9-2013

Investigating the Effects of Daily Inventory Record Inaccuracy in Multichannel Retailing

Thomas J. Kull
Arizona State University

Mark Barratt
Marquette University, mark.barratt@marquette.edu

Anníbal C. Sodero
University of Arkansas, Fayetteville

Elliot Rabinovich
Marquette University

Journal of Business Logistics, Vol. 34, No. 3 (September 20, 2013): 189-208. DOI.

Investigating the Effects of Daily Inventory Record Inaccuracy in Multichannel Retailing.

Thomas J. Kull
Arizona State University
Mark Barratt
Marquette University
Anníbal C. Sodero
University of Arkansas
Elliot Rabinovich
Arizona State University

Abstract

Inventory record inaccuracy (IRI) challenges multichannel retailers in fulfilling both brick-and-mortar and direct channel demands from their distribution centers. The nature and damaging effects of IRI largely go unnoticed because retailers assume daily IRI remains stable over time within the replenishment cycle. While research shows that a high level of IRI is damaging, in reality the level of IRI can change every day. We posit that daily IRI variation increases the uncertainty in the system to negatively affect inventory and service levels. Our research uses data collected daily from a multichannel retailer to ground a discrete-event simulation experiment. Going beyond testing just the level of IRI, we evaluate daily IRI variation's impact on operating performance. What we find in our empirical data challenges extant assumptions regarding the characteristics of IRI. In addition, our simulation results reveal that daily IRI variation has a paradoxical effect: it increases inventory levels while also decreasing service levels. Moreover, we also reveal that brick-and-mortar and direct channels are impacted differently. Our findings show that assumptions and practices that ignore daily IRI variation need revising. For managers, we demonstrate how periods of multiday counting help assess their daily IRI variation and indicate what the causes may be.

Keywords

inventory record inaccuracy, multichannel retailing, inventory management, distribution center

Introduction

To manage inventory complexities present across brick-and-mortar and direct (i.e., Internet-based) channels, retailers commonly use centralized decision support systems to serve both of these channels from a single distribution center (DC) (Blankley et al. [9] ; Galbreth and LeBlanc [28]). These systems assume that the retailer's system inventory record (SIR) is accurate, even though widespread inventory record inaccuracy (IRI) occurs in practice (Raman et al. [55] ; Rekik [56]). Research shows that IRI, which is the relative discrepancy between the SIR and the actual inventory on-hand of a stock-keeping unit (SKU), is damaging to operating performance (Waller et al. [70] ; DeHoratius and Raman [23]). However, because the level of IRI is continually changing, managers cannot make one-time adjustments to account for it, but must assess and manage the variability of IRI to maintain DC operating performance and predictability (Barratt et al. [6]).

In this study, we examine the operating performance effects of daily IRI variation within replenishment cycles (i.e., the interval between successive supplier deliveries) of a multichannel retailer DC. Daily IRI variation is the degree to which SIR and actual inventory differ on a day-to-day basis. In practice, because multichannel managers realize IRI damages channel performance (Metters and Walton [46]), they perform cycle counts and audits at single points in time to identify and correct IRI (Neeley [51]). This practice assumes that these occasional, periodic point estimates represent the true level of IRI, that is, there is stability during the replenishment cycle and between cycle counts. In this paper, we test this assumption and investigate daily IRI variation and its effects on operating performance under different demand patterns (i.e., order size and frequency) within multichannel retailing. Our study aims to answer the following main questions: In the context of a multichannel retailer's DC, to what extent does daily IRI variation affect performance? Do channel demand patterns alter the effects of daily IRI variation?

Our research employs a multimethod approach (Mentzer and Flint [45] ; Sanders and Wagner [59]), using both empirical and simulated daily IRI data to answer these questions. First, we review the impact that daily IRI variation has within the replenishment cycle and propose a paradoxical effect: that daily IRI variation increases inventory levels while also decreasing service levels (i.e., order fill rate). We also propose that direct channel demand patterns are more sensitive to the uncertainty created by this phenomenon.

Second, our research uses empirical, multiday IRI data from a multichannel retailer to investigate common assumptions of daily IRI variation. We also decouple SIR errors into those that are transaction-dependent (TD) and transaction-independent (TI) while accounting for the effects of inventory policies on performance. Finally, similar to recent logistics research (e.g., Shapiro and Wagner [63] ; Torres and Maltz [67]), we use our data to ground a simulation with daily IRI variation and test the hypothesized effects on DC performance (i.e., customer service and inventory level). Simulation is appropriate because it: (1) incorporates a high level of detail regarding the factors of interest, (2) accommodates nonlinearities essential to IRI research (e.g., frequencies in cycle counting and record corrections), and (3) accounts for stochastic elements in lead times, SIR errors, and demand across channels (Bowersox and Closs [11]).

We find that daily IRI variation increases inventory levels but decreases service levels, and that an interaction exists between daily IRI variation and channel demand patterns. Also, TI errors seem to be more damaging throughout the inventory system than TD errors. This reveals to managers the importance of knowing what type of errors induce daily IRI variation, so they may reduce the problem instead of using inventory to buffer against the problem.

Our study makes three primary contributions. First, it uses longitudinal evidence to extend the literature beyond current simplifying assumptions of daily IRI variation (cf. Kök and Shang [36] ; DeHoratius et al. [22]), giving a more accurate description for researchers to use. Second, it demonstrates how daily IRI variation is a detrimental phenomenon that managers should be aware of as they seek to avoid stock-outs and improve DC performance (Rabinovich [54]). Third, it presents an approach that managers can use to begin assessing and controlling the problems that cause daily IRI variation. Thus, our research leads to a better understanding of IRI behavior in DCs and suggests how managers can respond to further improve DC performance.

In the next two sections, we review the literature to propose direct and interaction hypotheses about daily IRI variation's effect on DC operating performance. Next, we present our empirical method and results, followed by our simulation method and results. We conclude with a discussion of our findings, contributions, and future research opportunities. Finally, we refer the reader to the Appendix.

Literature Review

Investigations concerning IRI

Since the early works of Rinehart ([58]), Schrady ([62]), and Iglehart and Morey ([32]), the extant literature has focused on the conspicuous presence of IRI in distinct contexts: manufacturing (Graff [30] ; Brown et al. [14]), retailing (Morey [47] ; Raman et al. [55] ; Corsten and Gruen [20] ; Nachtmann et al. [49]), and the extended supply chain (Delen et al. [24] ; Heese [31] ; Ukun et al. [68]). Common across this literature is the attempt by researchers to map the causes and consequences of IRI and to provide managerial guidance regarding ideal or optimal approaches for tackling the problem with inventory management practices.

This literature on IRI may be segregated into two streams of research: empirical and analytical. The empirical stream, which primarily investigates correlations between operating conditions and the presence of IRI, is best represented by DeHoratius and Raman ([23]). In that study, data from a retailer are used to develop a framework relating several factors that mitigate or exacerbate IRI, such as auditing practices, product variety, sales velocity, price, retail stores environment, and distribution structure. DeHoratius and Raman ([23]) argue that these factors should be incorporated into inventory planning tools to account for the presence of IRI. While empirical research has shown the influence and variability of IRI, there remains a lack of empirical characterizations of how IRI varies over time and of the errors that drive IRI variation, and the impact that this variation has on performance.

The analytical stream of IRI literature, by contrast, primarily considers auditing policies and base-stock levels to minimize inspection and inventory holding costs when IRI is present (e.g., Fleisch and Tellkamp [27] ; Delen et al. [24]). Kang and Gershwin ([33]) use simulation to demonstrate how even small levels (1%–3%) of IRI during replenishment cycles lead to severe stock-outs. Camdereli and Swaminathan ([15]) describe how IRI influences optimal replenishment policy decisions and coordinating contracts in a single-period, single-location system. Kök and Shang ([36]) develop a joint inspection and replenishment policy that minimizes total costs in a finite horizon, while DeHoratius et al. ([22]) propose replenishment policies that account for errors using a Bayesian updating of error distribution.

Taken as a whole, analytical research has reinforced the notion that the replenishment cycle is a crucial process in scheduling periodic inventory counts to manage IRI. However, it has yet to show how measuring and managing daily IRI variations within replenishment cycles can impact DC performance. In addition, the above research commonly makes several simplifying assumptions. First, it assumes that errors inducing daily IRI variation are identically distributed and independent of demand and of the channels through which demand arrives. Therefore, this research assumes that IRI follows a simple random-walk pattern over time.[1] Second, the research assumes that managers will only know of inaccuracies as a result of scheduled cycle counts. Third, it assumes that IRI variation involves only a single SKU, as opposed to a wider range of SKUs, which is more common in practice. By contrast, our research uses an empirically grounded simulation to examine these assumptions while showing how DC performance is affected by daily IRI variation across multiple channels. In addition, because managers may look to increase inventory to protect against IRI, we examine how replenishment policies interact with IRI-inducing errors to exacerbate the problem.

Drivers of IRI

Since IRI is the logistics equivalent of a manufacturing defect (Ernst et al. [26]), the manifestation of IRI is linked to SIR errors that are akin to unscheduled downtimes in manufacturing and material handling systems (Banks et al. [5]). Such SIR errors may be classified into two groups: (1) TD errors and (2) TI errors (Lee and Özer [40] ; Nachtmann et al. [49]). TD errors are changes in IRI as triggered by replenishments, demand orders, or product returns (Lee and Özer [40]). Errors in this category can be due to: (1) incorrect deliveries, (2) misplaced items, or (3) incorrect picking (Kang and Gershwin [33]). TI errors are changes in IRI that occur irrespective of transaction and are influenced by the amount of inventory on-hand (DeHoratius and Raman [23]). Such errors are related to: (1) internal materials movement, and (2) shrinkage from theft, spoilage, or damage (Kang and Gershwin [33]).

Channel characteristics also affect IRI. In some cases, multichannel retailers fulfill both their brick-and-mortar and direct channel demands from a single-location DC. While this allows for the pooling of inventory to reduce stock-outs (Ton and Raman [66]) and facilitates the coordination of operations (Metters and Walton [46]), there is speculation that the inherent differences in both channels may affect how IRI occurs and influences DC operating performance (Agatz et al. [1]). In particular, because consumer search costs are lower in direct channels than in brick-and-mortar channels, consumers in direct channels will have easier, more regular access to wider product variety (Bakos [3]). This will induce smaller, more frequent consumer orders in direct channels (Metters and Walton [46] ; Agatz et al. [1]) than in brick-and-mortar channels that can create a wider variety in the sources of IRI within the DC (DeHoratius and Raman [23]). Thus, relative to brick-and-mortar channels, the economics of direct channels create more opportunities for TD errors in the DC.

Inventory management policies may also influence the levels of IRI (Kök and Shang [36] ; DeHoratius et al. [22]). Safety stock policies that hold more inventory create complexity and tracking difficulties because of the sheer volume and variety of items (Lee and Billington [41] ; DeHoratius [21]). Replenishment policies that prescribe highly frequent order deliveries create more opportunities for TD errors (Bonney et al. [10]). Even the common method of using a single-day, periodic cycle count to reconcile SIR with physical inventory can present IRI-management problems (Brooks and Wilson [13]). Not only is cycle counting costly (Graff [30]), it is also prone to human error that can induce IRI rather than correct it (Neeley [51]). Moreover, it only offers snapshots of IRI that mask the TD and TI errors that influence daily IRI variation; items observed with low (high) IRI during a cycle count may spike (drop) in IRI after the count (Kang and Gershwin [33] ; Lee and Özer [40]). As such, managers can adopt policies based on unrealistic assumptions of SIR accuracy. Thus, understanding how channel characteristics and inventory policies interact with the underlying SIR errors that drive IRI variation will benefit both managers and academics.

