**GUILHERME CHINEN**

# IMPACT ON SALES OF INVENTORY CONTROL ERRORS AND ACTION PLANS: A SIMULATION STUDY

**SÃO PAULO**

**2023**

**GUILHERME CHINEN**

# IMPACTO NAS VENDAS DE ERROS DE CONTROLE DE ESTOQUE E PLANOS DE AÇÃO: UM ESTUDO DE SIMULAÇÃO

Trabalho de formatura apresentado à Escola Politécnica da Universidade de São Paulo para a obtenção do diploma de Engenheiro de Produção.

Orientador: Prof. Dr. Marco Aurélio de Mesquita

**SÃO PAULO**

**2023**

## FICHA CATALOGRÁFICA

Catalogação-na-publicação

Chinen, Guilherme
IMPACT ON SALES OF INVENTORY CONTROL ERRORS AND ACTION PLANS: A SIMULATION STUDY / G. Chinen -- São Paulo, 2023.
122 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Inventory Record Inaccuracy 2.Retail 3.Discrete-Event Simulation 4.Factorial Design I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

I thank my parents for all the support

in my journey up to this point.

# ACKNOWLEDGEMENTS

I thank, first of all, my parents for always having supported me with hard work and teachings and for being the one behind all the success I had, am having, and will still have ahead of me.

I also thank my dear friends, who have accompanied me throughout all these years and have been a fundamental source of inspiration and support on this journey. May I be able to reciprocate by being a source of comfort and inspiration while I'm out there.

I thank the Escola Politécnica da Universidade de São Paulo and the Technische Universität Darmstadt for making me the professional I am today. I am committed to use this knowledge to make the world a better place.

Finally, I thank Professor Dr. Marco Aurélio de Mesquita, mainly for supporting me in this final stage of my degree, which was as challenging as all the other years of my academic life.

## ABSTRACT

This study focuses on the relevance of the topic of inventory control for retailers and aims to validate the empirical findings of previous studies on the negative impact on sales of the phenomenon of positive inventory record inaccuracies (IRI), a scenario in which the retailer has more actual inventory than shown in the inventory management system. The research also seeks to identify the causes of these positive IRIs that most significantly affect sales results and possible mitigation actions against them.

The research begins by identifying the commonly mentioned causes of IRI. Next, a discrete event simulation model is developed that considers a single SKU in isolation to analyze these factors and confirm the negative impact on sales caused by positive discrepancies. Additionally, the simulation model aims to determine the factors, or causes of IRI, that most negatively impact sales. To this end, a full factorial design is performed to identify the factors that significantly impact sales.

The main findings of this study confirm previous research findings that positive discrepancies in IRI have a negative impact on sales. Specifically, the study highlights that unaccounted devolutions of products, one of the causes of IRI, stand out over other factors such as excessive outbound counting and excessive inbound shipping in terms of impact on sales. However, the author recognizes several limitations that should be addressed in future research. These limitations include the adoption of a single SKU simulation model that does not take into account the high level of product variety in stores, the setting of some parameters of the simulation model based on empirical observations, and the assumption of linear relationships between the levels of the factors simulated in the $2^k$ factorial design.



**Keywords:** Inventory Record Inaccuracy, Retail, Discrete-Event Simulation, Factorial Design.

**RESUMO**

Este estudo se concentra na relevância do tema controle de estoques para os varejistas e tem como objetivo validar as descobertas empíricas de estudos anteriores sobre o impacto negativo nas vendas do fenômeno de imprecisões positivas de registro de estoque (IRI), um cenário em que o varejista tem mais estoque real do que o mostrado no sistema de gerenciamento de estoque. A pesquisa também busca identificar as causas desses IRIs positivos que afetam de forma mais significativa os resultados de vendas e as possíveis ações de mitigação contra eles.

A investigação começa identificando as causas comumente mencionadas do IRI. Em seguida, é desenvolvida um modelo de simulação de eventos discretos que considera isoladamente um único SKU para analisar esses fatores e confirmar o impacto negativo nas vendas causado por discrepâncias positivas. Adicionalmente, o modelo de simulação visa determinar os fatores, ou causas de IRI, que mais impactam negativamente as vendas. Para isso, é realizado um design fatorial completo para identificar os fatores que impactam significativamente as vendas.

As principais conclusões deste estudo confirmam as descobertas de pesquisas anteriores de que discrepâncias positivas no IRI têm um impacto negativo sobre as vendas. Especificamente, o estudo destaca que devoluções não registradas, uma das causas de IRI, se sobressaem em relação aos outros fatores, como a contagem excessiva de saída e o envio excessivo de entrada, em termos de impacto sobre as vendas. No entanto, o autor reconhece várias limitações que devem ser abordadas em pesquisas futuras. Essas limitações incluem a adoção de um modelo de simulação de um único SKU que não leva em conta o alto nível de variedade de produtos nas lojas, a definição de alguns parâmetros do modelo de simulação com base em observações empíricas e a premissa de relações lineares entre os níveis dos fatores simulados no design fatorial $2^k$.

**Palavras-chave**: Imprecisão do registro de estoque, varejo, simulação de eventos discretos, projeto fatorial.

## LIST OF FIGURES

## LIST OF TABLES

**TABLE OF CONTENTS**

# 1 INTRODUCTION

Today's world is filled with innovations and changes in many different branches. For instance, from late February 2020 through April 2021, at the height of the first wave of the COVID-19 pan-demic, global shopping volumes increased significantly and led the retail sector to experience a 35% increase in market capitalization. Perhaps, the best perceived boost within the retail industry during this period was the acceleration of e-commerce growth due to pandemic-related lockdowns, individual self-isolation initiatives, and imposition of quarantine measures for infected citizens, all of which kept people at home and, therefore, shopping online. To give some examples, many people used the time indoors to renovate and redecorate their homes, start new hobbies, try new clothing styles, change their beauty routines, care for their families and pets, and so on (Bradley et al., 2021).

However, new trends experienced recently by the retail industry bring along many extremely complicated challenges to be dealt with. For instance, a study carried out by Boston Consulting Group (2022) shows that 80% of customers are more likely to do business with companies that offer personalized experiences, 64% of them expect responses and interactions in real time from the company side, around 19,5% of total retail sales are already originated from e-commerce (in comparison to just 13,6% in 2019), there is a 37% projected compound annual growth rate (CAGR) of intelligent virtual-assistant market size between 2020 and 2027, and the metaverse is expected to have a market worth of 1,3 trillion US dollars by 2030 (i.e., an over 40% CAGR if calculated from 2022).

In order to better visualize the conclusions of Boston Consulting Group (2022) and Bradley et al. (2021), an analysis of retail giant Amazon.com's (or simply Amazon) net sales between FY 2018 and FY 2021 was carried out and compared to the number of COVID-19 new cases in Europe and in the United States of America (USA). The data for the former was obtained from Amazon's investors relation website (Amazon.com, 2019a, 2019b, 2019c, 2020, 2021a, 2021b, 2021c, 2022a, 2022b), whereas the data for the latter was obtained from the ECDC (European Centre for Disease Prevention and Control) and CDC (Centers for Disease Control and Prevention) organizations respectively (Centers for Disease Control and Prevention, 2022; European Centre for Disease Prevention and Control, 2022). Moreover, the result of this comparison can be seen in figure 1 and, with the assistance of table 1, which provides Amazon's net sales data utilized in figure 1, it is possible to see that the CAGR-value of the company's total net sales between 2018 and 2019 (i.e., before the pandemic arrived in Europe and in the USA) amounted 21%, whereas the same indicator comparing the pre-pandemic year of 2019 with the mid-pandemic year of 2020 showed a much greater value of 37%. This surge in sales amidst the peak of the first wave apparently confirms the findings presented in the previous two paragraphs.

Table 1: Amazon's net sales between FY 2018 and FY 2021 (Amazon.com, 2019a, 2019b, 2019c, 2020, 2021a, 2021b, 2021c, 2022a, 2022b).

| | Amazon's Annual Sales | | | |
|---|---|---|---|---|
| Year | Net Sales Amazon - Total | Net Sales Amazon - North America | Net Sales Amazon - International | Net Sales Amazon - AWS |
| 2018 | 232887 | 141366 | 65865 | 25656 |
| 2019 | 282522 | 170773 | 74723 | 35026 |
| 2020 | 386064 | 236282 | 104412 | 45370 |
| 2021 | 469822 | 279833 | 127787 | 62202 |
| CAGR (18-19) | 21% | 21% | 13% | 37% |
| CAGR (19-20) | 37% | 38% | 40% | 30% |
| CAGR (20-21) | 22% | 18% | 22% | 37% |
| CAGR (18-21) | 26% | 26% | 25% | 34% |

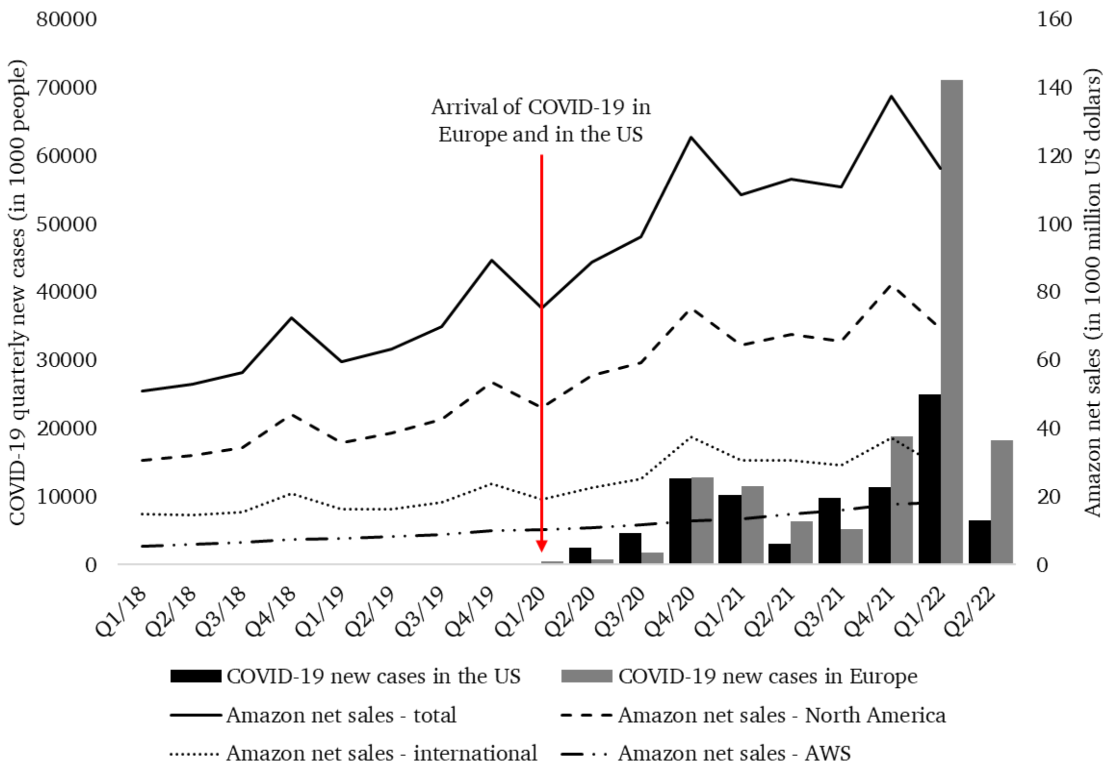


Figure 1: Comparison between Amazon’s net sales and Covid-19 new cases in the USA and Europe (Amazon.com, 2019a, 2019b, 2019c, 2020, 2021a, 2021b, 2021c, 2022a, 2022b; Centers for Disease Control and Prevention, 2022; European Centre for Disease Prevention and Control, 2022).

Although the already mentioned studies carried out by Bradley et al. (2021) and Boston Consulting Group (2022), as well as the Amazon brief case study, provide a good understanding of the recent challenges and developments of the retail industry, they are not enough to create a holistic view of its current stand. Therefore, in order to achieve this broader comprehension of how retailing is overall changing, Mostaghel et al. (2022) performed a systematic literature review around the topic and identified some activities that pointed towards a digitalization-driven business model innovation (BMI) experienced overall by the retail industry. These activities were divided into four aggregated dimensions, “digitalization enabled retail business model (RBM) innovation characteristics”, “digitalization enabled retail value creation”, “digitalization enabled retail value delivery”, and “digitalization enabled retail value capture”.

An example of RBM innovation enabled by digitalization is the creation of ecosystem partnerships and strategic alliances. This can be seen in the already-mentioned retailer Amazon, which runs a hybrid B2C business model complementing the inventories owned by the company with sales from independent suppliers on their platform. To sustain this hybrid approach, Amazon owns an international logistics network along with key strategic partners such as FedEx, UPS, and USPS, all of which play a decisive role in delivering the products to customers (Hänninen et al., 2018).

When it comes to digitalization driven value creation, one can mention the creation of digital services as a good example of it. This can be well portraited by Coop’s initiative to turn their customers’ experience more convenient. The European grocery retail chain has introduced a robot to their leading stores to help customers find receipts and the whereabouts of some products (Mostaghel et al., 2022).

With respect to digitalization-driven value creation, we have the enhancement of the omni-channel customer experience. Alexander & Blazquez Cano (2020) identify how a mixture of physical stores and virtual channels can enhance customer experience, proposing shifts to the functions and characteristics of the first.

At last, when it comes to digitalization enabled retail value capture, Mostaghel et al. (2022) argue that the digitalization has enabled parallel data computing, data transparency, permanent recording data authenticity, and data security, which led to effective and efficient decision making by enabling knowledge collection and reducing costs through optimized processes.

All these examples show clearly that data collection and processing is utterly important for retailers to run their business models. In addition to it, the high complexity of operating such involuted retail chains means that possessing reliable data is more critical than ever. Thus, in this context of data reliability, a pervasive problem faced by retailers is the inventory record inaccuracy (IRI), which refers to the discrepancy between physical and recorded inventory levels (Chuang & Oliva, 2015). The concept of IRI can be broken down into positive discrepancies, when the actual physical stock is greater than the recorded one, and negative discrepancies, when the physical stock is smaller than the recorded one (Rekik et al., 2019). Both of them are problematic because if the recorded inventory quantity does not match the quantity present on the store shelf, this system will either make a new order when such is rather unnecessary or fail to do so when an order should have been placed (DeHoratius & Raman, 2008). With these definitions, it is possible to finally comprehend the statement in the title of this work, which is to analyze what are the causes of positive discrepancies that most significantly have a negative impact on sales.

## 1.1 An overview of retailing

Before moving forward into the detailed analysis of the toll of positive discrepancies on sales for retailers, a brief contextualization of the retail sector is important. The term "retail" originates from the French word "retailer", which refers to the act of cutting-off a small piece of something bigger (Amit & Kameshvari, 2012). Nowadays, the concept of retailing could be understood as the sum of all activities for selling goods or services directly to the ultimate consumer for their personal non-business-related use (Kotler & Armstrong, 2012). Kotler & Armstrong (2012) also identify the 7 most important types of retail stores. Highlighting the differences between them is important because even though the conclusions of model presented in this work attempt to be as holistic as possible, it is unrealistic to believe that they accurately contemplate every retailer category.

- **Specialty Store**. A specialty store is a kind of store that provides to its customers a relatively narrow product line when compared with other kinds of retail, but with a deep assortment. They are typically portraited as apparel stores (i.e., stores specialized on selling clothing articles), sporting-goods stores, furniture stores, florists, and bookstores. An example of specialty store would be Sephora.

- **Department Store**. When compared to the specialty store the department store carries a lot more product lines, usually being clothing, home furnishings, and household goods. In this case, each line is typically operated as a separate department and therefore managed by specialist vendors or merchandisers. An example of Department Store would be Sears.

- **Supermarket**. Supermarkets are some of the most well-known retail types since they primarily sell essential goods that all human beings need to survive. They are designed in a way that makes them relatively large, low-cost, low-margin, high-volume, and self-service, serving the purpose of satisfying consumer's total needs for grocery and house-hold products. Supermarkets are, this way, the most frequently visited type of retail store and examples of it would be Rewe, Aldi, Netto, Penny, and TeGut.

- **Convenience Store**. Convenience stores are sometimes also referred to as convenience shops and they refer to relatively small stores that are usually located near residential are-as, with opening hours "24/7" (meaning they operate 24 hours per day along all 7 days of week) and carrying a limited line of high-turnover convenience products at slightly higher prices. An example of Convenience Store would be 7-Eleven.

- **Discount Store**. As the name suggests, discount stores are stores that offer to the general public standard merchandise sold at lower or discounted prices. The main strategy of business that opt to go this way is to operate under lower margins and, in exchange for that, enjoy higher sales volumes. An example of discount store would be Walmart.

- **Off-Price Retailer**. The off-price retailers are similar to discount stores in terms of pur-suing a low-price strategy to attract customers. In such stores, merchandise bought at less-than-regular wholesale prices are offered and sold at less-than-retail prices. These include facto-ry outlets owned and operated by manufacturers; independent off-price retailers owned and run by entrepreneurs or by divisions of larger retail corporations; and warehouse (or wholesale) clubs selling a limited selection of goods at deep discounts to consumers who pay membership fees. An example of Off-Price Retailer would be Costco.

- **Superstore**. The last category of retailer to be mentioned is the superstore. It usually refers to a very large store that has the objective of meeting consumers' total needs for routinely purchased food and non-food products. Under this category we can list supercenters, combined supermarket and discount stores, and category killers, which basically offer a relatively deep assortment of a particular product line. An example of superstore would be Walmart's Supercenter.

In addition to differentiating between the previously mentioned categories of retailers, Kotler & Armstrong (2012) also go further and present another four main characteristics which can be used to differentiate these retailers from one another. This way, each of the retailers can be individually analyzed under one or more of these four optics.

- **Amount of Service**. Basically, different types of customers and products require different amounts of service. The amount of service most adequate for a customer segment of product line varies according to other parameters such as the income of the customers and the level of sophistication of the products. With this in mind, in order to meet the many service needs, retailers may offer one of three service levels. The first one is the self-service, in which customers conduct their own locate-compare-select process to save time and money. The second service level is the limited service, which is adopted by retailers that carry a relatively large amount of shopping goods, about which customers often need more information and therefore usually enjoy the availability of a staff member for FAQs. At last, there are the full-service retailers, that operate high-end specialty stores and first-class department stores, providing customers with highly customized services.

- **Product Line**. When analyzing the retail sector, one can also classify it according to the length and width of the product assortment. To make it clear with an example, a specialty store can carry a narrow number of product lines, but with deep assortment within those lines (e.g., selling only a handful of beer brands but having a lot of them in stock to meet the customer demand). Here it is also worth mentioning that the product line of retailers does not restrict itself to consumer goods, being also possible to offer services that are equally classified according to their service lines. As examples of service retailers, there are hotels, motels, banks, airlines, restaurants, and movie theaters.

- **Relative Prices**. The relative prices could also be referred to as price categories and most re-tailers charge regular prices for their normal-quality goods and customer service. However, depending on their strategy, other retailers could and do also provide higher-quality goods and services for a premium on the price charged to the customer. On the other side of the spectrum, there are those retailers that opt to lower prices as much as they can in order to have a gain on the volume sold (e.g., typically discount stores and "off-price" sellers).

- **Organizational Approach**. When dealing with legal entities, the classification according to the organizational approach is good specially to have a comparison of, for example, the size of two different businesses. This way, just like in other sectors of the economy, retail-ers can also join together in forms of corporate or contractual organizations. Under this category, there are four major types of retail organizations that can be mentioned: corpo-rate chains, voluntary chains, retailer cooperatives and franchise organizations.

With respect to the size of the retail market, Sabanoglu (2022) estimates that in 2021 it generated over 26 trillion U.S. dollars globally, with a forecast to reach over 30 trillion U.S. dollars by 2024. As already mentioned in the introduction, this huge market branch is going through significant changes (Mostaghel et al., 2022). Consequently, when it comes to economic agents that operate in such branch and run inventory-carrying facilities, the provision of high product availability to customers at minimal operation costs is one of the key factors that determines the success of businesses Kang & Gershwin (2005). In this scenario in which the competition is fierce and profit margins are thin, many companies have decided to automate the inventory management processes to better meet customer demand and reduce operational costs (Kang & Gershwin, 2005). For example, many retailers use an automatic replenishment system which tracks the number of products in the store and place an order to the supplier in a timely fashion with minimal human intervention (Kang & Gershwin, 2005). With this aim of identifying inefficiencies, Hausruckinger et al. (2003) studied how product availability deteriorates along the supply chain from the manufacturer to the retailer. They came to the realization that while there are high service levels from the manufacturer's warehouse to the retailer's warehouse, and similarly high service levels from there to the retailer's stockroom (98–99%), this performance drops sharply over the final meters from the stockroom to the shelf (90–93%). In other words, on average, almost 1 out of 10 final customers does not get all the products they were looking for. Figure 2 provides a visual indication of how the service level decay along the supply chain.

In addition to this decay in service level, they also concluded that retailer in-store operations cause over 85% of all out-of-stocks (OOSt). When analyzing the main causes of these shortages, store ordering, shelf replenishment and inventory inaccuracy were pointed out as the key issues to be tackled (Hausruckinger et al., 2003). Given the proven relevance of inventory record inaccuracies (IRIs), the following section will provide a brief overview regarding what they are, and their most common causes, as well as presenting the basis for the assumptions used throughout the development of the current work.

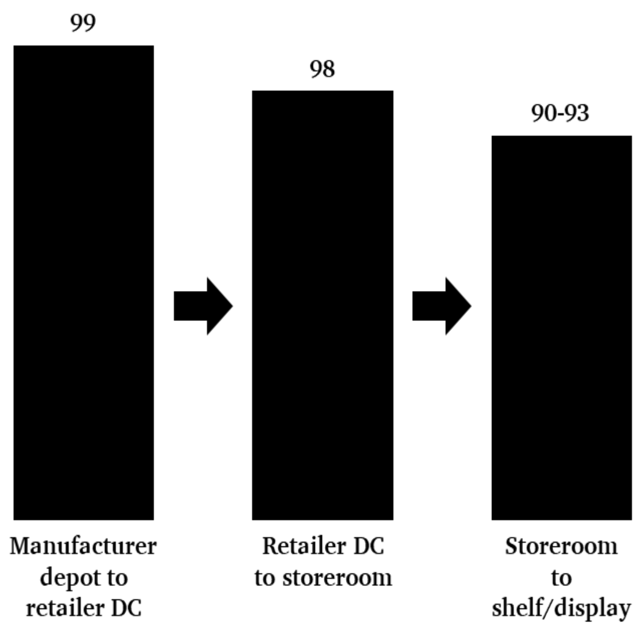


Figure 2: Service level along the supply chain in percentual points (Hausruckinger et al., 2003).

## 1.2 Overview of the Inventory Record Inaccuracy Problem

According to DeHoratius & Raman (2008), the first researcher to identify IRI as a potential obstacle was R. F. Rinehart in his 1960 paper entitled “Effects and Causes of Discrepancies in Supply Operations”. In the mentioned paper, Rinehart (1960) works on a project aimed at the quantitative identification of discrepancies causes between actual and recorded inventories at an American Federal Government agency. Back then, he already stressed that it was common sense among operators of supply systems that an out-of-stock position of an item, not reflected on the records, would lead to significant delays in the orders fulfillment of that item. On the other hand, reaching the point of order in the recorded inventory when the actual inventory does not require a refill would result in unrequired replenishment of the item in question and therefore unnecessary procurement and inventory-carrying costs. When framing the problem, Rinehart (1960) found that during the yearly inventory counts carried out in the FYs of 1957, 1958, and 1959, discrepancy rates from 20% to 50% could be seen among secondary items (e.g. repair and replacement parts for major items, tools, and general supplies).

The already more than 60 years of studies of this topic are not unjustified. The literature has many examples of how significant this problem is. When analyzing a large subset of stores from a global retailer, Kang & Gershwin (2005) realized that the best performing store is the one in which only 70%-75% of its inventory records matched the actual inventory. In a certain store, about 67% of its inventory records were inaccurate and, on average, the inventory accuracy of this global retailer’s stores was just 51%.

Similar numbers were obtained by Rekik et al. (2019). After analyzing approximately 1 million stock keeping units (SKUs) from about 100 stores of 7 European retailers operating in across 4 European countries, they found that about 60% of the SKUs analyzed contained IRIs. The average magnitude of IRIs for the affected SKUs was approximately +6.6 and -6.0 units for positive and negative discrepancies, respectively. The authors also came to the conclusion that correcting such inaccuracies could lead to approximately 4% to 8% of increased sales in the participating retailers. In a market that operate with tight profit margins, as already mentioned before, such increase on sales ought not to be neglected. Another curious observation made by Rekik et al. (2019) suggests that negative inaccuracies tend to accumulate over time, whereas positive discrepancies apparently remain on relatively stable levels. One possible explanation for this result could be that positive discrepancies are caused, for example, by misplaced items that are removed from inventory records during gap scans, but later on found again and placed back on the shelf for customers. These temporary losses in inventory due to product misplacement could somehow balance each other over time (such would be the case, for example, when a particular item is lost and another one is found again within a short period of time between these two events). On the other hand, as later presented on section 2.1, negative discrepancies are oftentimes caused by inventory shrinkage, which in turn translates into a permanent loss of items, leading to an accumulation of discrepancies over time. Finally, Rekik et al. (2019) have also discovered that positive discrepancies are just as common as negative ones.

Interestingly, retailers in general already know quite well the importance of IRI problems and do see it as one of the major obstacles when it comes to their operations. However, often they do not know when and where the causes of IRI occur and what is their magnitude (Kang & Gershwin, 2005). This way, for retailers to improve both the quality of their operations and the data integrity of their inventory management systems (IMSs), it is important for managers to comprehensively understand the causes of IRI and identify the policy levers that they can make use of in order to reduce it (Chuang & Oliva, 2015).

## 1.3 Objectives of the thesis

With the unquestionable relevance of IRI for retailers in mind, the current work was developed with two main objectives. The first one was to validate the findings of the study of Rekik et al. (2019) regarding the fact that positive inventory record inaccuracies can lead to negative impacts on sales in a typical retail store, as per represented in figure 3. If the first objective was validated, the second objective aimed at identifying the causes of these positive IRI that most significantly affect sales in a negative way.

