its performance nearly matches that of the optimal policy and it offers the relative simplicity of fixed inspection schedules. However, in the cases with high error levels, high inspection costs, and nonstationary data, the performance of the CCABS heuristic deteriorates slightly. In such cases, the IABS policy may be preferred. Third, our study indicates that an effective inventory inspection and replenishment policy, such as the IABS heuristic, can secure a significant portion of the value of accurate inventory information that may be provided by RFID systems. Finally, we discuss several extensions and the limitations of our model. These discussions point out the challenges for more general models and lay the groundwork for future research.

Acknowledgments
The authors gratefully acknowledge Gérard Cachon, Nicole DeHoratius, Jing-Sheng Song, Paul Zipkin, the associate editor, and the reviewers for their helpful comments.

Appendix A

Algorithm for Constructing the Lower-Bound Functions in Period t.

1. Set y^{∞} to a very large positive number. Set $y^{-\infty} = -y^{\infty}$.
2. For $j = 0$ to $T - t + 1$:
   (a) Solve for $\alpha_{j+1} = \min_y \mathsf{E}_D[\underline{V}_{t-1}(y - D, j+1)]$.[5]
   (b) Set $\beta_{j+1} = \min\{y\colon \alpha_{j+1} \ge \mathsf{E}_D[\underline{V}_{t-1}(y^{-\infty} - D, j+1)] - c(y - y^{-\infty})\}$.
   (c) Set $\eta_{j+1} = \max\{\max\{y\colon \alpha_{j+1} \ge \mathsf{E}_D[\underline{V}_{t-1}(y^{\infty} - D, j+1)] + \bar{h}_{t-1}(y - y^{\infty})\}, s_{10}\}$.
3. Set $\beta_1 = \min\{\beta_1, \beta_2, \ldots, \beta_{T-t+2}\}$.
4. For $j = 0$ to $T - t + 1$:

$$B_{t-1}(y, j+1) = \begin{cases} \alpha_{j+1} - c(y - \beta_{j+1}), & \text{if } y \le \beta_{j+1}, \\ \alpha_{j+1}, \quad \text{if } \beta_{j+1} < y \le \eta_{j+1}, \\ \alpha_{j+1} + \bar{h}_{t-1}(y - \eta_{j+1}) & \text{if } \eta_{j+1} < y. \end{cases}$$

[5] We perform a numerical search for finding the global minimizer in Step 2(a). In most of our numerical examples, the $\mathsf{E}_D[\underline{V}_t(y - D, j)]$ functions are unimodal.

Appendix B

Proof of Proposition 1. We first show Part (1). Note that

$$H(y, j) = cy + hy - h\mathsf{E}[D] + (b+h)\int_y^{\infty} \mathsf{P}(D + \varepsilon(j) > t)\, dt.$$

Because $(D + \varepsilon(j)) \ge_{var} (D + \varepsilon(i))$, by Definition 1 we have $\int_y^{\infty} \mathsf{P}(D + \varepsilon(j) > t)\, dt \ge \int_y^{\infty} \mathsf{P}(D + \varepsilon(i) > t)\, dt$. The result follows immediately. Part (2) follows because $H(s_i, i) < H(s_j, i) \le H(s_j, j)$.

Proof of Proposition 2. Part (1) holds because of the definition. For Part (2), note that $I^0(x, j) = \mathsf{E}_{\varepsilon(j)}[H(\max\{x - \varepsilon(j), s_0\}, 0)]$. Because $H(\max\{x, s_0\}, 0)$ is linear for $x \le s_0$ and convex increasing in x for $x > s_0$, $\mathsf{E}_{\varepsilon(j)}[H(\max\{x - \varepsilon(j), s_0\}, 0)]$ is convex.