The replenishment cycle as related to IRI

During the replenishment cycle, decision support system records (including the SIR) are reviewed either continuously or periodically to determine reorder needs. Periodic reviews are used in most real-world environments so to jointly replenish products and leverage economies of scale (Waller et al. [71]). In addition, (s, S) policies, also known as min–max, are seen as efficient in retail distribution (Scarf et al. [61] ; Ballou [4]). In fact, Viswanathan ([69]) shows that periodic (s, S) policies are optimal in many DC-related environments. Not surprisingly, such a policy has been adopted in retail organizations (Caro and Gallien [17]), including the company we discuss in the empirical portion of this paper. As such, without loss of generalizability, we use the period (s, S) structure as a framework to discuss IRI variation effects.

A primary concern of the replenishment cycle is when and how much to reorder, so that inventories and backorders are kept optimally low. When the system determines replenishment is needed, a reorder is placed that is to be received after a specific lead time but before inventory is insufficient to fill customer orders. While partial order fulfillment may be allowed, customer service is highest when orders are filled “on-time/in-full” or OTIF (Livingstone [43] ; Braithwaite and Samakh [12]). We therefore adopt in our paper the concept of a potential backorder point, defined as the point where inventory depletes just below the average customer order size (cf. Li et al. [42] , 411); this is the point where backorders not only become highly probable, but also have the greatest detriment to service levels. While the literature has focused on incorporating demand or supply uncertainties to prevent backorders (cf. Silver [64] ; LeBlanc et al. [39]), only recent work has examined uncertainties in the inventory record itself (DeHoratius et al. [22]). This is crucial because the time at which the reorder point and potential backorder point occur will change depending on the characteristics of IRI and the average customer order size in the distribution channel.

Figure 1 depicts an inventory profile with a periodic (s, S) policy and IRI during the reorder interval of a replenishment cycle. Actual and recorded inventory are distinguished, as are the time-ordered periodic reviews (in gray). Early in the replenishment cycle, inventory levels are high and IRI does not instigate performance problems. However, during the reorder interval of the replenishment cycle (shown in Figure 1 as beginning around the first periodic review), IRI creates problems that have both customer service and inventory performance ramifications. Although the literature has shown IRI to affect replenishment cycle performance, how daily IRI variation impacts performance has not been studied. In the next section, we hypothesize how daily IRI variation directly affects inventory and service levels, and how daily IRI variation during the reorder interval of the replenishment cycle interacts with channel demand patterns (shaped by the demand characteristics we discussed earlier for direct and brick-and-mortar channels) to impact inventory and service levels.

Hypotheses: How Daily IRI Variation Affects Performance

Consistent with literature (Angulo et al. [2] ; Kang and Gershwin [33]), we operationalize IRI as the relative discrepancy between the SIR amount (J) less the actual amount on-hand (I). When IRI is negative, the amount of inventory that managers believe they have is less than actual; when IRI is positive, the amount of inventory that managers believe they have is more than actual.

If IRI is negative, which the retail literature has referred to as “inflating” (Barratt et al. [6]), managers will begin a premature reorder interval at point “A” in Figure 1 at the third review instead of the more appropriate fourth review. This means early receipts and, thus, higher inventory levels than required. Moreover, while awaiting resupply, inventory depletes toward the average customer order size. If the system reaches the potential backorder point “B” in Figure 1, managers are more likely to believe they have insufficient stock to fill the order in full. The likelihood of a backorder increases, even though inventory is available, thus impacting service levels negatively. Under this scenario, managers will not receive indications that IRI exists and so IRI will remain negative.

If IRI is positive, which in a retail setting leads to “freezing” (Kang and Gershwin [33] ; DeHoratius and Raman [23]), managers will begin the reorder interval at point “C” at the sixth review in Figure 1 that is later than required, meaning late receipts and lower inventory levels. However, when actual inventory reaches point “D” in Figure 1, managers will likely accept orders that cannot be fulfilled, thus potentially hurting service levels. In contrast to negative IRI, under this scenario DC picking personnel will observe that stock is unavailable, informing managers that IRI exists, correcting the positive IRI condition and repairing the “frozen” SKU (Barratt et al. [6]). This has been referred to as a “free inspection” (Kök and Shang [36] , 201) and has been largely ignored in the IRI literature. So long as IRI stays positive during the reorder interval, the situation will be corrected at point “D” and thereby eliminate the IRI condition.

Our first two hypotheses relate to the direct performance effects of daily IRI variation. High daily IRI variation means that extreme levels of both positive and negative IRI are likely to occur. That is, when daily IRI variation is high, the range of values to which SIR can deviate from actual inventory is greater than if daily IRI variation was low. Although previous studies do not examine daily IRI variation empirically, it has been shown analytically that IRI variation can increase requisite inventory or decrease service levels. For instance, Iglehart and Morey ([32]) show that error variation increases requisite buffer stock under a periodic review system, while Kök and Shang ([36]) show that maintaining requisite service levels becomes more costly as error variations increase. While these studies do not investigate daily IRI variation in a multichannel DC setting, we expect similar outcomes. We posit that an increase in daily IRI variation has the paradoxical effect of concurrently raising DC inventory levels and lowering DC service levels.

Regarding service levels, both negative and positive values of IRI will diminish performance. As daily IRI variation increases, the range between negative and positive IRI values increases, which increases the frequency that the potential backorder points “B” and “D” in Figure 1 will occur. That is, either point “B” will cause premature backordering earlier in the reorder interval, or point “D” will cause failed pick attempts more often because of later reorders. We therefore posit the following:

H_1: Daily IRI variation decreases DC service levels.

Regarding inventory, higher daily IRI variation causes an increase in actual levels. While negative IRI increases inventory levels because of early reordering at point “A,” positive IRI decreases inventory levels because of late reordering at point “C.” Yet higher daily IRI variation means more instability, increasing the degree that both negative and positive IRI states exist. Although, as stated above, picker-induced corrections will repair positive IRI states, these corrections will be effective only if the IRI state remains stable. Because daily IRI variation creates instability during the reorder interval, a correction to record inaccuracies will be only temporary. Therefore, as daily IRI variation increases, regardless if IRI is positive or negative, the damaging effects of IRI will persist and inventory levels will remain higher. The related hypothesis is stated:

H_2: Daily IRI variation increases DC inventory levels.

We next hypothesize how channel demand patterns interact with daily IRI variation to influence DC performance. Similar to daily IRI variation, demand order size and frequency increase the potential range of SIR values. Our earlier discussion about demand characteristics for brick-and-mortar and direct channels suggests that the infrequent, large orders that typify brick-and-mortar channels are “lumpier” than the frequent, small orders that characterize direct channels (Ward [72]). Lumpy, erratic demand generally induces more uncertainty into an inventory system (Bartezzaghi and Verganti [7] , 116), increasing the likelihood that brick-and-mortar channels will have a wider range of potential SIR values, and higher SIR variability, than direct channels. Moreover, because of this disparity in SIR variability, an increase in daily IRI variation will have a relatively different effect between the two channels. Specifically, brick-and-mortar channels, with relatively higher SIR variability, will experience smaller increases in SIR variability from daily IRI variation than direct channels. Thus, brick-and-mortar channels will be less sensitive to increases in daily IRI variation than direct channels.

This implies that daily IRI variation exacerbates the service level problems discussed in the development of H_1 more in direct channels than in brick-and-mortar channels. The relatively larger effect by IRI variation on SIR variability in a direct channel increases the frequency in which potential backorder points “B” and “D” occur. Therefore, for an equivalent level of IRI variation across direct and brick-and-mortar channels, there will be more premature backorders and order-pick failures in direct than in brick-and-mortar channels. This leads to the following hypothesis.

> H_{3a}: Daily IRI variation has a more adverse effect on DC service levels under direct channel demand than under brick-and-mortar channel demand.

Daily IRI variation also exacerbates problems regarding inventory levels to a greater degree in direct channels than in brick-and-mortar channels. Lumpy demand, as in a brick-and-mortar channel, is generally associated with higher inventory (Bartezzaghi et al. [8]). The relatively larger effect by IRI variation on SIR variability in direct channels causes more frequent early ordering and less permanent IRI corrections. Thus, the inventory effects caused by IRI variation will be more damaging in direct than in brick-and-mortar channels. This leads to our last hypothesis.

> H_{3b}: Increases in DC inventory levels due to daily IRI variation are higher under direct channel demand than under brick-and-mortar channel demand.

Contextual factors

Beyond the above hypotheses, we discuss other factors that interact in the DC system for which we control (see in Figure 2). As we discussed in the literature review section, these factors and their effects on performance are based on whether daily IRI variations are caused by TD or TI errors. Detailed explanations are given in Table 1, but are summarized as follows: (1) because opportunities for TI errors are more frequent than TD errors, we expect that TI errors exert a greater interactive and direct influence than TD errors on DC operating performance; (2) because of more frequent customer orders, we expect that direct channel demand patterns interact more with the negative effects of TD errors than with the negative effects of TI errors on DC operating performance; (3) because inventory variations are proportional to the inventory policy reorder point (little s), we expect that little s interacts more with the negative effects of TI errors than with the negative effects of TD errors on DC operating performance; (4) because of the increased opportunity for TD errors, we expect that the inventory policy's reorder quantity (i.e., S – little s = Δ) interacts more with the negative effects of TD errors than with the negative effects of TI errors on DC inventory and service levels; and finally, (5) because of the increased opportunity for TI error to accumulate, we expect that cycle count frequency (time between counts) interacts more with the negative effects of TI errors than with the negative effects of TD errors on DC operating performance.

Table 1. Contextual factors interacting with record errors and affecting DC operating performance

Expected contextual factor	**Justification**
(1) TI errors exert a greater interactive and direct influence than TD errors on DC inventory levels and service levels	The literature suggests differing causes for TD and TI errors (Iglehart and Morey 32; Ernst et al. 26). TD errors occur when too many or too few units are actually picked than are recorded during a shipping transaction, and when receipt documents are inaccurate during a receiving or returning transaction. TI errors, by contrast, occur regardless of incoming or outgoing product, and are the result of unrecorded spoilage, damage, theft, and the like. Opportunities for TI errors are more frequent than TD errors because inventory exists nearly every day, while transactions do not
(2) Direct channel demand patterns interact more with the negative effects of TD errors than with the negative effects of TI errors on DC inventory levels and service levels	TD and TI errors will also interact with the demand patterns observed in different channels. Given an equal total demand in a period across channels, a direct channel will have smaller order sizes and higher order frequencies than a brick-and-mortar channel. If we represent the relationship between order frequencies and order sizes in a channel by a ratio r, a higher r represents a direct channel and translates into more frequent customer orders and, therefore, more opportunities for TD errors. This would compound the record errors in the system and reduce system performance
(3) Inventory policy reorder point (little s) interacts more with the negative effects of TI errors than with the negative effects of TD errors on DC inventory levels and service levels	TD and TI errors will also interact with the inventory levels used as reorder points. In a periodic review reorder policy (s, S) regime, the reorder point (little s)—which managers base on various costs, policies, and lead times—influences the average amount of stock carried: ceteris paribus, a higher reorder point will increase the average amount of inventory on hand. As the amount of inventory on hand increases, the opportunity for TI errors will increase. This is because as more units are available in stock, there will be more units to miscount, misplace, or experience shrinkage. In this way, as the reorder point increases, the exposure to the effects of TI errors will increase and damage DC performance. On the other hand, the reorder point has little bearing on the size or frequency of transactions and, thus, will have a relatively minor interaction with the negative effects of TD errors on DC inventory levels and service levels
(4) Inventory policy reorder quantity (Δ) interacts more with the negative effects of TD errors than with the negative effects of TI errors on DC inventory and service levels	The size of reorder quantities (Δ)—which managers base on various costs and policies—influence the occurrence of TD errors because these errors occur, in part, when inventory is replenished. The frequency of replenishment is inversely related to the average reorder quantity: ceteris paribus, as reorder quantity decreases, the frequency of reorder transactions increases. In this way, a smaller Δ increases the

	opportunities for possible TD errors and for replenishment cycle problems in general. Although Δ may also interact with TI because Δ influences the amount of inventory on hand, the constant amount of inventory on hand as influenced by reorder point (little s) is expected to dominate interactions with TI
(5) Cycle count frequencies (time between counts) interact more with the negative effects of TI errors than with the negative effects of TD errors on DC inventory levels and service levels	TD and TI errors and their impact on inventory and service level performance will depend on the inventory audit policy in place. A common policy to control IRI and improve performance is cycle counting, in which frequency (i.e., time between counts) is the primary parameter necessary to define the policy. Counting occurs in longer cycles for those items that are considered less important in terms of sales volume or item value (Cantwell 16). Because cycle counting occurs infrequently for these items, there will be more opportunity for TI errors to accumulate (as explained in item (1) above). As such, longer audit frequencies will increase the opportunity for TI errors to occur

Notes : DC, distribution center; TI, transaction-independent; TD, transaction-dependent.