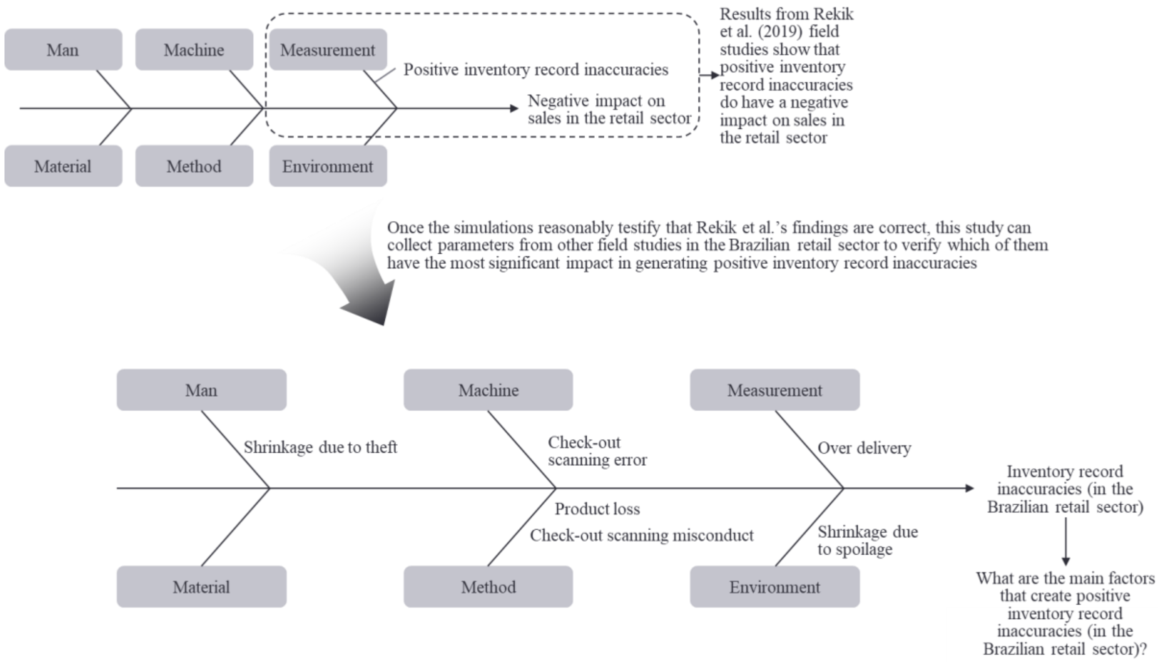


Figure 3: Schematic representation of the research framing (own authorship).

The idea of answering these two research questions comes primarily from the fact that, at first, it might seem counter intuitive to think that having more actual stock than recorded stock can adversely affect sales. However, since as demonstrated by Rekik et al. (2019) this scenario is not only possible, but it also happens quite often, it seemed reasonable to analyze it in further detail.

In this investigation, the first step was to identify what the commonly mentioned causes of IRI in the literature. Once they were identified, a separation into those that generate positive and those that generate negative discrepancies was made. The goal with such a separation was to allow for a specific analysis of the factors associated with positive discrepancies.

Next, a discrete-event simulation was developed to frame the problem in a way that these identified factors were translated into parameters that could be freely varied. The simulation had the aim of confirming if indeed the negative impact on sales claimed by Rekik et al. (2019) coming from positive discrepancies could be verified and, after that, to check which factors were the most impactful ones.

A full-factorial design was conducted subsequently with the hope to see if there were indeed factors that were significantly more impactful than others. The detailed step-by-step of the simulation and the full-factorial design are presented in section 3 of the current work.

## 1.4 Structure of the thesis

In this first section, the purpose of the current work was introduced together with a background on the topic encompassing the retail sector and why it is relevant and a brief overview of IRI, which will be discussed in more detail in the next section. Moving on to section 2, the results of a literature review are presented, where the main causes and consequences of IRI (both positive and negative) are investigated and possible approaches to answering the research question of this thesis are analyzed. Next, the methodology is presented in section 3 in which a combination of discrete-event and agent-based simulation in the software Plant Simulation followed by a full-factorial experiment were chosen tools to approach the problem. Moving on to section 4, the development of the simulation and its results are presented. Then, at last, section 5 presents the conclusion of the thesis together with some discussions regarding the further development of the topic.

# 2 LITERATURE REVIEW

The literature review had as its goal the consolidation of comprehensive reference material for developing the current research. Through the literature review, the decision of choosing a simulation approach to answer the research question was validated and the most important parameters for carrying on the simulations were identified, minimizing the need to resort to common knowledge and giving more consistency to the conclusions presented later on. The method utilized here was primarily a backward snowballing taking as reference the following six kick-off papers: Condea et al. (2012), Chuang & Oliva (2015), DeHoratius & Raman (2008), Kang & Gershwin (2005), Papakiriakopoulos (2012), and Rekik et al. (2019).

In addition to the backward snowballing, the database Scopus was used to further verify if other papers could purposefully contribute to this research. Scopus, along with Web of Science, EBSCO, and Science Direct, is taken as a reliable database for indexing and therefore identifying journals and authors with high-ranking scores (Mostaghel et al., 2022). For this reason, it was chosen as the reference platform for the searching the additional papers. Other minor contributions from reports, books and websites were obtained by simply searching for these individual materials in the Google website.

When looking for the additional material in the Scopus database, the query "“retail*” AND “simulation*” AND “inventory record inaccuracy”" was used as reference for the purpose of finding papers that could support the development of the simulation model. The Boolean operator “AND” acted combining the terms in a way that only papers that contain all the terms were filtered. The wild card “*” allowed variations of the terms “retail” and “simulation” (e.g., “retailer”, “retailing”, or simply “retail”) to be within the results of the search. Another query was also created to find papers related to IRI that were not necessarily related to simulation. This second query was simply "“retail*” AND “inventory record inaccuracy”". Additional material regarding simulation, retail and design of experiments (DOE) was searched, as mentioned before, in the Google website.

The following subsections are divided into the results that provided a better understanding of the inventory record inaccuracies in retailing in general and those related to developing simulation models and carrying out factorial designs to approach IRI in retailers.

## 2.1 Inventory Record Inaccuracy in Retailing

As already presented in section 1.2, IRI is a widely researched topic within the retail industry context and its relevance is by no means unjustified. Within the current section, the focus will be on expanding the basic knowledges of IRI, identifying its main causes and consequences for retailers, analyzing the most commonly adopted countermeasures against it, and finding literature parameters to support the methodology employed in this work. Each of these targets is approached in a dedicated subsection.

### 2.1.1 Findings on typical causes and consequences of IRI

Within the studies of inventory management, there are several already known causes for record inaccuracies of stock levels. Kang & Gershwin (2005) present some of the commonly observed ones, which are considered here as being stock loss (also known as shrinkage and associated to events such as theft of SKUs), transaction error (e.g., when customers bring similar products with equal prices to the POS and the cashier registers them as being two of the same product line, when they are not), inaccessible inventory (e.g., products that are in the store but cannot be found by customers or even staff), and incorrect product identification (e.g., when products are labeled wrongly by mistake).

Shrinkage, sometimes also referred to as stock loss, in this context encompasses every kind of loss of the products available for sale. One common example is theft, which according to Kang & Gershwin (2005), can be perpetrated either by shoppers, suppliers (both defined as external theft), or employees (defined as internal theft). To give a better picture of the relevance of theft for inventory shrinkage, Bamfield (2004) conducted a study with 476 major European retailers spread across 16 countries and found that the average rate of shrinkage was 1,46% of retail turnover, which is a measure for the retail total sales plus value-added taxes. An astonishing 82,4% of this loss was regarded as being crime-related (or simply theft). This percentage represented nothing less than a total cost of 24.763 million Euros annually for a single retailer. Internal theft, that happens, for example, when members of the staff steal products while in the store (either in backroom or on the sales floor) performing replenishment duties, was estimated to be on average 28,8% of total shrinkage among European retailers. In contrast to this number, external theft, at-tributed mostly to criminal customers and associated with cases in which, for example, products are hid amidst the clothing of the thief when leaving the store's facility, accounted for around 48,4% of shrinkage in the continent.

It is relevant to mention that Bamfield (2004)'s study deals only with thefts carried out by customers, staff, and suppliers (additionally to also studying process and administrative losses, which just as well constitute shrinkage cases). This means that other types of crime such as robbery, violence, arson, terrorism, virus attacks, and major frauds were not considered. Such framing of the research problem was considered to be acceptable, since approximately 93% of all crime losses suffered by retailers were contained within the first three categories of perpetrators.

When analyzing Bamfield (2004)'s findings in more detail, a discretization per country of the previously mentioned indicator (shrinkage rate as a percentage of annual turnover) was made possible. According to the research, this number had its highest value in the UK (1,77%) and its lowest value in Switzerland (0,85%). Germany found itself below the average, with a 1,19% shrinkage rate. These values are important because they were later adopted as parameters for the simulation (more details on how they were implemented are provided in section 3, where the methodology of the current work is discussed).

Another worth mentioning breakdown of Bamfield's (2004) conclusions differentiate the shrinkage numbers in Europe and in the USA into customer, employees, supplier theft and internal error contributions. Table 2 summarizes these indicators that also supported later on the development of simulation model used in the current work.

Table 2: Shrinkage breakdown according to customer theft, employee theft, supplier theft, and internal errors (Bamfield, 2004).

| | **Europe** | | **USA** | |
|---|---|---|---|---|
| | Percentage shrinkage | Estimated cost (in millions of Euros) | Percentage shrinkage | Estimated cost (in millions of Euros) |
| Customer theft | 48,4 | 14.670 | 30,8 | 10.230 |
| Employee theft | 28,8 | 7.820 | 45,9 | 15.243 |
| Supplier theft | 7,5 | 2.273 | 5,9 | 1.959 |
| Internal errors | 18,3 | 5.547 | 17,4 | 5.789 |
| Total | 100,0 | 30.310 | 100,0 | 33.221 |

Other two studies that focused on inventory shrinkage within the retail sector were carried out by Moraca et al. (2015) and Hollinger & Davis (2002). Both of them contributed to the American National Retail Security Survey (NRSS) report, where the former corresponds to the

2015 edition and the latter to the 2002 edition of these publications. According to Moraca et al. (2015), the NRSS is recognized as a key benchmark for loss prevention in the retail industry and provides a holistic view of this topic. The report aims at studying workplace-related criminality, as well as identifying successful security countermeasures (which are presented in section 2.1.2) to protect people, assets, and brands among retailers.

Moraca et al. (2015) found that inventory shrinkage in the US averaged 1,38% of revenues in FY 2014, where approximately 30% of surveyed retailers reported slightly higher levels than this, and 94,4% had losses below 3% of revenue. In absolute terms, this translated into an average loss of around US$ 44,02 billion in that fiscal year. In comparison Hollinger & Davis (2002) found this number to be 1,70% of revenue on average for FY 2001, accounting in absolute terms for US$ 31,3 billion in annual losses. When comparing these values to Bamfield's (2004) findings, there is a consistency between the numbers for shrinkage as a percentage of revenue/turnover presented by these three studies, which reinforces their individual credibility as a reliable source.

When seeking to analyze the main culprits of inventory shrinkage, Hollinger & Davis (2002) identified employee theft, shoplifting, administrative and paperwork error, vendor fraud, and check, cash, and credit card losses as being responsible for most of the shrinkage cases. These results are very close to those reported by Kang & Gershwin (2005). The following paragraphs will provide a more detailed explanation of what constitute “administrative and paperwork error”, “vendor fraud”, and “check, cash, and credit card losses”. “Employee theft” and “shoplifting” were already discussed previously in this section and, for this reason, were excluded from further explanation.

Administrative and paperwork errors occur, for instance, when pricing activity for a certain SKU is carried out wrongly. This situation can lead to a markup (product is attributed a price higher than it should) or a markdown (product is rather attributed a price lower than it should) errors. Such kind of errors accounted for around 15% of total shrinkage in Hollinger & Davis (2002)’s study and were reported to be on average 16,5% by Moraca et al. (2015).

Vendors fraud was understood by both Moraca et al. (2015) and Hollinger & Davis (2002) as happening when vendors steal merchandise while stocking it. It’s important to em-phasize here that such a distinguishment between employee theft and vendors fraud is not made in any other study analyzed in the current work.

Finally, check, cash, and credit card losses account for another kind of shrinkage not related to inventory shortage. To properly understand it, it is important to know that Hollinger & Davis (2002) refer to inventory shrinkage as a financial loss incurred by the store rather than just a subtraction of physical items. In this definition, for in the FY 2001, bad check losses totaled 0,55% of annual sales. Following it came cash shortage losses, amounting for 0,37% of annual sales. Lastly, credit card charge-back losses accounted for 0,27% of annual sales in that year (Hollinger & Davis, 2002).

Figure 3 provides a graphical view of the contribution of each shrinkage factor mentioned in the study of Moraca et al. (2015), which corresponds to the most recent data from the NRSS reports of the two mentioned publications.

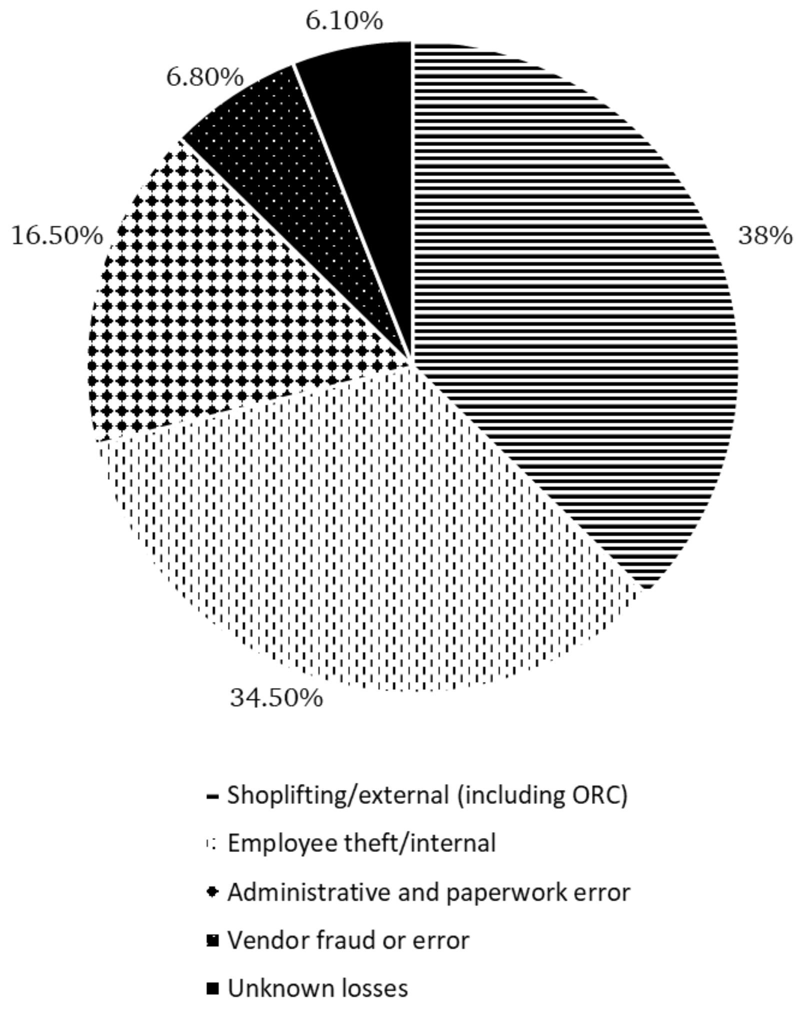


Figure 4: Sources of inventory shrinkage (Moraca et al., 2015).

In connection to the research target of this work and coherently with the results of Moraca et al. (2015), Hollinger & Davis (2002), and Bamfield (2004), Chuang & Oliva (2015) found that these backroom and shelf shrinkage losses are actually the dominant drivers of IRI in comparison to others. This way, additionally to these dominant drivers, they also concluded that “under-shelving”, “erroneous checkout” and “erroneous data capture” likewise created IRI, although they accounted for a relatively negligible impact and were clearly overshadowed by shrinkage losses.

Other minor causes of inventory shrinkage mentioned by Kang & Gershwin (2005) can also be attributed to collusions between customers and staff as well as the unauthorized in-store consumption of no “free sample” products by either shoppers or even employees (such as eating a package of chips in the facility aisle or trying a lipstick without buying it later). Finally, shrinkage can also be present when products are deemed unsuitable for sale by becoming out of date (i.e., expired), damaged, or even spoiled (e.g., due to inadequate keeping conditions).

Ending the considerations regarding shrinkage as a generator of IRI and going into further details on behalf of the transaction errors, Kang & Gershwin (2005) state that these kinds of errors are typically present at the inbound (where products arrive from the upstream sup-ply chain entity to resupply the backroom inventory) and outbound (checkout area where customers leave) sides of the store.

Typical inbound problems arise from the fact that the shipments that arrive from the supplier are ought to be registered into the IMS of the retailer and in some cases, there might be a discrepancy (positive or negative) between the shipment recorded and the actual one. Since it is impossible for the retailer staff to carry out a 100% inspection of the arriving lots of new products, such errors not unusually pass unnoticed. DeHoratius & Raman (2008) confirm these kinds of error by stating that, typically, store employees do not scan every single item delivered from the distribution center (DC). What actually happens is that the staff just verifies if the ordered amount of pallets and cases has been delivered by the supplier. DeHoratius & Raman (2008) also say that the store inventory record is automatically updated upon the placement of the replenishment order and, by doing so, the store management oftentimes relies on the assumption that the order has been filled correctly. Hence, if the situation arises in which a DC employee incorrectly fills the store order by accidentally picking and shipping the wrong item or the wrong quantity of the correct item, the amount received by the retail store will not match the quantity ordered. As a direct consequence of it, a IRI situation (either positive or negative) will take place.

When analyzing another point of view of the distribution structure of a product, DeHoratius & Raman (2008) found out that retailers usually adopt one of two different shipping methods when distributing merchandise lots to their stores. In the first scenario, retailers operate their own DC and, therefore, are solely in charge of replenishing their stores. Alternatively, retailers could also opt to have external vendors shipping the merchandise directly to the stores. The researchers concluded that variations that arose from these different ways in which the inventory could be managed during the store resupplying process usually created, due to contrasting governance styles, variations in IRI levels.

One hypothesis formulated by DeHoratius & Raman (2008) suggested that these variations from one distribution structure to the other could be explained by means of transaction cost economics (in short, this theory states that organizations generally aim to minimize the overall costs associated with transactions, such as searching costs, bargaining costs, and enforcement costs). According to the researchers, the transaction cost economics approach point to the conclusion that, in the absenteeism of efficiency distortions related to the internal organization of the retailer, a vertical consolidation of the supply chain leads to efficient and effective coordination of the resupplying process. Simply stated, it is more advantageous to internalize the resupplying process than to outsource it to a third party. The explanation to this conclusion lies on the fact that through a variety of mechanisms, including the increased likelihood of physical proximity between DC and store, well-established patterns of communication between supply chain entities, and greater willingness to cooperate with other members of the same group, a retailer-owned DC would outperform a vendor-owned one. Therefore, in this scenario, one could expect to encounter fewer inventory discrepancies among items shipped by retailer owned DCs.

On the other hand, DeHoratius & Raman (2008) present an alternative hypothesis to the one exposed in the last paragraph. In this second hypothesis, the presence of internal procurement bias, internal expansion bias, persistence on sub-optimal routine practices, and in-ternal communication distortions could potentially cause one to expect more errors among items shipped from a retailer-owned DC than from a vendor-owned one. Another supporting argument in favor of this hypothesis relies on the fact that vendors are in general highly motivated to spend more than the minimal amount of effort to ensure quality performance (e.g., progressively more accurate shipments), given that the store can usually resort to switching suppliers if it wishes to. Since stores do not enjoy such power of rewarding (or penalizing) retailer-owned DCs for accurate (or inaccurate) fulfillment of resupplying orders, it would be reasonable to assume that

IRI-related discrepancies should be greater in this latter case. The misalignment of incentives due to the inability of the store to resort to alternative sourcing options qualify a classical monopoly problem.

When comparing these two conflicting hypotheses in their paper, DeHoratius & Raman (2008) found that vendor-owned DCs outperform retailer-owned DCs where IRI is concerned. For a typical retail chain, the researchers found that the IRI of a retailer-owned DC was on average 25% higher than a comparable vendor-owned DC. In addition to this, the researchers also discovered that the impact of different distribution structures on IRI is not the same across product lines. In-stead, they claimed that the more expensive a certain product category is, the less difference in IRI magnitude exists between the two distribution structures. By running some simulations, DeHoratius & Raman (2008) saw that, for expensive items, a retailer-shipping structure had a 19% greater IRI than a vendor-shipping one. This gap increased to 47% when making the very same comparison for inexpensive items.

Zooming in into the analysis of the outbound side of the retail store, it is important to understand that a common practice among retailers is to reward the cashier staff according to the speed of checkouts. Simply put, the faster the cashiers process the customer products, the better the re-wards. On the one hand, such rewarding policy may increase customer satisfaction by reducing the time spent in the checkout queue; however, on the other hand, it also increases the chances of checkout errors, that in turn leads to IRI. A representative example of a failure mode that can take place was, to other purposes, already mentioned before. In such cases, the shopper brings two similar products with identical prices to the POS and the retail staff opts to scan only one of these products, processing both of them as being two of the same SKU. This obviously leads to IRI for both products, where one of them will face a positive discrepancy and the other will face a negative discrepancy (Kang & Gershwin, 2005). DeHoratius & Raman (2008) go further in this analysis and quantify their findings for an average retailer. They came to the conclusion that record inaccuracy outstanding from transaction errors for fast-moving consumer goods (FMCG), sometimes also regarded as items with high transaction frequency or even high dollar-volume items, are around 2,1 times greater than for their slow-moving counterparts.

Proceeding forward to framing the problem of inaccessible inventory, Kang & Gershwin (2005) depict it as being the situation in which the product is somewhere in the retailer, although it simply cannot be found (sometimes the product can be considered lost, since it is found only after no longer being suitable for sale) or cannot be reached (when store staff places the product in a place of difficult access). This factor could potentially lead to other problems such as OOSh and shrinkage due to product expiration or spoilage, as already mentioned in the current work. A classical situation that leads to such problem happens when the customer picks up a prod-uct from the shelf and, for some reason, leaves it in another point of the store without buying it. Another mirrored situation can take place in the backroom when the staff picks a product and leaves it somewhere else where it should not be (Kang & Gershwin, 2005). A feature for simulation purposes of inaccessible products is that they are no different from being nonexistent as far as revenue is concerned. In other words, when modeling inaccessible products, one can simply consider them as lost and, therefore, categorize it as shrinkage.

The next mentioned issue by Kang & Gershwin (2005) is the incorrect product identification and is a problem that can happen both in the retailer or upstream in the supply chain. In this case, wrong labels wrong labels are assigned to products and when these labels have their barcode scanned, the IMS deducts the wrong item from its records. This kind of error also happens during manual inventory counts.

A last finding of the paper of Kang & Gershwin (2005) regarding the main causes of IRI highlights a rather interesting phenomenon. After conducting the simulation and analytical work, they observed that, if no corrective/mitigative measures are taken, even the smallest rates of stock loss (i.e., stock that goes missing) is capable of disrupting the entire replenishment process and consequently generate impactful OOSt. This is showed in their simulation, where inventory record levels move constantly upwards while actual inventory levels get every time lower (this fact was also verified in the simulation developed here when only negative discrepancies were taken into account). The increasing gap between inventory record and actual inventory can in turn result in loss of sales and all other sorts of damages related to OOSt. One final remark to be made on this behalf is that the harmful impact of such kind of stock loss is expected to have a heavier toll in lean environments, usually characterized by short lead times, small order quantities, and small levels of safety stocks.

Before moving on to presenting the main consequences of IRI, it is important to discuss some of the last findings presented by DeHoratius & Raman (2008). First, the researchers found that inventory record inaccuracy is negatively associated with the unit cost of an item. They also found that IRI is negatively associated with the dollar-volume of an item. Another finding suggests that high (low) levels of inventory density are associated with high (low) levels of IRI. And finally, high (low) levels of product variety are associated with high (low) levels of IRI. Each of these conclusions will be discussed in detail in the next two paragraphs.

The intuitive approach to verify that the IRI is negatively associated with the cost and dollar-volume of an item relies on observing the fact that retailers usually pay more attention to high-value items. To give an example, items with high value are commonly monitored more closely throughout the supply chain in comparison to those with lower value. The retailers pay such a special attention that they often allocate theft-prevention efforts to protect the former class of merchandise. Not rarely they may be shipped in special cartons, placed into locked cages in the backroom prior to replenishing the sales floor, or audited at multiple stages of the distribution process. Since the dollar-value of a product usually increases with its cost, both behave in similar ways regarding IRI. DeHoratius & Raman (2008) further showed that, in a typical retailer, the IRI of expensive items (defined here as the 90th percentile of the retailer's portfolio ordered according to item cost) was about 57% less than that of inexpensive items (defined as being the 10th percentile). The researchers also showed that, when it comes to a quantitative analysis of the dollar-value parameter, the annual selling quantity of the product is a decisive variable. On average, the record inaccuracy of fast-moving inexpensive items is 150% greater than of fast-moving expensive items. Similarly, the record inaccuracy of slow-moving inexpensive items is around 56% greater than that of slow-moving expensive ones.

Moving on, the relationship between inventory density and IRI, just like the relationship be-tween degree of product variety and IRI, is contained in the field of environmental complexity. To better understand it, a first introduction to information processing theory is needed. The topic of information processing has been widely studied at the cognitive level, within organizations, and across organizations. This way, the information processing theory states that whenever people carry out a task, they are making trade-offs between speed and accuracy. Properties of the task itself, the individual executing it, and the environment surrounding it influence the probability of error occurrence for a given level of employed effort (e.g., the more complex a task is, the more prone to errors it is). When bringing these properties to a retailer's reality, the environmental complexity could be framed in two different dimensions: inventory density and product variety. There-fore, stores with greater inventory density (defined as the total number of selling items divided by the facility area) and with greater product variety (defined as the number of different merchandise categories sold by a store) have a more complex task environment than other stores. On behalf of this issue, DeHoratius & Raman (2008) posit that items kept in stores with high levels of inventory density (defined as the 90th percentile of studied retailers classified according to their inventory density) exhibit an inaccuracy level 25% greater than items kept in stores with low levels of inventory density (defined as being the 10th percentile). In a similar manner, stores with high levels of product variety showed a record inaccuracy around 13% greater than items in stores with low levels of product variety.