Let ψ_j and Ψ_j denote the pdf and cdf of $\varepsilon(j)$, respectively. For Part (3),

$$\begin{aligned} &\lim_{x\to\infty} N^0(x, j) - \lim_{x\to\infty} I^0(x, j) \\ &\quad = \lim_{x\to\infty} H(\max\{x, s_j\}, j) - \lim_{x\to\infty} \mathsf{E}_{\varepsilon(j)}[H(\max\{x - \varepsilon(j), s_0\}, 0)] \\ &\quad = \lim_{x\to\infty} (H(\max\{x, s_j\}, j) - (h+c)x) \\ &\qquad - \lim_{x\to\infty} \mathsf{E}_{\varepsilon(j)}[H(\max\{x - \varepsilon(j), s_0\}, 0) - (h+c)x]. \end{aligned} \tag{A.1}$$

Consider the last term in (A.1). For fixed $\varepsilon(j)$, $H(\max\{x - \varepsilon(j), s_0\}, 0) - (h+c)x$ is monotone decreasing in x and converges to $\kappa + (h+c)\varepsilon(j)$ as $x \to \infty$, where κ is some constant. Thus, by monotone convergence theorem (MCT), Equation (A.1) is equal to

$$\begin{aligned} &\mathsf{E}_{\varepsilon(j)}\Big[\lim_{x\to\infty} (H(\max\{x, s_j\}, j) - (h+c)x) \\ &\qquad - \lim_{x\to\infty} (H(\max\{x - \varepsilon(j), s_0\}, 0) - (h+c)x)\Big] \\ &\quad = \mathsf{E}_{\varepsilon(j)}\Big[\lim_{x\to\infty} (H(\max\{x, s_j\}, j) - H(\max\{x - \varepsilon(j), s_0\}, 0))\Big] \\ &\quad = \mathsf{E}_{\varepsilon(j)}[(c+h)\varepsilon(j)] = 0. \end{aligned}$$

The second-to-last inequality is due to $\lim_{x\to\infty} H(x, j) = \lim_{x\to\infty} H(x, 0)$.

For Part (4),

$$\begin{aligned}\lim_{x\to-\infty} \mathsf{E}_{\varepsilon(j)}[H(\max\{x-\varepsilon(j), s_0\}, 0)] &= \mathsf{E}_{\varepsilon(j)}\Big[\lim_{x\to-\infty} H(\max\{x-\varepsilon(j), s_0\}, 0)\Big] \\ &= \mathsf{E}_{\varepsilon(j)}[H(s_0, 0)] = H(s_0, 0).\end{aligned}$$

The exchange of the limit and the expectation is justified again by the MCT, as $H(\max\{x-\varepsilon(j), s_0\}, 0)$ is monotone in x and converges to $H(s_0, 0)$ almost everywhere. Thus,

$$\lim_{x\to-\infty}(N^0(x, j) - I^0(x, j)) = H(s_j, j) - H(s_0, 0) > 0.$$

Finally, we show Part (5). Because $N^0(x, j)$ is constant for $x < s_j$ and $I^0(x, j)$ is increasing convex for all x, the result holds for $x < s_j$. For $x \ge s_j$, $N^0(x, j) = H(x, j)$. Recall that the error and demand distributions are continuous. Differentiating N^0,

$$\begin{aligned}\frac{\partial N^0(x, j)}{\partial x} &= c + h + (b+h)\frac{\partial}{\partial x}(\mathsf{E}_{\varepsilon(j), D}[(x - D - \varepsilon(j))^-]) \\ &= c + h - (b+h)\mathsf{P}(D + \varepsilon(j) \ge x) \\ &= c + h - (b+h)\int_{-\infty}^{\infty} \mathsf{P}(D \ge x - u)\psi_j(u)\,du.\end{aligned}$$

On the other hand, $I^0(x, j) = \mathsf{E}_{\varepsilon(j)}[H(\max\{s_0, x - \varepsilon(j)\}, 0)]$. Thus,

$$\begin{aligned}\frac{\partial I^0(x, j)}{\partial x} &= \int_{-\infty}^{x-s_0} \frac{\partial}{\partial x} H(x-u, 0)\psi_j(u)\,du \\ &= \int_{-\infty}^{x-s_0} [c + h - (b+h)\mathsf{P}(D \ge x - u)]\psi_j(u)\,du.\end{aligned}$$

