ORCA – Online Research @ Cardiff

On the causes of positive inventory discrepancies in retail stores

Julian Best and Christoph H. Glock
Institute of Production and Supply Chain Management, Technical University of Darmstadt, Germany
Eric H. Grosse
Digital Transformation in Operations Management, Saarland University, Germany
Yacine Rekik
DISP Lab, EM LYON Business School, France,
Aris Syntetos
PARC Institute of Manufacturing, Logistics and Inventory, Cardiff University, UK

Abstract

Purpose

Ensuring high on-shelf availability at low inventory costs remains an important challenge in retailing. Inaccurate inventory records, i.e., discrepancies between the stock records displayed in the inventory system and the stock quantity actually found in the retail store, have been identified as one of the most important drivers of retail stockouts in the past. The purpose of this work is to investigate the causes of positive inventory discrepancies in retailing, i.e., where there is more inventory on-hand than identified by the inventory system.

Design/methodology/approach

Based on input from retailers, we develop a simulation model of a retail store that considers various error-prone processes and study in a full factorial test design how the different operational errors may drive inventory discrepancies, paying special attention to the sources of positive inventory record inaccuracies.

Findings

This enables us to gain insights into the process parameters retailers need to adjust to avoid inventory records becoming inaccurate. In addition, we analyze how positive inventory discrepancies relate to stockouts to further our understanding of the role so-called phantom products may play.

Originality

While negative inventory discrepancies (where the stock that is available in the store is less than what the system displays) and their sources (theft, shrinkage, etc.) have been discussed quite frequently in the literature, the causes of positive inventory discrepancies (where the available inventory exceeds the system inventory) have received much less attention.

Keywords: Inventory record inaccuracies, Inventory discrepancy, Retailing, Simulation, Stockcount

Introduction

On-shelf out-of-stocks continue to pose a major challenge in retailing. The literature has reported average out-of-stock rates of around 7-10% (Corsten and Gruen, 2003; Aastrup and Kotzab, 2009; Avlijas *et al.*, 2015), with this share remaining surprisingly constant over the last decades (Aastrup and Kotzab, 2010). Stockouts can have severe negative effects on the company. While some customers may decide to buy a substitute product during a stockout, other customers may prefer to buy the item at another store, postpone the purchase, or decide not to buy the item at all, which reduces sales (Corsten and Gruen, 2003). In addition to an immediate loss in sales, stockouts can also impact future sales since customers experiencing a stockout are less likely to purchase from the same retailer in the future (Goyal *et al.*, 2016). Besides affecting sales, stockouts may influence various indirect costs, for example the cost of staff attending shoppers who are unable to find an out-of-stock item (Ehrenthal and Stölzle, 2010).

The literature has discussed various causes of stockouts, including forecasting errors, ordering errors, distribution center or manufacturer problems, transport delays, or shelf replenishment issues (Bertolini *et al.*, 2015; Ehrenthal and Stölzle, 2010). Some researchers hypothesized that stockouts are predominantly caused by internal drivers, with external drivers playing a much smaller role (Corsten and Gruen, 2003). One internal driver of stockouts that is at the core of this work are inventory record inaccuracies (IRI).

IRI can be defined as the discrepancy between the recorded and the actual (physically available) inventory in a store or warehouse (DeHoratius and Raman, 2008; Chuang *et al.*, 2016; Goyal *et al.*, 2016). Inaccurate inventory records may be caused by selling and restocking errors, replenishment errors, the misplacement of items, transaction errors, or theft (Kang and Gershwin, 2005; DeHoratius and Raman, 2008; Su *et al.*, 2021). IRI can also be induced by the dynamics of the inventory system itself, e.g., due to time-consuming processes that modify

the inventory system and that occur simultaneously, creating a potential for mismatches in the system (Kull *et al.*, 2013; Barratt *et al.*, 2018). IRI are a severe problem in retailing, with earlier empirical research indicating that the share of stock keeping units (SKUs) affected by IRI may reach 65% (Kang and Gershwin, 2005; DeHoratius and Raman, 2008). SKUs suffering from IRI may display both negative and positive inventory discrepancies (Rekik *et al.*, 2019a; 2019b). In the case of negative inventory discrepancies, the number of units of a particular SKU available in the store is less than the quantity recorded in the stock management system, whereas in the case of positive inventory discrepancies, the actually available quantity exceeds the recorded one. If a SKU suffers from negative inventory discrepancies, the SKU may get ordered too late, triggering stockouts later in the ordering cycle. If positive inventory discrepancies prevail for a particular SKU, the SKU may get ordered too early, leading to unnecessarily high inventory levels at the time of the replenishment. Surprisingly though, such unnecessarily high stock levels do not necessarily prevent out-of-stock situations (Rekik *et al.*, 2019a). Amongst other reasons, this is because the inventory system is not aware of where such extra stock is (say in backrooms where the customer is unable to access it), preventing the store from selling these units.

While IRI has been discussed quite frequently in the literature especially in recent years, prior research either focused on negative inventory discrepancies (Rekik and Sahin, 2012; Cannella *et al.*, 2015; Rekik *et al.*, 2015) or on absolute discrepancies that do not differentiate between positive and negative deviations (Fleisch and Tellkamp, 2005). Positive inventory discrepancies, despite being very common in retailing (Barratt *et al.*, 2018; Rekik et al., 2019a), have received much less attention. In particular, there is a lack of research on the drivers of positive inventory discrepancies. Intuitively, one would expect that extra stock in the retail store leads to additional units that can be sold. Rekik *et al.* (2019a), however, have shown empirically that removing positive inventory discrepancies from the records may lead to a sales

increase, which may point towards the fact that surplus inventory occurs somewhere in the store or backroom where the customer is unable to access it. Rekik *et al.* (2019a) therefore hypothesized that positive IRI are as important as negative, and impact sales to a similar extent. Ton and Raman (2010) referred to such 'hidden' products as 'phantom products'. A more detailed investigation of the factors driving positive IRI, where surplus inventory occurs, and how this interacts with stockouts, has not been conducted yet.