Out of all of these mentioned causes of inventory record inaccuracies, many impactful consequences arise for retailers and other economic agents upstream in the supply chain. Perhaps the most notorious consequence of IRI is the situation in which the customer does not find the desired product in the retailer, also known as out-of-stock, that in turn leads to a whole set of other impacts Ernst et al. (1993). In the following paragraph, the relationship between inventory record inaccuracy and out-of-stock encountered by Ernst et al. (1993) gets better clarified.

When developing a quality control approach for monitoring inventory stock levels and improving the performance of inventory management systems, Ernst et al. (1993) realized that many retailers are faced with the problem of having to keeping track of thousands of SKUs. This way, in order to obtain competitive and economic advantages from an efficient control of these inventories, several inventory control models were developed throughout the years. Here it is important to emphasize that the inventory management is based on recorded, and not actual, stock levels to take actions such as ordering new lots and determining the parameters that aim at optimizing the inventory control. These parameters usually have an influence on both operational and financial decisions in a way that, for example, the replenishment of a shelf can be triggered too late or not at all, causing customers to be dissatisfied by not finding the products they were originally looking for. This OOSt situation in turn leads to loss of goodwill for the retailer, loss of production time for the manufacturer, and loss of sales for all parts involved. It is also worth mentioning the mirrored situation is equally detrimental for businesses, in which over-stock conditions may increase the capitalization and carrying costs for the retailer.

Gruen et al. (2002) also attempted to examine the extent, causes, and consumer responses towards OOSt. They found that 70% to 75% of out-of-stock situations are a direct result of retail store practices (i.e., either by underestimating customer demand for a certain item or by carrying out too lengthy ordering processes/cycles) and shelf-restocking practices (i.e., when the item is actually within the limits of the store, for example, in the backroom, but is unavailable on the shelf). When further breaking down these numbers, the researchers found that store ordering practices accounted for 34% of OOSt situations, store shelving itself accounted for 25% of OOSh, and store forecasting accounted for 13% of OOSt (Gruen et al., 2002).

The researchers further concluded that, typically, causes of OOSt tend to be assigned to one of the following three general processes: ordering, replenishing, and planning. The ordering practices cover two general categories. First, the retailer may have placed a replenishment order of an insufficient amount or even a reasonable amount, but too late, so that the warehouse could not deliver the order before the retailer ran out of the SKU. Second, the retailer forecast happens to have mis-judged demand for a certain SKU and ended up ordering an insufficient amount of resupply. For the case in which an item is promoted, usually the ordered amount of supply intended to meet the demand is inadequate.

When framing the case of replenishment practices, one comes across the situation in which the product is in the store, usually in the backroom, although sometimes also in another area of the store, but it is not on the shelf when the consumer arrives to purchase the product. One event that might lead to such situation is the inadequate shelf space allocation for a certain SKU so that the item runs out of the shelf before the regular restocking occurs. Other causes like the lack of an adequate signal to the store management that the product is no longer on the shelf, or poor backroom inventory handling procedures that create obstacles for store personnel to move products from the backroom onto the shelf, are also typically found among retailers. These sorts of problems also affect upstream players of the supply chain. On a warehouse level, for example, the warehouse may be faced with insufficient inventory levels to meet final demand and this “scratches” the retailer’s order.

At last, the planning practices. This category of processes that lead to OOSt encompasses several possible causes. One typical example can be seen when the item was discontinued but such information was not communicated to the retail management, which can result in the retailer even still promoting the item and creating expectations among the customers to find this item on the shelf. In other cases, the manufacturer simply may not have shipped the adequate amount to properly refill the store inventory, or there may be a so-called product “drought”, in which the manufacturer finds itself unable to produce enough quantities to meet customer demand (Gruen et al., 2002).

Other causes that also contribute to generating OOSt scenarios and are not contemplated within the three root causes mentioned before are: inadequate shelf capacity, inverse effect of inventory, ad and price changes, new product phase in and out, and manufacturer minimum order sizes. Each of these secondary causes will be briefly explained in the next paragraph.

Sometimes an adequate shelf space is not designate to a SKU that requires a higher shelf capacity to properly supply its demand. This problem emphasizes the importance of identifying slow and fast-moving items, since oftentimes slow-moving products occupy an inordinate amount of shelf space, crowding the space available for fast-moving items. A typical solution for this problem would be the shelf reallocation of products; however, retailers usually face the constraint of having to consider case-pack sizes, which sets a limit to how much space fast-moving products can gain over slow-moving counterparts. In dealing with the inverse effect of inventory, the results might be to some extent counter intuitive. It was shown that the greater the inventory warehoused in the backroom of a certain SKU, the higher the OOSh (i.e., the product is somehow in the store but unavailable on the shelf for the customers) for that SKU. An explanation to this observation could be that the excess of carried inventories create congestion in the supply chain and, this way, reduce the degree of synchronization between the many replenishment processes. Moving on to ad and price changes, it was identified that advertisement campaigns and price changes, when coupled with inadequate communications with warehouse and store logistics and purchasing managers, lead to major OOSt rates among advertised items. Perhaps more intuitively understood than the previously mentioned causes is the problem that arises from new product's phase in and old product's phase out. Such events require changes in the IMS and a number of communication break-downs are prone to occur. Finally, there is the problem with minimum order sizes with manufacturers. This upstream issue can lead to delaying an order placement to a future moment past the ordering point, which increases the risk of incurring OOSt.

Additionally to the general processes that lead to OOSt situations, Gruen et al. (2002) also analyzed the OOSt magnitude for different countries and regions, different product categories, different days of the week, different kinds of promotion, different rates of product movement, different brands, and different durations of the OOSt (i.e., how long, on average, did the products stay unavailable for customers). The findings show that the average worldwide value for OOSt ex-tent was 8,3%, with small regional deviations around this mean. When it comes to the product lines, the differences between categories were more expressive, as presented in figure 4. An interesting consistent pattern could be perceived among the studied retailers when breaking down the analysis between the different days of the week, in which the rate of OOSt decreased throughout the days of the week, starting with a large rate on Sunday and ending with a below-average rate on Saturday (figure 5). Two main hypotheses could explain this phenomenon: first, assuming that the heaviest shopping volumes coincide with the weekend days, reordering and deliveries occur on Monday and Tuesday; second, its common knowledge that in some countries (e.g., Germany) stores are closed on Sunday and restocking does not occur until Monday, which increases the OOSt rate of the first weekday. Regarding the promotional effects, the conclusions showed that promoted items systematically presented higher OOSt rates when compared to non-promoted items. With respect to the velocity of product movement, studies suggest that OOSt rates for fast-moving items are around 13% and 15%, whereas this number for slow-moving items averages 8,3%. Moving on to the brand effects, the researchers could come to a solid conclusion about how specifically different brands influence out-of-stock rates; however, they could attest that the velocity of product movement impacted OOSt rates regardless of their brand. Finally, when analyzing the duration of the out-of-stock events, the numbers showed that around 20% of the items stayed in a OOSt status for less than 8 hours, whereas some 36% remained in the status between 1 and 3 days, with 19% of the items exceeding the 3-days mark.

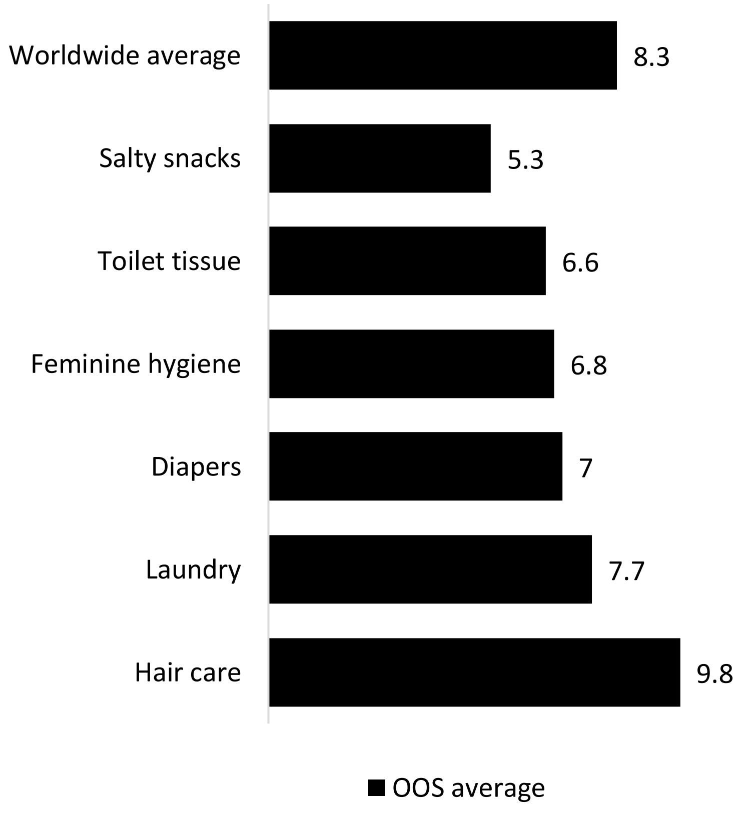


Figure 5: Out-of-stock averages by product category (%) (Gruen et al., 2002).

The researchers stressed that, due to three factors observed in the past few years, tackling the issue of OOSt proved itself to be more important than ever. The first of these factors states that consumers are becoming less tolerant to OOSt events because they are empowered with more information and variety of available channels for purchasing their necessities, meaning they will increasingly switch to an alternative outlet to find the items they need. Secondly, the benefit for retailers coming from solving issues related to OOSt as overall increased, since this market segment continues to consolidate worldwide and solutions to operational issues have the potential to be globally implemented. The third factor is related to the new technologies available for retailers, that provide them with new-found solutions to address this problem (e.g., RFID technology), rather than relying on traditional and costly countermeasures (Gruen et al., 2002).

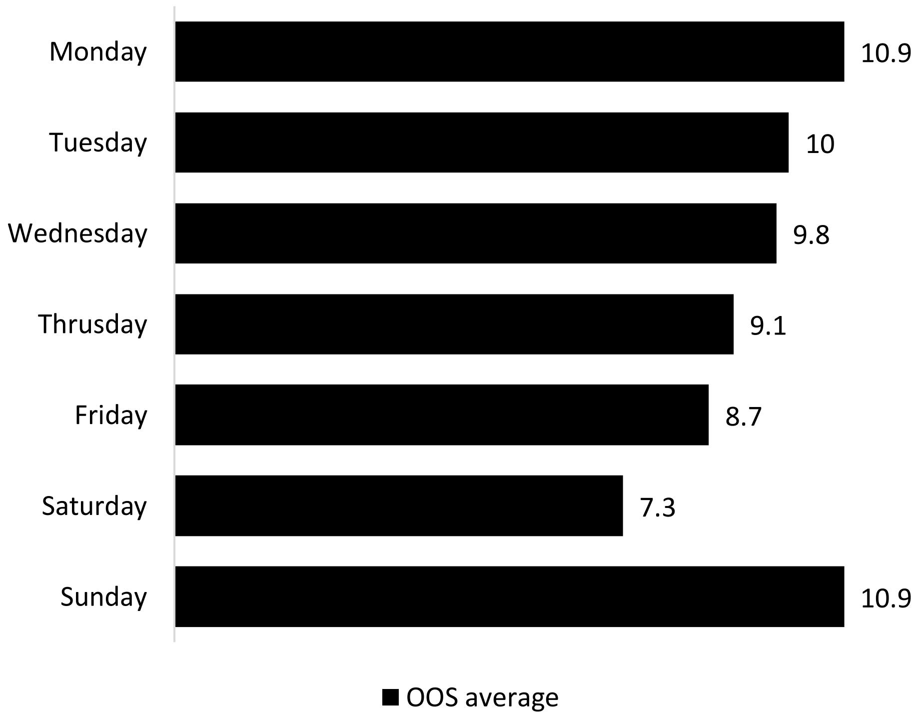


Figure 6: Out-of-stock averages per day of the week (%) (Gruen et al., 2002).

In another study that the German consultancy company Roland Berger carried out together with the Efficient Consumer Response (ECR) community (which, according to Gruen et al. (2002), was started in 1993 in the United States by grocery retailers, distributors, and manufacturers of FMCGs with the aim of reducing the many inefficiencies that can be found throughout the supply chain), they analyzed how product availability impacts not only sales, but also consumer behavior. Before discussing their findings, it is important to define what they refer to as out-of-stock (OOSt) situations: the authors basically state that OOSt assumes one of three forms. A classic OOSt occurs when there is a shelf-edge ticket, but no product (i.e., the product was supposed to be available on the shelf for sale but is missing instead). The second form of OOSt is called dual placement out-of-stock and happens when the product should be found on the shelf and in a second placement site but for some reason cannot be found in one or even both of them. The third OOSt situation is referred to as delisting out-of-stock and describes the situation in which the product is listed in the catalog but taken away from the shelves by the store staff (Hausruckinger et al., 2003).

After going through 18 different studies, across 11 countries and involving over 20,000 consumers, they came to some important conclusions about consumer behavior. The first one was regarding the consumer needs, where “shorter queues” ranked as a top necessity for consumers (being mentioned by roughly 67% of them), followed by “more promotions” (mentioned by 52% of responders) and “fewer items out-of-stock” (mentioned by 30% of consumers). This reveals that the average range between 7% to 10% of OOSt levels in Europe is found to be a too high of number for such an indicator under customers’ perspective.

The second conclusion to which the researchers arrived was related to the customer loyalty. When confronted with an OOSt situation, clients usually take one of the following courses of action: buying a different size or type of the desired product, buying a different brand of that product, not buying at all, buying the desired brand of the product in another retailer, or returning to the same retailer later to see if the product can be found. With this in mind, the findings show that 37% of consumers would buy a different brand, 21% would buy the brand elsewhere, 17% would return later, 16% would buy a different size of the product and 9% would not buy anything. When analyzing these data into more detail, the researchers came to the conclusion that Europeans are more likely to switch brands than retailers, which gives the latter group a leverage over manufacturers in situations in which the OOSt is caused by supplying issues from the manufacturer. Figure 6 depicts how consumer loyalty relates to brand switching and store switching.

The third conclusion regarding consumer behavior took into account the recurrence of OOSt situations experienced by the customer in a certain store throughout the time. It showed that although 69% of customers choose to substitute brands and 31% choose to change to a different re-tailer in the first OOSt situation, by the third OOSt situation this ratio is exact the opposite. By then, 31% of customers still choose to substitute the desired brand whereas 69% of them opt to change to a different store.

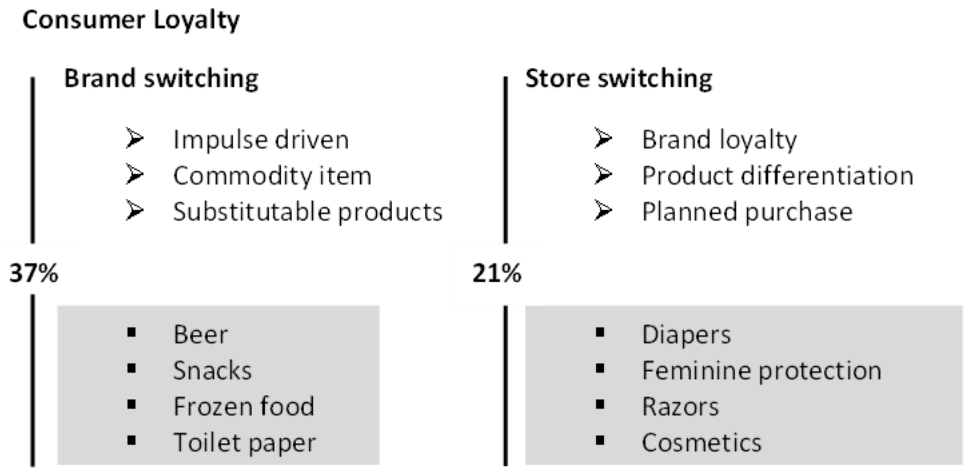


Figure 7: Customer Loyalty (Hausruckinger et al., 2003).

When studying the modus operandi of retailers and alternative solutions that could minimize the already mentioned problems of OOSt and OOSh, Yaw Wong & McFarlane (2007) analyzed in detail the shelf replenishment process and raised a list of the different factors that could potentially influence the effectiveness of its performance. The 7 major culprits that the re-searchers identified were the observation delay, checkout delay, delays in picking list preparation, delays in locating products in the backroom or on the shelf, delays in moving the products from the backroom onto the shelf, the time needed to fill up the shelf to its maximum capacity, and the product removal pattern. To understand each of these factors proved to be important for the later developed simulation and therefore a brief explanation of each of them is provided ahead.

Starting with the observation delay, in the "pull" replenishment policy it depends on the frequency of inspections (sometimes also referred to as audits) to verify whether a certain product type on the shelf appears low on its quantity or even off-sale (OOsh situation). When formulated in another way, such delay is essentially the time between inspections. On the other side of the spectrum, in the "push" policy the observation delay is defined as the time interval between inspections carried out in the backroom. What both policies share in common, however, is the fact that the observation delay is the trigger that starts the replenishment of shelf inventory levels. Hence, it is comprehensible that the frequency of audits affects directly the frequency of shelf replenishments. In other words, less frequent observations may result in lower service level, or also depicted by a lower shelf availability of products.

Moving on to the next factor, comes the checkout delay. It is defined as the total time elapsed between the moment a product was taken from the shelf and the future moment when it was processed in the POS by the cashier (or by the customer in the case of a self-checkout POS). Accounted in here are the customer walking time around the store facility, that can prove to be significant if the size of the store is large (e.g., department stores), and the customer queuing time at the check-out, that is usually expressive during peak hours. So next, by the time the product is finally registered at the POS, the same is deducted from the IMS (in an ideal situation, this process occurs flawlessly); in other words, the barcode-based data capture feeds the EPOS information into the inventory system. This is the information that is later on used to estimate the quantity of product on the shelf and, because of it, it is natural to conclude that the longer the checkout delay, the more inaccurate this estimate will be (sometimes this delay is decisive in generating OOSh situations and, therefore, the management of the retail must counter it somehow).

The delay in the preparation of the picking list comes subsequently and, before any analysis, requires first a clarification as to what the purpose of such a management tool is. This way, a pa-per-based or digital picking list, whether generated by the IMS or simply manually recorded by the retail employees, serves the impetus of recording the quantity to be replenished from the backroom to the shelf, with the goal of orienting the people carrying on this task. Relevant is the fact that this picking list may take a relatively long time to prepare if elaborated manually using a pen and paper approach. In comparison, with the aid of a handheld barcode reader, this process has the potential to be vastly simplified, since in this case the product is automatically identified (apart from rare exceptions in which the employee has to input the ID of the product manually due to scanning issues) and only the number of outstanding items on the shelf needs to be recorded. However, if the execution of the upcoming shelf replenishment process takes too long to be accomplished, there is the risk that the picking list might become obsolete (a situation that is comparable to the problem encountered by the checkout delay with the inaccuracy of shelf availability), especially when dealing with FMCG.

Delays in locating products in the backroom or on the shelf come next. They are relevant to consider since the ability of a retailer to systematically locate different SKUs within its store (be it in the backroom or on the shelf) is imperative for a quick and efficient shelf replenishment process. Although at first glance mentioning it might seem trivial when analyzing the retail shop floor's performance, a cluttered backroom is no exception in this sector and it can transform the shelf replenishment task into an extremely challenging, as well as equally time-consuming and laborious, activity. Another aspect worth pointing out in this context is the fact that shelf displays at different locations can potentially confuse the employees. Even in the ideal scenario in which the IMS is operating under a perfectly accurate regime by the time of the shelf replenishment process, employees who are working under time pressure oftentimes can "write-off" (i.e., unadvisedly neglect) unobserved or not rarely misplaced items, leading to the exact inaccuracies that are targeted in the cur-rent work. More experienced members of the staff take advantage of their experience and some-times rely on their memory and skills to quickly locate products. It is of considerable relevance in this topic to reiterate the fact that neither the manual nor the barcode data capture approach pro-vides an instantaneous product location information (out this fact comes the literature support that some products do get lost in the backroom and on the shelf leading to unreported shrinkage and, henceforth, IRI).

Succeeding the list of influencing factors come the delays in moving products from the backroom to the shelves. So, following the steps that compose the replenishment process, once the product to be replenished is correctly located in the backroom storage, it must be transported to the shelf (usually with the support of a cage or dolly that allow the employees to carry multiple items at a time). Depending on how extent and diverse in terms of SKUs the picking list might be, there could be a whole variety of different product categories that require replenishment at a certain time and, consequently, need to be put into the cage/dolly before they are rolled out to be displayed for the customers. The correct quantity should be replenished over a single run to avoid as much as possible double handling situations. This fact is important because it will be later on presented as an assumption for the simulation (at every replenishment cycle, the shelf is filled up until its maximum capacity). A direct consequence of this procedure of moving items from the backroom to the main area of the store lies on the fact that the detailed layout (arrangement of aisles, shelves, fridges and-so-forth) of the retailer has a direct influence on the delay time of this step.

The next factor in line comprehends the total time needed in filling the shelf up until its maxi-mum allocation. To shed light on this process, one must first clarify align the concept of maximum shelf allocation: it refers to the utmost number of products allocated on a shelf according to a dis-play arrangement (that in turn can change depending on, for example, promotion campaigns). A specific allocation of product quantity in a retailer normally takes into account the previous sales patterns of that product line (if available), the demand forecasts for that product line (which in itself is a topic of profound complexity, but not the focus of the current work), and it usually changes in a weekly basis (this fact was not accounted for in the simulation since the model simulated just a single item). A phenomenon that arises from a high shelf allocation and can be seen in retailers in general is the increase in the time spent on product presentations, as well as the amount of products that need to be replenished at any one time from the backroom. Another circumstance that is often observed is the fact that products that are relatively small in size or, on the very oppo-site side, products that are heavy and bulky are usually difficult to organize neatly on the shelves, taking time from the staff and, by doing so, being somewhat more susceptible to errors.

Finally, as a last identified factor that influences the replenishment of the shelves in retail establishments, and also the last step of this cycle, there is the impact on replenishment occasioned by the product removal pattern. The definition of product removal pattern sets it as being the rate at which products are removed from the shelf by the customers. Although it might not seem intuitive at first, this pattern is different from the demand pattern, which is generated when customers pay for their products during the checkout. To illustrate it better, it is important to picture that the demand pattern is set as a function of four factors: the removal pattern itself, the customer walking time, the customer queuing time at checkout counters, and the possible theft (either internal or external) after removal. To add to the relevant features of product removal patterns, it must be mentioned that it also depends on the characteristics of the product (in other words, different SKUs will present different removal patterns, which can seem intuitive to many people). This way, high demand product lines (and, hence, FMCGs) such as detergents, spirits, carbonated soft drinks, ice cream, confectionery and fresh ready meals have higher off-sales levels (which translate to higher OOSh situations or a lower service level for these categories). In a similar logic, findings suggested that promoted products behaved similarly and surveyed to have relatively higher than average off-sales levels. Moreover, although this factor does not present any specific delay like other factors mentioned before did, it faces another kind of problem: products with high removal rates also demand the shelf to be replenished within a shorter period of time in order to maintain the same on-shelf-availability (or service level) before the customers. At last, in both “pull” and “push” policies (explained in the following paragraph), product removal patterns cannot be continuously monitored. Instead, their monitoring can only be carried out on a periodic basis by the workers. This fact is one of the arguments later presented in the methodology (section 3) as to why a system dynamics simulation model was discarded when choosing how to simulate a retail store (Yaw Wong & McFarlane, 2007).

In addition to the detailed explanation of the factors influencing the replenishment policy, the authors also distinguished between two broad classes of replenishment policies in the retail sector (Yaw Wong & McFarlane, 2007).

The first one is a so-called “pull” replenishment policy, in which the act of monitoring the shelf inventory levels triggers the replenishment from the backroom. For this feature, the “pull” policy is also sometimes referred to as Kanban system.

The second policy is referred to as a “push” replenishment and is described as the situation in which the store staff monitors rather the backroom in order to estimate shelf inventory levels and this is the act that in turn eventually triggers the shelf replenishment.

As noted, the product observation point – in other words, the point at which the product inventory in general is directly monitored – is a decisive distinguishing factor between the “pull” and the “push” policies. With regards to this aspect, the authors claim that a simultaneous real time direct physical observation (as opposed to inferred) of product inventory both on the shelf and in the backroom is impossible. Another observation worth mentioning is the fact that the “pull” policy is the most common of the two policies among retailers and, for this reason, was presented in section 3 as the replenishment policy for the developed model. Table 3 presents a comparison be-tween the “push” and “pull” replenishment policies.

Table 3: Comparison between "pull" and "push" replenishment policies (Yaw Wong & McFarlane 2007).

| **“Pull” replenishment policy** | **“Push” replenishment policy** |
|---|---|
| Visual inspection starts on the shelf. | Visual inspection starts in the backroom. |
| Products are replenished according to shelf arrangements. | Products are replenished according to backroom arrangements. |
| For barcode approach, amount of product in the backroom is estimated. | Amount of product on the shelf is estimated. |
| Time to move product from the backroom to shelf is affected by the time to find the product in the backroom. | Time to move product from the backroom to shelf is affected by the time to find the product on the shelf. |

The authors also provided a schematic view of each of the policies, identifying every step of both “pull” and “push” policies, as well as how the information and product flows behave in each case. Such portrait can be found in figure 7.