Finally, it is important to note that an influential characteristic of IRI is bias, meaning that over time the system may consistently be in a positive IRI or a negative IRI state. As stated above, the effects of these conditions on inventory levels are well known. As such, although no hypotheses are given for this characteristic, in order to test our hypotheses regarding daily IRI variation, we control for the IRI bias condition in our analysis.

Methodology

To test our hypotheses with a simulation model, we first collected data to represent essential DC characteristics while controlling for spurious sources of variation. This enabled us to stringently focus on DC fulfillment operations and daily IRI variation. We chose a DC that fulfilled demand from both direct and brick-and-mortar retail channels of a midsize apparel retailer in the United States. This research setting was ideal because the retailer's DC faced high IRI levels across many SKUs even though efforts to mitigate IRI were in place—for example, the implementation of continuous improvement initiatives and the adoption of a modern warehouse management system (WMS). Moreover, given the nature of the apparel retailer's business, there was relatively less variation in the physical size of the SKUs than in most industrial settings. This mitigated the effects of a factor that has been found elsewhere to significantly induce IRI (DeHoratius and Raman [23]). In addition, although the retailer held and managed separate inventories for its direct and brick-and-mortar channels, operational processes, equipment, and personnel were shared between the fulfillment operations of both channels within the DC, thus preventing such factors to induce biases between channels. By focusing on a single DC location of a multichannel retailer, we avoided the introduction of bias that might otherwise compromise the comparability of measures and establishment of relationships between concepts. The homogeneity of factors across both channels also controlled for extraneous sources of IRI, which might introduce unobservable variability.

The retailer followed a periodic review policy based on an (s, S) approach for reordering and replenishing: when the periodic reviews showed that inventory (i.e., system record + in-transit replenishment orders – backordered demand orders) reached a position equal to or approaching the reorder point, little s, the DC reordered from suppliers an amount necessary to take the level back to a preset maximum S. While little s performs the function of a reorder point with some element of safety stock, the difference Δ (i.e., S – little s) accounts for demand patterns, replenishment lead time, and reordering costs, so it serves to control reorder quantity.

Data collection

The retailer carried approximately 15,000 SKUs across seven product categories. Of these, 100 SKUs were common to both channels, representing just four of the seven product categories. SKUs for each channel were in separate locations within the DC. We ranked these 100 common SKUs by unit sales volume within each of the four product-channel categories to identify fast-, medium-, and slow-velocity products. The two products best representing each type of velocity were selected, resulting in 24 matched pairs (i.e., 2 [SKUs] × 3 [velocity types] × 4 [product categories]) across the two channels to allow for comparability of findings (Metters and Walton [46]) and to cover a broad spectrum of prices and popularities. Because three of the 24 sampled SKUs had been scheduled to be discontinued during our data collection process, we also selected three additional SKUs as possible substitutes. The three additional SKUs were chosen based on their similarities (category, velocity, price) to the SKUs that had been scheduled to be discontinued. As we continued to see inventory and product receipts for these three SKUs during the data collection, we decided to maximize our sample by retaining the three “discontinued” as well as the three additional SKUs in the final sample. The final sample of 27 SKUs across two channels (for a total sample of 54) is presented in Table 2.

Table 2. IRI parameters among SKU characteristics and channels

				Estimates of IRI parameters (% on-hand inventory)			
				Bias (mean μ_{IRI})		**Daily variability (SD σ_{IRI})**	
Item	**SKU**	**Price ($)**	**Rank in category**	**Direct channel**	**Brick-and-mortar channel**	**Direct channel**	**Brick-and-mortar channel**
1	2446300	4	Slow	0.00	−1.04	0.00*	2.37
2	2315968	4	Medium	−1.77	2.35	3.38	15.16
3	2315927	4	Fast	0.03	2.14	0.19	8.52
4	2100402	4.75	Unique†	0.04	−0.14	1.88	1.36
5	2083954	4.75	Slow	−1.37	−1.08	1.56	2.31
6	2083947	4.75	Medium	0.00	−0.04	0.00*	0.12
7	2008985	4.75	Fast	−2.00	−16.11	1.19	21.99
8	2225977	5	Unique†	−10.00	−10.00	31.62	31.62
9	2268167	5	Medium	−1.12	−2.21	0.69	12.26
10	2268100	5	Fast	−10.00	−10.00	31.62	31.62
11	2268118	5	Fast	−0.09	−3.57	0.38	11.29
12	2237402	14	Slow	−0.13	39.26	0.41	64.69
13	2199966	14	Medium	0.26	0.27	0.83	0.43
14	2199982	14	Medium	−0.26	0.56	0.46	0.48
15	2237410	14	Fast	0.57	−2.85	1.92	9.35
16	1577527	27	Slow	−1.29	0.00	0.36	0.00*
17	1603919	27	Slow	−9.90	−10.00	31.66	31.62
18	1740356	27	Medium	−1.64	−18.79	0.62	31.47
19	1372960	27	Fast	0.66	0.00	0.43	0.00*
20	1862911	28	Slow	−7.17	0.00	16.12	0.00*
21	1785328	28	Medium	−1.81	−2.45	4.86	7.75
22	1785419	28	Fast	0.23	0.00	0.42	0.00*
23	2099356	30	Medium	−2.73	2.49	1.27	4.10
24	2099612	30	Fast	−0.40	2.03	1.26	4.91
25	2259570	44	Slow	−0.47	0.00	3.89	0.00*
26	2259588	44	Fast	−0.68	0.00	1.01	0.00*
27	2374106	60	Unique†	−10.23	10.00	17.93	31.62
	Average (SD)			−2.27% (3.63%)	−0.71% (9.99%)	5.78% (10.29%)	12.04% (15.85%)

Notes : IRI, inventory record inaccuracy; SKU, stock-keeping unit; SD, standard deviation.

*SKUs with zero variability, that is, accurate over the period, were excluded from error estimations.
†Item was announced as “discontinued” but continued to be monitored.

To assess IRI, we focused on tracking the physical inventory held for the 27 sampled SKUs at the DC for both the direct and brick-and-mortar channels during a period of 10 consecutive business days. Simultaneously, we contrasted this information with the data in the retailer's SIR. We did not count the physical inventory during the weekend because the DC operated only from Monday to Friday. The 10 business days corresponding to two calendar weeks were chosen to avoid the interference of seasonality in demand. The company carried out a complete physical audit of every SKU in the DC two weeks before our data collection.

To track the physical inventory and minimize the likelihood of counting error, at least two of the researchers counted the number of items in storage for each SKU and for each channel. Every day, the counters were randomly assigned to 27 SKUs (13 or 14 SKUs for each channel), to avoid counting bias due to fatigue and knowledge of previous inventory levels. It has been shown that repetitive counting of many items of the same SKU might decrease the counters’ attention, or increase their expectations about a “correct” count (Neeley [51] , [52]). At the beginning of every counting day, for all selected SKUs, we printed the SIR balance and the locations from the WMS onto a counting sheet. The items in the counting sheet were then sorted by nearby locations within each channel to ensure that we counted the SKUs in the same order every day. Each SKU had a fixed channel location in the DC—that is, one direct channel location and one brick-and-mortar channel location. All counts were performed at the start of each day before operations started to ensure that items were not in temporary storage or shipping/packing locations. The average on-hand amount per SKU ranged from near zero to 9,000 units, with an overall average of about 140, 430, and 770 units for slow-, medium-, and fast-velocity SKUs, respectively.

From 540 observations of SIR and actual inventory across 27 SKUs in both channels over 10 consecutive business days, a relative measure of IRI was computed as shown in Equation that follows from the IRI literature (Angulo et al. [2] ; Fleisch and Tellkamp [27]).

(1)

$$IRI_{\mathrm{pqt}} = \begin{cases} d_{\mathrm{pqt}}/I_{\mathrm{pqt}}, & \text{if } I_{\mathrm{pqt}} > 0 \\ 0, & \text{otherwise} \end{cases}$$

where d_{pqt} is the difference between the SIR balance ($\mathrm{J}_{\mathrm{pqt}}$) and physical inventory (I_{pqt}) of SKU p in channel q at counting day t ($p = 1, \dots, 27;\ q = 1, 2;\ t = 1, \dots, 10$). Our measure of IRI departs from some previous studies because we measure the relative rather than the absolute discrepancy (cf. DeHoratius and Raman [23]). A relative measure enables comparability of IRI across different SKUs and is the common approach used in practice (Brown et al. [14]). Although the possibility exists for d_{pqt} to be nonzero when I_{pqt} is zero (which we observed 1.1% of the time), our relative measure provides better generalizability than an absolute measure. In addition, a relative discrepancy (i.e., percentage of actual inventory on hand) is consistent with a percentage representation of TD and TI errors. Finally, because in our analysis, we statistically assess the relative impact of daily IRI variation, use of a percentage versus an absolute measure is common among all scenarios and is thus controlled.[2]

As shown in Table 2, daily IRI variability (measured as standard deviation σ_{IRI}) remained equal to zero across the ten counting days in eight out of the 54 series. In the 46 series that presented some variation, IRI showed large variability relative to the bias (measured as mean magnitude μ_{IRI}). Although, on average, the direct channel had a more negative bias than the brick-and-mortar channel, a between-subjects multivariate analysis of variance revealed no significant influence of price, sales velocity, or channel type on μ_{IRI} or σ_{IRI}. This suggests a latent tendency toward a negative IRI bias, possibly because picker corrections are occurring when IRI is positive.

To further assess literature assumptions of IRI behavior, we fit the 46 IRI series that presented IRI variability to various auto-regressive integrated moving average (ARIMA) models. We found it most likely (at the.05 level) that 37 time series have an integrated and moving average term—that is, an ARIMA (0, 1, 1). This departs from the literature's assumption that IRI moves in a random-walk pattern—that is, ARIMA (0, 1, 0)—and is likely a result of TD errors that are demand-dependent and intermittent in nature. In addition, IRI variation across time was found to be 35% lower than IRI variation across SKUs, showing that it is erroneous to assume IRI varies similarly across time versus across SKU. Examples of daily d_{pqt} values are shown in Figure 3 and were observed to have random shocks, followed by periods of stability. These observations challenge some of the simplifying assumptions made in the literature.

We also examined the issue of IRI patterns more closely by modeling the probability of occurrence and the magnitude of the errors—the result of underlying, unobserved incidents that generate IRI (DeHoratius et al. [22]). An error occurs in the multichannel retailer's inventory system when the difference (d_{pqt}) between $\mathrm{J}_{\mathrm{pqt}}$ and I_{pqt} of a particular SKU in a given channel changes between two consecutive days. That is:

(2)

$$e_{\mathrm{pqt}} = d_{\mathrm{pqt}} - d_{\mathrm{pq,t\text{-}1}}$$

where $\mathrm{e}_{\mathrm{pqt}}$ is an error that changes the IRI of SKU p in channel q at counting day t ($p = 1, \dots, 27;\ q = 1, 2;\ t = 1, \dots, 10$). We model error differently from the unobservable process; others have termed stock loss (Kang and Gershwin [33]), random error (Kök and Shang [36]), invisible demand (DeHoratius et al. [22]), or demand error (Nachtmann et al. [49]). This literature assumes that an independent and identically distributed random variable following a continuous distribution impacts inventory in each period, and that this random variable is independent of any receipts, returns, or shipments over time; thus leading to a random-walk pattern. We see from the empirical data that this is not the case. Therefore, we make no such assumptions and, instead, we associate errors closer to IRI's generating mechanisms (cf. Rinehart [58] ; Lee and Özer [40]). Following early work on IRI (Iglehart and Morey [32] ; Ernst et al. [26]), we seek distinct estimates of TD and TI errors. As we could not directly observe every possible error over the study time frame, we adopt a heuristic to estimate TD and TI errors: Errors during no SIR change are assumed to be TI errors while those during SIR change are divided into TI and TD errors based on TI error rates observed when there was no SIR change (see the Appendix).