The work at hand therefore aims to expand the knowledge on the causes of IRI in a retailing context with a special focus on positive inventory discrepancies. We develop a simulation model of a retail store that considers various error-prone processes occurring in the store and investigate how the different errors may drive (in particular: positive) inventory discrepancies. This enables us to gain insights into the process parameters retailers need to adjust to avoid stock records becoming inaccurate. In addition, we include different types of stocktakes performed at the retail store into the simulation model to further our understanding of how counting stock contributes to lowering (positive) IRI over time.

The remainder of the paper is structured as follows. The next section discusses the related literature, and Section 3 outlines the details of our simulation methodology. Section 4 statistically analyses the data obtained during the simulation runs and presents managerial insights on the causes of positive IRI. Section 5 summarizes our work and presents an outlook of future research opportunities.

Literature review

IRI has attracted the attention of researchers for many decades. One of the first works to investigate IRI is the one of Rinehart (1960), who studied the extent and the causes of inventory discrepancies at a supply facility of the US Federal Government. The researcher found that up to 50% of the items displayed some kind of inventory record error, and that the by far most

important driver of IRI were processes whose actual function was to correct the discrepancies, such as adjustments of the stock records. Kang and Gershwin (2005) performed weekly stock counts at a retailer over an eight-week time horizon and found that on average, 49% of the stock records of the retailer suffered from IRI. DeHoratius and Raman (2008) collected stocktake data at 37 stores of a major US retailer and analyzed the distribution of errors, both within and across stores. They found that 65% of the inventory records they analyzed were inaccurate. The absolute difference between the stock on the shelf and the inventory record was 35% of the average number of units found on the shelf for a particular SKU. Rekik et al. (2019a) conducted an experiment at seven major European retailers in which they counted the retailers' inventories at specific points in time over a 24-week time horizon. The researchers again found that about 60% of the stock records they analyzed were inaccurate at the time of the stocktakes, with about half of these SKUs suffering from negative and the other half suffering from positive inventory discrepancies. The authors also investigated the sales increase that could be achieved by reconciling stock records during a stocktake and found that removing both negative and positive discrepancies can lead to a major sales increase.

To protect themselves against the negative effects of IRI, companies can identify and eliminate the sources of errors, correct the errors, for example during stocktakes, or reflect the existence of errors in the (design of the) stock replenishment policies (Hardgrave et al., 2009; Rekik et al., 2019a). One topic that has attracted some attention in the past is the use of RFID in preventing and correcting IRI. Goyal *et al.* (2016) and Bertolini *et al.* (2015) compared RFID-supported counts to manual and barcode-supported counts and found that RFID led to a sharp decrease in both IRI and stockouts. Morey and Dittman (1986) proposed methods for calculating the optimal frequency of stock audits needed to meet a given accuracy level. Stock control policies that take account of inaccurate inventory data were proposed by Kang and Gershwin (2005) and Rekik *et al.* (2019b). Kang and Gershwin (2005) developed a continuous

5

review inventory control model that takes account of inventory shrinkage. The authors then evaluated different methods for protecting the company against IRI, such as additional safety stocks, manual stock counts, or the use of RFID that was assumed to eliminate shrinkage. Rekik *et al.* (2019b) proposed a periodic order-up-to level inventory system subject to IRI and showed how the stock control policy should be adjusted to minimize the total expected inventory cost subject to a service level constraint. They also proposed a method for estimating the inventory error from data the company may collect over time.

Works that are particularly relevant to the manuscript at hand developed simulation models to explain how IRI occur in retail supply chains or to investigate how managerial actions can contribute to reducing IRI. Fleisch and Tellkamp (2005) developed a simulation model for a three-stage retail supply chain and considered several factors causing IRI in their model, namely theft, incorrect deliveries, misplaced items, and unsalable items. The authors compared the case where IRI accumulate over time to a situation where inventory records are reconciled at the end of each period. The results of the study show that removing inaccuracies from stock records can leverage the performance of the entire supply chain. Gumrukcu *et al.* (2008) also simulated a supply chain consisting of a retailer, a distribution center, and a supplier, and compared the cost of carrying additional inventory to the case where cycle counting is used to detect IRI. They found that cycle counting is particularly effective for low-demand-high-cost items. Condea *et al.* (2012) investigated how RFID can improve shelf replenishments in a retail store suffering from shrinkage. The authors considered a scenario where RFID is used to detect cases moving from the backroom to the salesroom; the information collected this way is used for scheduling shelf replenishments. The performance of this system was compared to a system where inventory is checked on shelves in periodic manual stocktakes. The results of the study imply that depending on data (reading) quality, RFID has the potential to lower total cost and increase service levels. If reading mistakes occur, then backroom replenishments that rely

exclusively on RFID data may trigger too many replenishments; a heuristic developed by the authors that helps to eliminate unnecessary replenishments improves the performance of the system in this case. Thiesse and Buckel (2015) also compared an RFID-based store replenishment system to a periodic review system and extended earlier works by considering two different RFID tagging strategies, namely item level and case level tagging, and by considering constraints on shelf space and service levels. Their results show that case level tagging is much more susceptible to reading errors; item level tagging, in turn, does not suffer that much from reading errors. Kull *et al.* (2013) studied the influence of high daily IRI variation on inventory and service levels of a multi-channel retailer. Their simulation model considered two different types of IRI drivers, namely transaction-dependent errors and transaction-independent errors. The results of their study imply that using across-SKU IRI variation to infer actual IRI variation is an overestimate that may lead to excessive on-hand inventory. In addition, they showed that referring just to the mean IRI in replenishing an inventory system is not sufficient, as the variance of IRI can exceed the mean in many cases. Finally, Barratt *et al.* (2018) developed a simulation model to gain insights into how employee-system interaction causes IRI. One main result was that visual complexity caused by higher inventory levels complicate IRI correction. Albeit focusing on distribution centers, their results are relevant for IRI research in retail stores as well.