We have

$$\begin{aligned}&\frac{\partial N^0(x, j)}{\partial x} - \frac{\partial I^0(x, j)}{\partial x} \\ &= (c+h)(1 - \Psi_j(x - s_0)) - (b+h)\int_{x-s_0}^{\infty} \mathsf{P}(D \ge x - u)\psi_j(u)\,du \\ &< (c+h)(1-\Psi_j(x-s_0)) - (b+h)\mathsf{P}(D \ge s_0)(1-\Psi_j(x-s_0)) = 0.\end{aligned}$$

The last equality holds because $\mathsf{P}(D \ge s_0) = (c+h)/(b+h)$.

For Part (6),

$$\lim_{x\to-\infty} \mathsf{E}_{\varepsilon(j)}[k + g(x - \varepsilon(j))^+] = k + g\mathsf{E}_{\varepsilon(j)}\Big[\lim_{x\to-\infty}(x - \varepsilon(j))^+\Big] = k.$$

The exchange of the limit and the expectation is again justified by the MCT, as $G(x - \varepsilon(j))^+$ is monotone in x and converges to zero almost everywhere.

Proof of Proposition 3. Part (1) follows from Lemma 1. Part (2) is a result of the observations that $N(x, j)$ is independent of k and g and $I(x, j)$ increases with k and g. For Parts (3), note that $H(s_0, 0)$ is independent of σ_ε and j. From Proposition 1, it can be inferred that $H(y, j)$ is increasing j for all y.

For Part (4) note that $\bar{k}_j = H(s_j, j) - H(s_0, 0) = (b+h)\cdot \phi(\Phi^{-1}((b-c)/(b+h)))(\sqrt{\sigma^2 + j\sigma_\varepsilon^2} - \sigma)$ where ϕ and Φ are the density and distribution functions of the standard normal distribution. $\sqrt{\sigma^2 + j\sigma_\varepsilon^2} - \sigma$ is decreasing in σ and increasing in σ_ε. $\phi(\Phi^{-1}((b-c)/(b+h)))$ is increasing in c and h for $(b-c)/(b+h) > 0.5$. For the result with b, we need to show that $(b+h)\phi(\Phi^{-1}((b-c)/(b+h)))$ increases in b. Define z as the solution to $\Phi(z) = (b-c)/(b+h)$. By the implicit function theorem, $\partial z/\partial b = (h+c)/[(b+h)^2\phi(z)]$. Then $\partial[(b+h)\phi(z)]/\partial b = \phi(z) + \phi'(z)(b+h)\partial z/\partial b = \phi(z) - z\phi(z)((b-c)/(b+h)\phi(z)) = \phi(z) - z(1 - \Phi(z)) > 0$ for all z.

Proof of Proposition 4. Analogously to Part (1) of Proposition 1, we can show that $H(y, j) \ge H(y, i)$ for all y and for $j > i$. The results are implied by this observation and the definition of the functions.

Proof of Proposition 5. Part (1) holds because $\underline{H}_t(y, j)$ is the sum of two convex functions.

The proof of Part (2) is by induction. We have that $B_1(y, j+1) > B_1(y, 1)$ by $V_1(y, j+1) > V_1(y, 1)$. We also have $L(y, j) > L(y, 0)$. Hence, $\underline{H}_2(y, j) \ge \underline{H}_2(y, 0)$ for all y, which leads to $\underline{V}_2(y, j) > \underline{V}_2(y, 0)$ and $B_2(y, j) > B_2(y, 0)$ and so on.