In Table 3, we examine the frequency and magnitude of TD and TI errors across channels, price, and sales velocity in accordance with literature (DeHoratius and Raman [23]). Because high skewness exists in portions of the data, we focus on the median values. Regarding TD and TI frequencies, we find mostly no significant differences ($p > .05$), except for lower price SKUs having significantly higher TD rates that could be due to larger pack sizes. Regarding TD and TI magnitudes, we find mostly no significant differences ($p > .05$), an exception being that channels show significant TD difference. Further investigation reveals that because error rates are percentage values, a series of uncharacteristically small retail transaction quantities induced large TD values—if these are removed then no significant TD difference is found ($p > .10$). In sum, we conclude that assuming similar TD and TI frequencies and magnitudes across channels and sales velocities is acceptable, but that price may have some effect on IRI through differences in TD frequencies.

Table 3. Observed inventory error frequencies and magnitudes across product characteristics[‡,§],

	Frequency		Magnitude	
Product characteristics	**TD errors**	**TI errors**	**TD errors**	**TI errors**
Price				
High				
N	21	24	52	18
Mean (%)	51.4	25.8	−102.9	−8.4
Median (%)	57.1	15.5	−47.7	−1.0
Low				
N	25	29	76	54
Mean (%)	69.0	34.9	9.2	−1.7
Median (%)	77.8	33.3	−8.0	−0.1
p-value of differences				
Mean*	.218	.213	.248	.365
Median†	.038	.179	.280	.276
Channel				
Direct				
N	26	26	98	29
Mean (%)	61.3	39.3	−33.4	−5.3
Median (%)	70.8	29.2	−11.8	−0.6
Brick-and-mortar				
N	20	27	30	43
Mean (%)	60.5	22.6	−45.9	−2.1
Median (%)	55.0	14.3	−99.0	−0.1
p-value of differences				
Mean*	.968	.115	.912	.630
Median†	.552	.341	.012	.471
Sales velocity				
Fast				
N	16	17	50	22

Mean (%)	54.3	39.5	98.5	−1.9
Median (%)	56.3	33.3	18.3	0.1
Medium				
N	13	16	32	29
Mean (%)	60.1	27.9	−241.4	−1.1
Median (%)	60.0	25.0	−99.9	−0.2
Slow				
N	11	14	30	17
Mean (%)	57.6	30.4	−33.6	1.7
Median (%)	50.0	15.5	−16.3	−0.7
Unique				
N	6	6	16	4
Mean (%)	86.9	14.6	−52.8	−50.0
Median (%)	92.9	12.5	−85.7	−50.3
p-value of differences				
Mean	.247	.284	.048	.003
Median	.379	.292	.054	.610

Notes : TD, transaction-dependent, TI, transaction-independent.
*p-value from univariate ANOVA tests.
†p-value from nonparametric median difference tests.
‡Distribution values across groups are found to be equivalent except for TD magnitudes between velocity and channels ($p < .01$), revealing the influence that order patterns have (based on nonparametric Kolmogorov–Smirnov two-sample tests for price and channel, and Kruskal–Wallis k-sample tests for sale velocity).
§The estimation procedure used is found in the Appendix.

Simulation model

On the basis of our empirical analysis, we used a discrete-event simulation (Kelton et al. [34] , 197) to test our hypotheses and to observe outcomes for the relationships among the contextual factors (see Figure 2). The simulation methodology accommodates nonlinearities (e.g., cycle counting and picker corrections) that are key elements of IRI research and that cannot be assessed accurately using analytical techniques that require limiting assumptions (Bowersox and Closs [11]).

In order to explore the IRI dynamics in the DC of a multichannel retailer, a simple inventory management system is considered with a single SKU available in a retail channel: either direct or brick-and-mortar. The retailer follows an (s, S) inventory policy at the DC that is consistent with our empirical setting and extant literature, making the simulation generalizable to other DCs (Nahmias and Smith [50] ; cf. Ballou [4]). The amount of reordered inventory, Δ, corresponds to the difference between S and the inventory position. Under such review policies, both the predetermined minimum level, little s and the Δ (i.e., S – little s) quantity are expected to influence IRI and inventory performance as described in earlier sections. The replenishment of inventory at the DC involves supply lead times given by a random variable with mean μ_{leadtime} and standard deviation $\sigma_{\mathrm{leadtime}}$ from an uncapacitated supply source.

Demand input corresponds to the arrival of batch orders in the brick-and-mortar channel and the arrival of individual, end-customer orders in the direct channel. We do not allow multiple SKUs in the orders to preserve the internal validity of the results; doing so allows the simulation to consider DC performance in an unbiased fashion, across all demand orders, irrespective of their product composition. Orders in the simulation may induce multiple items for an SKU. These order sizes follow a known distribution with mean μ_d and standard deviation σ_d. They arrive according to a Poisson distribution with mean λ_d and standard deviation $\sqrt{\lambda_d}$.

Demand in the orders is fulfilled by the WMS when the SIR shows inventory availability; otherwise, they are backordered. No partial shipments are allowed (Kumar et al. [37]). This ensures consistency with OTIF measures of customer service that can be found in practice (Godsell and Van Hoek [29]), and also ensures an equal treatment of small versus large order sizes that are common in direct and brick-and-mortar channels, respectively. If SIR shows enough inventory availability to fulfill an order but physical inventory is too low to fill the order in full (i.e., point "D" in Figure 1), the nonlinearity of a picker correction (i.e., an adjustment made by picking employees) is modeled by making the SIR equal to the actual inventory while backlogging the order.

A fundamental distinction between an inventory management system in the direct channel and one in the brick-and-mortar channel is the difference in demand patterns—that is, customer order size and frequency (Agatz et al. [1]). To assure comparability of results between channels, we assume that the total volume of demand (in units) is equal in both channels. This allows holding constant both s and Δ while varying a ratio, r, which is the relative change from a base brick-and-mortar order frequency. Higher values of r represent more frequent and, thus, smaller orders, according to the demand pattern in the direct channel than those based on the demand pattern in the base brick-and-mortar channel.

Therefore, in our simulation, a lower ratio r represents the brick-and-mortar channel and a higher ratio r represents the direct channel. For example, if the average monthly demand is 100 units in each channel, while the base brick-and-mortar channel (with r = 1) will receive an average of 20 orders of 5 items each, the direct channel (with r = 2.5) will receive 50 orders of 2 items each. By dividing μ_d, σ_d, and λ_d by the ratio r, we modify the differences in magnitudes and interarrival times of demand to represent channel characteristics, while holding the coefficients of variation $\sigma_d \mu_d$ and $\lambda_d \lambda_d$ and total demand constant.

The simulation also takes into account the sources of IRI variation. Rather than assuming that IRI varies from an identically distributed, demand-independent random integer variable, we seek to replicate our empirical results by allowing IRI to vary from two sources: TD errors and TI errors. The magnitude of a TD error is defined by the realization of a random percentage variable, normally distributed (cf. Fleisch and Tellkamp [27]) with mean zero and standard deviation σ_{TD} multiplied by the size of the transaction. TD errors occur according to a conditional probability P_{TD}. Conversely, the magnitude of a TI error is defined by the realization of a random percentage variable, normally distributed with mean zero and standard deviation σ_{TI} times the amount of physical inventory. TI errors occur according to a probability P_{TI}. The occurrence of either TD or TI errors will change daily IRI dependent on demand and inventory levels. Finally, in the simulation, we correct IRI by accounting for the possibility that IRI will be reconciled with SIR every time cycle counting is performed at f time intervals—larger f values mean more infrequent (less often) cycle counting.

Holding the channel fixed, our simulation not only takes into account a daily sequence of events in the inventory system that follows past simulation studies addressing IRI (Kök and Shang [36]), but also incorporates important nonlinearities. First, TD errors may occur only on days with replenishments or order picking, while TI errors are possible every day. Second, if fulfillment orders are released without enough stock, order pickers will notice this and the SIR level will be corrected. Third, physical inventory is not allowed to become negative. Fourth, cycle counting corrects SIR balances at discrete time intervals. Finally, IRI parameters, mean μ_{IRI} and standard deviation σ_{IRI}, are computed from the daily IRI values. Further details about our simulation model development (Sargent [60]) are in the Appendix.

Results

To test our hypotheses, we designed a model in line with our empirical observations shown in Tables 2 and 3 for a full factorial experiment across eight factors: little s (2 levels: 15 and 20), Δ (2 levels: 35 and 50), r (3 levels: 1, 2, and 3), f (3 levels: 30, 180, and 360), σ_{TD} (2 levels: 0.03 and 0.06), P_{TD} (2 levels: 0.40 and 0.60), σ_{TI} (2 levels: 0.03 and 0.06), and P_{TI} (2 levels: 0.10 and 0.20). To assure proper statistical power across factor combinations,[3] 35 replications were conducted for each of the 576 factorial combinations to generate a total of 20,160 observations. Further background information relating to parameters, demand, errors, cycle counting, and model operations in our model is summarized in Table 4.

Table 4. Summary of simulation background information

Replenishment lead time and service level
Replenishment lead time: normally distributed with mean (μ_{leadtime}) 5 and SD ($\sigma_{\mathrm{leadtime}}$) 1 (in days)
Target order fill rate (frequency of nonbackordering) determines the reordering policy parameters: 99%
Demand
Representative customer originates all demand
Demand orders are independent and identically distributed
Base† demand order size is normally distributed with mean (μ_{d}) 4.5 and SD (σ_{d}) 1 (in units)*
Base† demand order interarrival times follow a Poisson distribution with mean (λ_{d}) 3 (in days)*
Errors
TI errors are independent and identically distributed
TI errors arrival times follow a Poisson distribution
TD error magnitude is a percentage of transaction size. The percentage is normally distributed with mean zero and an SD
TD errors occur according to a conditional probability with given mean and an SD
Cycle counting
Physical inventory levels are correctly assessed
IRI levels are correctly assessed
SIR levels are correctly adjusted at the end of cycle counting
Model operations
DC operates 24 hr, seven days per week
DC stock levels preloaded with typical days-on-hand inventory
Reordering policy determined based upon strategic goals and no IRI assumption

Reordering policy is min–max, or (s, S), following a seven-day periodic review
Reordering policy parameters (little s and Δ) are stagnant throughout simulation run

Notes : DC, distribution center; TD, transaction-dependent, TI, transaction-independent; IRI, inventory record inaccuracy; SIR, system inventory record; SD, standard deviation.

*Additional information regarding the estimation of these parameters is included in the Appendix.

†Base demand represents the brick-and-mortar channel.

To test the hypotheses, we used regressions analysis based on the output from 20,160 simulation runs. Because we expected interactions among experimental factors, we included, in the regression analysis, the factors' main effects and two-way interactions along with the direct effects of IRI bias μ_{IRI} and daily IRI variation σ_{IRI}. To avoid multicollinearity, we centered all factor variables before we formed interaction terms (Cohen et al. [19]). We estimated variance inflation factors for the full model and found values ranging from 1 to 3.914. These values are below the threshold value (10) that would signal a multicollinearity problem.