Our overview of the literature shows that IRI has been analyzed from different perspectives and using a variety of methodologies in the past, including simulation. Positive inventory discrepancies and their causes have, despite their importance in practice (Rekik *et al.* 2019a), not attracted much attention so far. The work at hand contributes to closing this research gap by developing a simulation model that helps to explain how different decisions and errors occurring in retail stores (e.g., shelf sizes, selling of items in case packs) may cause positive inventory discrepancies or modulate their intensity. We extend prior research that investigated

causes of IRI (Raman *et al.*, 2001; Fleisch and Tellkamp, 2005; Kull *et al.*, 2013; Barratt *et al.*, 2018). The main contribution of our work is the investigation of new influence factors of positive IRI and the analysis of interaction effects between them in a single simulation.

Simulation model

Practitioner workshop

To practically ground our simulation model and to improve the interpretation of its results, we conducted a workshop with professionals from eight different retail companies and associations. The objective of the workshop was to motivate our research questions, highlight the importance of considering positive IRI, and verify our research design and the parameter values used in the simulation runs. In the workshop, the participants discussed possible influence factors and the relevance of positive IRI in retail stores, and quantified examples of order-delivery-discrepancies, misplacements of items and checkout errors. It was found that practitioners generally focus more on negative IRI because it tends to create more availability issues. However, also positive discrepancies are considered relevant due to potential losses in sales. The participants also stated that the checkout quality has a direct influence on IRI and provided multiple examples for possible error causes (e.g., product recognition errors, incorrect barcodes on the items, scanning errors like over- and underscanning). Especially the use of case packs poses a problem at the checkout. These packs increase the overall complexity of handling items as there are multiple "levels of aggregation" that can lead to problems with stock taking and inventory management on the shopfloor. For example, multiple single items entering the store could get registered as case packs, creating positive IRI. One of the retailers pointed out that positive IRI also originate from items being misplaced while the affected product is counted in a partial stocktake.

The workshop clearly highlighted the importance of the topic investigated in this paper. All error classes mentioned during the workshop were considered in the simulation model. The empirical grounding of the simulation model enables us to reflect the results and verify the outcome integrity in light of the practitioner feedback we received.

General framework

We develop a discrete event-controlled simulation model that is roughly based on the retail store model of Condea *et al.* (2012) that consists of a backroom and a salesroom, an automatic inventory system that handles reorders and shelf replenishments, and basic inventory management processes to simulate item and information flow. The model was implemented in the software Tecnomatix Plant Simulation. The retail store investigated in this paper is illustrated in Figure 1. Solid arrows represent the flow of physical goods, whereas dotted arrows represent information that is exchanged between the different areas, linking IRI sources to the product and information flow (see subsection "sources of IRI" below).

Items arriving at the retail store from a supplier or a distribution center of the retailer are stored in the backroom without being scanned individually, and the inventory system is updated directly based on the ordered amount. The backroom inventory is used to replenish the shelves in the salesroom, where each item sold by the store has a dedicated shelf area with a certain storage capacity. While the "regular" flow of products is from the backroom to the salesroom, items may also flow in the opposite direction, e.g., if misplaced items are found again during a stocktake.

Figure 1: Conceptual model of the retail store including IRI sources

The inventory control system monitors the replenishment of the store for all items. We assume that the retailer uses a reorder threshold, order quantity (r, Q) continuous review inventory control policy (Kull *et al.*, 2013), as this is a policy that is, according to the authors' experience,

frequently used in retailing. In this case, an order quantity Q is placed at the suppliers once the total available inventory of an item reaches the reorder threshold r. We assume that the quantity actually delivered to the store can differ from the order quantity Q due to errors (Q + error), with the error being normally distributed. The error has an expected value of zero, and its variance depends on Q, which means that higher order quantities may lead to higher deviations of the delivery quantity from Q.

Once a customer enters the salesroom, a shopping list is generated. A binomial distribution determines the item classes that will be bought, and then the number of requested units of each item class is generated according to a geometric distribution to account for the fact that retail customers often buy small item quantities.

Each time an item has been sold with the sale registered at the checkout counter, the stock record quantity of that item is reduced. Once a shelf replenishment threshold of an item has been reached (note that this is different from the store reorder threshold that initiates store reordering), a replenishment order is triggered at the backroom and the shelf is fully replenished (Condea *et al.*, 2012). The simulation model also tracks the "real" inventory in the system that is subject to errors; the "real" inventory, however, is not known to the stock control system that only has access to information that a real-world retail store would have as well.