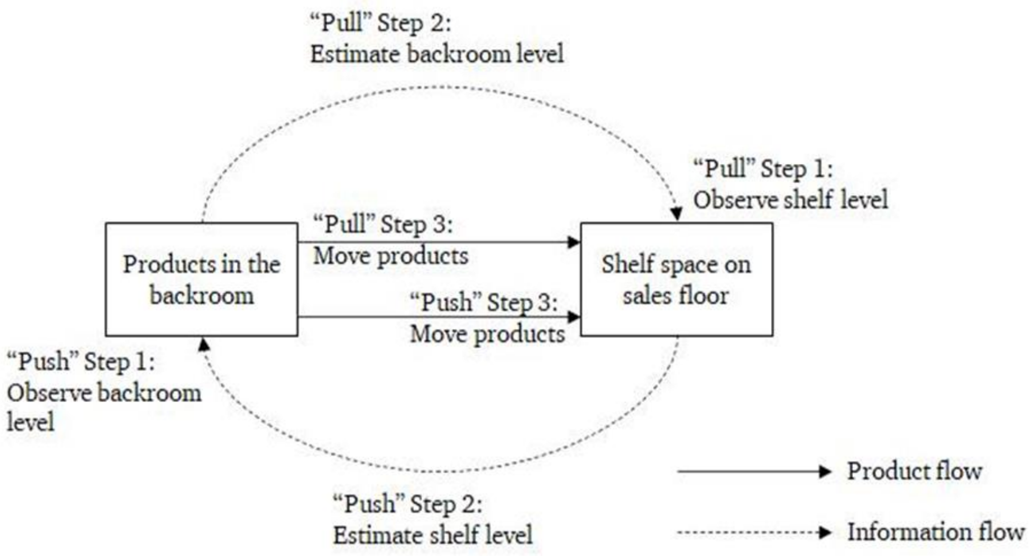


Figure 8: Representation of "pull" and "push" replenishment policies (Yaw Wong & McFarlane, 2007).

### 2.1.2 Countermeasures against IRI

Oftentimes, the suggested countermeasures against IRI were usually presented in the very same papers that were dedicated to study their causes. These countermeasures proved to be an important piece of the puzzle when constructing a holistic understanding of the IRI problem for retailers. Therefore, it was judged as relevant to dedicate a subsection to present the main findings regarding such countermeasures. The insights taken out of this section, just like those from the previous one, were considered when developing the simulation model presented in section 3.

The next LP prevention strategy are the LP prevention awareness and training programs. According to Hollinger & Davis (2002), the most widely implemented awareness program was the discussion about shrinkage during new hire orientation (89,8% of surveyed companies reported doing it). Following it there were bulletin boards notices/posters (84,7%) and the use of anonymous telephone hotlines (84,7%) as second most implemented techniques. Next, periodic programs/lectures (78%) and dissemination of an employee “code of conduct” (75,4%) appeared as above-average mentioned initiatives. The last group of measures mentioned by over half of inter-viewed retailers was comprised of training/instructing videos (62,7%), newsletters (60,2%), and honesty incentives (58,5%). Finally, the least mentioned initiatives were in-store employee loss prevention committees (29,7%), informational paycheck stuffers (29,7%), internet communications (28%), employee surveys (22,7%), and training audio tapes and announcements (10,2%). Moraca et al. (2015), just like in the last paragraph, separated the analysis into retail category. Their findings were the following: men's and women's specialty apparel retailers focused primarily on training videos about LP for their employees (83,3% against the average of 71,6%), with further support from anonymous online/email notification (75% versus the average of 55,4%), and online communications (75% versus the average of 55,4% as well). Department stores directed their efforts into hotline, bulletin board notices, code of conduct, and new hire orientation (100% each), and, to a lesser extent, Active Shooter training programs, honesty incentives, and in-store LP committees were also in place quite often. The final category, grocery stores and supermarkets, more often than average implemented internet-based interactive or CD-ROM training methods. A comparative view of the findings of both NRSS editions is provided in figure 9.

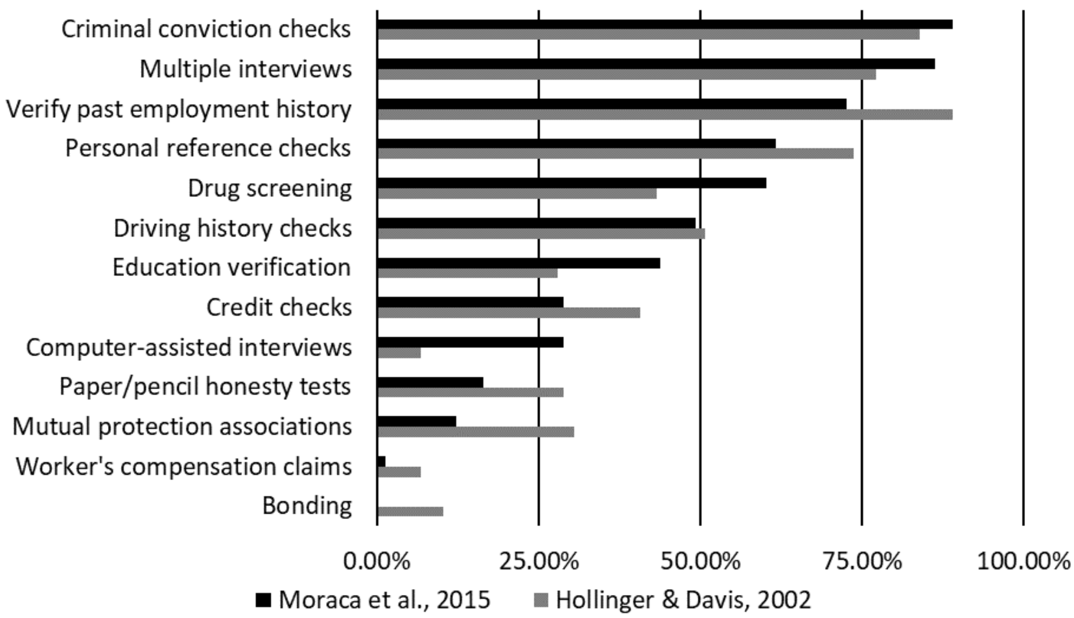


Figure 9: Most used integrity screening measures (Moraca et al., 2015; Hollinger & Davis, 2002).

Asset control policies come in sequence as a common shrinkage prevention strategy. It is important to note here that Moraca et al. (2015) do not refer to this category in their study because their focus was mostly directed to criminality-associated shrinkage and so, in other words, there are no related data available in the 2015 edition of the NRSS. Diving into details, Hollinger & Davis (2002) state that this category was with no doubt the most broadly adopted LP strategy among retailers. Topping the list of most frequently used asset control policies is the re-fund control (mentioned by 93,2% of surveyed retailers), followed by void controls (88,1%), employee package checks (80,5%), POS exception-based reporting (79,5%), trash re-moval controls (76,3%), inter-store transfer controls (73,7%), POS bar coding/scanning (73,7%), price change controls (70,3%), unobserved exit door controls (67,8%), inventory bar coding/scanning (66,9%), de-tailed merchandise receiving controls (55,1%), and use of newly developed, electronically con-trolled access to cash handling areas (17,8%). Figure 10 shows these asset control policies in a summarized graph.

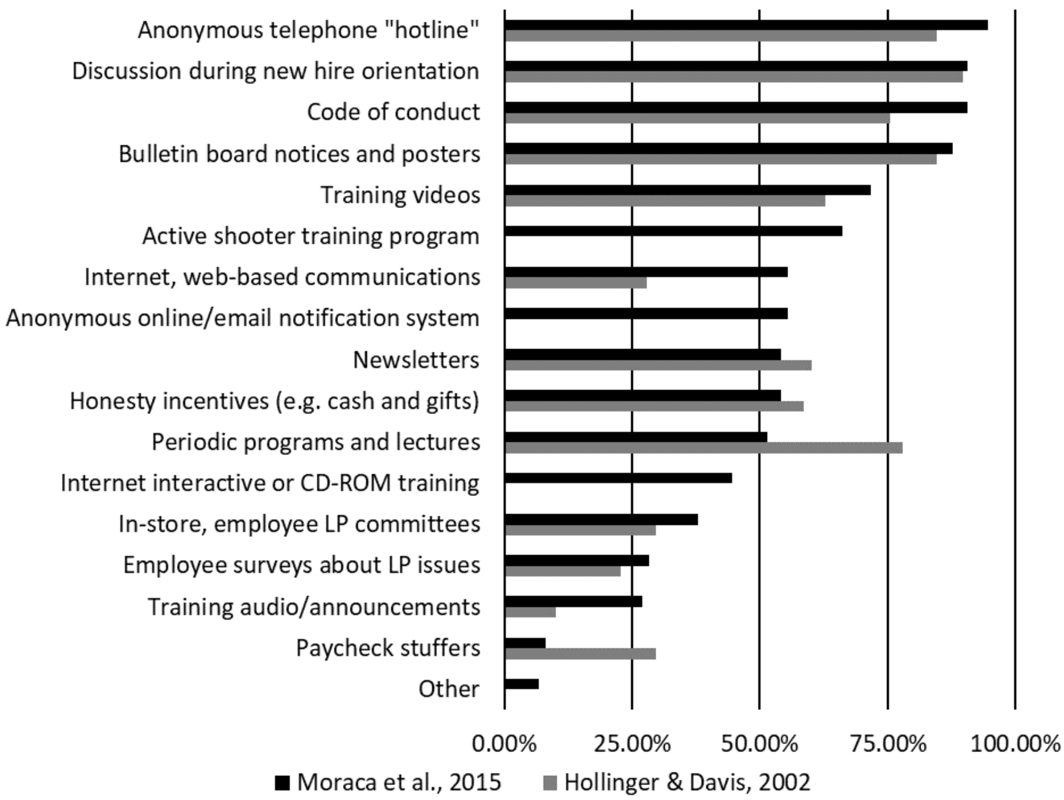


Figure 10: Most used LP awareness and training programs (Moraca et al., 2015; Hollinger & Davis, 2002).

The last LP prevention strategy mentioned in both 2002 and 2015 NRSS editions is the loss prevention systems and personnel. Hollinger & Davis (2002) referred to it as being the largest category of their study and comprehending the use of mechanical, electronic, video, and personnel to monitor not only shoppers, but also employees and, of course, the store's merchandise. At the top of the list of the study of Hollinger & Davis (2002) are the burglar alarms (present in 94,4% of all participant retailers), followed by visible CCTV cameras (73,3%), check approval databases (61,9%), armored car deposit pick-ups (56,8%), cable/locks/chains (51,7%), live hidden CCTV (50,8%), digital video recording systems (50,8%), observation mirrors (49,2%), POS data mining software (49,2%), secured display fixtures (48,3%), drop safes (46,6%), mystery/honesty shoppers (45,8%), acousto-magnetic, electronic security tags, or EAS (42,4%), shoplifting deterrence signage (41,5%), uniformed guards (37,3%), ink/dye benefit denial tags (36,4%), silent alarms (32,2%), visible CCTV (29,7%), radio-frequency EAS tags (28,0%), plain clothes detectives (27,1%), merchandise alarms (26,3%), fitting room attendants (18,6%), microwaves EAS tags (17,8%), POS exception based CCTV interface recording (17,8%), timed entry safes (14,4%), acousto-magnetic EAS source source-tagged merchandise (13,6%), radio-frequency source tagged merchandise (10,2%), observation booths (10,2%), and RFID tags (1,7%). As shown in figure 11, these numbers varied a lot in comparison to those presented by (Moraca et al., 2015).

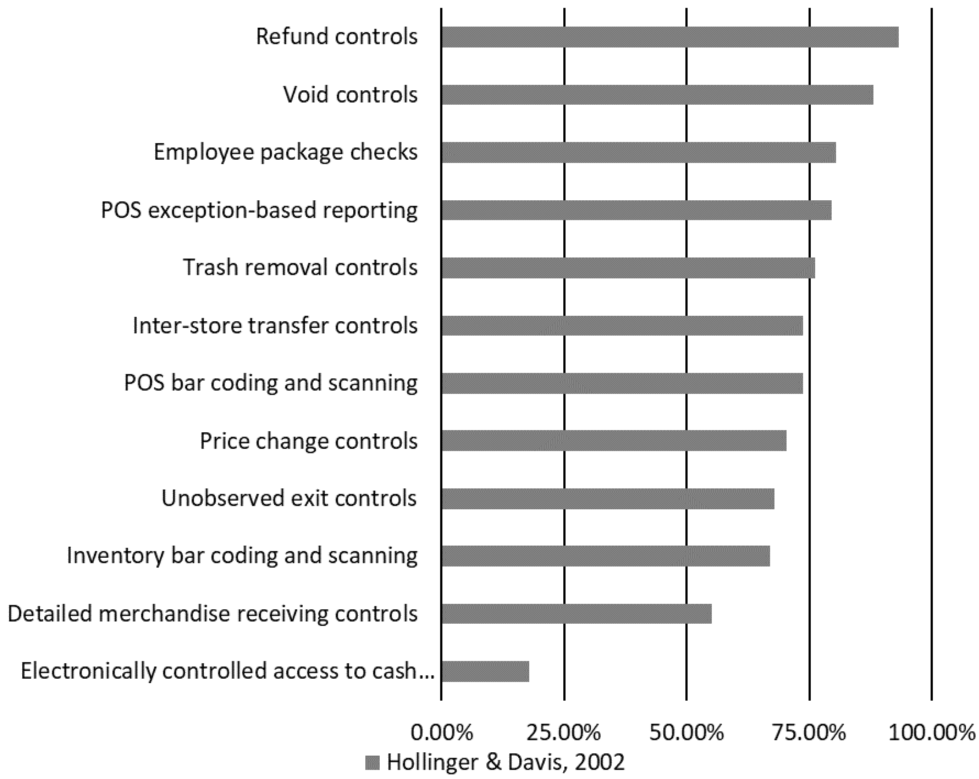


Figure 11: Most used asset control policies (Hollinger & Davis, 2002).

The study of Moraca et al. (2015) also presents additional insights into the strategy mentioned in the last paragraph that are worth mentioning. The first one is that 100% of retailers across all categories responded that they had burglar alarms in operation to assist against LP. Another finding shows that all surveyed retailers that runed men’s and women's specialty apparel and sporting goods stores used acousto-magnetic, electronic security tags. In comparison, only one-third of department stores implemented them. When it came to ink/dye benefit denial tags, all the addressed sporting goods retailers used such tags, whereas just half of department stores and specialty apparel retailers did so. Subsequently, all retailers that operated grocery stores and super-markets used POS data mining where, by contrast, only two-thirds of department stores and half of men’s and women's specialty apparel stores did so. At last, when comparing the numbers regarding shoplifting deterrence signage, two-thirds of department stores and one-third of men’s and women's specialty apparel reported posting any kind of warning signs exposing the consequences of shoplifting. Even fewer goods retailers reported adopting such countermeasure (just one-fifth of them) and absolutely none of the grocery stores and supermarkets did so.

After applying these countermeasures, retailers reported a significant drop in overall perceived shrinkage (Hollinger & Davis, 2002). However, they are not collectively exhaustive and, because of it, further results from other literature were also brought to this section. For instance, Moraca et al. (2015) mention that only 16,4% of participating retailers reported using RFID technology for inventory control. This topic alone is analyzed in deeper details by Condea et al. (2012) and Kang & Gershwin (2005) in their papers and, therefore, will be presented here as well.

Starting with Condea et al. (2012), their study aimed at verifying the feasibility of RFID technology as a cost-efficient inventory control system that could improve the overall service level in the retailing industry. The researchers identified in the RFID a potential for further improvement of process efficiency in retail stores and improvement of overall supply chain visibility of material flow. The reason behind the high expectations around this technology lies on the fact that it differentiates itself from the widely adopted barcodes due to the fact that it can carry out bulk registrations, identify items out sight in an unambiguous manner, store data in the objects themselves (i.e., products), and provide robustness against environmental influences and disturbances (i.e., it is more resilient to damage than the traditional barcode). Some case studies that make use of these mentioned potentials of the RFID technology are already available to be seen in the market. For instance, in order to take advantage of the traits of RFID within their supply chains, the retailers Walmart, Tesco, and Metro (among others) required that some of their suppliers attached RFID tags into their products for better monitoring them downstream in the supply chain all the way to the customer checkout (Condea et al., 2012).

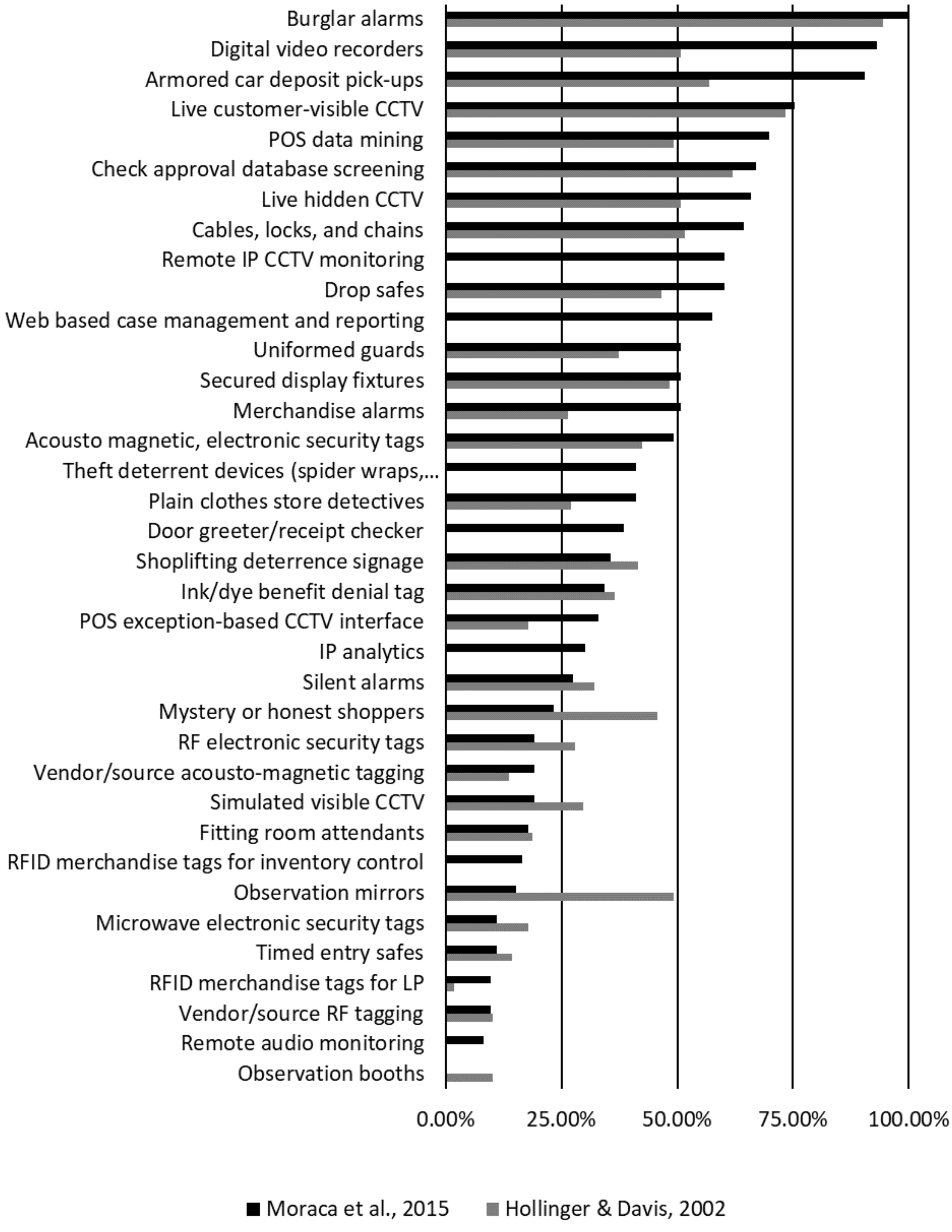


Figure 12: Most used LP systems and personnel (Moraca et al., 2015; Hollinger & Davis, 2002).

Another identified potential for the RFID technology highlighted in the same paper is the improvement of in-store processes, specially the critical one of shelf replenishment. In such case, the RFID technology would act as a way to automate the monitoring of inventory levels and, by doing so, would also trigger the replenishment from the backroom to the sales floor in real time and with higher accuracy than current standards whenever it was needed. Such an approach to the shelf replenishment process would be a promising alternative to having the store staff visually inspect shelf by shelf through recurrent inspections throughout the store's aisles (Condea et al., 2012).

Although the RFID technology clearly has advantages when compared to currently used methods of inventory control, it also has one clear downside that prevents it from being broadly accept-ed: its costs. RFID is no cheap solution, and the researchers recognize that. They state that, for ex-ample, solutions such as "smart shelves", which suggest installing RFID reading devices in every shelf, could be unrealistic at present due to its high costs. Instead, they propose a far more realistic alternative way by implementing one single RFID reading point between the backroom and the sales floor that would be responsible for tracking the movement of some cases instead of many single items (Condea et al., 2012).

With the goal of verifying if a RFID-enabled replenishment process would outperform a traditional periodic review (PR) one, the researchers carried out a series of simulations with both replenishment policies and monitored their results. Their conclusions confirmed the previous beliefs and the RFID-based replenishment policy indeed led to better responses to customer demand than the traditional PR.

Moving on to the study of Kang & Gershwin (2005), they also approached the IRI problem via simulation. The assumptions and parameters adopted by the researchers are presented in section 3 and they were of great importance for the development of the simulation model used in the current work. However, relevant for this section are the countermeasures that they suggested to tack-ling the problem of discrepancies between actual and recorded inventories. This way the following paragraphs will be dedicated to presenting the so-called "compensation methods for inventory inaccuracy" (Kang & Gershwin, 2005).

It is important to note that the methods to counter IRI problems are many and these suggestions are regarded by Kang & Gershwin (2005) as not being collective exhaustive, meaning that other countermeasures that are perhaps even more effective might not be listed here. However, the solutions presented are representative and aligned with the best practices among retailers.

The first compensation method is the safety stock. Retailers usually resort to this type of stock to shield themselves against uncertainties at an operational level. Examples of such uncertainties would be variations in demand consumption rate and supplier/replenishment lead time. However, the safety stock can just as well serve as a buffer against variations in the inventory record of the store if, of course, the retailer is aware of such problem. In a (Q,R) inventory policy, in which a fixed reorder quantity Q is placed whenever stock level reaches a certain value R, the level of safety stock is adjusted according to the parameter R. In other words, if the store management wishes to in-crease their safety stock, they adopt a higher value for R, likewise, if they wish to reduce inventory carrying costs, they can reduce the safety stock by lowering the R-value. Applying the same logic to IRI situation, a higher uncertainty regarding inventory accuracy could be buffered by increasing R.

The second compensation method is the manual inventory verification. This compensation method is to this date still widely adopted among retailers in general as a countermeasure to mitigate the harmful effects of IRI. It consists simply of having store personnel manually counting the SKUs in the facility (backroom and shop floor) in order to later compare it to inventory levels in the IMS and, in turn, correct any discrepancies between actual and recorded values. The retailer’s management can opt to prioritize the verification of just part of the entirety of SKUs that compose its portfolio. In such case, the retail staff would carry out more audits of these selected items than the yearly minimum required for fiscal purposes. One example would be prioritizing the stock counts of FMCGs, since, as already mentioned in the last section, they are reportedly more susceptible to IRI (DeHoratius & Raman, 2008). Another example would be carrying out audits of high-valuable items according to an ABC classification, that ranks SKUs in decreasing order according to their annual sales values and splits this array into an A-class (e.g., the more or less 20% of items that account to around 70% of total turnover), a C-class (e.g., the 50% or so items that correspond to approximately 5% of total turnover), and a B-class (the remaining items in between the A- and C-classes) (Rekik et al., 2019). Naturally, retailers are not capable of conducting as many stock counts as they wish since these processes are constrained by labor and time availability.

DeHoratius & Raman (2008) complement the recommendation of Kang & Gershwin (2005) by confirming that, in an average retailer, a higher frequency of audits indeed decrease IRI numbers and that items audited once a year had around 12% greater record inaccuracies than items audited twice in the year.

Another important observation to be made here is that, according to Iglehart & Morey (1972), an inventory count does not completely eliminated discrepancies between the recorded and physical stock. In practice, non-trivial inaccuracies usually remain in the inventory records be-cause of human error in the counting procedure itself. Moreover, transactions that take place during the count make this task extremely difficult and costly to assure 100% accuracy.

Following with the list of Kang & Gershwin (2005) of compensation methods, there is the manual reset of the inventory record. To better understand this countermeasure, it is important to understand that sometimes the simulation developed by the researchers reached a point called “replenishment freeze”, in which the actual inventory level of the simulated product line was zero, but the recorded level was greater than the replenishment point R. This situation led to a stalemate on that item’s sales, since no replenishment order was sent to the supplier anymore and no products were available on shelf for customers. In the rare occasions that the store management could be faced with such situation, the researchers suggested simply setting up the recorded inventory of the affected product line to zero manually, which in turn would allow the auto-mated replenishment system to start placing orders again.

The fourth compensation method is the constant decrement of the inventory record. This countermeasure requires that the inventory manager knows that there is a problem of stock loss within a certain product line, and it also requires that the stochastic behavior the product is known beforehand. If these two conditions are fulfilled, then the manager could simply decrement the inventory record by the average stock loss demand in each period (e.g., every three months). (Kang & Gershwin, 2005) recognize that this measure most likely would not eliminate the inaccuracy between actual and recorded inventory. However, it would be better than leaving the system untouched, since in this case the system would eventually reach the “replenishment freeze” point and OOSt situations would arise together with all of its consequences.

The last proposed compensation method is the implementation of the Auto-ID product identification system. At the time the paper was published, this technology was under development by the Auto-ID Center, which was founded in 1999 at the Massachusetts Institute of Technology and is sponsored by over 100 global companies. The basic idea behind the Auto-ID system is to have RFID tags placed on physical products such as, for example, food packages, clothing, and machinery spare parts. These tags would then be tracked by RFID readers installed in key locations along the supply chain, which would enable managers to track intermediate and final products all the way from the upstream-positioned manufacturers down to the customer check-out at the POS, and per-haps even beyond (Kang & Gershwin, 2005). The researchers adopted the premise that Auto-ID could provide a perfectly accurate measurement of the real physical inventory level of a certain product line and, therefore, in terms of performance it would rank as the best available countermeasure against IRI (Kang & Gershwin, 2005).