The regression results appear in Table 5. There are two models for each DC performance variable: service level (as measured by the percentage of orders accepted on-time and in-full) and average actual inventory level. The first model (model 1) includes all controls, while the second model (model 2) includes all control as well as daily IRI variation (σ_{IRI}) and its interaction with channel demand pattern (r). Both models list standardized regression coefficients. Because model 2 controls for IRI effects, substantial changes in the coefficients indicate IRI sensitivity to particular variables. Overall, the R^2 values range from 0.847 to 0.981 and show an improvement from model 1 to model 2 for each dependent variable. In addition, the second model's $\Delta\chi^2$(df) values for service level and average inventory show a significant improvement in variance explained [151(2) and 2,992(2), respectively]. Thus, the importance of understanding dynamic IRI characteristics is supported.

Table 5. Regression results

Dependent variable main effects	Service level		Average inventory	
	Model 1†	Model 2†	Model 1†	Model 2†
little s	0.900 (0.001)***	0.884 (0.001)*	0.552 (0.002)**	0.564 (0.002) ***
Δ	0.103 (0.000) ***	0.118 (0.000) ***	0.807 (0.001) ***	0.796 (0.001) ***
r	−0.053 (0.003) ***	−0.053 (0.003)**	−0.024 (0.007) ***	−0.024 (0.006) ***
F	−0.211 (0.000) ***	−0.060 (0.000) ***	0.112 (0.000) ***	0.002 (0.000)
P_{TD}	−0.012 (0.022) ***	−0.004 (0.022)	0.008 (0.056) ***	0.002 (0.049)*
σ_{TD}	−0.050 (0.150) ***	−0.027 (0.148) ***	0.026 (0.375) ***	0.008 (0.337) ***
P_{TI}	−0.118 (0.045) ***	−0.068 (0.048) ***	0.069 (0.113) ***	0.034 (0.109) ***
σ_{TI}	−0.206 (0.150) ***	−0.109 (0.196) ***	0.124 (0.375) ***	0.054 (0.448) ***
μ_{IRI}	−0.183 (0.052) ***	−0.102 (0.058) ***	−0.309 (0.129) ***	−0.369 (0.132) ***
σ_{IRI} (H_1, H_2)		−0.207 (0.026) ***		0.149 (0.058) ***
σ_{IRI}Xr (H_{3a}, H_{3b})		−0.052 (0.025) ***		0.001 (0.057)
Two-way control interactions of experimental factors				
sXΔ	−0.111 (0.000) ***	−0.105 (0.000) ***	−0.008 (0.000) ***	−0.012 (0.000) ***
sXr	0.067 (0.001) ***	0.068 (0.001) ***	0.009 (0.003) ***	0.009 (0.002) ***
sXf	0.039 (0.000) ***	0.045 (0.000) ***	0.045 (0.000) ***	0.041 (0.000) ***
sXP_{TD}	0.005 (0.009)	0.005 (0.009)	0.000 (0.023)	0.001 (0.020)
sXσ_{TD}	0.017 (0.060) ***	0.015 (0.058) ***	−0.001 (0.150)	0.001 (0.131)
sXP_{TI}	0.036 (0.018) ***	0.039 (0.017) ***	0.008 (0.045) ***	0.006 (0.039) ***
sXσ_{TI}	0.056 (0.060) ***	0.062 (0.058) ***	0.014 (0.150) ***	0.010 (0.132) ***
ΔXr	0.009 (0.000)**	0.015 (0.000) ***	0.009 (0.001) ***	0.008 (0.001) ***
ΔXf	−0.028 (0.000) ***	−0.014 (0.000) ***	0.045 (0.000) ***	0.035 (0.000) ***
ΔXP_{TD}	0.006 (0.003)*	0.006 (0.003)*	0.002 (0.008)	0.002 (0.007)*
ΔXσ_{TD}	0.003 (0.020)	0.002 (0.019)	0.001 (0.050)	0.002 (0.044)*
ΔXP_{TI}	−0.008 (0.006)**	−0.004 (0.006)	0.014 (0.015) ***	0.011 (0.013) ***
ΔXσ_{TI}	−0.022 (0.020) ***	−0.013 (0.019) ***	0.023 (0.050) ***	0.017 (0.044) ***
rXf	0.000 (0.000)	0.028 (0.000) ***	−0.003 (0.000)**	−0.004 (0.000)**
rXP_{TD}	0.000 (0.028)	0.001 (0.027)	0.000 (0.069)	0.000 (0.060)
rXσ_{TD}	0.003 (0.183) ***	0.008 (0.179)**	0.000 (0.459)	0.001 (0.409)
rXP_{TI}	−0.001 (0.055)*	0.012 (0.057) ***	−0.001 (0.138)	−0.002 (0.129)*
rXσ_{TI}	−0.019 (0.183) ***	0.007 (0.221)*	0.001 (0.459)	−0.001 (0.503)

fXP_{TD}	−0.006 (0.000) ***	−0.002 (0.000)	0.004 (0.000) ***	0.001 (0.000)
$fX\sigma_{TD}$	−0.016 (0.001) ***	−0.006 (0.001)*	0.012 (0.003) ***	0.006 (0.002) ***
fXP_{TI}	−0.014 (0.000)**	−0.006 (0.000)*	0.030 (0.001) ***	0.023 (0.001) ***
$fX\sigma_{TI}$	−0.032 (0.001) ***	−0.014 (0.001) ***	0.054 (0.003) ***	0.041 (0.002) ***
$P_{TD}X\sigma_{TD}$	−0.007 (1.498)**	−0.003 (1.443)	0.005 (3.751) ***	0.002 (3.287)*
$P_{TI}X\sigma_{TI}$	−0.051 (2.996) ***	−0.034 (2.922) ***	0.039 (7.503) ***	0.027 (6.660) ***
$P_{TD}XP_{TI}$	0.006 (0.449)*	0.005 (0.432)*	0.001 (1.125)	0.002 (0.986)
$P_{TD}X\sigma_{TI}$	0.007 (1.498)*	0.006 (1.442) *	0.001 (3.751)	0.001 (3.285)
$P_{TI}X\sigma_{TD}$	0.003 (2.996)	0.001 (2.884)	0.000 (7.502)	0.002 (6.573)
$\sigma_{TD}X\sigma_{TI}$	0.007 (9.986)*	0.000 (9.633)	−0.001 (25.007)	0.004 (21.954) ***
N	20160	20160	20160	20160
R^2	0.847	0.858	0.918	0.981
$\Delta\chi^2$(df)		151(2)		2,992(2)
% χ^2 improvement		7.4%		23.3%

Notes : †Standardized coefficients are shown along with standard errors.

*p < .05.

**p < .01.

***p < .001.

We first examine the results regarding the impact of daily IRI variation (σ_{IRI}) on DC performance and their support for H_1 and H_2. As shown in model 2 for service level as dependent variable, the standardized coefficient for σ_{IRI} is −0.207 ($p < .001$), in support of H_1. Moreover, as shown in model 2 for average inventory level as dependent variable, the standardized coefficient for σ_{IRI} is 0.149 ($p < .001$), in support of H_2. These coefficients show that daily IRI variation is quite damaging to DC performance. This is because daily IRI variation is shown to decrease service levels and increase average inventory level.

We also examined how channel demand patterns r exacerbate the damaging effects of daily IRI variation σ_{IRI}, in order to test H_{3a} and H_{3b}. As shown in model 2 for service level as dependent variable, the standardized coefficient for σ_{IRI}Xr is −0.052 ($p < .001$), which supports H_{3a}. However, as shown in model 2 with average inventory level as dependent variable, the standardized coefficient for σ_{IRI}Xr is 0.001 ($p = .374$), which provides no evidence to support H_{3b}. Thus, channel demand patterns are shown to only interact negatively with daily IRI variation with respect to service level, meaning that customer satisfaction should be especially a concern for direct channel managers when confronted with daily IRI variability.

We also evaluated the effect of IRI bias, μ_{IRI}, on DC performance. As shown in the results for model 2 for service level and average inventory as dependent variables, the standardized coefficients for μ_{IRI} are −0.102 ($p < .001$) and −0.369 ($p < .001$). Thus, as expected, a positive IRI bias (SIR > actual inventory) is unfavorable to service level but favorable to average inventory, and a negative IRI bias (SIR < actual inventory) has the opposite effect. This result comports with the literature's traditional perspective about the effects of IRI, but also contrasts with this study's findings about the effects of daily IRI variation.

While not of primary theoretical interest, we also had expectations as to how the various experimental factors interacted with the TD and TI errors that drive IRI and influence performance. Regarding channel effects, the direct channel (with higher r) has worse service levels under IRI conditions (a negative direct effect) and interacts with the magnitude of TI in reducing service levels—smaller orders are sensitive to TI. This is contrary to our expectations. Also unexpected is the favorable (i.e., negative) direct effect r has on average inventory, showing that direct channels use more inventories during replenishment cycles and that larger reorder points are needed when direct channels have IRI. Regarding the order policy factors (little s and Δ), both benefit service levels as expected when in the presence of IRI. Yet, unexpectedly, little s interacts favorably with TI and TD for service level, while Δ has a stronger unfavorable interaction with TI than with TD for both performance variables. This shows that little s is the more useful factor in the (s, S) regime as errors increase. Regarding audit policy, as expected a larger time between cycle counts (f) is generally unfavorable to both performance variables, but f does not affect inventory when daily IRI variability is accounted. Interactions between f and TI are more damaging because of the longer exposure of SIR to error opportunities. Finally, regarding TD and TI, the results overall indicate that TI is more damaging to operating performance than TD, both in the direct effects of P_{TI} and σ_{TI} and in the interaction P_{TI}Xσ_{TI}. This was expected because opportunities for these types of errors exist more frequently. Interestingly, the TI magnitude (σ_{TI}) is more damaging than the TI probability (P_{TI}), suggesting that reducing the size of the errors is more beneficial than reducing the likelihood of the errors. The issues with TI also suggest that managers should not increase inventory through larger order sizes to compensate for IRI because more on-hand inventory creates more opportunities for error.

Conclusion

Inventory management theory and practice has perpetuated the assumption that static, periodic measures of IRI provide adequate information on which to base inventory reorder and audit policies. In this study, we examine the unexplored effects of daily IRI variation as an important influence on the operating performance of DCs in retail channels. We used continuous, multiday observations of IRI from a retailer's DC that uncovered problematic literature assumptions of IRI. We then used this data as a basis to build a simulation and validate hypotheses regarding the effect that daily IRI variation and the interaction between this variation and channel demand patterns have on DC performance. Our study also uncovers how operating conditions—inventory control policies and the frequency of cycle counts—interact with IRI to impact DC performance in favorable and unfavorable ways.

Contributions to theory

This is the first study to both describe empirically the characteristics of daily IRI variation and, with a grounded simulation, give its performance implications. As such, this paper breaks new ground into how IRI behaves over time and how DC operations are affected. Our data collection at a retailer's DC shows that substantial variability exists over time in ways that do not comport with common literature assumptions (see Williams and Tokar [73]). We find that (1) IRI follows a moving average pattern likely tied to demand rather than a random-walk pattern, and (2) IRI variability is better understood by SKU-specific observations across days than single-day observations across SKUs. In particular, our data collection finds that using across-SKU IRI variation to infer actual IRI variation is an overestimate that will lead to excessive inventory on-hand. Our results also underscore the fallacy of using single-point estimates to gauge SKU IRI. As Table 2 shows, an item's IRI variation is often two or three times an item's IRI mean. Thus, our study shows empirically that there are better ways to estimate the true extent to which an item is exposed to IRI problems.

Our study is also the first to use empirical data to both describe the characteristics and explore the effects of two key drivers of daily IRI variation—TD and TI errors. Potentially surprising are the high frequencies and magnitudes of both TD and TI errors that occur. Those familiar with the many opportunities for errors in DC operations, however, will likely attest that order fulfillment and inventory replenishment are quite susceptible to such problems (Lee and Özer [40]). Our study gives evidence to support such an intuition. Yet, we push the boundary of understanding further in our simulation experiment, showing in Table 5 that the prevalence of TI errors interacts with inventory policies and impairs DC performance more than TD errors. This suggests that drivers of TI errors, such as storage practices and human behavior (DeHoratius [21] ; Rekik et al. [57]), are at the root of daily IRI variation and could be a useful avenue for future behavior research in logistics (Tokar [65] ; Knemeyer and Naylor [35]). In addition, the different interaction effects of TI and TD reveal the importance of using a multiday counting method as we have to not only assess daily IRI variation, but also to indicate whether TI or TD errors are most prevalent (as we do in Table 3).