Sources of IRI

The simulation model considers various sources of IRI. For each IRI source, an individual probability distribution is used. The implementation of the IRI sources (in terms of type and probability distribution) was taken from the related literature (Raman *et al.*, 2001; Fleisch and Tellkamp, 2005; Condea *et al.*, 2012) as well as from initial individual discussions with loss protection managers from a group of retailers. All considered sources of IRI were

compressively discussed, validated and approved in the subsequent practitioner workshop. The following error classes and sources of IRI are considered in the simulation model:

- *Wrong deliveries*: A store may receive more or fewer items than ordered, for example because of errors that occurred during the ordering process (modelled using a normal distribution) or due to damage during transport (modelled using a binomial distribution);

- *Checkout errors*: At the checkout counter, errors may lead to wrong adjustments of the stock records. We consider transaction errors (an item is not scanned or scanned multiple times; modelled using a geometric distribution), customer theft (an item leaves the store without being scanned, modelled using a binomial distribution), and case pack errors (an individual item is scanned as a case pack or vice versa, modelled using a binomial distribution);

- *Shrinkage*: Items disappear in the store due to employee theft or spoilage not registered in the inventory system (modelled using an Erlang distribution for the event to occur and a combination of a binomial and a geometric distribution to determine the affected products and the number of stolen/spoiled items);

- *Misplacement errors*: Customers or employees put items into wrong shelf locations, either in the salesroom or the backroom (modelled using a binomial distribution for the products and a uniform distribution for the shelf).

The literature also discusses *stock count errors* as a possible source of IRI where employees count the wrong number of items during stocktakes (Barratt *et al.*, 2018; DeHoratius and Raman, 2008). Stock count errors were, however, not included in the simulation model due to their obvious effects on IRI that would only compromise statistical analysis without providing meaningful insights. For example, systematically counting more items than actually available

in the store will inevitably lead to positive IRI. Note that if no systematic errors occur during stocktakes (i.e., employees do not have a tendency to count too much or too little) and the result of each stocktake fluctuates evenly around the real value, the statistical results are not affected as these errors cancel out in the long run.

Formulation of research questions

To analyze the influence of a set of model parameters that were identified as especially important in the literature review and the practitioner workshop on the behavior of the system, we formulate research questions (*RQs*) and test them using a statistical model:

- *RQ 1*: Do case packs have higher positive IRI than individual units?

- *RQ 2*: How do shelf sizes (backroom and salesroom) and thresholds (store reordering and salesroom replenishing) affect positive IRI?

- *RQ 3*: How do sales affect positive IRI?

- *RQ 4*: How do different rates of misplacement in the backroom and salesroom affect positive IRI?

- *RQ 5*: How does time affect positive IRI (within and across stocktake cycles)?

- *RQ 6*: Is there a relationship between stocktake frequency and positive IRI?

Note that the RQs are not only investigated independently but also in combination to allow testing of interesting interaction effects between them (e.g., do case packs introduce positive IRI in conjunction with higher misplacement rates etc.). This allows us to gain more granular insights into the sources of positive IRI than the extant literature. Also note that not only the existence of a link is investigated, but also the extent of the effect. In formulating the *RQs*, we excluded parameters that have a direct effect on IRI. For example, if the theft probability is increased in the model, then shrinkage will increase as well, which reduces sales.

Implementation of stock counts

Stock counts are mandatory under commercial law for most companies, and they are also an effective measure for removing errors from the stock records of a company. Given that stock counts are very labor-intensive and, hence, expensive, companies aim at optimizing the trade-off between the cost of stock counts and the reduction of IRI-induced errors due to improved inventory information (Kök and Shang, 2007). In this paper, we assume that the retailer may schedule stock counts to correct inventory data in the stock control system. The simulation model considers the following stock count policies:

1. *Bi-annual stock counts*: Every six months, the entire store is counted (complete stock count).

2. *Stock counts during failed replenishments*: If the system wants to replenish a shelf in the salesroom but fails to find sufficient items in the backroom (despite the system indicating that sufficient items are available), and if no orders are outstanding, then a stock count is triggered for the affected product class(es) (partial stock count). This strategy avoids inventory freezing.

3. *Stock counts during backroom overflows*: If the store receives a shipment but is not able to store the full charge of items in their dedicated zone(s) in the backroom, as this zone is already full, then a stock count is triggered for the affected product class(es) (partial stock count). This strategy avoids the very rare situation where inventory for a particular SKU becomes excessive. During normal operations where only a limited number of items exceeds the capacity of a backroom shelf, no stock count is triggered.

In addition to verifying and eventually correcting stock records, the purpose of stock counts in our model is also to identify misplaced products and to bring them back to their correct storage location in the backroom. Wrong storage locations may also be corrected

- when a new shipment arrives at the backroom of the store. With a given probability (binomial distribution) the misplaced items in the respective backroom shelves are detected and brought back to their correct location;

- when a shelf is replenished. With a given probability (binomial distribution) the misplaced items in the respective salesroom shelves are detected and brought back to their correct location.

Simulation times

The simulation model assumes that almost all considered processes require time to be completed. This applies to delivery, receiving, store reordering, shelf replenishment, shopping, paying, stock counts, correcting errors, and customer arrivals. These random times may leverage the effect of inaccurate data. For example, replenishing a shelf takes a certain time, and during this replenishment process, further sales may be lost. All times are stochastic and drawn from an Erlang distribution, except for the bi-annual stock count which is assumed to be completed instantly without affecting other store functions. Also, we assume that all chosen process times are the same for all products/product categories.