Now that all compensation methods were presented, it is possible to proceed to the last major finding regarding countermeasures against IRI. In their paper, Chuang & Oliva (2015) conducted a comparison between part-time (PTE) and full-time employee (FTE) performance when executing inventory audits (a factor that the researchers added to the "labor availability" category). The researchers' conclusions suggested employing PTEs does not improve the quality of the audits (it actually has no effect on it) and, therefore, it is not a valid countermeasure against IRI. In other words, what the store management should do to cover staff gaps or simply to improve audit quality is to hire FTEs instead. Some reasons behind the worse performance of PTEs in comparison to FTEs are based on the facts that PTEs are less committed and less trained and, this way, more prone to commit mistakes.

Finally, to mention two countermeasures that are very particular of the papers they originated from, there is the system to identify OOSh products developed by Papakiriakopoulos (2012) and the approach to monitor the performance of IMSs by using control charts, which was developed by Ernst et al. (1993). From the literature collected and presented in this section, it was possible to conclude that both countermeasures against IRI are relatively uncommon and, for this reason, will not be further covered in this thesis.

## 2.2 Methodology-supporting literature

This section is dedicated to presenting the literature supporting the choice of simulation and factorial design as ways to answer the research questions presented in section 1.3 and, furthermore, to presenting the literature supporting some of the parameters used as input in the simulation. With this in mind, section 2.2 was further subdivided into 4 subsections: "Inventory management policies", "Simulation: in-depth approach and useful takeaways", "Design of Experiments: in-depth approach and key takeaways", and "Simulation and factorial design as ways to approach IRI".

### 2.2.1 Inventory management policies

The first step when developing the simulation model presented in section 3 was to understand the dynamics of different replenishment policies. For this purpose, the focus of the current work was briefly shifted from studying the causes, consequences, and countermeasures against IRI to dive into the vast field of Operations Research. Within the area of Operations Research, a special focus was given to the topic of Scientific Inventory Management and the guiding literature used here was the book of Hillier & Lieberman (2015), "Introduction to Operations Research".

Starting with a quick background, the importance of efficiently managing inventories is not restricted to retailers. Holding inventories is an activity inherent in any company that deals with physical products. For instance, manufacturers and wholesalers also pay very close attention to inventory carrying costs, which are all but negligible. The total value of all inventories (here including finished goods, partially finished goods, and raw material) in the US at the time the 10th edition of book was published totaled over a trillion dollars, which corresponded to more than US$ 4.000,00 per living person in the country. On the other side of the globe, Japanese companies like the Toyota Motor Corporation (or simply Toyota), were pioneers in introducing the concept of "just-in-time (JIT) inventory system" in alignment with their worldwide known lean production model. In this inventory system, an emphasis is placed in planning and scheduling so that no more than the required materials arrive exactly at (or, more realistically, close to) the time they are needed, which then result in large amounts of savings due to reduced inventory levels, but also face the trade-off leaving the supply chain more vulnerable to, for example, variabilities on demand. Within this broader context, one can define the concept of Scientific Inventory Management (SIM). In just a few words, it refers to the application of Operations Research techniques to control inventory policies with the goal of gaining a competitive edge by doing so.

The application of SIM techniques comprises of four steps: formulation of a mathematical model describing the behavior of the IMS, search of an optimal inventory policy with respect to the previously formulated model, usage of a computerized information processing system (or just IS) to maintain a record of the current inventory levels, and the usage of recorded inventory levels to apply the optimal inventory policy developed in step 2.

The mathematical inventory models mentioned in step 1 can be subdivided into deterministic models and stochastic models according to how predictable the demand is. Demand here is defined as the number of units that are taken from the inventory at any given period. If this demand can be accurately predicted, in other words, if the future demand is known beforehand, a deterministic inventory model could be used. However, in cases where the future demand simply cannot be predicted very well, a stochastic inventory model would be more recommended since it primarily treats the demand as being a random variable rather than a known constant (Hillier & Lieberman, 2015).

Now that the basic idea behind SIM was presented, a discussion of the main components of inventory models can take place. The defining parameter when choosing an inventory model is its relative profitability in comparison to others. Profitability in turn is basically directly or indirectly influenced by two other parameters: costs and revenues. Finally, these two can be further subdivided into six categories: ordering costs, holding costs, shortage costs, revenues, salvage costs, and discount rates. Each of these six will be approached in detail in the next paragraphs (Hillier & Lieberman, 2015).

The ordering costs can be represented by a function c(z) where z is the amount of purchased or produced items. In its simplest form, this function is directly proportional the ordered amount.

$$c(z) = c \cdot z \tag{1}$$

Another common form of the ordering cost function is given by equation 2, in which a term that is directly proportional to the amount of items ordered and a setup term (K) that is constant build up the function. The setup cost includes, for example, administrative costs of ordering or, in case of producing the items, the costs of setting up the machinery to start the production.

$$\begin{cases} 0 & if\ z = 0 \\ K + cz & if\ z > 0 \end{cases} \tag{2}$$

Coming next are the holding costs, which sometimes are referred to as storage costs. It represents the entirety of costs associated with the storage of the inventory up until it is eventually sold or used.

The shortage costs, also known as costs of unsatisfied demand, is present whenever the demand exceeds the available stock. Such cost category can be subdivided into a backlogging case and a no backlogging case. In the backlogging case, the exceeding demand is unsatisfied but not lost. This means that the demand is put on hold until it can be satisfied. In the no backlogging case, the exceeding demand is either somehow served by a priority shipment or lost forever. Analyzing these two cases, one can finally picture the shortage costs as being, for example, the cost of customers’ goodwill decrement (and subsequent loss of willingness to do business with the company), cost of postponed revenue inflow, extra administrative costs, cost of organizing a priority shipment, and cost of loss of potential revenue.

The next parameter to be discussed is the revenue itself. Since it is considered to be a parameter of secondary priority for inventory management, it may or may not be included in the model. This happens because the revenue from sales is independent of the firm’s inventory policy (assuming a scenario in which demand is met). If the store management opts for neglecting the revenue, any “loss in revenue” must be accounted for in the already discussed shortage costs.

Subsequently there is the salvage cost of an item. Salvage value is defined as the value of a leftover item in a situation where no further inventory is desired. In such cases, the item might be disposed through a discounted sale. This generates a negative impact in the firm’s profitability that is called salvage cost.  If, for example, there is a cost arising from the disposal of the undesired product, the salvage cost will be attributed to a positive value.

At last, the discount rate here is the same as in Financial Mathematics. It takes into account the time value of money (e.g., in an inflationary scenario, the value of a certain quantity of money decreases with time). Whenever the firm ties up its capital in inventory, the firm is prevented from using this capital in alternative investments and, therefore, opportunity costs arise. The discount factor is especially important to be considered when modeling long time horizons.

Now that the profitability parameter was extensively discussed, it is possible to move on the next one: the lead time. In inventory management, lead time is defined as being the amount of time between the placement of a replenishment order and the receipt of these ordered goods in a future time. The lead time can be fixed or variable depending on the assumptions made when developing the model (Hillier & Lieberman).

According to Hillier & Lieberman (2015), the studies around SIM also take into consideration the way the inventory is monitored. This leads to the next modeling parameter, which verifies if the monitoring policy is a continuous review (i.e., a replenishment order is placed as soon as the stock level reaches a point equal to or lower than a reorder point R) or a periodic review (i.e., the inventory level is assessed at discrete time intervals).

With the distinction between continuous review and periodic review in the last paragraph, all the most important simulating parameters were already discussed. Now, a detailed distinction between the deterministic continuous-review, deterministic periodic-review, stochastic continuous-review, and stochastic periodic-review will be presented.

Kicking-off with the deterministic continuous-review policy (Hillier & Lieberman, 2015). This policy will be introduced via the well-known Economic Order Quantity (EOQ) model, oftentimes also referred to as Economic Lot-Size model. Here, units of the analyzed product are considered to be consumed from the inventory at a known and constant rate a. Moreover, it is also assumed that the inventory is instantly resupplied with a fixed batch size of Q units. The last new variable to be defined is h, that represents the unit inventory holding cost of an item. Note that since there is no lead time and a is constant at any given time, there is no need for a safety stock. Figure 12 illustrates how such a model would look like.

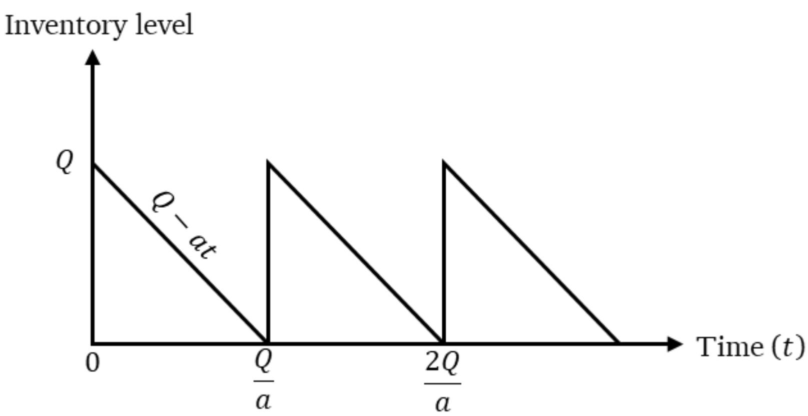


Figure 13: Diagram of inventory level as a function of time for the basic EOQ model (Hillier & Lieberman, 2015).

The objective of a model such as the one presented in figure 12 is to determine when to replenish the inventory and with how many units, so as to minimize the total costs T per unit time. The following equations 3 to 8 indicate the step-by-step to reach the optimal order quantity Q* that minimizes T.

$$Production\ or\ ordering\ cost\ per\ cycle = K + cQ \tag{3}$$

$$Holding\ cost\ per\ cycle = \frac{hQ^2}{2a} \tag{4}$$

$$Total\ cost\ per\ cycle = K + cQ + \frac{hQ^2}{2a} \tag{5}$$

$$T = \frac{K + cQ + \frac{hQ^2}{2a}}{\frac{Q}{a}} = \frac{aK}{Q} + ac + \frac{hQ}{2} \tag{6}$$

$$\frac{dT}{dQ} := 0 = -\frac{aK}{Q^{*2}} + \frac{h}{2} \tag{7}$$

$$\therefore Q^* = \sqrt{\frac{2aK}{h}} \tag{8}$$

Equation 8 is a very well-known EOQ formula and the cycle time that corresponds to it is given by equation 9.

$$t^* = \frac{Q^*}{a} = \sqrt{\frac{2K}{ah}} \tag{9}$$

Now that the most basic deterministic continuous-review model was presented, it is possible to introduce the EOQ model with planned shortages, which considers OOSt situations. Simulating OOSt allow managers to verify if allowing for shortage costs is a better strategy than sustaining inventory holding costs. For this simulation, one needs new parameters such as the shortage cost per unit short per unit of time short p, the inventory level just after a batch of Q units is added to the inventory S, and the shortage in inventory just before a batch of Q units is added Q-S. Equations 10 to 20 allow for an understanding of how to get to optimal batch size Q*, optimal cycle length t*, and maximum allowed shortage $Q^* - S^*$.

$$Holding\ cost\ per\ cycle = \frac{hS^2}{2a} \tag{10}$$

$$Shortage\ cost\ per\ cycle = \frac{p(Q-S)}{2} \cdot \frac{Q-S}{a} = \frac{p(Q-S)^2}{2a} \tag{11}$$

$$Total\ cost\ per\ cycle = K + cQ + \frac{hS^2}{2a} + \frac{p(Q-S)^2}{2a} \tag{12}$$

$$T = \frac{K + cQ + \frac{hS^2}{2a} + \frac{p(Q-S)^2}{2a}}{\frac{Q}{a}} = \frac{aK}{Q} + ac + \frac{hS^2}{2Q} + \frac{p(Q-S)^2}{2Q} \tag{13}$$

$$\frac{\partial T}{\partial S} := 0 = \frac{hS}{Q} - \frac{p(Q-S)}{Q} \tag{14}$$

$$\frac{\partial T}{\partial Q} := 0 = -\frac{aK}{Q^2} - \frac{hS^2}{2Q^2} + \frac{p(Q-S)}{Q} - \frac{p(Q-S)^2}{2Q^2} \tag{15}$$

$$\therefore S^* = \sqrt{\frac{2aK}{h}} \cdot \sqrt{\frac{p}{p+h}} \tag{16}$$

$$\therefore Q^* = \sqrt{\frac{2aK}{h}} \cdot \sqrt{\frac{p+h}{p}} \tag{17}$$

$$t^* = \frac{Q^*}{a} = \sqrt{\frac{2K}{ah}} \cdot \sqrt{\frac{p+h}{p}} \tag{18}$$

$$Q^* - S^* = \sqrt{\frac{2aK}{p}} \cdot \sqrt{\frac{h}{p+h}} \tag{19}$$

$$\therefore \frac{\frac{S^*}{a}}{\frac{Q^*}{a}} = \frac{p}{p+h} \tag{20}$$

Figure 13 allow for a visual analysis of the behavior depicted by equations 10 to 20. Note that when p→∞ and h is kept constant (in other words, a situation in which shortage costs dominate inventory holding costs), Q*-S*→0. In a similar manner, when h→∞ and p is kept constant (a situation in which inventory holding costs dominate the shortage costs), S*→0. In this last case, it is uneconomical to keep any kind of inventory and Q* just plays the role of eliminating the demand shortages.

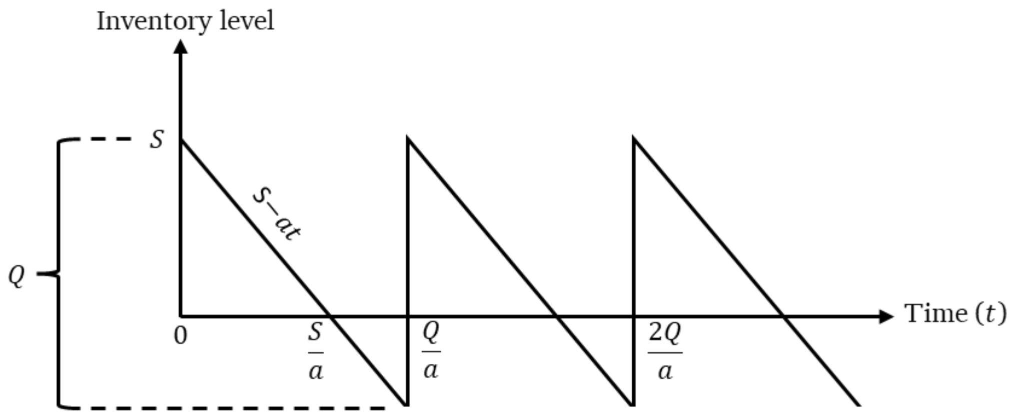


Figure 14: Diagram of inventory level as a function of time for the EOQ model with planned shortage (Hillier & Lieberman, 2015).

Finally, there is the EOQ model with quantity discounts. In this case it is assumed that the unit cost of an item depends on the different values that Q can assume. Moreover, an additional incentive is given to make large orders more advantageous. This is done by reducing the unit cost of large batches in comparison to the unit cost of small ones. After making all necessary adjustments, the results were be depicted in figure 14. In this figure, $T_{(i-1)}<T_i<T_{(i+1)}$ and

the total cost per unit for Q<Q_1 is T_(i+1), whereas the total cost per unit for Q_1≤Q≤Q_2 is T_i, and for Q>Q_2 is T_(i-1).

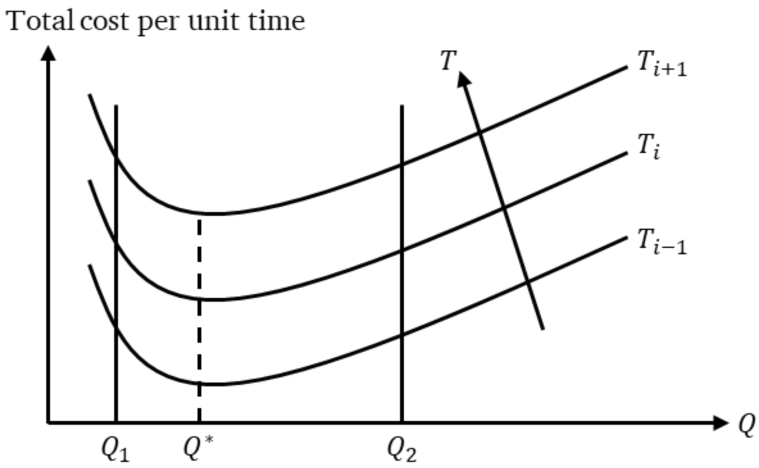


Figure 15: Total cost per unit as a function of lot size in the EOQ model with quantity discounts (Hillier & Lieberman, 2015).

Before proceeding to discuss the deterministic periodic-review models, this first introduction given by the deterministic continuous-review ones already provide enough background on the topic to allow for some further discussions regarding SIM in general. This way, a first important distinction to be made differentiates the concepts of dependent and independent demand. Independent demand occurs when the demand for a product does not depend on the demand of any other one. On the other hand, a dependent demand refers to a situation in which the demand for a product depends directly on the demand of another one. To better illustrate it with an example, the demand for cars could be considered to a certain extent independent (i.e., the demand for cars is not bound to the demand of any other product such as fridges or laptops); however, the demand for car components (e.g., steering wheel, tires, engine, and so on) are directly bound to the car demand (Hillier & Lieberman, 2015).

Understanding the concepts of dependent and independent demands is important to prepare the ground for what comes next: the ideas behind the material requirements planning (MRP) and JIT (Hillier & Lieberman, 2015). Briefly introducing these two notions is important to have a picture of just how complex the supply chain is before products reach the retail store and, by extent, the final customer.

MRP can be seen as a computer-based system that fulfills the purpose of planning, scheduling, and controlling (here also included the task of monitoring) the production of a final product. The first step taken by this system is to break down the final product into all its sub-assemblies (this procedure is sometimes referred to as “exploding” the final product and its result is the so-called bill of materials). It then organizes a production schedule by taking as inputs the demand for each component, as well as the lead times to receive/produce the missing components and to assemble the intermediary and final products. If more units of a certain component are lacking in the inventory system, the MRP automatically generates a purchase/production order.

Bringing a closure to the MRP discussion and framing the JIT concept now, Kang & Gershwin (2005) already mentioned in section 2.1.1 that lean environments are more vulnerable to IRI issues because they usually operate with lower inventory levels, shorter lead times, and smaller order quantities, all of which hit the core of the JIT inventory system. Before entering in more details about JIT itself, a brief historical background could be helpful to bring more clarity to this topic. The lean philosophy was first developed in Japan by Toyota in the late 1950s and was given much of the credit for the outstanding improvements in Japanese car automakers’ productivity through much of the late 20th century. One of the key pillars of the lean philosophy in turn is the JIT system, which places emphasis on reducing inventory levels to a bare minimum so that items are delivered only when needed, where needed, and in the exact quantity needed (Hillier & Lieberman, 2015). These features do not make a JIT inventory system incompatible to an EOQ model, since both of them are rather complementary to each other.

After this brief pause to discuss the concepts of independent and dependent products, MRP, and JIT, it is possible to give continuity to the SIM discussion and present the deterministic periodic-review model (Hillier & Lieberman, 2015).

In a general periodic-review model, the inventory manager has to plan for the upcoming n periods how many units ought to be produced/purchased, with the goal of efficiently replenishing the depleted stock level and minimizing the total cost over these n periods. Since the discussed model is a deterministic one, the demands r_i for each period are known beforehand.

A suitable algorithm for finding an optimal inventory policy could be developed to exploit the deterministic periodic-review model in its full potential. In this case, one could define C_i as being the total variable cost of an optimal policy for all the periods considered in the simulation, in the case where period i starts with zero inventory. By solving this problem using a dynamic programming approach, which analyzes the array of periods backwards, one first finds the values for C_n, C_(n-1), C_(n-2), and so on up until C_(i+1). After this, C_i is obtained out of the recursive relationship. Equation 21 shows how such problem can be formulated.

$$C_i = \min_{j=i,i+1,\ldots,n} \{C_{j+1} + K + h[r_{i+1} + 2r_{i+2} + 3r_{i+3} + \ldots + (j-1)r_j]\} \tag{21}$$

In equation 21, j is an index that denotes the moment at which the inventory level reaches zero for the first time since the beginning of period i. With this in mind, h is defined here as being the inventory holding cost during the interval of time between i and j. Hence, for j=n, C_(n+1)=0. Another property of equation 21 states that, since it is being solved for a minimal value of j, if the inventory level does go down to zero in period i, then the produced/purchased number of products in i must cover the entirety of demand up until period j.

With the discussion contained in the last three paragraphs, one can have an idea of how general deterministic periodic-review models work. This way, it is possible to move forward and analyze the dynamics of stochastic continuous-review models (Hillier & Lieberman, 2015). They are basically designed with the goal of analyzing inventory systems operating under conditions of relatively uncertain future demands, which in turn prevents any kind of mathematical modelling of the demand behavior from being developed and allows for the introduction of randomness to the simulation.

The traditional way of simulating continuous-review inventory systems was basically using a so-called two-bin system. In this case, all the items of a certain product were ought to be kept in one of two bins, A and B. The capacity of one of them, say bin A, would represent the already-discussed reorder point. The units would then first be removed from bin B and, logically, the emptying of this bin would trigger a replenishment order. During the lead time to refill bin B, products would be consumed from bin A.

As time went on, the two-bin approach was progressively replaced by computerized inventory systems, which were already referred to in the current work by the name of IMS. In this case, inventory replenishment and product sales are recorded electronically, so that ideally the current inventory levels are always displayed in the computer. The evidence presented in section 2.1.1 show that, in the real world, sustaining a perfect match between actual and recorded inventories is not an easy task and IRIs might emerge.

Due to the widespread use of computers in the contemporary inventory management, continuous-review inventory systems for products that are deemed important enough to warrant a formal inventory policy. Not only this, but simulation approaches to the retail industry such as the one of Chuang & Oliva (2015) were also allowed to resort to continuous-review methods, like the system dynamics convention, with no significant loss of accuracy in comparison to the real world.

The simplest continuous-review inventory system for a particular simulated item has two fundamental pillars: the reorder point R and the order quantity Q. The idea behind it is likewise simple: if the stock level reaches R, a replenishment order of Q must be placed. This policy is referred to as "reorder-point policy", "order-quantity policy", of just "(Q,R) policy". Kang & Gershwin (2005) happen to use such a (Q,R) policy in their paper when studying the impacts of IRI on a typical retail store.

Although the (Q,R) policy in the continuous-review context is extremely simple, there are ten important assumptions behind it that must be taken into consideration:

1. The first assumption is rather very intuitive and states that each application of the (Q,R) policy involves one, and only one, product. Although this may seem a simple statement at first, it can prove to be a rather complicated constraint for retailers in the real world, since some batches that are shipped from suppliers contain several product types.
2. The inventory level is continuously reviewed, which implies that its current level is always known.
3. In such a (Q,R) policy, the only controllable parameters that managers are able to influence are the reorder point R and the order quantity Q.
4. In this model, a lead time between the order placement and receipt must be taken into account. It is also worth noting here that this lead time can be either fixed or subjected to random variability.
5. The demand for the simulated product during the lead time mentioned in "4" is uncertain. However, its probability distribution can be taken as known. This allows for an estimation of the demand during the lead time, but not for a determination of its exact value.
6. If, for any reason, a stockout situation arises before the replenishment products arrive in the store, two outcomes can follow. In the case where clients wait for the new products instead of leaving the store, the "excess demand" should be backlogged and served once the

resupply arrives. In the alternative scenario in which clients give up on the purchase decision, the "excess demand" is simply lost.

7. The inventory manager incurs a fixed setup cost K whenever a replenishment order is placed.
8. In addition to the setup cost K mentioned in "7", the cost of the order is proportional to the order quantity Q.
9. A holding cost h must be taken into account for each unit kept in the inventory during the simulated time.
10. A shortage cost p in incurred by the store whenever stockout situations happen. This cost is attributed to each item backordered per unit of time.

When choosing the order quantity Q in the (Q,R) policy, the most straightforward approach is to adopt as reference the equation 17, first introduced in the EOQ model with planned shortages. From this equation, one can come up with equation 22.

$$Q = \sqrt{\frac{2AK}{h}} \cdot \sqrt{\frac{p+h}{p}} \tag{22}$$

Note that when comparing equations 17 to 22, some adjustments were made. First of all, the consumption rate a was replaced by the average demand per unit A. The remaining parameters of the new equation were already defined in assumptions 7, 9, and 10.

Moving on to determine the reorder point R, it is important first to identify the management's desired service level for the simulated product. There are in turn five most usual ways to define service level:

1. The first definition for service level of a product understands it as the probability that a OOSt situation will not occur in the time between the order placement and the order receipt.
2. The second definition sees it as the average number of stockouts per period (e.g., per month or per year).
3. An alternative third approach to service level sets it as the average percentage of, for example, annual demand that can be satisfied immediately (i.e., without incurring in a OOSt situation).

4. Next, service level can also be understood as the average delay in filling backorders when a stockout happens.
5. Finally, the last definition for service level sees it as the delay in filling orders for a certain product. In this case, when there is no OOSt, the delay is considered as being equal to zero.

The first definition of service level is perhaps the most convenient one and, therefore, was chosen to be discussed in more detail. For this purpose, L was denoted as being the management's desired probability that a stockout will not occur between the time a replenishment order is placed and later on received. Additionally, a parameter D was also created and defined as the demand during the lead time in filling the order.

For instance, in the case where D uniformly distributed in a time interval [t_1,t_2], the reorder point would be given by:

$$R = t_1 + L(t_2 - t_1) \tag{23}$$

Where:

$$P(D \leq R) = L \tag{24}$$

Since the mean of a uniform distribution is given by:

$$E(D) = \frac{t_2 + t_1}{2} \tag{25}$$

Here, the safety stock, understood as the inventory level just before the order quantity is received, can be obtained through equation 26.

$$Safety\ stock = R - E(D) = t_1 + L(t_2 - t_1) - \frac{t_2 + t_1}{2} = (L - \frac{1}{2})(t_2 - t_1) \tag{26}$$

Now that the procedure to obtain R and the safety stock for a uniform distribution was explained, similar approaches can be adopted for other distributions following the general step-by-step:

1. Choose L
2. Solve the problem for R in a way that $P(D \leq R) = L$

The exposition of these two steps finishes the discussion of stochastic continuous-review models. This way, it is possible now to present the stochastic periodic-review models and, once it is done, to put an end to the analysis of SIM and move on to section 2.2.2.