A crucial contribution of this study is the introduction of a new approach toward modeling IRI's daily variation. We analytically replicated the empirically observed patterns and parameters of daily IRI variation. In addition, the damaging effects of daily IRI variation on DC performance were shown to be acute. An example of this is given in Figure 4; although the impact varies, daily IRI variation σ_{IRI} decreases service levels regardless of the experimental factor. This result is important because research has yet to recognize day-to-day IRI variation as critical, instead of being focused on static observations to infer variation (DeHoratius and Raman [23]) or to detect biases (Morey and Dittman [48]). Without an accounting for daily IRI variation, such an important feature of the IRI phenomena is not recognized. Daily IRI variation also has a pernicious cognitive effect: record confidence degradation. Management of increasingly complex, multichannel DCs is reliant on decision support systems for ordering and product replenishment. As the informational foundation for managerial judgment erodes, inefficient triple-checking of records and buffering of time and material likely grow, driving down operating performance. These secondary, negative effects of daily IRI variation may be worse than the primary effects we show; however, the message is the same that the variability characteristics of IRI need further study.

A final theoretical contribution stems from our regression analysis where a more nuanced understanding is given as to the role that inventory control policies play with respect to IRI. As Table 5 shows, higher reorder points (little s) and larger reorder quantities (Δ) have the expected effect of buffering inventory to protect service levels. However, our study shows these policy parameters have differing interactions with the TD and TI record errors that drive IRI variations. For instance, little s is shown in Table 5 to be particularly beneficial to service levels in its interaction with record errors; meaning that even though more inventory during the replenishment cycle can exacerbate TI errors, the earlier reorder point from a larger little s is more beneficial in the presence of larger TI errors. Yet, this is not the case for the reorder size Δ. Like little s, a larger Δ increases inventory that exacerbates TI errors, but Δ has no benefits for the replenishment cycle and, thus, the interaction of Δ and TI error is detrimental. This suggests that reducing the reorder size is a potential IRI-management strategy. Thus, by examining in detail these interactions with both type of errors, we come closer to understanding the role various managerial choices play in the operation of DCs in the presence of IRI.

Contributions to practice

DC managers must choose where to allocate limited resources, such as employee time or technology investment, in their effort to diminish the negative impact of IRI. Our research suggests that DC managers should focus their attention not only on IRI bias, but also on daily IRI variability. Indeed, the paradoxical effect of daily IRI variation increasing inventory while decreasing service levels should be of keen interest to managers. Yet, this means devoting resources to multiday cycle counts without corrections in order to estimate the degree to which IRI variation exists. If managers correct on a daily basis, then IRI's inherently dynamic nature is not captured and crucial information may be lost. Future research should provide guidance as to how many days are required.

Another recommendation from our research is that critical SKUs with perplexingly high inventory levels and low service levels be targeted for multiday counts. Following the approach described in our methodology, managers can gain rough estimates as to the likelihood that TD and TI errors exist. Items demonstrating high daily IRI variation and high TI errors should be the first priority, with particular attention being given to the magnitude of TI errors. Our results show that large, infrequent TI errors—which may be associated with large-scale spoilage—are more detrimental than small, frequent TI errors—which may be associated with small item breakage or theft. Plus, as Table 5 shows, TI errors correlate with daily IRI variation and interact with inventory policies more than TD errors. Therefore, resources dedicated to lessen large magnitude stock loss, such as assuring a proper storage environment to avoid spoilage, may be more beneficial than resources dedicated to reduce small, frequent errors, such as petty-theft prevention. Of course, reducing both the magnitude and frequency of errors are important. Yet, understanding the broader performance effects of daily IRI variation will help managers make more informed decisions under resource constraints, and help avoid the naive reaction and potentially damaging strategy of simply increasing inventory levels.

The observations we make with respect to the mixed channel effects will likely give confidence to managers that significant changes will not be required in managing IRI in multichannel DCs. That is, the demand pattern difference between direct channels as opposed to brick-and-mortar channels has minor direct and interacting influences on operating performance. The high frequency of customer orders that are typically observed in a direct channel may create the impression to managers that even minor TD errors will lead to large problems, but it appears that the smaller sizes that are common in these orders negate these consequences. However, it should be noted that because order sizes are lower in a direct channel, the potential backorder point is lower and, thus, the likelihood for picker correction diminishes. We find this to be of significant consequence to service levels, and managers should develop more intentional methods for ad-hoc corrections. For example, early warning indicators within the WMS of paradoxically higher inventory levels and lower service levels than expected could perhaps be used as an initial trigger. Another important note for direct channels is that behavioral influences such as cognitive fatigue from repeated transactions (cf. Perrey et al. [53]) or overconfidence from frequent experience (Einhorn and Hogarth [25]) are more likely and should be examined in future research. As our empirical observations show, although not statistically significant, a higher median frequency of TD errors is possible in direct channels as opposed to brick-and-mortar channels. These issues should be of interest to researchers and managers alike.

A final note to managers should be the importance of the DC employees. The context of a DC is substantially different than a retail store environment because DC employees see inventory storage for every sale. This increases the opportunity for DC employees to monitor the status of IRI. In particular, the frequency with which items are “frozen” is substantially less because orders that are accepted will not get filled and will therefore be rejected back to customer service. Crucial assumptions are that rejections will be expedient and proper stock corrections will be made. In our study, these assumptions are made and the benefits to operating performance are significant. As can be seen in Table 5, a positive, “freeze-inducing” IRI bias still has a negative effect on service, but the impact is not as strong as daily IRI variation. In a retail store environment, this would not be the case; a stable IRI bias would create long periods of no replenishment and unsatisfied customers. Our results, therefore, emphasize the crucial role that employees have on the ability of DCs to diminish the damaging effects of IRI on performance. Having DC employees solely focused on productivity may seem to be beneficial, but a loss of accuracy for the sake of increased productivity may not be in the best interest of the operation.

Limitations and future research opportunities

One of our study's limitations is that we did not observe TD or TI errors as they occurred. However, doing so would have created a “Hawthorne effect” (Mayo [44]), with employees artificially curtailing

IRI due to our presence in the DC. Therefore, we ensured that employees were unaware of the SKUs that we used for our study. Also, observing all stock locations and transactions for all 27 items simultaneously over 10 days is highly resource intensive. However, future studies could focus on one or two items in fewer locations to examine closely how TD and TI errors translate into IRI variation. Modeling work could also ascertain costs and benefits of daily IRI variation knowledge. While there is valuable information from multiday counts, these studies could determine how many days and under what conditions such an activity is beneficial.

Our simulation experiment assumed a zero error mean for both TD and TI errors. This designed the experiment to have no inherent IRI drift. While our empirical data support this assumption, it is possible that certain SKUs have an inherent negative or positive IRI tendency. For example, food items with short shelf-lives likely have positive TI error averages; meaning actual inventory depletes faster than recorded and leads to a positive IRI bias. Our simulation study also did not vary the total velocity (i.e., sales volume) of SKUs. The ramifications of an inherent drift or a higher volume will most likely exacerbate of the effects we already observed, but these should be investigated further.

Because we found empirically that errors do not occur in a random-walk fashion (largely due to TD errors), future research should also look into advancing reorder policies that account for this—particularly those items with higher transaction frequencies. Moreover, TI errors may be incorporated as indicated by more frequent stock relocations within the WMS system (Chen et al. [18]). Internal stock movement increases the opportunities for errors and such information could be used for updating the inventory policies.

Finally, our simulation experiment did not include actual behaviors of managers or employees during extreme levels of IRI. Managers will likely change inventory and monitoring policies as problems with order fulfillment occur, thus affecting IRI. Repeated failed pick attempts by employees will likely diminish confidence in the WMS, perhaps leading to communication problems or delays between departments and further exacerbating the problem. Studying the implications of such factors is beyond the scope of this paper. However, examining the social and behavioral processes that affect and interact with daily IRI variation would be interesting and useful for logistics research (Knemeyer and Naylor [35]).

Appendix

The Appendix provides additional information regarding the empirical analysis and simulation design. We organize the information by section, detail into how transaction-independent (TI) and transaction-dependent (TD) error estimates were computed. It also gives specifics on the simulation model development.

Computation of TI and TD error estimates

We use changes in the system inventory record (SIR) balance from the previous day $J_{pqt} - J_{pq,t-1}$ as a proxy for transaction amount. We infer that a receipt/return takes place when the SIR balance increases, $\Delta J_{pq} > 0$, and that a shipment takes place when the SIR balance decreases $\Delta J_{pq} < 0$. We assume that an error at day t is TI if a transaction is not observed at day t. However, if a transaction is observed at day t, we make an assumption that the error may be a combination of TI and TD. We explain the procedure for computing TI and TD below. Note the magnitude of the TD error, $\mathrm{TD}_{\mathrm{pqt}}$, is measured as the ratio between the estimated TD error quantity (TDe) with respect to the transaction amount in day t. Conversely, the magnitude of the TI error, TI_{pqt}, is measured as the ratio between the estimated TI error quantity (TIe) with respect to the available physical inventory amount in day t (Ernst et al. [26] , 1993). The frequency of TD is the sum of the number of simultaneous errors and transactions divided by the total number of transaction days, while the frequency of TI is the sum of the number of nontransaction errors divided by the number of nontransaction days. Thus, we demonstrate how periods of multiday counting can assess daily IRI variations and can indicate what the underlying causes may be.

Procedure

Recall that

$$\mathrm{IRI}_{pqt} = \begin{cases} \dfrac{d_{pqt}}{I_{pqt}}, & \text{if}_{\ pqt} > 0 \\ 0, & \text{otherwise} \end{cases}$$

where d_{pqt} is the difference between the SIR balance ($\mathrm{J}_{\mathrm{pqt}}$) and physical inventory (I_{pqt}); and that $e_{pqt} \;=\; d_{pqt} \;-\; d_{pq,t-1}$

We compute day counts D according to Table A1.

We note the proportion of days with errors $= \frac{D^e}{D} = \frac{D_{NT}^e + D_T^e}{D_{NT} + D_T}$

Let $\dot{e}_{pqt} = \dot{e}_{pqt}$ for all D_{NT} (i.e., $J_{pqt} \;-\; J_{pq,t-1} \;=\; 0$) = independent error TIe_{pqt} for that instance

Average $\overline{TIe} = \frac{\Sigma \dot{e}_{pqt}}{D_{NT}^e}$, which estimates $\mu(TIe)$. Median $M(TIe)$ = middle of $\{TIe\}$ series

$$P(TIe) = \frac{D_{NT}^e}{D_{NT}}$$

Let $\ddot{e}_{pqt} = \dot{e}_{pqt}$ for all D_T (i.e. $\mathrm{J}_{\mathrm{pqt}} \;-\; \mathrm{J}_{\mathrm{pq,t-1}} \neq\; 0$) = overall error for that instance.

Table A1. Notation to classify day counting

Day characteristics	With errors	Without errors	Total
Without transaction	D_{NT}^e	D_{NT}^{ne}	D_{NT}
With transaction	D_T^e	D_T^{ne}	D^T
Total	D^e	D^{ne}	D

We assume that $\ddot{e}_{pqt}$ can be a combination of a transaction independent error (TIe_{pqt}) and a transaction dependent error (TDe_{pqt}). That is, $\ddot{e}_{pqt} = f[TIe, TDe, P(TIe), P(TDe)]$.

We assume no difference exists in transaction versus nontransaction days with respect to independent errors. Because nonparametric analyses show a difference in error quantities (not percentiles) only by velocity (i.e., unique[u][1] , slow[s], medium[m], fast[f]); we assume accurate $\mu(TIe)$ and $P(TIe)$

estimates can be made from a "sample" of the nontransaction days (with $\dot{e}_{pqt}$ values) for the sample of instances pqt within each velocity category.