For Part (3), we have, by construction,

$$\frac{\partial B_{t-1}(y, 1)}{\partial y} \ge \frac{\partial B_{t-1}(y, j+1)}{\partial y} \quad \text{for all } y \text{ and } j, \tag{A.2}$$

where $\partial/\partial y$ operator denotes the left derivative if the derivative does not exist. Note that in Step 2(c) of the algorithm of constructing the lower-bound functions $B_{t-1}(y, j+1)$, we restrict η_{j+1} to be no less than s_{10}. Thus, by construction, we have $\partial B_{t-1}(y, 1)/\partial y \le 0$ at $y = s_{10}$. Because s_{10} is the minimizer of $L(y, 0)$, $\partial L(y, 0)/\partial y = 0$ for $y = s_{10}$. Together, we have

$$\frac{\partial L(y, 0)}{\partial y}\bigg|_{s_{10}} + \gamma\frac{\partial B_{t-1}(y, 1)}{\partial y}\bigg|_{s_{10}} = \frac{\partial \underline{H}_t(y, 0)}{\partial y}\bigg|_{s_{10}} \le 0,$$

which implies $\underline{s}_{t0} \ge s_{10}$ because $\underline{s}_{t0}$ is the minimizer of $\underline{H}_t(y, 0)$.

Note that $\partial L(y, 0)/\partial y \ge \partial L(y, j)/\partial y$ for $y \ge s_{10}$. Therefore, $\partial L(y, 0)/\partial y|_{\underline{s}_{t0}} \ge \partial L(y, j)/\partial y|_{\underline{s}_{t0}}$. Recall Equation (A.2), $\partial B_{t-1}(y, 1)/\partial y \ge \partial B_{t-1}(y, j+1)/\partial y$ for all y. Thus,

$$\frac{\partial L(y,0)}{\partial y}\bigg|_{\underline{s}_{t0}} + \gamma\frac{\partial B_{t-1}(y,1)}{\partial y}\bigg|_{\underline{s}_{t0}} \ge \frac{\partial L(y,j)}{\partial y}\bigg|_{\underline{s}_{t0}} + \gamma\frac{\partial B_{t-1}(y,j+1)}{\partial y}\bigg|_{\underline{s}_{t0}}.$$

That is, $\partial \underline{H}_t(y, 0)/\partial y|_{\underline{s}_{t0}} \ge \partial \underline{H}_t(y, j)/\partial y|_{\underline{s}_{t0}}$, which implies $\underline{s}_{tj} \ge \underline{s}_{t0}$.

Part (4) follows from Part (3), because $\partial L(y, 0)/\partial y \ge \partial L(y, j)/\partial y$ for all $y \ge \underline{s}_{tj}$ and $\partial B_{t-1}(y, 1)/\partial y \ge \partial B_{t-1}(y, j+1)/\partial y$ for all y.

Proof of Proposition 6. Parts (1) and (2) are by definition of these functions. Part (3) is by induction. We show the induction step for $t = 1$. We have that $L(y, j)$ converges to $L(y, 0)$ as $y \to \infty$. Hence, $H_1(y, j)$ converges to $H_1(y, 0)$ as $y \to \infty$. Because it is never optimal to inspect as $x \to \infty$, $V_1(x, j)$ converges to $H(x, j)$ for any j. Therefore, $\mathsf{E}V_1(y - D, j+1)$ converges to $\mathsf{E}V_1(y - D, 1)$. By construction, $B_1(y, j+1)$ converges to $B_1(y, 1)$ converge as $y \to \infty$.

Part (4) is due to Part (2) of Proposition 5. Part (5) follows from Part (4) of Proposition 5. Part (6) follows from Parts (1)–(5).

Proof of Proposition 7. The proof is similar to the proof of Proposition 3, Parts (2)–(4). It is possible to show that $B_{t-1}(y, j+1) \geq B_{t-1}(y, i+1)$ and $\underline{H}_t(y, j) > \underline{H}_t(y, i)$ for $j > i$ using induction on t, and then the results follow.

References

Atali, A., H. L. Lee, Ö. Özer. 2005. If the inventory manager knew: Value of RFID under imperfect inventory information. Working paper, Stanford University, Stanford, CA.

Bensoussan, A., M. Çakanyıldırım, S. Sethi. 2005. Partially observed inventory systems: The case of zero-balance walk. *SIAM J. Control Optimization.* Forthcoming.