According to Hillier & Lieberman (2015), the major change in these new models is that the inventory level is only periodically assessed. In other words, only at the end of each period the current inventory level is determined, and a replenishment decision is made.

To better understand this category of models, it is perhaps better to start with the simple stochastic two-period model with no setup cost and then gradually move on to the more complex and general stochastic multi-period model with setup cost. The defining difference between modelling a system that plans ahead only one period at a time and one that takes into consideration multiple periods into the future is the capability of reliably forecasting the demand for the simulated product. With this in mind, the assumptions for the stochastic two-period model can be presented:

1. Just like when the (Q,R) policy was presented, it is important to make it clear that the equations here are valid for a modelling a single stable product.
2. The planning horizon totals two periods, in which unsatisfied demand in period 1 is backlogged to be considered in period 2. The same does not happen in period 2 since it is the last modelled period.
3. The demands $D_1$ and $D_2$ for periods 1 and 2 are considered as being independent and identically distributed random variables. That is, they have the same probability distribution with a probability density $\phi_D(\xi)$ and a cumulative distribution function $\Phi(\xi)$.
4. The starting inventory level at the beginning of period 1 is represented by $x_1$ with ($x_1 \geq 0$).
5. The optimal inventory levels to be reached by replenishment processes at the beginning of periods 1 and 2 are represented by $y_1$ and $y_2$, respectively.
6. Finally, the overall goal is to minimize the expected total cost for both periods, considering the purchasing/producing unit cost c, the holding cost per unit remaining in the inventory by the end of each period h, and the shortage cost per unit of unsatisfied demand by the end of each period p.

Now that these six assumptions were presented, the analysis of the problem can be developed. For this purpose, y_i^0 is defined as the optimal value of y_i (i=1,2), C_1 (x_1) is the expected total cost for both periods when applying an optimal policy, and C_2 (x_2) is the expected total cost for period 2 specifically, with x_2 being the inventory level before replenishment at the beginning of period 2.

The procedure goes as follows: first, a solution is found for C_2 (x_2) and y_2^0, where there is only one period to analyze. After that the encountered results are used to find C_1 (x_1) and, in turn, y_1^0. Hence, y_2^0 is found by solving equation 27.

$$\Phi(y_2^0) = \frac{p-c}{p+h} \tag{27}$$

By the definition of x_2, the optimal policy for equation 27 is:

$$\begin{cases} x_2 < y_2^0 & \mathit{order}\ y_2^0 - x_2\ \mathit{to\ bring\ the\ inventory\ level\ to}\ y_2^0 \\ x_2 \geq y_2^0 & \mathit{do\ not\ order\ anything} \end{cases} \tag{28}$$

The cost associated with the policy of equation 28 is given by equation 29:

$$C_2(x_2) = \begin{cases} L(x_2) & if\ x_2 \geq y_2^0 \\ c(y_2^0 - x_2) + L(y_2^0) & if\ x_2 < y_2^0 \end{cases} \tag{29}$$

With L(z) being the expected shortage plus the holding cost for a single period of inventory level z, as depicted in equation 30.

$$L(z) = \int_z^\infty p(\xi - z)\varphi_D(\xi)d\xi + \int_0^z h(z-\xi)\varphi_D(\xi)d\xi \tag{30}$$

Finally, when both periods are considered, the optimal cost for this two-period policy is portraited in equation 31.

$$C_1(x_1) = \min_{y_1 \geq x_1} \{c(y_1 - x_1) + L(y_1) + E[C_2(x_2)]\} \tag{31}$$

Where:

$$E[C_2(x_2)] = \int_z^{\infty} C_2(y_1 - \xi)\varphi_D(\xi)d\xi \tag{32}$$

At last, it is possible to demonstrate that $C_1(x_1)$ has a unique minimum and that the optimal value of $y_1$, denoted by $y_1^0$, is given by equation 33. Here it is important to note that the exact solution of this equation depends on the probability distribution adopted (e.g., normal distribution, uniform distribution, exponential distribution, and so on).

$$E[C_2(x_2)] = \int_z^{\infty} C_2(y_1 - \xi)\varphi_D(\xi)d\xi \tag{33}$$

Which leads to the optimal policy for period 1 depicted in equation 34:

$$(p+h)\Phi(y_1^0) + (c-p)\Phi(y_1^0 - y_2^0) + (p+h)\int_0^{y_1^0 - y_2^0} \Phi(y_1^0 - \xi)\varphi_D(\xi)d\xi = 0 \tag{34}$$

Equation 34 in turn result in the optimal policy for period 1, stated in equation 35 in a similar manner to equation 28.

$$\begin{cases} x_1 < y_1^0 & order\ y_1^0 - x_1\ to\ bring\ the\ inventory\ level\ to\ y_1^0 \\ x_1 \geq y_1^0 & do\ not\ order\ anything \end{cases} \tag{35}$$

Now that the simplest stochastic periodic-review model was presented, it is possible to discuss the more complex cases, such as the multiperiod model with no setup cost and the multiperiod model with setup cost. Before doing so, however, it is worth mentioning that the adaptations made to equations 27 to 35 are minor ones and very few additions are actually made, so much so that the assumptions remain the same.

Starting then with the multiperiod model with no setup cost, the main difference here in comparison to the last presented model is that a total of n periods, with $n>2$, are considered. In addition to it, a discount factor $\alpha$ is introduced to calculate the expected total cost for the n periods, with $0<\alpha<1$. The first similarity to point out is that the problem still remains to find the $y_1^0, y_2^0, \ldots, y_n^0$ that best describe the optimal inventory policy.

Similarly to equations 28 and 35, for each period i (i=1,2,…,n), that has $x_i$ as the inventory level before replenishment, the optimal policy is given by equation 36. Note that, in this case, $y_n^0 \leq y_{(n-1)}^0 \leq \ldots \leq y_1^0$.

$$\begin{cases} x_i < y_i^0 & \mathit{order}\ y_i^0 - x_i\ \mathit{to\ bring\ the\ inventory\ level\ to}\ y_i^0 \\ x_i \geq y_i^0 & \mathit{do\ not\ order\ anything} \end{cases} \tag{36}$$

For the particular case when n→∞, $y_n^0 = y_{(n-1)}^0 = \ldots = y_1^0 := y^0 = \text{constant}$. The solution is then obtained from equation 37.

$$\Phi(y^0) = \frac{p - c(1-\alpha)}{p+h} \tag{37}$$

The last variation of the stochastic periodic-review model is the multiperiod model with setup cost. Here, K will be needed to analyze the situation, which as already defined before represents the fixed setup costs incurred when ordering replenishment items. In addition to it the introduction of a (s,S) policy is needed, where s indicates when to order (i.e., if the inventory level drops below a s point, a replenishment order is placed) and S indicates how much to order (i.e., how many new items are needed to bring the inventory level up to S again).

Since many periods are being considered now, it is necessary to make adjustments and define $s_i$ and $S_i$ for every period i. This way, the optimal policy could be represented by equation 38.

$$\begin{cases} x_i < s_i & \mathit{order}\ S_i - x_i\ \mathit{to\ bring\ the\ inventory\ level\ to}\ S_i \\ x_i \geq s_i & \mathit{do\ not\ order\ anything} \end{cases} \tag{38}$$

Equation 38 concludes section 2.2.1 that discusses different inventory management policies that could be used in the simulation developed in the current work and presented in section 3.

### 2.2.2 Simulation: in-depth approach and useful takeaways

Section 2.2.1 presented a fundamental piece of this thesis by expanding the knowledge on SIM. The inventory replenishment models exposed by Hillier & Lieberman (2015) were thoroughly considered when developing the simulation depicted in section 3. However, they did not provide enough background to properly answer the research questions of the current work.

Therefore, this section presents an in-depth investigation of simulation as a method to solve problems.

Starting then with a definition of simulation. According to Banks et al. (1996), a simulation could be interpreted as an imitation of a certain real-world process or system over a certain period of time. Interestingly, a simulation is not necessarily done through a computer. It could be carried out, for example, by hand. Of course, computer-assisted simulations are far more capable than their counterparts done by hand and, therefore, the former has way more practical applications than the latter. Another, perhaps more poetic, way to understand a simulation is by seeing it as a generation of an artificial history of a target system, in addition to the observation of this artificial history to draw conclusions and inferences regarding the operating characteristics of the real-world system.

The evolution of the behavior of the mentioned system through time is studied by means of a so-called simulation model. Such a model is created based on a set of assumptions about the real-world system. These assumptions in turn are the key building blocks that define the mathematical, logical, symbolic, and behavioral relationships between the entities of the simulation, which correspond to the objects of interest of the real-world system represented in the simulation.

The logical question to be asked now is "when is simulation an appropriate tool for answering a, for example, research question such as the ones presented in section 1.3 of this thesis?". The answer for this question is also provided and expanded by Banks et al. (1996). According to the authors, there are seven main scenarios in which a simulation could be a suitable choice to approach a problem.

1. Whenever the researchers want to carry out experimentations or simply study the internal interactions of a system that is considered to be complex in the real world, simulation could be an option.

2. Informal, organizational, and environmental changes in the real world could also be simulated with reason to believe that the conclusions obtained by the researchers would be valid in the real world. The effects of these changes could be just as well simulated if they were also part of the scope.

3. It is also important to observe that the knowledge gained when designing a simulation model has the potential to significantly improve the quality of suggestions when the aim is to improve a real-world system's performance.

4. Next, by changing a simulation's inputs and observing how it influences the outputs, very important insights can be acquired about the real-world system's behavior under similar circumstances. This situation depicts exactly why a simulation approach was chosen to investigate the relationship between positive IRI and a negative impact on sales in retail stores.

5. When it comes to instruction assistance, simulations can and are used as pedagogical de-vices to, for instance, reinforce analytic solution methodologies.

6. Not-yet developed system could also make use of simulations. In other words, simulation models could be used with new designs (e.g., products, services, or processes) or policies (e.g., how citizens would react to a new monetary policy form a country's central bank) be-fore their implementation in other to make adjustments in the planning phase and not in the execution phase (or even in the aftermarket phase, through mitigating or corrective actions).

7. Finally, but not less important the previous six purposes, simulation could be used to simp-ly verify analytic solutions.

Now that these main usages of simulation models were presented it is also worth noting that nowadays there is a wide range of special-purpose simulation languages and software with massive computing capabilities. For example, the current thesis was developed in a computer specially designated to carry out simulations at the TU Darmstadt.

Like any other tool, however, simulation has its advantages and disadvantages. An example of a commonly perceived advantage is that simulations are intuitively appealing to clients due to their capability of mimicking what happens (or better said, could happen) in real-world systems. On the other hand, a limitation of simulations when compared to, for example, optimization models lies on the fact that the former are simply "run', whereas the latter are actually solved and give the researchers a clearly stated recommendation in their solutions. To better comprehend these pros and cons of simulation it is better to list its attributes and limitations.

Advantages of simulation models:

1. The first advantage to be mentioned is the ability to explore new policies, operating procedures, decision rules, information flows, organizational procedures, and so on without completely disrupting ongoing operations (e.g., in a factory) of the real-world system.

2. Next, one can say that testing new hardware designs, physical layouts, transportation systems, and so on without committing precious and scarce resources in doing so is a great advantage of simulation models.

3. Allowing for hypothesis testing is also widely known advantage of simulating systems. An example could be seen once again in the current work. As stated in section 1.3, the re-search questions led to the pre-simulation hypothesis that positive IRI, just like negative IRI leads to decrease in sales. This hypothesis, as presented in section 4, was later validated via a simulation model of a retail store developed in this thesis.

4. The time lapse tool enjoyed by computer-assisted simulations is a great advantage of simulations. In other words, the ability to expand or compress time by allowing for "speed ups" and "slowdowns" of the events being simulated is of incredible usage for researchers.

5. Taking the example of carrying out a DOE (explained in more detail in section 2.2.3), valuable information can be obtained from studying the interaction of variables, since the effect of combined variables not necessarily represents the sum of the individual effects of each variable.

6. More insights could also be obtained regarding how exactly the simulated variables influence the performance of the simulated systems and, consequently, the real-world one as well.

7. Going beyond the assistance in insight development, simulations can also carry out bottle-neck analysis, indicating with high accuracy where work in progress, information, materials, and so on are being delayed more than tolerable (e.g., in a production plant).

8. A simulation study could help in eliminating biases and exposing how the real-world system operates instead of how stakeholders think it operates. In other words, it reduces the human factor in conclusion-drawing processes.

9. Finally, the nineth remarkable trait of simulations is their capability of answering the famous "what if" questions. This ability to simulate different scenarios by, for example, carrying out a sensitivity analyses is a particularly powerful tool when design new systems.

Opposing these advantages just mentioned, there are four well-known disadvantages associated with simulations as well:

1. Perhaps the most intuitive disadvantage of simulations is that they require special training for the developers in order to properly simulate the target system. Simulation building takes

relatively long time to master in comparison to other comparable competences and its quality depends a lot on the experience of the developer. Moreover, two models with the same purpose but developed by two different people are highly unlikely to be the same.

2. The second disadvantage on the list is the fact that simulations can oftentimes be very difficult to interpret depending on the complexity of the framed problem, the capability of the developers, and the simulation approach adopted. For instance, since most simulation outputs are random variables, in other words, since most simulation models are stochastic, then it can be quite hard to determine if an observation is due to randomness or directed interactions between entities.

3. As a third factor to be considered, simulation modeling can be verry time-consuming and, therefore, expensive. However, limiting the allocation of the adequate amount of resources into the development of a simulation may result in a model that is not suitable or simply not enough to answer the questions it was meant to answer.

4. As a last factor, sometimes simulation is chosen as the tool of choice to solve problems that rather require analytical solutions, or that should preferably be solved by analytical means. For instance, this situation can be seen when a simulation approach is chosen to simulate waiting lines where closed-form queueing models are already available.

Fortunately, since these disadvantages are quite well-known, some countermeasures against them are also available. For example, against the first mentioned disadvantage, many companies that develop simulation software have created packages that contain models that only need input data for their operation, gaining the name of "simulators" or "templates". Such is the case of Plant Simulation, developed by Siemens, and AnyLogic, developed by The AnyLogic Company.

Since the developed simulation presented in section 3 was a stochastic one, the proper management of random variables proved to be relevant. To this end, further literature providing guidelines to handling randomness in simulations was collected. The reference material for this topic was Law's (2007) book "Simulation Modeling & Analysis", one of the most referenced literatures when it comes to simulation.

Law (2007) state that any system or process in the real world that in one way or another contains inherent random components requires a method of somehow generating or, in other cases, obtaining random numbers. A good way to do it is to make use of the interval [0,1] of a uniform distribution. Such distribution will be denoted U(0,1). Before continuing with the development of U(0,1), it is important to mention that a similar process to handle randomness was also used in the simulation depicted in chapter 3.

The first concept to be introduced here is the idea of using an arithmetic generator to obtain random numbers. This way, there are some properties of "good" arithmetic random-number generators that must be taken into account.

1. First of and above all, the numbers generated by this method should appear, at first glance, to be uniformly distributed in the interval [0;1] and should likewise not exhibit any correlation with each other (e.g., trends are not desired). If this requirement is not fulfilled, the results of the simulation may incur in biases and other errors.

2. When we consider the computational memory allocated for the simulation, it is important (but in most cases not critical) to avoid the need for a lot of storage for random numbers and to prioritize speed.

3. It is important that the random numbers are generated in a way that a stream or array of them can be exactly replicated. In other words, if the same seed is used, the same random numbers should be obtained. This property is important for debugging the simulation and to make more precise comparisons.

4. The generator of random numbers should be developed in such a way that it is easy to separate different streams of random numbers. For this purpose, a "stream of random numbers" can be simply understood as a subsegment of all the numbers produced by the generator. In this sense, one stream begins where the previous ends and it is possible then to dedicate particular streams to particular sources of randomness in the simulation.

5. A good generator of random numbers should be portable. This means that it should be able to produce the same sequence of random numbers in every compiler and computer.

In practice, the most common generators used in simulations are actually quite fast and require very little storage. They also can relatively easily reproduce a given sequence of random numbers, which fulfill the requirements 2 and 3 just presented. Moreover, most generators, especially in simulation software like Plan Simulation and AnyLogic (both already referenced in this the-sis), have the facility for multiple streams at least to a certain extent, which satisfies point 4. One criterium that was still not quite often satisfied at the time the book was published was point 1. This is somehow critical because it hinders the acquirement of correct simulation results.

Now that the general and fundamental idea behind simulation and random numbers was presented, it is possible to move on and discuss the methodology behind the second pillar of this the-sis, the Design of Experiments.

### 2.2.3 Design of Experiments: in-depth approach and key takeaways

Now that the general and fundamental idea behind simulation and random numbers was presented, it is possible to move on and discuss the methodology behind the second pillar of this the-sis, the Design of Experiments.

Section 2.2.3 will occupy itself with the provision of enough theoretical background on DOE to draw the necessary conclusion to the problem stated in section 1.3. To this purpose, only a brief introduction in the vast field of Experimental Design will be carried out. The top reference literature to this section was the book "Design and Analysis of Experiments" from Montgomery (2009).

According to Montgomery (2009), an experiment is nothing more than a test to discover something about a process or a system. For instance, the purpose of adopting DOE in this thesis was to try to determine which factors had the most negative impact on sales among those that generated positive IRI. This way, an experiment could be defined as follows:

"An experiment is a test or series of tests in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in the output response." (Montgomery, 2009)

As mentioned in the previous paragraph, in most cases the goal of experiments is to analyze the performance of processes and systems. Such process or system could be represented like in the well-known diagram of figure 15.

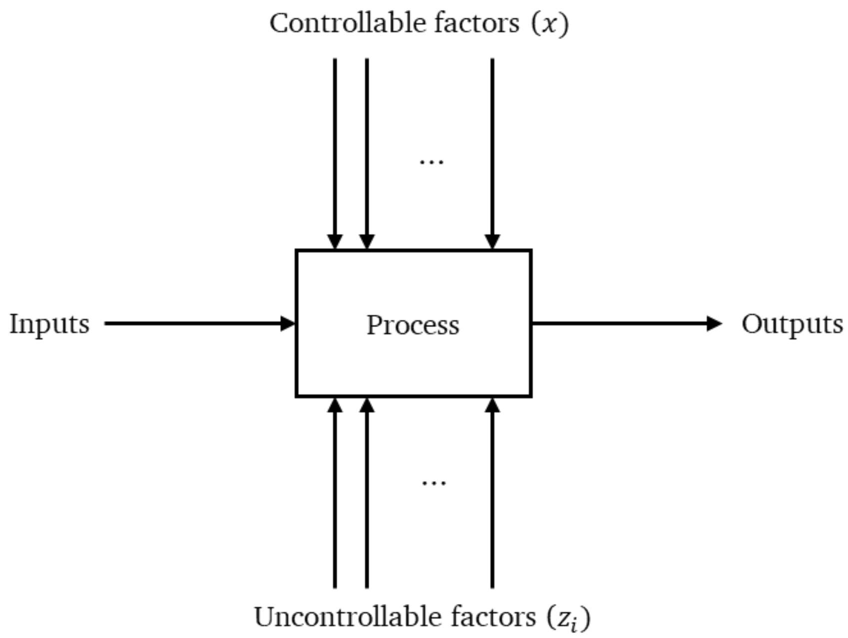


Figure 16: Diagram of relevant factors to a process (Montgomery, 2009).

As shown in figure 15, a process could be seen as a combination of operations, machines, methods, people, and some other resources that serve the purpose of transforming the inputs (e.g., raw materials, information, work in progress) into an output that has one or more observable response variables. It is also possible to see that some variables that influence the process are controllable whereas other ones are uncontrollable (from the perspective of the researchers).

Now that a general idea of what constitutes an experiment was presented, it is possible to list some of its most common goals:

1. Topping the list is the goal of determining which of the input variables are the most influential ones on the response variables. This also happens to the goal of the DOE of the current work, explained in more details in section 3.
2. Another common goal is to try to control the response variable in a way that it gets closer to a desired nominal value. This mentioned controlling process takes place through the adjustment of the input variables.
3. Moving on, researchers could also try to reduce the variability of the response variable by controlling the input variables.

4. Finally, the last commonly set goal is to set up the controllable variables in a way that the undesired effects of the uncontrollable variables are minimized.

With this background, the next step is to define the concept of strategy of experimentation, which basically corresponds to the general approach to planning and conducting the experiment. From now up until the end of this section, some most commonly adopted strategies will be presented and, in section 3, the chosen strategy to this thesis will be presented in closer detail.

Starting the list of strategy is the one called “best-guess approach”. To better illustrate this strategy, a practical example is presented. If one was to analyze what factors influence the performance of a golf player, the following parameters could be raised:

1. The type of driver used (in this sport, it could be, for instance, an oversized or a regular sized driver).
2. The kind of ball (balata ball or three-piece-ball).
3. The wear and tear on the player measured by walking and carrying his own golf equipment or riding in a golf cart.
4. The impact of drinking water to hydrate himself or beer instead.
5. The environmental conditions such as visibility measured by playing in the morning time or in the afternoon.
6. The influence of the weather, captured by playing when it is hot and playing when it is cool.
7. The type of golf shoes, either spike (metal) or worn (soft), could also influence the player’s performance.
8. Finally, the wind speed and direction are great influencing factors in this sport and, therefore, playing on a windy day or calm day is also a valid experiment.

After raising the these most influencing factors, the researchers decide, out of expertise, that factors 5, 6, 7, and 8 have a small effect in comparison to the first four and, because of that, could be removed from the experimentation phase. These types of qualitative screening decisions are very common among engineers and researchers. Then, according to the best-guess approach, the researchers carrying out the experiment should now, under consideration of resource constraints (e.g., maximum amount of rounds that can be played), arbitrarily choose the combination of factors that should be analyzed and check the results of these experiments. At first glance, this approach might seem flawed and lacking framework background, but Montgomery (2009) believes it works well mainly due to the technical and theoretical background of the research team involved in the project. In other words, a key requirement when choosing the best-guess strategy of experimentation is to assure that the people involved have enough expertise in the topic.

Another worth-mentioning approach is the “one-factor-at-a-time”, or simply OFAT. This method starts by determining a so-called baseline set of levels or, in other words, a starting point for each analyzed factor. After it, the next step is to successively vary one factor at a time in order to see how each of them influence the response variable. From this procedure it is possible already to determine the perhaps biggest disadvantage of this method: it does not consider the interactions between factors. However, a good advantage of the OFAT approach Is that it is of straight forward interpretation of results, requires less experiments than the other methods that will be presented next, and it already allows for a first selection of an optimal combination of factors when the researchers want the response variable to present, for instance, its maximal (or minimal) possible value.

The next approach tackles exactly the OFAT deficiency of not considering the interactions between factors. It is based on a statistical approach to design the experiment and is often referred to as factorial experiment. In this experimental strategy, the selected factors are varied together. This factorial approach can be further broken down into, to mention just the most important ones, a full-factorial or a fractional factorial experiment. A full-factorial experiment analyzes all the possible combinations of the different levels of the factors, whereas the fractional factorial experiments only analyze a fraction of the total due to resource constraints.

Since the best-guess and OFAT approaches were not adopted in this thesis, they will not be covered in more details in order to give more space to the more relevant approaches: the factorial designs. To discuss them in detail, however, the basic principles of DOE must be first presented (Montgomery, 2009).

The three basic principles of DOE are randomization, replication, and blocking. In deeper understanding, randomization is the most important principle to be followed and it basically states that every experimentation must be carried out in a random order since the statistical methods involved require that the observations (or errors) are independent from one another. Another desired property of randomizing the observations is that the effects of external uncontrollable factors is, in a way, “averaged out”. This means that their influence is more or less equally distributed in the experiment, minimizing the chances of reaching out biased conclusions. Sometimes experimenters encounter difficulties in randomizing experiments due to external complexities (e.g., setting up a production line with a certain combination of factors). There are methods to counter this problem, but they will not be covered here because the simulation presented in section 3 showed no barriers to completely randomize the experiments.

The second principle, replication, is translated into the independent repeat of each factor combination. It has two major properties that are decisive for the DOE. First, replication enables the researcher to have a good understanding of the experimental error, which in turn becomes the basic unit of measurement for verifying if the collected data are indeed statistically different from one another. The second worth-mentioning property states that if the sample mean $\bar{y}$ is applied to estimate the true mean response for one of the factor levels in the experiment, then the replication would allow to a more precise estimation of this particular parameter. To understand it better, let us observe equation 39. A simple analysis of equation 39 already allows for the conclusion that the variance of the sample $\sigma_y^2$ decreases as the number of replications n increases.

$$\sigma_y^2 = \frac{\sigma^2}{n} \tag{39}$$

The third and last basic principle surrounding the DOE is the blocking principle. Blocking itself is a design technique that is often used to improve the precision with which comparisons are made.

Now that these fundamental principles were presented, the next procedure is to structure the guidelines for correctly designing experiments (Montgomery, 2009). There are basically seven guiding principles to carry out a good experiment, which are presented in table 4.

The first principle, named recognition of and statement of the problem, is by no means a simple task. Here, it is absolutely necessary that all ideas about the objectives of the experiment are stated. For instance, in some cases the inputs for the experiment involve shutting down the production line of a product or requiring information that is hard to collect for a certain department of a company and this requires a lot of resources and, therefore, is very expensive. A clearly stated objective helps developing a better structured plan and, this way, avoiding costly rework later in the project.

Table 4: Guidelines for designing an experiment (Montgomery, 2009).

| **Guidelines for Designing an Experiment** |
| --- |
| 1. Recognition of and statement of the problem |
| 2. Selection of the response variable |
| 3. Choice of factors, levels, and ranges |
| 4. Choice of experimental design |
| 5. Performing the experiment |
| 6. Statistical analysis of the data |
| 7. Conclusions and recommendations |

Next in table 4, there is the selection of the response variable. Once the problem is clearly stated, it is possible to define the response variable (or variables) upon which the controllable factors will exercise an effect, which in turn will enable the researchers to draw important conclusions.