To compute $\mathrm{TDe_{pqt}}$ for each velocity (u, s, m, f), we randomly sample N values of $\ddot{e}_{pqt}$, where $N \tilde{} P(TIe) \times (D_T)$, through a random number b and estimate $TDe_{pqt}^{u,s,m,f}$ as follows:

$$TDe_{pqt} = \begin{cases} \ddot{e}_{pqt}, & \text{if } b > P(TIe) \\ \ddot{e}_{pqt} - \overline{TIe}^{u,s,m,f}, & \text{if } b \leq P(TIe) \end{cases} \forall \text{velocities}(u, s, m, f)$$

Therefore, $\mu(TDe)$ is estimated by $\overline{TDe} = \frac{\Sigma TDe_{pqt}}{\text{count of } TDe}$, median $M(TDe)$ = middle of $\{TDe\}$ series, and $P(TDe) = \frac{\text{count of } TDe}{D_T}$

Simulation model development

This section provides additional information into how our simulation model was set up and verified. A daily sequence of events in our inventory system is shown in Figure A1 and closely follows past simulation studies addressing IRI (Kök and Shang [36] ; DeHoratius et al. [22]). Figure A1 reveals important nonlinearities that underscore the necessity of using a discrete-event simulation.

In our model, the baseline demand values (μ_d, σ_d, and λ_d) represent an inventoried item having a medium sales velocity with demand and variability levels within the range of values from our empirical data. Demand order quantities follow a normal distribution with μ_d = 4.0, σ_d = 1, while demand orders' interarrival times follow a Poisson distribution with mean λ_d = 3. Based upon previous simulation studies in DC contexts (Ballou 2005; Rabinovich [54]), the replenishment lead time from an uncapacitated supply source is modeled as being normally distributed with a mean μ_{leadtime} = 5 days and a standard deviation 20% of the mean (σ_{leadtime} = 1 day). Lead-time represents one day for order fulfilling and four days in transit.

Initial physical inventory is set based on a managerial expectation of 10 days of inventory on hand. Inventory is continuously reviewed following the (s, S) regime, with an order of size S minus the inventory position being placed at little s. Parameters little s and Δ are calculated based on demand and replenishment lead times characteristics ignoring IRI. Little s is calculated based upon a common reorder point formula. The Δ is set based upon expected replenishment lead times and a customer service target of a 99% order fill rate. Thus, the parameters little s and Δ are set at five and 10 days of inventory on hand, respectively. These values are within the range of values for little s and Δ in our empirical data set.

We tested the simulation program written in Arena 5.0, following the prescriptions of Kelton et al. ([34]) and Sargent (2005). Model validation occurred through visual tracing, assessing output behavior, and comparing results with our empirical results. Random number seeds demonstrated low coefficient of variability across replications (<1%). Sensitivity analyses with model input parameters showed predictable patterns—for example, customer service levels decreased with longer lead times and smaller little s values. Evidence of "freezing" and "inflating" was also noticeable when higher IRI levels were induced. For each simulation factorial combination, we assessed the daily IRI time series and found a nonrandom walk ARIMA (0, 1, 1) pattern in most configurations at the 0.05 level. The typical range of values for daily IRI variability σ_{IRI} and IRI bias μ_{IRI} were 0.05–1.2 and −0.29 to 0.26, respectively. The results were consistent with our empirical data, indicating that our model structure replicates more realistically than previous studies on the underlying processes of IRI. Each scenario, with its customer service levels under the condition of SIR accuracy, is shown in Table A2, and the resultant correlations among our experimental factors, IRI parameters, and DC performance variables are shown in Table A3.

Table A2. Planned supply chain configuration

Little s (in days on hand)	**Δ (in days on hand)**	**Channel r**	**Service level (%)**
5	5	1	98.73
5	5	2	98.74
5	5	3	98.53
5	10	1	99.04
5	10	2	99.08
5	10	3	99.08
10	5	1	99.85
10	5	2	99.92
10	5	3	99.90
10	10	1	99.85
10	10	2	99.92
10	10	3	99.95

Table A3. Correlations among experimental factors, IRI parameters, and DC performance variables

	s	**Δ**	**r**	P_{TD}	P_{TI}	σ_{TD}	σ_{TI}	**f**	μ_{IRI}	σ_{IRI}	**Service level**
μ_{IRI}	0.291	0.099	−0.012	0.000	−0.006	−0.010	0.016	−0.480	1		
σ_{IRI}	0.034	0.111	−0.008	0.040	0.234	0.112	0.473	0.546	0.030	1	
Service level	0.847	0.085	−0.050	−0.012	−0.117	−0.049	−0.209	−0.123	0.207	−0.266	1
Average inventory level	0.462	0.776	−0.020	0.009	0.071	0.029	0.119	0.260	−0.115	0.295	0.421

Notes : *p <.05, **p <.01.

Footnotes

1 This can be represented by the designation ARIMA (0, 1, 0).

2 While use of a percentage measure is appropriate for our purpose, we note that it does induce a low, base level of variation: regular fluctuations in on-hand inventory vary IRI even if d remains constant. However, because such background noise is common among cases, it is accounted for through our statistical analysis.

3 To determine the number of replications, an approach based on Law and Kelton's (38, p.512) $n^*(\gamma) = \min\left\{i \geq n: t_{i-1,1-\frac{\alpha}{2}}\sqrt{S^2(n)/i}/|\bar{X}(n)| \leq \gamma'\right\}$ was used. The average variance and mean among the outcome variables for various scenarios were computed for 10 replications.

Then, we incremented i from 1 to n^*, at which point an error below $\gamma = .05$ was attained at $\alpha = .01$ level. Law and Kelton (38) recommends at most $\gamma = .15$ and $\alpha = .05$. It was determined that while at least 10 replications were appropriate, using 35 replications would achieve an error below 2.5% (i.e., $\gamma = .025$).

4 Unique items were announced as “discontinued” but continued to be monitored.

References

1 Agatz, N., Fleischmann, M., and Van Nunen, J. 2008. “ E-Fulfillment and Multi-Channel Distribution—A Review.” *European Journal of Operational Research* 187 (2): 339 – 56.

2 Angulo, A., Nachtmann, H., and Waller, M.A. 2004. “ Supply Chain Information Sharing in a Vendor Managed Inventory Partnership.” ***Journal of Business Logistics*** 25 (1): 101 – 20.

3 Bakos, J.Y. 1997. “ Reducing Buyer Search Costs: Implications for Electronic Marketplaces.” *Management Science* 43 (12): 1676 – 92.

4 Ballou, R.H. 2004. *Business Logistics Management*. Englewood Cliffs, NJ : Prentice Hall.

5 Banks, J., Carson, J.S., Nelson, B.L., and Nicol, D.M. 2000. *Discrete Event System Simulation*. 3rd ed. Paramus, NJ : Pearson Education.

6 Barratt, M., Rabinovich, E., and Sodero, A. 2010. “ Inventory Accuracy: Essential, but Often Overlooked.” *Supply Chain Management Review* 14 (2): 36 – 46.

7 Bartezzaghi, E., and Verganti, R. 1995. “ Managing Demand Uncertainty Through Order Overplanning.” *International Journal of Production Economics* 40 (2–3): 107 – 20.

8 Bartezzaghi, E., Verganti, R., and Zotteri, G. 1999. “ Measuring the Impact of Asymmetric Demand Distributions on Inventories.” *International Journal of Production Economics* 60 (1): 395 – 404.

9 Blankley, A.I., Khouja, M., and Wiggins, C.E. 2008. “ An Investigation Into the Effect of Full-Scale Supply Chain Management Software Adoptions on Inventory Balances and Turns.” ***Journal of Business Logistics*** 29 (1): 201 – 23.

10 Bonney, M., Popplewell, K., and Matoug, M. 1994. “ Effect of Errors and Delays in Inventory Reporting on Production System Performance.” *International Journal of Production Economics* 35 (1–3): 93 – 105.

11 Bowersox, D., and Closs, D. 1989. “ Simulation in Logistics: A Review of Present Practice and a Look to the Future.” ***Journal of Business Logistics*** 10 (1): 133 – 48.

12 Braithwaite, A., and Samakh, E. 1998. “ The Cost-to-Serve Method.” *International Journal of Logistics Management* 9 (1): 69 – 84.

13 Brooks, R.B., and Wilson, L.W. 1993. *Inventory Record Accuracy: Unleashing the Power of Cycle Counting*. Essex Junction, VT : Oliver Wight Publications.

14 Brown, K.L., Inman, R.A., and Calloway, J.A. 2001. “ Measuring the Effects of Inventory Inaccuracy in MRP Inventory and Delivery Performance.” *Production Planning & Control* 12 (1): 46 – 57.

15 Camdereli, A., and Swaminathan, J. 2005. “ Misplaced Inventory at the Retailer: Information and Coordination.”In *Electronic Proceedings of MSOM Conference*. Chicago, IL : INFORMS.

16 Cantwell, J. 1985. “ The How and Why of Cycle Counting: The ABC Method.” *Production and Inventory Management* 26 (2): 50 – 54.

17 Caro, F., and Gallien, J. 2010. “ Inventory Management of a Fast-Fashion Retail Network.” *Operations Research* 58 (2): 257 – 73.

18 Chen, L., Langevin, A., and Riopel, D. 2011. “ A Tabu Search Algorithm for the Relocation Problem in a Warehousing System.” *International Journal of Production Economics* 129 (1): 147 – 56.

19 Cohen, J., Cohen, P., West, S., and Aiken, L. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. Mahwah, NJ : Lawrence Erlbaum Associates.
20 Corsten, D., and Gruen, T. 2003. " Desperataly Seeking Shelf Availability: An Examination of the Extent, the Causes, and the Efforts to Address Retail out-Of-Stocks." *International Journal of Retail & Distribution Management* 31 (11/12): 605 – 17.
21 DeHoratius, N. 2002. *Critical Determinants of Retail Execution*. DBA Dissertation 3046095. Cambridge, MA : Harvard University.
22 DeHoratius, N., Mersereau, A.J., and Schrage, L. 2008. " Retail Inventory Management When Records Are Inaccurate." *Manufacturing Service Operations Management* 10 (2): 257 – 77.
23 DeHoratius, N., and Raman, A. 2008. " *Inventory Record Inaccuracy: An Empirical Analysis." Management Science* 54 (4): 627 – 41.
24 Delen, D., Hardgrave, B.C., and Sharda, R. 2007. " RFID for Better Supply-Chain Management Through Enhanced Information Visibility." *Production and Operations Management* 16 (5): 613 – 24.
25 Einhorn, H.J., and Hogarth, R.M. 1978. " Confidence in Judgment—Persistence of Illusion of Validity." *Psychological Review* 85 (5): 395 – 416.
26 Ernst, R., Guerrero, J.L., and Roshwalb, A. 1992. " Maintaining Inventory System Accuracy." *International Journal of Purchasing and Materials Management* 28 (3): 33 – 37.
27 Fleisch, E., and Tellkamp, C. 2005. " Inventory Inaccuracy and Supply Chain Performance: A Simulation Study of a Retail Supply Chain." *International Journal of Production Economics* 95 (3): 373 – 85.
28 Galbreth, M.R., and LeBlanc, L.J. 2010. " Overcoming Spreadsheet Risk in Supply Chain Modeling." ***Journal of Business Logistics*** 31 (2): 21 – 33.
29 Godsell, J., and Van Hoek, R. 2009. " Fudging the Supply Chain to Hit the Number: Five Common Practices That Sacrifice the Supply Chain and What Financial Analysts Should Ask About Them." *Supply Chain Management* 14 (3): 171 – 76.
30 Graff, R.T. 1987. " The Limitations of Cycle Counting." *Production and Inventory Management Journal* 28 (4): 39 – 43.
31 Heese, H.S. 2007. " Inventory Record Inaccuracy, Double Marginalization, and RFID Adoption." *Production and Operations Management* 16 (5): 542 – 53.
32 Iglehart, D.L., and Morey, R.C. 1972. " Inventory Systems With Imperfect Asset Information." *Management Science Series B-Application* 18 (8): B388 – 94.
33 Kang, Y., and Gershwin, S.B. 2005. " Information Inaccuracy in Inventory Systems: Stock Loss and Stockout." *IIE Transactions* 37 (9): 843 – 59.
34 Kelton, W.D., Sadowski, R.P., and Sturrock, D.T. 2004. *Simulation With Arena*. New York, NY : McGraw-Hill Science Engineering.
35 Knemeyer, A.M., and Naylor, R.W. 2011. " Using Behavioral Experiments to Expand Our Horizons and Deepen Our Understanding of Logistics and Supply Chain Decision Making." ***Journal of Business Logistics*** 32 (4): 296 – 302.
36 Kök, A.G., and Shang, K.H. 2007. " Inspection and Replenishment Policies for Systems With Inventory Record Inaccuracy." *Manufacturing Service Operations Management* 9 (2): 185 – 205.
37 Kumar, S., Nottestad, D.A., and Murphy, E.E. 2009. " Effects of Product Postponement on the Distribution Network: A Case Study." *Journal of the Operational Research Society* 60 (4): 471 – 80.