Experiment design

In the simulation, we study nine model parameters (listed in Table 1) that are necessary to answer the RQs. The impact of these parameters on IRI is tested in a full factorial design (see e.g., Fleisch and Tellkamp, 2005). For each parameter, we assume that it can either adopt a high or a low value, or that the respective model parameter is implemented or not implemented (e.g., a high vs. a low probability of misplacement, or the SKU is stored in a case pack vs. it is stored as individual units on the shelf). Considering nine parameters with two possible values each leads to $2^9 = 512$ scenarios we study in the simulation. The experiment was set up to consider 512 different products, with each product corresponding to one of the scenarios, and

to study how the different products behave over time. This allows gaining insights into how the different parameters impact the behavior of the system and thus IRI. The simulation model captures two types of inventories over time: the inventory recorded in the stock management system and the "real" inventory that corresponds to the number of items physically available in the store. The difference between both is the inventory discrepancy for product i on day t (that can be both positive or negative):

$$Inventory\ discrepancy_{it} = real\ inventory_{it} - recorded\ stock_{it}$$

Table 1 summarizes the different model parameters that are varied for the full factorial test design. Please note that there are more model parameters that are kept constant during all simulation runs. The inventory discrepancy for every product is sampled on a daily basis and then averaged for a time period of 30 days to smooth the high volatility. Note that the results of the IRI sampling will be used exclusively for the statistical analysis and do not enter the simulated inventory system of the retail store, which operates unaffected by this sampling. Every 180 days, a complete stocktake (bi-annual stock count) is carried out to update the records and find misplaced items, but not to restore the initial stock level, leaving us with six observations for every product in one stock-taking cycle. This facilitates the statistical analysis by filtering high-frequency components from the time series and thus reducing complexity while still retaining data points between stocktakes to represent the overall IRI trend. The averaging may lead to the cancelling out of positive and negative IRI, but it preserves the overall trend in the long run. Since the simulation model relies heavily on statistical distributions to model processes, the number of model repetitions for every time step needs to be selected carefully. To derive a feasible number of repetitions for the simulation run, a short test run of 30 days was evaluated. The results suggest that with only 100 repetitions, the confidence interval of the mean IRI can already be reduced significantly, while more repetitions lead to quickly decreasing benefits compared to the computational effort. Based on

15

this test run, the number of repetitions was set to 500 to get the confidence interval below 0.2 while also providing some headroom. Next, the overall simulation time was determined after analyzing a short test run with only 32 products. For the test run, the confidence interval change for the mean IRI was calculated for an increasing number of stocktake cycles (mean over all products; the last 30 days prior to the stock count were used due to the highest variance of the sample). The results show that the mean confidence interval quickly drops below 0.9 within just six cycles before the improvement decelerates drastically compared to the computational effort. Therefore, the simulation time was set to 10 stocktake cycles to provide reasonable data input while also enabling some additional headroom.

Other assumptions made in developing the simulation model are the following:

- The dimensions of the items on stock (such as size of cartons) are not considered. Shelf capacities are therefore formulated just as a function of the number of items.

- To avoid initialization biases, the model starts after a complete stock count with a fully replenished store and ends after another complete stock count.

- The model considers a temporary storage area where items are kept in case the shelf capacity for an item class in the backroom and the salesroom has been reached. During the validation phase of the simulation model, we verified that the temporary storage area did not exceed its capacity during any run and that it did not hold unrealistic amounts of items at the same time.

Table 1: Model parameters for full factorial test design

Results

Descriptive results

Figure 2a plots the mean inventory discrepancy for every month and every item individually during the simulation timeframe. As can be seen, inventory discrepancies are almost centered

on 0 (mean = -0.7, median -2.07) with an asymmetric variance. In addition, it can be observed that the overall variance increases almost linearly within each stocktake cycle until it is reset after the bi-annual stock count. The result is a seasonal pattern of inventory discrepancies, where discrepancies are the highest just before the stock count and the lowest after the stock count has taken place. Another interesting observation from Figure 2a is that discrepancies are larger in the positive direction. This could be the result of misplaced items that can potentially drive positive discrepancies if the inventory level is updated incorrectly. Figure 2b illustrates the example of a misplaced shipment that leads to large positive discrepancies. After the shipment has been misplaced, the associated products are no longer found in the store, and consequently the system orders anew after the error is detected. This leads to a high positive inventory discrepancy. The missing units are then discovered towards the end of the experiment, and the records are reconciled.

Figure 2: a) Mean inventory discrepancies; b) Misplaced item causing high positive inventory discrepancies

Additionally, an analysis of individual item classes in our data reveals that the inventory discrepancy develops almost linearly (increasing and decreasing) between the bi-annual stock counts. The general behavior of IRI observed in the results of the simulation study confirm the empirical findings of Rekik *et al.* (2019a), who also found that IRI increase linearly after a stock count (as opposed to Rinehart (1960), who observed a degressive increase).

Statistical model

We now use a linear model for analyzing how the different model parameters influence the dependent variable (inventory discrepancies). Due to the experiment design, we use all model parameters as well as all interaction terms as independent variables. In addition, variables that identify the stocktake cycles as well as the observations within each cycle were added to

17

account for autocorrelation of the time series. The first and last observation within each cycle were discarded due to undesirable inconsistencies caused by the bi-annual stock count (e.g., settling phase of the simulation model after the stocktake, imprecise assignment of some variables to a specific stocktake cycle close to the stocktake, etc.). To be able to answer *RQ 6*, we later also expand the model with the number of registered stocktakes as an independent variable to predict inventory discrepancies.

Results of the regression analysis

The results of the basic regression analysis (interaction effects are not displayed) are summarized in Table 2. We observed that about 75.5% of the variance of the inventory discrepancies can be explained by the independent variables ($R^2 \sim 0.755$). The p-values (heteroscedasticity-consistent standard errors were used) of several variables indicate that the results are highly significant. The encoding of the model treats all "low" characteristics (see Table 1) as the control group. Therefore, the intercept β_0 = -4.505 shows the mean IRI in case all simulation parameters are set to their "low" characteristic. All other estimates describe the relative impact of the independent variables on IRI when switched to the "high" characteristic. The analysis thus investigates the case where the independent variables change from "low" to "high". Tables that show the results of the regression analysis for each individual research question can be obtained from the authors upon request.