The choice of factors, levels, and range is then required to proceed with the experiment. The factors that the experimenters judge to be relevant for the experiment because they have a potentially big influence on the response variable are called design factors. These design factors are the ones that the researchers will alter to verify how this alteration influences the response variable. Apart from this category of factors, there are also the held-constant factors and the allowed-to-vary factors, both equally important for the experiment. Held-constant factors are variables that could undesirably influence the value of the response variable and, therefore, are kept constant. On the other hand, allowed-to-vary factors are oftentimes the uncontrollable variables that are countered by randomizing the experiment. A good way to determine the factors that may influence the experiment is to use a Ishikawa diagram, like the one represented in figure 16. The identification of levels and ranges is a more complex process, and it relies a lot on the process knowledge of the experimenters.

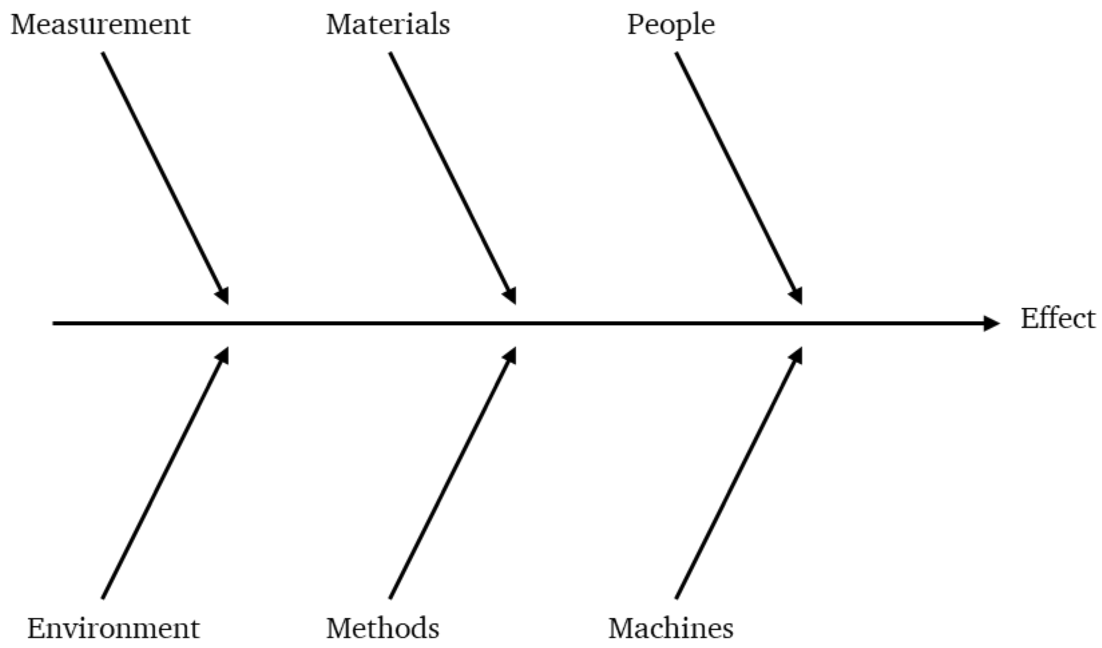


Figure 17: Diagram of relevant factors to a process (Montgomery, 2009).

Once the factors, levels, and ranges were determined, it is possible to choose the experimental design. To take some of the already mentioned ones, the experimenters should decide between, for example, carrying out an OFAT, full factorial experiment, fractional factorial experiment, or even a best-guess approach.

With the experimental design chosen, it is then time to perform it. This execution phase is just as crucial as the planning phase, since any errors in carrying out the experiment could potentially destroy it and result in immense costs for the sponsors of the project.

After finishing the experiment, it is necessary to statistically interpret its results. The statistical trait of this analysis allows for the conclusions to be objective instead of judgmental. In this phase, it is important to once again verify the quality of the collected data and to pay attention to correctly carry out the statistical methods, since they are by no means trivial and, therefore, error prone. Graphical methods here are widely applied and in this thesis, it was no different, as better illustrated in section 4, which presents the results of the current work.

Finally, there are the conclusions and recommendations. This last step summarizes all the other steps and is the value generating deliverable for the client. If possible, a good practice here is to conduct follow-up runs and confirmation tests to double-check if the conclusions are correct.

The explanation of the guidelines for correctly designing experiments puts an end to all introduction needed regarding the DOE, allowing to move on and expose the literature behind the 2^k factorial design, one of the underlying methodologies supporting the conclusions of this thesis. Here, the key reference was chapter 6 of the same book introduced at the beginning of this section (Montgomery, 2009).

The 2^k design is very useful in the early stages of an experimental work where a great number of factors are probably going to be analyzed. The first trait that makes such a design so important is the fact that it provides the smallest amount of runs with which k factors can be studied in a complete factorial design. For this reason, the 2^k designs are very useful in factor screening experiments. One of their biggest downsides lies on the required assumption that the response is approximately linear over the range of the factor levels that were chosen, which sometimes is not the case.

An interesting but not so useful property of the 2^k designs is that they contain 2^k-1 effects plus interactions of effects. Take for instance a 2^3 factorial design. As one can see from table 5, it contains 7 effects in total, namely the effects from the main factors A, B, and C, and the ones from the interactions AB, AC, BC, and ABC. Important to note in table 5 as well is that the lowercase letters a, b, ab, and so on represent the total of the response observation at all n replicates of each treatment combination.

Table 5: Design matrix for a 2^3 factorial design (Montgomery, 2009).

| **Run** | **A** | **B** | **C** | **Labels** | **A** | **B** | **C** |
|---|---|---|---|---|---|---|---|
| 1 | - | - | - | (1) | 0 | 0 | 0 |
| 2 | + | - | - | a | 1 | 0 | 0 |
| 3 | - | + | - | b | 0 | 1 | 0 |
| 4 | + | + | - | ab | 1 | 1 | 0 |
| 5 | - | - | + | c | 0 | 0 | 1 |
| 6 | + | - | + | ac | 1 | 0 | 1 |
| 7 | - | + | + | bc | 0 | 1 | 1 |
| 8 | + | + | + | abc | 1 | 1 | 1 |

After establishing the design matrix as presented in table 5, the next step is to carry out the analysis of variance (ANOVA). The summarized view of the ANOVA is presented in table 6, where, once again, k is the number of factors and n is the number of times each run was replicated.

Table 6: ANOVA for a 2^k factorial design (Montgomery, 2009).

| Source of variation | Sum of Squares | Degrees of Freedom |
|---|---|---|
| $k$ main effects | | |
| $A$ | $SS_A$ | 1 |
| $B$ | $SS_B$ | 1 |
| $\cdots$ | $\cdots$ | $\cdots$ |
| $K$ | $SS_K$ | 1 |
| $\binom{k}{2}$ two-factor interactions | | |
| $AB$ | $SS_{AB}$ | 1 |
| $AC$ | $SS_{AC}$ | 1 |
| $\cdots$ | ... | $\cdots$ |
| $JK$ | $SS_{JK}$ | 1 |
| $\binom{k}{3}$ two-factor interactions | | |
| $ABC$ | $SS_{ABC}$ | 1 |
| $ABD$ | $SS_{ABD}$ | 1 |
| $\cdots$ | $\cdots$ | $\cdots$ |
| $IJK$ | $SS_{IJK}$ | 1 |
| $\binom{k}{k}$ two-factor interactions | | |
| $ABC \ldots K$ | $SS_{ABC\ldots K}$ | 1 |
| Error | $SS_E$ | $2^k(n-1)$ |
| Total | $SS_T$ | $n \cdot 2^k - 1$ |

To obtain the sum of squares of each effect and interactions of effects, one first has to calculate their contrasts. For this purpose, equation 40 shows how to calculate the contrast of interaction AB…K and a similar procedure is carried out to calculate all other contrasts.

$$Contrast_{AB\ldots K} = (a \pm 1)(b \pm 1) \ldots (k \pm 1) \quad (40)$$

Since equation 40 is very generic, two examples will be provided for the contrast of AB in a 2^3 factorial design and for the contrast of ABCD in a 2^5 factorial design, respectively.

$$Contrast_{AB} = (a-1)(b-1)(c+1) \tag{41}$$
$$Contrast_{ABCD} = (a-1)(b-1)(c-1)(d-1)(e+1) \tag{42}$$

Finally, the effect of a generic factor or interaction of factors is given by equation 43.

$$Effect_{AB\ldots K} = \frac{2}{n \cdot 2^k}(Contrast_{ab\ldots k}) \tag{43}$$

If one pays close attention to equations 40, 41, 42 and 43, it is possible to derive a simplified way of calculating the contrasts of each effect and interaction of effects. A generic formulation for this simplified approach is given in equation 44.

$$Effect_{AB\ldots K} = 2 \cdot \frac{(Effect_{AB\ldots K}(+) - Effect_{AB\ldots K}(-))}{k} \tag{44}$$

Where the effect of a certain factor or interaction of factors is obtained through the simple average of the replicates of each run. Also, the symbols (+) and (-) represent the same codes as in the design matrix (see the example in table 5).

At last, it is possible to calculate the sum of squares 〖SS〗_X of each factor and interaction of factors (as represented in table 6) by adopting equation 45.

$$SS_{AB\ldots K} = \frac{1}{n2^k}(Contrast_{AB\ldots K})^2 \tag{45}$$

From equation 45 and with the help of the degrees of freedom (DF) of each factor and interaction of factors, it is possible to carry out a F-test (named after Ronald Fisher) by calculating first the mean squares of each sum of squares, and then their respective F-values. Equations 46 and 47 show how this procedure is done.

$$MS_{AB\ldots K} = \frac{SS_{AB\ldots K}}{DF_{AB\ldots K}} \tag{46}$$

$$F_{calc} = \frac{s^2_{AB\ldots K}}{s^2_{Error}} \tag{47}$$

To end the analysis, each F_calc is compared to a predetermined critical value named F_crit, that in turn depends on the probability α of incurring an error type I (rejecting the null hypothesis when it is in reality true).

The universe of DOE is vast and encompasses many other considerations and further discussions that, due to lack of relevance to this work, were not presented. The next section is the last of the literature review and will discuss how simulation and DOE could be applied together to produce meaningful insights.

### 2.2.4 Simulation and factorial design as ways to approach IRI

This last subsection of the literature review is shorter than the previous ones and will basically present the main inspirations for the methodology developed in section 3. The idea of a simulation approach combined with a factorial design to investigate IRI in a typical retail store came from Chuang & Oliva (2015). Of course, this paper is not the only one to present such combination as a suitable methodology to approach the research question described in section 1.3. Actually, it rather served as a very well-structured guide that provided valuable ideas to this thesis.

Some differences between the methodology adopted by Chuang & Oliva (2015) and in the current work can be seen, for instance, already in the simulation model. While the former researchers developed a system dynamics simulation model, in section 3 it will be shown that the simulation model of the current work is discrete-event one.

Another key different lies in the factorial design. Chuang & Oliva (2015) conducted a 3^k factorial design, whilst the current thesis focused on a 2^k one. It is perhaps easy to realize that a 3^k design produces more accurate results than a 2^k one; however, the marginal benefits of the former over the latter were considered to be not enough when compared to the extra resources that it would cost to adopt it.

# 3 METHODOLOGY

So far, section 1 introduced in a broad way what is this thesis about, discussing for example how big the retail industry is and how impactful IRI is for retailers. Section 2 then presented a whole set of papers, reports, and books that together composed the literature review of the current work. There, it was possible to understand, for instance, what are the most common causes and consequences of IRI for retailers, as well as the most commonly adopted countermeasures against it. Section 2 also analyzed in detail the suitability of using simulation and DOE approaches to answer the research questions stated in section 1.

All of it leads to section 3. Here, the methodology adopted in the thesis will be exposed in de-tail. For this purpose, the section was further divided into two subsections named "Simulation model" and "Design of Experiments". Each of them explains an important part of the procedure that led to the results presented in section 4 and the summarized conclusions presented in section 5.

Finally, before getting into the details within each subsection, it is important to reaffirm the re-search question that was first introduced in section 1.3, since it will be referred to a lot in this sec-tion and is the core of all the work developed in this thesis. Hence, one goal of the current work is was to validate the findings of Rekik et al. (2019), which suggested that positive inventory rec-ord inaccuracies, just like their negative IRI counterparts, also have an adverse impact on sales. If the prior goal was achieved, the other goal of the thesis was to identify what causes of positive IRI most significantly had this adverse impact on sales.

## 3.1 Simulation Model

This section was solely dedicated to the presentation of the simulation model. Since this task alone is already quite extensive, it was further broken down into three subsections, "Overall simu-lation setup assumptions and parameters", "Execution of the simulation model without IRI", and "Implementation of IRIs into the simulation". Each of these subsections constitute a fundamental part of understanding how the simulation model was developed and how it helped answering the research question of this thesis.

### 3.1.1 Overall simulation setup assumptions and parameters

The first step to answering the research question was choosing a simulation approach. The reason for this decision was heavily based on the benchmarking done in section 2.2.4 with the paper of Chuang & Oliva (2015), that combines a simulation model with DOE. However, other literature references also brought up simulation initiatives as valid ways to investigate IRI. Examples of it can be found, for instance, in the papers of Kang & Gershwin (2005) and Condea et al. (2012). The final argument that settled the decision to carry out a simulation of a retail store was provided by Banks et al. (1996) in section 2.2.2, when the authors of the book "Discrete-event system simulation" mentioned that resorting to simulation models is appropriate when the research involves changing a simulation's inputs and observing how it influences the outputs, which is exactly what was done in this thesis.

Once the development of a simulation model was validated as a viable way to answer the research question, the next step was to choose which simulation model to adopt. The choice between a stochastic or deterministic approach was mainly based on the benchmarking with other authors that simulated retail stores. Once again, Kang & Gershwin (2005), Condea et al. (2012), and Chuang & Oliva (2015) mentioned that they used stochastic methods to simulate a retail system. Here it is important to mention that Kang & Gershwin (2005) actually compared a stochastic and a deterministic model in their paper, so in reality the researchers developed both models.

Next, the choice to opt for a continuous approach (e.g., through system dynamics) or a discrete-event approach was mostly based on the previous experience of the student. In other words, since the student had no previous background with continuous simulation models and time was a decisive constraint in the development of the thesis, the discrete-event approach was chosen. However, another factor also weighted in favor of a discrete-event simulation. As exposed by Yaw Wong & McFarlane (2007) in section 2.1.1, the product removal patterns cannot be continuously monitored in most cases. Instead, their monitoring is carried out manually on a periodic basis by the workers.

Similarly to the decision criterium of which simulation model to select, the software of choice was equally chosen based on the previous experience of the student. The familiarity to the software Plant Simulation, developed by Siemens, immediately qualified it as a favorite in comparison to others. It is important to note, however, that there is no right choice for software picking as it can be seen, for example, with Condea et al. (2012), that used C# to draw their conclusions.

Now that the technical background was finally defined, the theoretical pillars could be established. Since the real-world system that the simulation was meant to mimic was a retail store, decisions regarding number of SKUs, product lines, replenishment policies, arrival pattern of clients, and so on needed to be defined. To better organize these parameters to be defined, table 7 provides a more structured presentation.

Starting with the number of simulated product lines, the benchmarking paper from Chuang & Oliva (2015), as well as Kang & Gershwin (2005) and Condea et al. (2012), developed single-item simulation models. In such cases, retailers are modeled as selling only one type. This simplification was considered to be acceptable when developing the model of the current work because if the negative impact on sales could be verified for one product line, then the findings of Rekik et al. (2019) could already be considered validated.

Table 7: Setup parameters to be defined in the simulation (own authorship).

| **Setup parameters to be defined in the simulation** |
|---|
| 1. Number of simulated product lines |
| 2. Inventory capacity |
| 3. Replenishment policy |
| 4. Arrival quantity and pattern of clients to the store |
| 5. Inventory costs |
| 6. Product price to calculate sales |
| 7. Opening and closing times |

Secondly, the inventory capacity had to be determined. For this specific case, the absolute amount of SKUs carried by the retailer was not too relevant. In other words, the focus here was to measure the variation of a set of controllable variables and see how they impacted the sales of the store, which made the absolute sales of each treatment not so relevant. However, Chuang

& Oliva (2015) concluded that a typical retail store carries on average 33000 SKUs, so this number was adopted as reference.

Still concerning the topic of inventory capacity, the decision of simulating a store with or without backroom had to be made. According to Yaw Wong & McFarlane (2007) and as already exposed in section 2.1.1, the choice of keeping a backroom is more space- and cost-effective for retailers than resorting exclusively to shelves. In addition to it, Condea et al. (2012) also highlighted the careful modeling of a backroom-shelf replenishment policy. Finally, Yaw Wong & McFarlane (2007) also mentioned in section 2.1.1 that IRI could also happen due to delays in the replenishment activities from the backroom to the shelves and from products getting lost in the backroom due to environmental complexity. This way, although implementing a backroom made the simulation more laborious to be developed, it proved to be necessary to simulate the real-world system of a retail store with greater precision.

Moving on to discussing the replenishment policy, this topic was further broken down into two other questions to be answered: “how would the shelf be replenished?” and “how would the backroom be replenished”.

In order to answer the first question, a consideration between simulating with a “pull” or a “push” replenishment policy as presented by Yaw Wong & McFarlane (2007) had to take place. According to the researchers, the “pull” policy is more widely adopted than the “push” one. This way, to keep the simulation as close to reality as possible, a “pull” policy was implemented. The only adjustment made was that the monitoring of inventory levels was also done at the POS, instead of relying exclusively on periodic audits. In other words, as customers checked out, the IMS deducted the purchased products from its inventory. This replenishment activity from the backroom to the shelf was set to last one hour (the value used by Condea et al. (2012) was 30 minutes, which makes the 1 hour approximation a reasonable one). Additionally to it, with the goal of increasing record accuracy and avoid unnecessary OOSh situations, physical stock counts were carried out periodically.

When it comes to the second question, the literature presented in section 2.2.1 regarding SIM played a key role in the decision made. To explain it better, Hillier & Lieberman (2015) provided a valuable background when presenting the different inventory management policies that could be adopted, such as the EOQ and the (Q,R) policies. Then when analyzing the simulation models developed by Kang & Gershwin (2005, p. 846) and Chuang & Oliva (2015), it was possible to verify that both of them adopted the (Q,R) replenishment system. This benchmarking in turn was the last piece of evidence needed to also adopt a (Q,R) replenishment system to the backroom of this thesis.

Last paragraph ends the discussion about the replenishment policies, allowing to move forward and analyze the arrival quantity and pattern of clients in the simulated store. For this purpose, the assumptions of Kang & Gershwin (2005), Chuang & Oliva (2015), and Condea et al. (2012) when building their model was a decisive benchmark. According to all these researchers, it is reasonable to adopt a Poisson distribution to determine the arrival pattern of clients to the store in each simulated day, which consequently settled the simulation of this thesis as being a stochastic one. When it comes to the absolute quantity of arriving clients in each day NC, it was concluded that, for this study, this number played a minor role under the same argument as used when establishing the absolute inventory levels. This way, equation 48 shows how NC was determined for every simulated day. Also regarding the customers and with the aim of getting closer results to reality, it was defined that the amount of products that each customer intended to buy NP was defined by a normal distribution with mean 4 and standard deviation 1, according to equation 49.

$$NC \sim Pois(\lambda)\ , \qquad \lambda = 1000 \tag{48}$$

$$NP \sim |N(\mu, \sigma^2)|\ , \qquad \mu = 4\ and\ \sigma = 1 \tag{49}$$

Next come the inventory costs, which are necessary to determine the value of the order quantity Q in a (Q, R) replenishment policy. When analyzing equation 8, the cost-associated parameters that needed to be defined were the setup cost K and the inventory keeping cost h per SKU. Since both costs were of minor importance, an estimated value was attributed to them, resulting in the values K=\$14.000,00 and h=\$0,50. This best-guess approach was also adopted by Condea et al. (2012) when they determined the very same parameters. Still regarding the (Q, R) policy, Kang & Gershwin (2005) define the lead time to resupply the backroom with new items as LT=3 days. This way, with the help of equation 50 presented by Kang & Gershwin (2005), it was possible to determine the reorder point R.

$$R = (expected\ demand\ during\ lead\ time) + (safety\ stock) \tag{50}$$

The determination of the safety stock is usually judgmental and, because of it, this parameter was arbitrarily chosen to be 5.000. So, for the simulation of this thesis, Q and R were calculated according to equations 51 and 52, respectively.

$$Q^{*} = \sqrt{\frac{2 \cdot NP \cdot NC \cdot K}{h}} \sim \sqrt{\frac{2 \cdot 4 \cdot 1000 \cdot 14.000}{0{,}50}} = 15.000 \tag{51}$$

$$R = LT \cdot NP \cdot NC + (safety\ stock) \sim 3 \cdot 4 \cdot 1.000 + 5.000 \tag{52}$$

Regarding the product price, since the simulation model is a single-item one, the price is simply a constant that multiplies the sales quantity. This way, once the aim is to compare the variation on sales (as better explained in section 3.2, when the factorial design is introduced), the price of each product was set as being P=$1,00.

Finally, the opening and closing times of the retail store also played a minor role in the simulation and, because of it, the opening hours from 8:00 a.m. until 8:00 p.m. were adopted based on common sense. To verify if these numbers were not absurd, a comparison with Condea et al. (2012) was carried out, where the researchers simulated a store that stays open for 10 hours per day.

In the last paragraph, the last of the most important setup parameters was defined. This opened the path to move on and proceed to the detailed modeling and execution of the simulation.

### 3.1.2 Execution of the simulation model without IRI

The parameters defined in the last section were very important to allow for the execution of a first version of the simulation model without IRI. The objective of simulating this first version of the model was to check if the replenishment activities were working appropriately and also to check if the logic behind customers arrival and their interaction with the store was happening as planned. However, before presenting the results of these first trials, it is important now to introduce the simulation model, explaining in detail each of its components. To this end, figure 17 presents a screenshot of the developed model in the software Plant Simulation that opens this discussion.

In figure 17, there are 5 major blocks, a green one named “Setup Parameters”, a light blue one named “Controlling Parameters”, two red ones named “Negative IRI Parameters” and “Positive IRI Parameters”, and a last purple one named “Retailer”. The “Retailer” block contains the core of the simulation and represents the retail store. All the other blocks contain parameters used in the simulation.

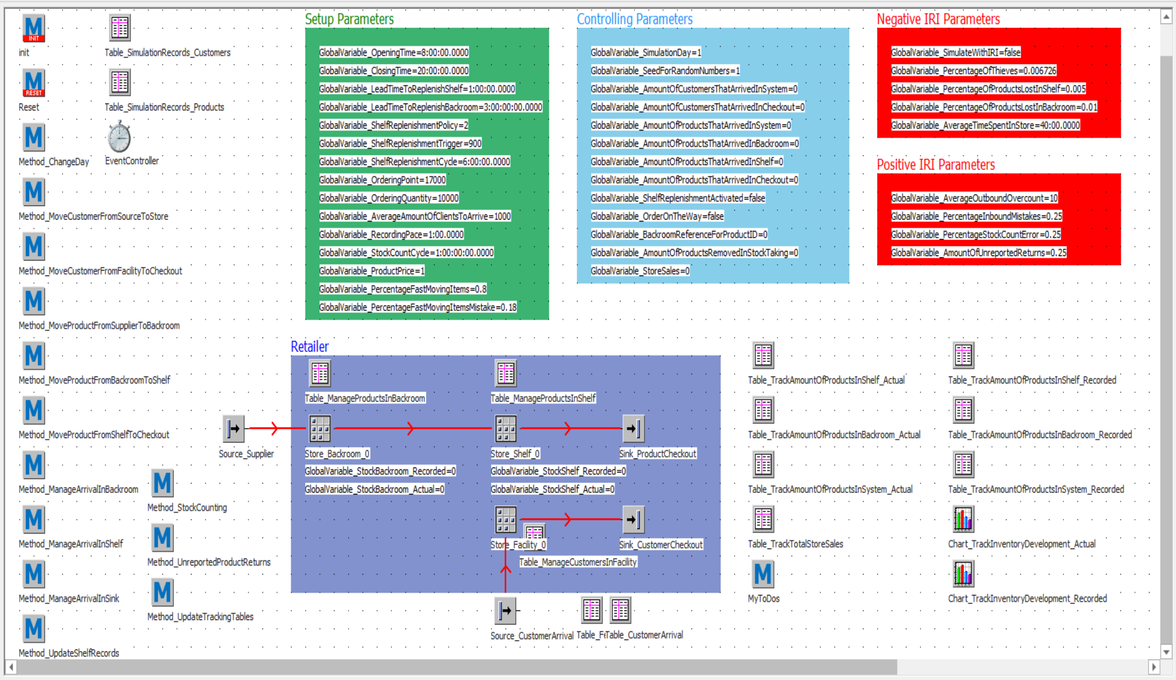


Figure 18: Overview of the developed model in the software Plant Simulation (own authorship).

A closer look at the “Retailer” block will provide further information regarding how the simulation model was developed and how each individual element interacts with each other. For this purpose, figure 18 will be used as reference.