Bensoussan, A., M. Çakanyıldırım, S. Sethi. 2006. Optimality of base-stock and (s, S) policies for inventory problems with information delays. *J. Optimization Theory Applications*. **130**(2) 153–172.

Camdereli, A., J. M. Swaminathan. 2005. Coordination of a supply chain under misplaced inventory. Working paper, Kenan-Flagler Business School, University of North Carolina, Chapel Hill, NC.

Chen, J., D. D. Yao, S. Zheng. 2001. Optimal replenishment and rework with multiple unreliable supply sources. *Oper. Res.* **49**(3) 430–443.

DeCroix, G. A., V. S. Mookerjee. 1997. Purchasing demand information in a stochastic-demand inventory system. *Eur. J. Oper. Res.* **102** 36–57.

DeHoratius, N., A. Raman. 2004. Inventory record inaccuracy: An empirical analysis. Working paper, Graduate School of Business, University of Chicago, Chicago, IL.

DeHoratius, N., A. Mersereau, L. Schrage. 2005. Retail inventory management when records are inaccurate. Working paper, Graduate School of Business, University of Chicago, Chicago, IL.

Derman, C., G. J. Lieberman. 1967. A Markovian decision model for a joint replacement and stocking problem. *Management Sci.* **13** 609–617.

Iglehart, D. L., R. C. Morey. 1972. Inventory systems with imperfect asset information. *Management Sci.* **18** B388–B394.

Kang, Y., S. B. Gershwin. 2003. Information inaccuracy in inventory systems—stock loss and stockout. *IIE Trans.* **37**(9) 843–859.

Lee, L. H., Ö. Özer. 2005. Unlocking the value of RFID. Working paper, Stanford University, Stanford, CA.

Lee, L. H., M. J. Rosenblatt. 1985. Optimal inspection and ordering policies for products with imperfect quality. *IIE Trans.* **17**(3) 284–289.

Peters, M. H., H. Schneider, K. Tang. 1988. Joint determination of optimal inventory and quality control policy. *Management Sci.* **34**(8) 991–1004.

Piasecki, D. J. 2003. *Inventory Accuracy: People, Processes and Technology*. Ops Publishing, Kenosha, WI.

Pinkerman, J. 2005. EMC2. Private communication.

Raman, A., Z. Ton. 2004. Operational execution at Arrow Electronics. Harvard Business School Case. N9-603-127, Harvard University, Cambridge, MA.

Raman, A., N. DeHoratius, Z. Ton. 2001. Execution: The missing link in retail operations. *California Management Rev.* **43** 136–152.

Rosenfield, D. 1976. Markovian deterioration with uncertain information. *Oper. Res.* **24** 141–155.

Ross, S. 1971. Quality control under Markovian deterioration. *Management Sci.* **17** 587–596.

Shaked, M., J. G. Shanthikumar. 1993. *Stochastic Orders and Their Applications*. Academic Press, San Diego, CA.

Sheppard, G. M., K. A. Brown. 1993. Predicting inventory record-keeping errors with discriminant analysis: A field experiment. *Internat. J. Production Econom.* **32** 39–51.

Singh, M. R., J. S. Song, C. A. Yano, A. Moreno-Beltran. 2004. Production and repair decisions with time consuming repair and a deadline. Working paper, Columbia University, New York.

So, K. C., C. S. Tang, 1995. Optimal batch sizing and repair strategies for operations with repairable jobs. *Management Sci.* **41** 894–908.

Song, J. S. 1994. The effect of leadtime uncertainty in a simple stochastic inventory model. *Management Sci.* **40** 603–613.

Venkatesan, M. 1984. Production-inventory with equipment replacement. *Oper. Res.* **32** 1286–1295.

White, C. C., III. 1978. Optimal inspection and repair of a production process subject to deterioration. *J. Oper. Res. Soc.* **29** 235–243.

Yano, C. A., H. L. Lee. 1995. Lot sizing with random yields: A review. *Oper. Res.* **43** 311–334.

Zipkin, P. 2006. The best things in life were free: On the technology of transactions. *Manufacturing Service Oper. Management*. **8** 321–329.