Table 2: Partial results of the regression analysis

Answering the research questions

Using the results of the regression analysis, we are now able to answer the research questions:

RQ 1: Do case packs have higher positive IRI than individual units?

The model suggests that the casing of items is a major driver of positive IRI, with cased items increasing IRI by 10.761 units. This supports the results of our workshop where the participants

18

mentioned that case packs cause handling errors, such as checking out a single item as a case pack. Further supporting the strong positive link between items sold in case packs and positive IRI, we can see that in-between the bi-annual stock counts, IRI increases monotonically by about 2.7 units per time step if the item is sold in case packs.

Other interactions show significant influences as well. Shelf sizes (CB and CS) and both thresholds (RE and RO) show significant negative interaction estimates with our case pack parameter, indicating that items sold in case packs are more likely to result in more positive IRI when shelf sizes (CB, CS) are smaller and thresholds (RE, RO) are lower. Our understanding of this behavior is that small shelves and thresholds lead to more shelf stockouts, which matches the results of Eroglu *et al.* (2011). From the workshop with the retailers, we know that in case of stockouts, retailers tend to carry out partial stocktakes of the affected item classes to correct the inventory records. More stockouts lead to more stocktakes, which creates more opportunities to falsely update the inventory records in a positive direction when items from that class are misplaced.

The findings also suggest a strong positive relationship with a high misplacement rate of items, showing that cased items are more likely to suffer from higher positive IRI than uncased items. This result is intuitive, as misplacing a case pack directly leads to a higher count of misplaced items. Especially the misplacement of case packs during the replenishment of the salesroom shelves impacts IRI considerably (4.533). This could also be traced back to the mechanism outlined before, as misplacements of items in the wrong salesroom shelf can lead to a perceived stockout in addition. However, while the misplacement of case packs during the replenishment of the salesroom shelves impacts IRI considerably, case packs misplaced by customers and during the reorder process in the backroom do not seem to influence IRI substantially, most likely due to their smaller impact on the emergence of stockouts on the salesfloor. This result

is in line with the statements of the workshop participants, who predicted that misplacements by customers would only have a negligible effect overall.

RQ 2: How do shelf sizes (backroom and salesroom) and thresholds (store reordering and salesroom replenishing) affect positive IRI?

The main effects for shelf sizes and thresholds suggest that IRI increase monotonically over time for smaller shelf sizes (CB and CS) and lower store reorder thresholds (RO), indicating an overall positive effect on IRI. Additionally, there are important interactions with other influence factors that can alter the impact on IRI considerably. In this regard, we observed a strong relationship between the shelf sizes and thresholds and the misplacement rate of items during the replenishment of the salesroom shelves. The estimates suggest that small shelf sizes and thresholds lead to more positive IRI in combination with higher misplacement rates during the shelf replenishment process. This is most likely also due to the effect we discussed above already – more replenishments lead to more stockouts and a higher potential to falsely update inventory record in the positive direction. The cross-interaction effects between the four parameters (CB, CS, RE, RO) support this argument and were also estimated to have a positive impact on the overall IRI.

RQ 3: How do increased sales affect positive IRI?

Our results suggest that higher sales, altogether, induce more negative IRI (base effect = -1.504) that further decreases monotonically in-between the bi-annual stock counts. This outcome confirms the results of DeHoratius and Raman (2008), who showed that higher item sales lead to more IRI; we extend their analysis, however, by also considering the direction of IRI. Our model shows that the direction of IRI that result from higher customer demand depends on the accumulation of different errors within the store. In general, higher customer demand increases the frequency of processes in the store (e.g., more salesfloor replenishments, more store reorder

processes, etc.) and therefore boosts inherent errors. For example, our results suggest that a higher customer demand in combination with a higher misplacement rate during shelf replenishments induce more positive IRI (2.461). This result is intuitively appealing, as a larger customer purchase quantity leads to more frequent salesroom replenishments, creating more chances for item misplacements. However, the interaction estimate between cased items and customer demand (-1.181) indicates more negative IRI due to more negative errors during checkout.

RQ 4: How do different rates of misplacement (misplacement of items in the backroom, salesroom, and by customers) affect positive IRI?

Overall, only the main effects of the rate of misplaced items during the replenishment of salesroom shelves show a significant (positive) impact on IRI that increases about linearly in time in-between the bi-annual stock counts (~1.9 per time step). In contrast to that, the estimates for products misplaced by customers or during the reordering process in the backroom are, overall, not significant or of very small magnitude. This confirms the results of Fleisch and Tellkamp (2005), who found an improvement in IRI for lower misplacement rates. Extending their results, our model suggests multiple relevant interaction effects between other influence factors and higher misplacement rates (see RQ1, RQ2, and RQ3). Our model suggests that cased items as well as smaller shelf sizes (CB, CS) and lower replenishment and reordering thresholds (RO, RE) favor more positive IRI when combined with high misplacement rates during salesroom replenishments. Additionally, higher sales lead to higher IRI estimates in combination with more frequent item misplacement.

RQ 5: How does time affect positive IRI (within and across stocktake cycles)?

Across the bi-annual stocktake cycles, no significant IRI patterns could be identified. As Figure 2a shows, the time series is stationary and behaves almost identical for the different cycles.

21

However, our model emphasizes the importance of using the time dimension to describe the data within a stocktake cycle, as the interaction effects as well as the base time effects could be estimated with very high significance. The development of IRI shows a linear behavior over time after the bi-annual stocktake for most of the investigated factors. This shows that the induced errors are consistent over time (considering both impact strength and direction) and do not induce large gaps or leaps or even change effective direction. Previous works observed different development patterns of IRI after stocktakes. Our simulation supports the findings of Rekik *et al.* (2019a) and contributes to our understanding of how IRI develops over time.