38 Law, A.M., and Kelton, W.D. 2000. *Simulation Modelling and Analysis*. New York, NY : McGraw-Hill Higher Education.
39 LeBlanc, L.J., Hill, J.A., Harder, J., and Greenwell, G.W. 2009. " Modeling Uncertain Forecast Accuracy in Supply Chains With Postponement." ***Journal of Business Logistics*** 30 (1): 19 – 31.
40 Lee, H., and Özer, Ö. 2007. " Unlocking the Value of RFID." *Production and Operations Management* 16 (1): 40 – 64.
41 Lee, H.L., and Billington, C. 1992. " Managing Supply Chain Inventory: Pitfalls and Opportunities." *Sloan Management Review* 33 (3): 65 – 63.
42 Li, J., Sava, A., and Xie, X. 2009. " An Analytical Approach for Performance Evaluation and Optimization of a Two-Stage Production-Distribution System." *International Journal of Production Research* 47 (2): 403 – 14.
43 Livingstone, G. 1992. " Measuring Customer Service in Distribution." *International Journal of Physical Distribution & Logistics Management* 22 (6): 4 – 6.
44 Mayo, E. 1933. The Human Problems of an Industrial Civilisation. New York : MacMillan.
45 Mentzer, J.T., and Flint, D.J. 1997. " Validity in Logistics Research." ***Journal of Business Logistics*** 18 (1): 199 – 216.
46 Metters, R., and Walton, S. 2007. " Strategic Supply Chain Choices for Multi-Channel Internet Retailers." *Service Business* 1 (4): 317 – 31.
47 Morey, R.C. 1985. " Estimating Service Level Impacts From Changes in Cycle Count, Buffer Stock, or Corrective Action." *Journal of Operations Management* 5 (4): 411 – 7.
48 Morey, R.C., and Dittman, D.A. 1986. " Optimal Timing of Account Audits in Internal Control." *Management Science* 32 (3): 272 – 83.
49 Nachtmann, H., Waller, M.A., and Rieske, D.W. 2010. " The Impact of Point-of-Sale Data Inaccuracy and Inventory Record Data Errors." ***Journal of Business Logistics*** 31 (1): 149 – 58.
50 Nahmias, S., and Smith, S.A. 1993. " Mathematical Models of Retailer Inventory Systems: A Review." In *Perspectives in Operations Management: Essays in Honor of Elwood S. Buffa*, edited by R.K. Sarin, 249 – 78. Norwell, MA : Kluwer.
51 Neeley, P.S. 1983. " A Framework for Cycle Counting." *Production and Inventory Management* 24 (4): 23 – 32.
52 Neeley, P.S. 1987. " Simple Mathematics Applied to Inventory Accuracy." *Production and Inventory Management* 28 (3): 64 – 68.
53 Perrey, S., Thedon, T., and Rupp, T. 2010. " NIRS in Ergonomics: Its Application in Industry for Promotion of Health and Human Performance at Work." *International Journal of Industrial Ergonomics* 40 (2): 185 – 89.
54 Rabinovich, E. 2005. " Consumer Direct Fulfillment Performance in Internet Retailing: Emergency Transhipments and Demand Dispersion." ***Journal of Business Logistics*** 26 (1): 79 – 112.
55 Raman, A., DeHoratius, N., and Ton, Z. 2001. " The Achilles' Heel of Supply Chain Management." *Harvard Business Review* 79 (5): 25 – 28.
56 Rekik, Y. 2011. " Inventory Inaccuracies in the Wholesale Supply Chain." *International Journal of Production Economics* 133 (1): 172 – 81.
57 Rekik, Y., Sahin, E., and Dallery, Y. 2009. " Inventory Inaccuracy in Retail Stores Due to Theft: An Analysis of the Benefits of RFID." *International Journal of Production Economics* 118 (1): 189 – 98.
58 Rinehart, R.F. 1960. " Effects and Causes of Discrepancies in Supply Operations." *Operations Research* 8 (4): 543 – 64.

59 Sanders, N.R., and Wagner, S.M. 2011. " Multidisciplinary and Multimethod Research for Addressing Contemporary Supply Chain Challenges." ***Journal of Business Logistics*** 32 (4): 317 – 23.
60 Sargent, R. 2000. Verification, Validation, and Accreditation of Simulation Models. Paper read at *Proceedings of the 2000 Winter Simulation Conference. Orlando, FL, December 10–13*.
61 Scarf, H. 1962. " The Optimality of (S, s) Policies in the Dynamic Inventory Problem." In *Mathematical Methods in Social Sciences*, edited by K.J. Arrow, S. Karlin, and P. Suppes, 196 – 202. Stanford, CA : Stanford University Press.
62 Schrady, D. 1970. " Operational Definitions of Inventory Record Accuracy." *Naval Research Logistics Quarterly* 17 (1): 133 – 42.
63 Shapiro, J.F., and Wagner, S.N. 2009. " Strategic Inventory Optimization." ***Journal of Business Logistics*** 30 (2): 161 – 73.
64 Silver, E.A. 2008. " Inventory Management: An Overview, Canadian Publications, Practical Applications and Suggestions for Future Research." I*NFOR: Information Systems and Operational Systems* 46 (1): 15 – 27.
65 Tokar, T. 2010. " Behavioural Research in Logistics and Supply Chain Management." *International Journal of Logistics Management* 21 (1): 89 – 103.
66 Ton, Z., and Raman, A. 2010. " The Effect of Product Variety and Inventory Levels on Retail Store Sales: A Longitudinal Study." *Production and Operations Management* 19 (5): 546 – 60.
67 Torres, O.C., and Maltz, A.B. 2010. " Understanding the Financial Consequences of the Bullwhip Effect in a Multi-Echelon Supply Chain." ***Journal of Business Logistics*** 31 (1): 23 – 41.
68 Ukun, C., Karaesmen, F., and Savas, S. 2008. " Investment in Improved Inventory Accuracy in a Decentralized Supply Chain." *International Journal of Production Economics* 113 (2): 546 – 66.
69 Viswanathan, S. 1997. " Periodic Review (s, S) Policies for Joint Replenishment Inventory Systems." *Management Science* 43 (10): 1447 – 54.
70 Waller, M.A., Nachtmann, H., and Hunter, J. 2006. " Measuring the Impact of Inaccurate Inventory Information on a Retail Outlet." *International Journal of Logistics Management* 17 (3): 355 – 76.
71 Waller, M.A., Williams, B.D., and Eroglu, C. 2008. " Hidden Effects of Variable Order Review Intervals in Inventory Control." *International Journal of Physical Distribution & Logistics Management* 38 (3): 244 – 58.
72 Ward, J.B. 1978. " Determining Reorder Points When Demand is Lumpy." *Management Science* 24 (6): 623 – 32.
73 Williams, B.D., and Tokar, T. 2008. " A Review of Inventory Management Research in Major Logistics Journals." *International Journal of Logistics Management* 19 (2): 212 – 32.

Figures

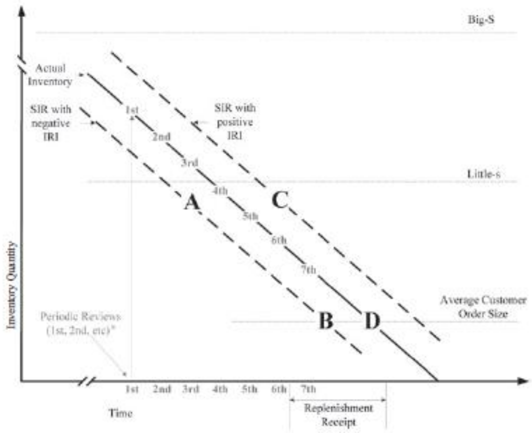

Figure 1. The reorder interval of an inventory profile with inventory record inaccuracy (IRI). Note: SIR, system inventory record.

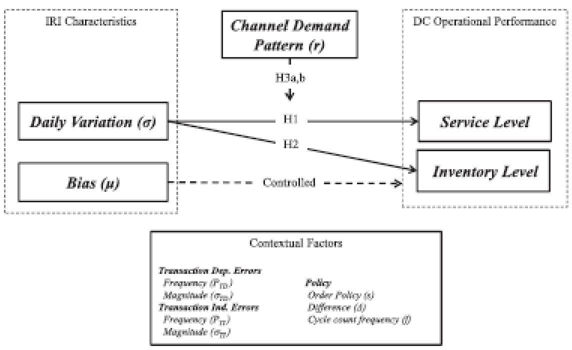

Figure 2. Influence of dynamic inventory record inaccuracy (IRI) characteristics on distribution center (DC) operating performance.

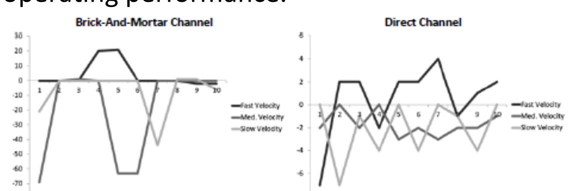

Figure 3. Example patterns of actual daily inventory record inaccuracy variations over 10 days.

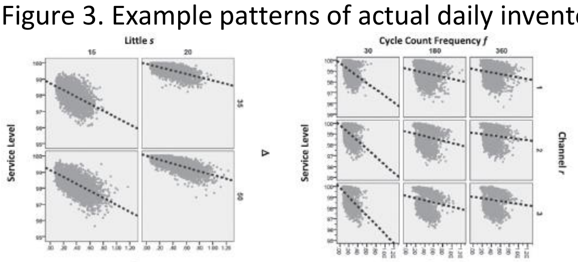

Figure 4. Impact of daily inventory record inaccuracy variation on service level over different conditions.

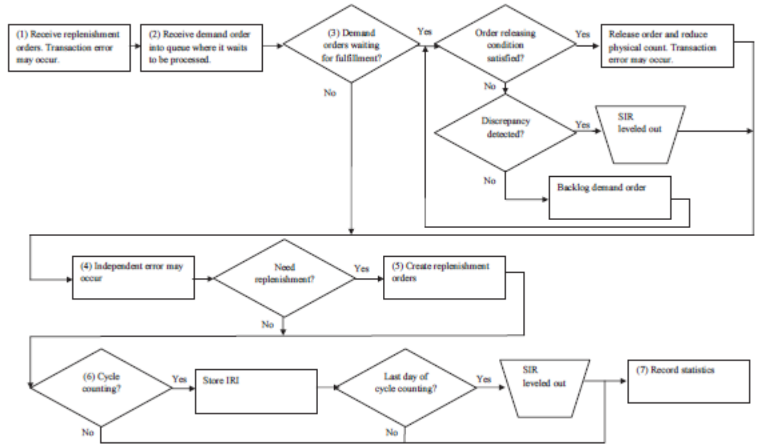

Figure 5. A1 List of events in the simulation.