RQ 6: Is there a relationship between stocktake frequency and positive IRI?

An alternative model that only investigates the relationship between the stocktake frequency and IRI reveals that the stocktake frequency is also a suitable predictor of IRI. The average number of stocktakes per time step alone was able to explain about 59.1% of the variance in IRI ($R^2 \sim 0.591$). A plot of this relationship (see Figure 3) also suggests a slight saturation effect. Therefore, an additional quadratic term was fitted, increasing the R^2 of the model to 60.1%. The plot also indicates that the estimated slopes differ substantially for the different time steps. To account for this behavior, a mixed model was fitted to allow for different slopes over the different points in time. A Likelihood-Ratio test calculated a Pseudo-R^2 of 70.6 %.

The very significant negative estimate (*p*-value $< $ 2e-16) of the quadratic term implies an inverse U-shape. In combination with the negative intercept estimate (-3.313), this supports the concern that was raised by the retailers examined by Rekik *et al.* (2019a) and the extant literature that a certain amount of stocktakes can true up the stock records effectively (DeHoratius and Raman, 2008; Hardgrave *et al.*, 2013; Chuang *et al.*, 2016), while too many stock counts can deteriorate the stock record quality (Rinehart, 1960). We confirm and extend these insights, as the quadratic estimate also suggests that this effect saturates as the number of stocktakes becomes large.

Figure 3: Relationship between stock taking per time step and (positive and negative) IRI

Conclusion

Main insights

This paper contributes to the literature on inventory record inaccuracies (IRI) in retailing by investigating the causes of positive inventory discrepancies in retail stores. A simulation model of a retail store was developed that considers a comprehensive set of IRI sources, different stock count policies as well as various error-prone processes occurring in the store. A workshop involving retail practitioners was conducted to validate the simulation model and to help interpret the results. The main insights obtained in this work can be summarized as follows.

First, the misplacement rate of items was shown to be one of the main drivers of positive IRI. While this result might seem obvious at first glance, it is also meaningful to compare the different types of misplacements and their impact on IRI for the retail store. The results indicate that misplacement of products during shelf replenishments has the largest effect on positive IRI, while item misplacement by customers or misplacement of incoming orders in the backroom did not show the same magnitude and significance. Possible reasons could be that the magnitude and/or frequency of misplacements is too low to be reliably observable in the data or the rectification of the errors happens too quickly before discrepancies can manifest themselves. In this context, it is also important to note that misplacements alone never lead to positive IRI, but only in interaction with processes that falsely update the inventory systems (e.g., partial stocktakes). The simulation showed an improvement of IRI with an increasing frequency of stocktakes, while too many stocktakes impose the risk of falsely updating the inventory records due to overlooking misplaced items, creating positive discrepancies. We also saw that case packs have a very significant positive impact on discrepancies. Especially in

interaction with the misplacement of items during shelf replenishments, a large positive effect could be observed in the simulation data. This indicates additional error processes for items in case packs that act primarily in a positive direction (e.g., wrong counting during checkout or easier misplacement of multiple items at once). The simulation also confirms the findings of Rekik *et al.* (2019a) that IRI develops linearly over time after a stocktake.

The results of our simulation should be helpful for managers to better understand the drivers of incorrect inventory records and to isolate factors leading to positive inventory discrepancies. The insights obtained from this work could educate the selection of countermeasures in retail stores to reduce IRI and ensure higher on-shelf availability.

Limitations

This work has limitations. First, we did not use empirical data in the simulation study. Parameterizing the model using an empirical dataset could help us gain further insights for a specific retail application. We note, however, that our simulation model generally supports such a parameterization. Secondly, the results of the regression model depend on the parameters and their characteristics assumed in the simulation model and changing them could lead to different results. In addition, we only considered a selection of errors in the model that frequently occur in retail stores in practice and neglected other types of errors that could also lead to inventory discrepancies (e.g., errors during stock counts or errors induced by visual complexity or high inventory density). We also note that the interpretation of the results required some amount of judgement, as for example a positive effect could have also resulted from a reduction of negative discrepancies still leading to negative IRI. Finally, the developed simulation model relies on some simplifications to reduce complexity. For example, we did not model individual employees with unique properties and we only considered a simple supply chain model to keep the paper focused on processes occurring in retail stores, which could be seen as a limitation as well.

Future research

Since this work only investigated the effects of different influence factors on positive IRI for a single retail store, future studies could investigate entire supply chains as the observed effects could differ along the supply chain stages. Future studies could also extend this work by investigating how technology (e.g., smart shelves and RFID) can support retailers in tracking stock movements that would otherwise remain unobserved. A simulation model could, for example, account for the fact that RFID-supported stock counts are not always accurate (Bertolini *et al.*, 2015) and investigate how stock record errors introduced by RFID influence the system. Apart from this, future studies could also explore the impact of frequent gap scans on the development of IRI over time and try to derive an 'optimal' number of gap scans that balances the cost of verifying stock records and those of positive or negative inventory discrepancies. The simulation model developed for this study could also be expanded to represent retail stores with integrated online shops. The model could be extended, for example, to account for a situation where online orders are satisfied from the backroom, such that an additional (error-prone) demand process depletes inventory in this area of the retail store. Another useful extension of the model would be the addition of errors during stock counts in combination with the simulation of opening hours as this allows for detailed insights in some dynamic behaviors of retail stores with regard to sources of IRI (e.g., the timing of stock counts, gap scans, etc.).

References

Aastrup, J.; Kotzab, H. (2009): Analyzing out-of-stock in independent grocery stores: an empirical study. *International Journal of Retail & Distribution Management*, 37(9), 765-789.

Aastrup, J.; Kotzab, H. (2010): Forty years of Out-of-Stock research – and shelves are still empty. *International Review of Retail, Distribution and Consumer Research*, 20(1), 147-164.

Avlijas, G., Simicevic, A., Avlijas, R., Prodanovic, M. (2015): Measuring the impact of stock-keeping unit attributes on retail stock-out performance. *Operations Management Research*, 8(3), 131-141.

Barratt, M.; Kull, T.J.; Sodero, A.C. (2018): Inventory record inaccuracy dynamics and the role of employees within multi-channel distribution center inventory systems. *Journal of Operations Management*, 63(11), 6-24.

Bertolini, M.; Bottani, E.; Romagnoli, G.; Vignali, G. (2015): The impact of RFID technologies on inventory accuracy in the apparel retailing: Evidence from the field. *International Journal of RF Technologies*,6(4),225-246.

Cannella, S.; Framinan, J.M.; Bruccoleri, M.; Barbosa-Póvoa, A.P.; Relvas, S. (2015): The effect of inventory record inaccuracy in information exchange supply chains. *European Journal of Operational Research*, 243(1), 120-129.

Chuang, H.H.-C.; Oliva, R.; Liu, S. (2016): On-shelf availability, retail performance, and external audits: a field experiment. *Production and Operations Management*, 25(5), 935-951.

Condea, C.; Thiesse, F.; Fleisch, E. (2012): RFID-enabled shelf replenishment with backroom monitoring in retail stores. *Decision Support Systems*, 52(4), 839-849.

Corsten, D.; Gruen, T. (2003): Desperately 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(12), 605-617.

DeHoratius, N.; Raman, A. (2008): Inventory record inaccuracy: an empirical analysis. *Management Science*, 54(4), 627-641.

Ehrenthal, J.C.F.; Stölzle, W. (2010): An examination of the causes for retail stockouts. *International Journal of Physical Distribution & Logistics Management*, 43(1), 54-69.

Fleisch, E.; Tellkamp, C. (2005): Inventory inaccuracy and supply chain performance: a simulation study of a retail supply chain. *International Journal of Production Economics*, 95(3), 373-385.

Goyal, S.; Hardgrave, B.C.; Aloysius, J.A.; DeHoratius, N. (2016): The effectiveness of RFID in backroom and sales floor inventory management. *International Journal of Logistics Management*, 27(3), 795-815.

Gumrukcu, S., Rossetti, M.D., Buyurgan, N. (2008): Quantifying the costs of cycle counting in a two-echelon supply chain with multiple items. *International Journal of Production Economics*, 116(2), 263-274.

Hardgrave, B.C.; Aloysius, J.; Goyal, S. (2009): Does RFID improve inventory accuracy? A preliminary analysis. *International Journal of RF Technologies: Research and Applications*, 1(1), 44-56.

Hardgrave, B.C.; Aloysius, J.; Goyal, S. (2013): RFID-enabled visibility and retail inventory record inaccuracy: experiments in the field. *Production and Operations Management*, 22(4), 843-856.

Kang, Y.; Gershwin, S.B. (2005): Information inaccuracy in inventory systems: stock loss and stockout. *IIE Transactions*, 37(9), 843-859.

Kök, A.G.; Shang, K.H. (2007): Inspection and replenishment policies for systems with inventory record inaccuracy. *Manufacturing & Service Operations Management*, 9(2), 185-205.

Kull, T.J.; Barratt, M.; Sodero, A.C.; Rabinovich, E. (2013): Investigating the effects of daily inventory record inaccuracy in multichannel retailing. *Journal of Business Logistics*, 34(3), 189-208.

Morey, R.C.; Dittman, D.A. (1986): Optimal timing of account audits in internal control. *Management Science*, 32(3), 272-282.

Raman, A.; DeHoratius, N.; Ton, Z. (2001): Execution: The Missing Link in Retail Operations. *California Management Review*, 43(3), 136-152.

Rekik, Y.; Sahin, E. (2012): Exploring inventory systems sensitive to shrinkage – analysis of a periodic review inventory under a service level constraint. *International Journal of Production Research*, 50(13), 3529-3546.

Rekik, Y.; Syntetos, A.; Jemai, Z. (2015): An e-retailing supply chain subject to inventory inaccuracies. *International Journal of Production Economics*,167,139-155.

Rekik, Y.; Syntetos, A.A.; Glock, C.H. (2019a): Inventory Inaccuracy in Retailing: Does it Matter? https://ecr-shrink-group.com/page/inventory-record-accuracy-research.

Rekik, Y.; Syntetos, A.A.; Glock, C.H. (2019b): Modeling (and learning from) inventory inaccuracies in e-retailing/B2B contexts. *Decision Sciences*, 50(6), 1184-1223.

Rinehart, R.F. (1960): Effects and causes of discrepancies in supply chains. *Operations Research*, 8(4), 543-564.

Su, H.C., Rungtusanatham, M.J.; Linderman, K. (2021): Retail inventory shrinkage, sensing weak security breach signals, and organizational structure. *Decision Sciences*, 1-21

Thiesse, F.; Buckel, T. (2015): A comparison of RFID-based shelf replenishment policies in retail stores under suboptimal read rates. *International Journal of Production Economics*, 159(1), 126-136.

Ton, Z.; Raman, A. (2010): The effect of product variety and inventory levels on retail store sales: a longitudinal study. *Production and Operations Management*, 19(5), 546-560.