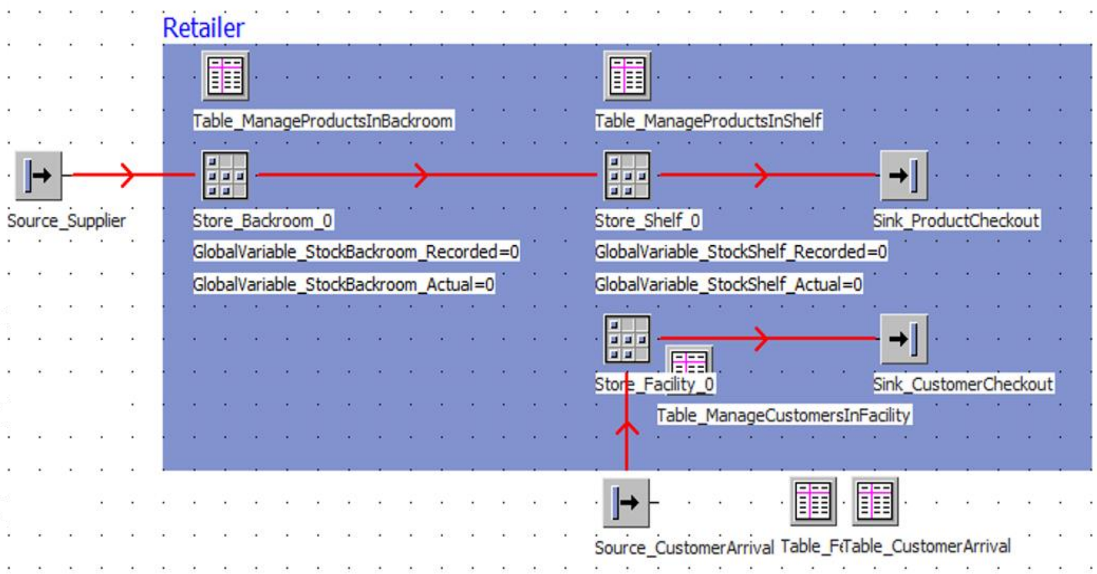


Figure 19: “Retailer” block with all its components (own authorship).

The analysis of figure 18 starts from the two sources. The bottom source, named “Source_CustoemrArrival” is the one responsible for generating the customers that will enter the retailer during a certain day. The moments in which each customer arrives in the store are defined beforehand and stored in the data table “Table_CustomerArrival”, to the right of the source “Source_CustomerArrival”.

As customers enter the retailer, they retrieve products from the shelves, identified as the building block “Store_Shelf_0” in the center of figure 18. “Store_Shelf_0” represents the entirety of shelves available in the store. The data table “Table_ManageCustomersInFacility” is responsible to keep track of which customers are in the store at any given time of the simulation and was only relevant for debugging purposes. Data tables “Table_ManageProductsInBackroom” and “Table_ManageProductsInShelf” also served debugging purposes and, therefore, will not be covered in further detail.

After taking the products from the shelves, if they were not empty, the customers would then proceed to the checkout area, where the items they intended to acquire would be processed by the IMS and the recorded inventory levels would be updated. This mentioned checkout area is represented by the building block “Sink_CustomerCheckout” and the red arrows connecting the “Source_CustomerArrival” to the “Sink_CustomerCheckout” represent the customers’ journey in the system.

Analyzing now the side of the products, their journey originates in the source “Source_Supplier” and ends in the drain “Sink_ProductCheckout”. While on the shelves, products are taken as customers arrive in the stores. These products removed from the shelves then proceed to the checkout, where the IMS acknowledges they are not part of the inventory anymore. The replenishment from the backroom to the shelf occurs as described by Yaw Wong & McFarlane (2007), where a replenishment cycle of 6 hours was set up (this parameter was captures by the variable “GlobalVariable_ShelfReplenishmentCycle”). Finally, the replenishment from the supplier to the backroom obeys a (Q,R) policy as it was depicted in the last section and captured in equations 51 and 52.

At last, since this simulation model is a stochastic one, it is important to mention how the random variables were generated. Taking the literature background presented by Law (2007) into account and the perks already embedded in the Plant Simulation software, the randomness of the events was assured. To better illustrate it, let us analyze in more detail the light blue block called “Controlling Parameters” exposed in figure 17 and zoomed in in figure 19. As shown in figure 19, there is a variable within that block called “GlobalVariable_SeedForRandomNumbers”. This variable is the most important one when it comes to ensuring the randomness of the events in the simulation.

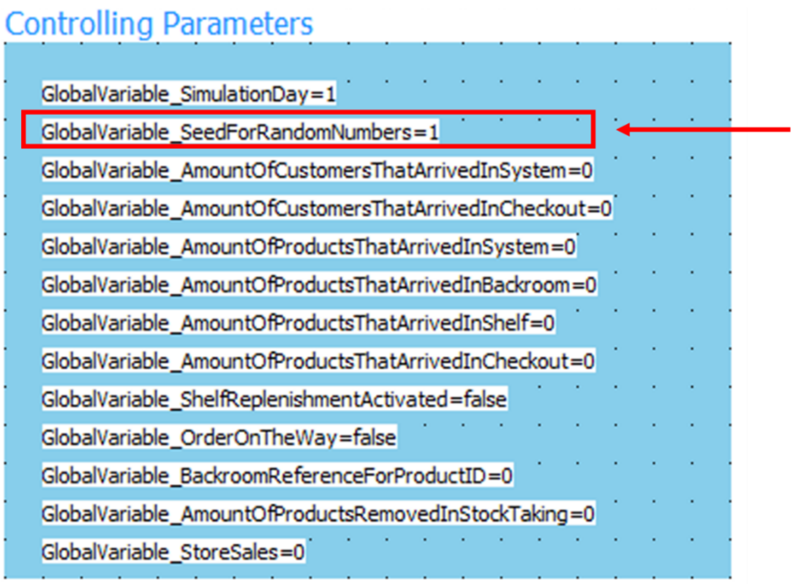


Figure 20: “Controlling Parameters” block with all its variables (own authorship).

To better understand how the variable "GlobalVariable_SeedForRandomNumbers" works, it is useful to analyze equations 48 and 49, that require a degree of randomness. Equation 48 defines the amount of customers that arrive in the retail store in a given day and it is captured in the simulation model by the code line presented in figure 20.

```
-- The following row defines how many customers will arrive in the store on the first day
intAmountOfCustomers := z_poisson(GlobalVariable_SeedForRandomNumbers,GlobalVariable_AverageAmountOfClientsToArrive)
```

Figure 21: Poisson distribution to determine the number of customers should arrive in the store at any given day (own authorship).

In figure 20, it is possible to see that the variable "GlobalVariable_SeedForRandomNumbers" is one of the inputs in the function "z_poisson". The other input, "GlobalVariable_AverageAmountofClientsToArrive", represents the average number of clients expected to arrive in the store at any given day, captured by λ in equation 48. In order to ensure that the seed will change from one simulation day to the next, its value was updated automatically by a method called "Method_ChangeDay", executed every time the simulation time reached a multiple of 24 hours. Figure 21 shows how the update of the seed took place.

```
-- The following blocks are very similar to the method init
-- Controlling parameters
intPositionInTable := 1
GlobalVariable_SimulationDay += 1
GlobalVariable_SeedForRandomNumbers += 1
```

Figure 22: Code lines that update the seed responsible to generate random numbers and some other variables (own authorship).

Now that the most important parameters were introduced, it is possible to execute the simulation and show the results. Note in figure 22 that the controlling variables "GlobalVariable_StockShelf_Recorded" and "GlobalVariable_StockShelf_Actual", that capture the recorded and actual inventory levels of the shelves respectively, present the same values (in this case, 9.685). The same thing happens to the backroom, that has 14.662 items in both recorded and actual stocks.

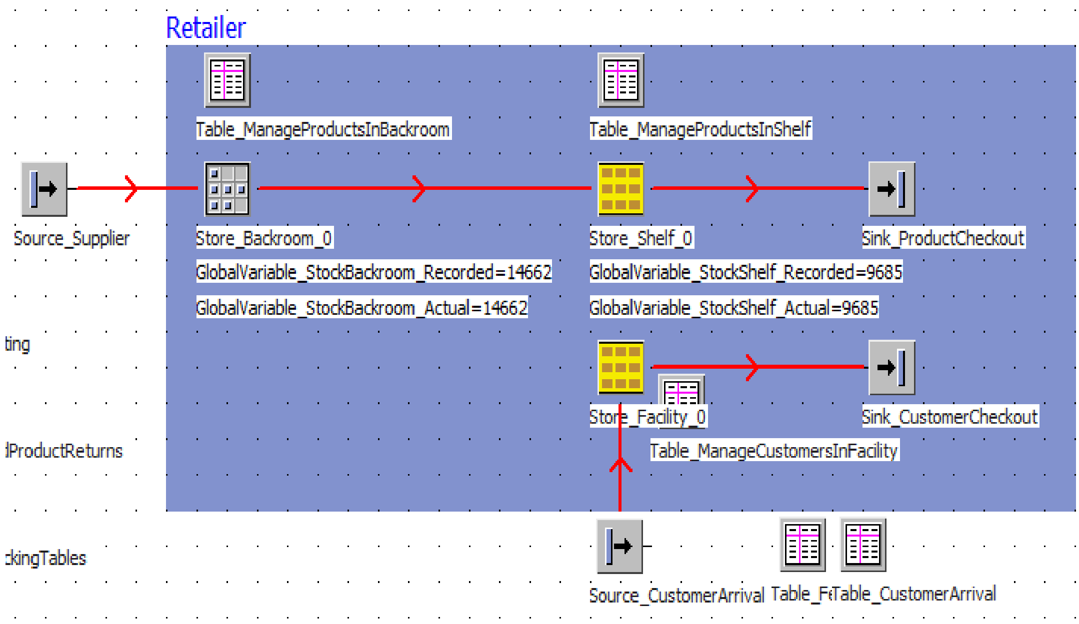


Figure 23: “Retailer” block in a simulation without IRI (own authorship).

The result of the simulation under these conditions is presented in figure 23, that shows the evolution of total inventory in the store (backroom plus shelf inventories) as time goes on up until day 40 of the simulation.

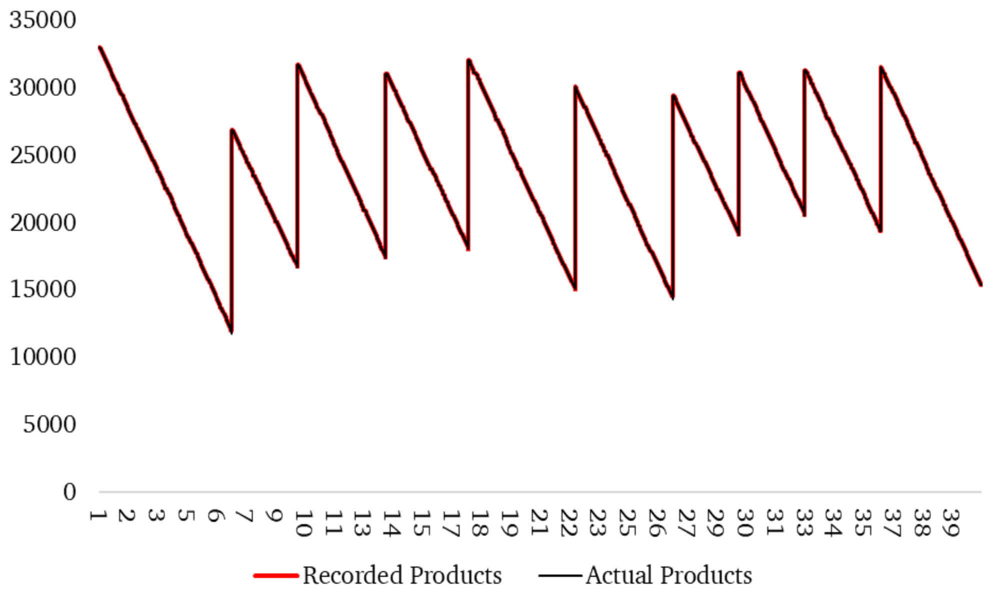


Figure 24: Simulation results without IRI (own authorship).

Some important conclusions can be taken from figure 23. The first of them is that the replenishment policy works well under perfect conditions (no IRI) and no OOSt situations arise. The second conclusion is directed to the fact that both actual and recorded inventory levels show the same results, indicating that the simulation was properly developed regarding its coding. The third and last conclusion needs help from figure 24, which amplifies a section of figure 23, allowing for a closer look at it.

In figure 24, it is possible to see that the curves are not linear, indicating that the depletion of the inventory levels is indeed stochastic and not deterministic. This is a good way to validate that the generator of randomness introduced in figure 21 is properly working.

With these verifications, section 3.1.2 is finally concluded, and it is possible to analyze how the simulation model behaves when failure modes that generate IRIs are implemented. This will be done in the following section 3.2.3.

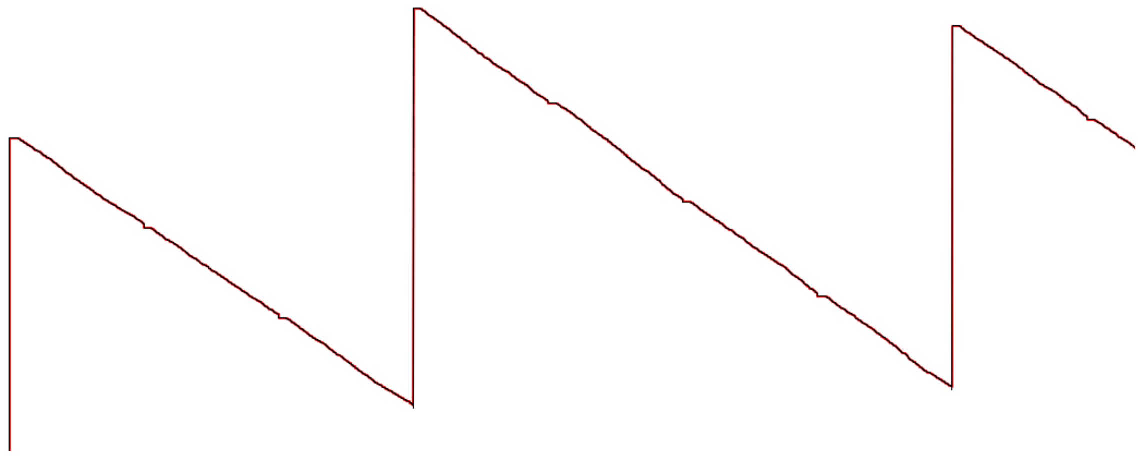

Figure 25: Zoom in Figure 23 to observe stochastic effects (own authorship).

### 3.1.3 Implementation of IRIs into the simulation model

As stated at the end of the last section, this one has as its primary goal the implementation of record inaccuracies to the simulation model. For this reason, there is a deep connection between this section and section 2.1.1, that in turn identifies the main causes and consequences of IRI.

The causes of IRI identified in section 2.2.1 were further subdivided here into those that generate positive discrepancies (actual inventory is greater than the recorded one) and those that generate negative discrepancies (actual inventory is smaller than the recorded one). Since the goal of this thesis was to study the impact of the positive discrepancies on sales, the negative ones were theoretically not necessary. However, to approximate the simulation as much as possible to the real-world system of a retail store, negative discrepancy generators were also introduced (they were later in section 3.2 considered as hold-constant parameters in order not to influence the results of the DOE).

Now that a brief introduction to this section was provided, it is possible to present the IRI parameters that were utilized in the simulation. Taking as reference figure 17, the IRI parameters are located in the red blocks at the top-right corner of the figure. In order to better visualize them, tables 8 and 9 provide a clearer representation of which IRI parameters were considered and their values.

Starting with the negative IRI, Kang & Gershwin (2005) stated that inventory shrinkage is one of the main causes of it. Section 2.1.1 approaches this topic in deeper detail, but, for the sake of the simulation, it was important to breakdown this shrinkage factor into its subcomponents and to quantify them so that reasonable input can be added to the model. With this in mind, Bamfield's (2004) contribution was decisive. The researcher pointed out that the main cause of shrinkage was theft (both internal and external). This fact was also confirmed by Moraca et al. (2015) and Hollinger & Davis (2002), who studied the shrinkage problem in US retail stores.

Table 8: Negative IRI parameters taken into account in the simulation (own authorship).

| **Negative IRI parameters** | |
|---|---|
| Thieves | 0,57596% |
| Products lost on the shelf | 5% |
| Products lost in the backroom | 10% |
| Average time spent in store | 40 minutes |

Table 9: Positive IRI parameters taken into account in the simulation (own authorship).

| **Positive IRI parameters** | |
|---|---|
| Average outbound overcount | 3 |
| Inbound mistakes | 5% |
| Stock count error | 5% |
| Amount of unreported returns | 5% |

The “Thieves” parameter in table 8 was captured in the simulation as the variable “GlobalVariable_PercentageOfThieves” and the idea behind it was to determine what percentage of customers were thieves. This would generate negative discrepancies because thieves would simply steal the merchandise instead of purchasing them, subtracting the actual inventory but not altering the recorded one. In the case of table 8, 0,57596% of all customers were considered to be thieves. The rationale for this number is based on can be obtained from Bamfield’s (2004) study, where the researcher found that, in Germany, shrinkage corresponded on average to 1,19% of retailer’s turnover. In addition to it, 48,4% of all shrinkage was attributed to external theft in Europe. With these two data, equation 53 illustrates how the percentage of thieves for the simulation was obtained.

$$Percentage\ of\ thieves = 1{,}19\% \cdot 48{,}4\% = 0{,}57596\% \tag{53}$$

Moving on to products lost on the shelf and in the backroom, these factors were also mentioned by Kang & Gershwin (2005) as one major contributor to negative discrepancies and, because of it, they were considered to be relevant to the simulation. The numbers in table 8 suggest that out of the total inventory carried by the store, 5% is lost on the shelf and another 10% is lost in the backroom. The mechanisms that controlled this in the simulation were attributes of the class “Product” named “GotLostInShelf” and “GotLostInBackroom”. Figure 25 shows these two attributes as well as all other ones used to control the products in the simulation model.

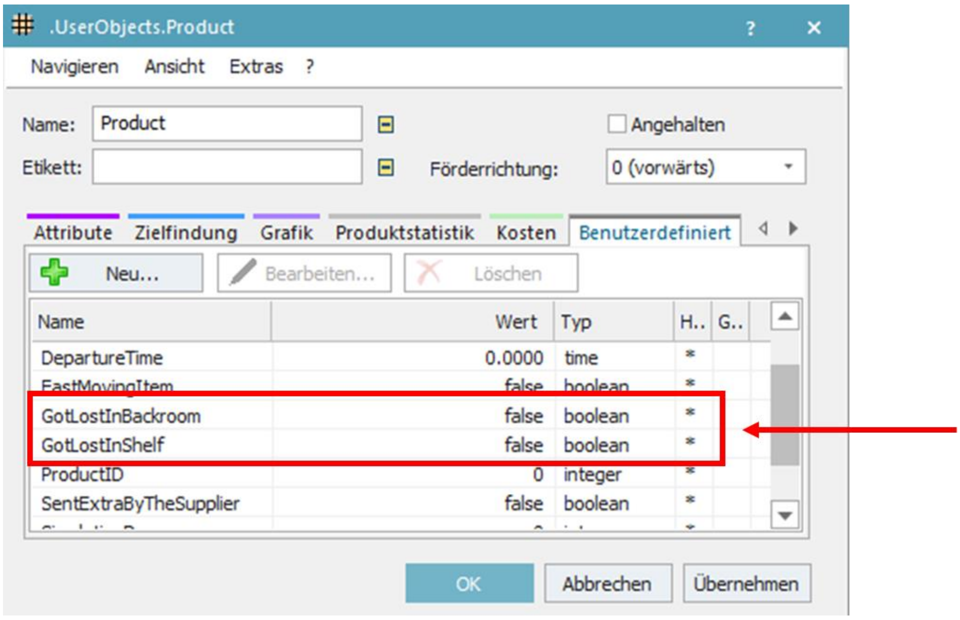


Figure 26: Attributes of the class "Product" that controlled product loss (own authorship).

It is important to note here the choice of setting the backroom loss as being greater than the shelf loss was not arbitrary. As presented in section 2.1.1, DeHoratius & Raman (2008) found that high (low) levels of inventory density are associated with high (low) levels of IRI. In other words, since the backroom usually has a far greater inventory density than the selves, the chances of generating IRI, for instance by product loss, is greater in the former than it is in the latter.

Finally, the average time of customers in the store is also a relevant factor to take into consideration, since the products are not available on the shelves during the time customers are in the store, but at the same time the POS did not yet register their checkout. This factor was put in the simulation because it was mentioned by Yaw Wong & McFarlane (2007).

Before moving on to introduce the parameters that controlled the positive discrepancies simulated in the model, it is important to show how the model behaved when simulated only with negative discrepancies. In section 2.1.2, it was shown that the simulation of (Kang & Gershwin (2005) eventually reached a point of "replenishment freeze", in which no new orders were placed because the recorded inventory levels were higher than the ordering point and, at the same time, no new sales happened as well because the actual inventory was zero. If a similar situation could be found in the model developed in the current work, it would be an indication that the negative IRI parameters were correctly simulated.

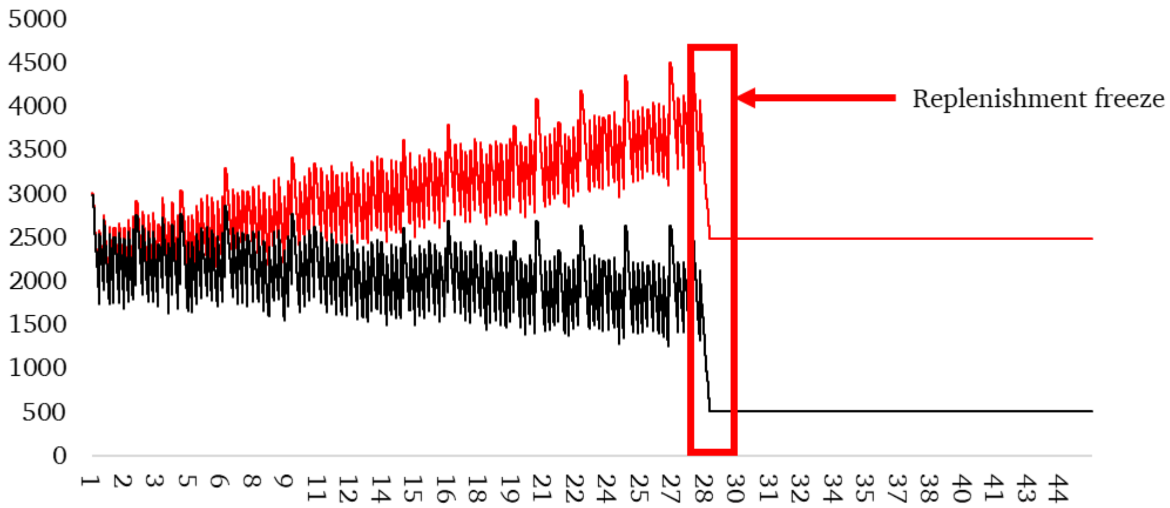


Figure 27: Replenishment freeze (own authorship).

As shown in figure 26, a replenishment freeze situation could be recreated when simulating the retail store with just negative-discrepancy parameters. However, an adjustment had to be made. The replenishment cycles were shortened, and customer demand was increased in order to be able to see the replenishment freeze happening within the first 30 days of simulation. If the conditions presented in tables 8 and 9 were kept, I would take over 200 days to get to the same result.

It is now possible to start discussing the parameters associated with the generation of positive discrepancies. The first parameter mentioned in table 9 was the "Average outbound overcount". To understand it, however, it is necessary to bring up DeHoratius & Raman's (2008) findings suggesting that FMCGs are more prone to transaction errors than their slow-moving counterparts (more details can be found in section 2.1.1). To illustrate it better with an example, this type of transaction error occurs when the cashier processes two or more SKUs of same price and similar labels as being the same product type, when they are actually different ones.

To capture this phenomenon in a way that it could be simulated as a positive discrepancy, the parameter "Average outbound overcount" was created. In the simulation model, it was represented by the variable "GlobalVariable_AverageOutboundOvercount". The idea here was to count a fraction of FMCGs n-fold, with n being, in the case of table 9, equal to 3. An example may help to understand it better. Figure 27 presents the variables "GlobalVariable_PercentageFastMovingItems" and "GlobalVariable_PercentageFastMovingItemsMistake". The first variable states that 80% of the products in the store are FMCGs, whereas the second variable states that 18% of the FMCGs are processed wrongly (find some kind of transaction error). This 18% share of the FMCGs will be the ones to be counted n-fold at the checkout and, therefore, generating positive IRIs.

```
GlobalVariable_PercentageFastMovingItems=0.8
GlobalVariable_PercentageFastMovingItemsMistake=0.18
```

Figure 28: FMCG simulating variables (own authorship).

The next positive discrepancy failure mode is the inbound mistake. In section 2.2.1, DeHoratius & Raman (2008) also say that the store inventory record is automatically updated upon the placement of the replenishment order and, by doing so, the store management oftentimes relies on the assumption that the order has been filled correctly. Hence, if the situation arises in which a DC employee incorrectly fills the store order by accidentally picking and shipping the wrong item or the wrong quantity of the correct item, the amount received by the retail store will not match the quantity ordered. This problem is a potential source of positive discrepancies, and, because of it, it was added to table 9. In the example of table 9, the "Inbound mistakes" is associated with a 5% value. This relationship means that, out of the total shipment received from the supplier at a certain replenishment order, an extra of 5% of the ordered quantity is added to the lot and not accounted for. This mistake ensured that a positive discrepancy was generated and that its effect could be later in detail studied.

Following the "Inbound mistakes" are the "Stock count errors". According to Iglehart & Morey (1972), an inventory count does not completely eliminate discrepancies between the recorded and physical stock. In practice, non-trivial inaccuracies usually remain in the inventory records because of human error in the counting procedure itself. Moreover, transactions that take place during the count make this task extremely difficult and costly to assure 100% accuracy. This fact paved the way to assume that there was a potential for positive discrepancies arising from stock counts and, therefore, "Stock count errors" was considered an IRI parameter. The logic behind this parameter as represented in table 9 is that, in every stock count, a 5% overcount of the total inventory mistakenly happens.

The last simulated positive-discrepancy parameter was the “Amount of unreported returns”. The literature that served as benchmark for this idea was Rekik's et al. (2019, p. 13) study. According to the researchers, in retail branches such as the fashion and apparel ones, returned products could be an important source of IRI, especially positive discrepancies. This is due to the fact that providing customers with the opportunity to return products create chances for errors like forgetting to update the IMS with the product is returned. This way, the data in table 9 can be understood as follows: an extra amount of 5% of the daily number of customers that shop in the retail store every day come to return an item that is not accounted for in the IMS and, therefore, leads to IRI.

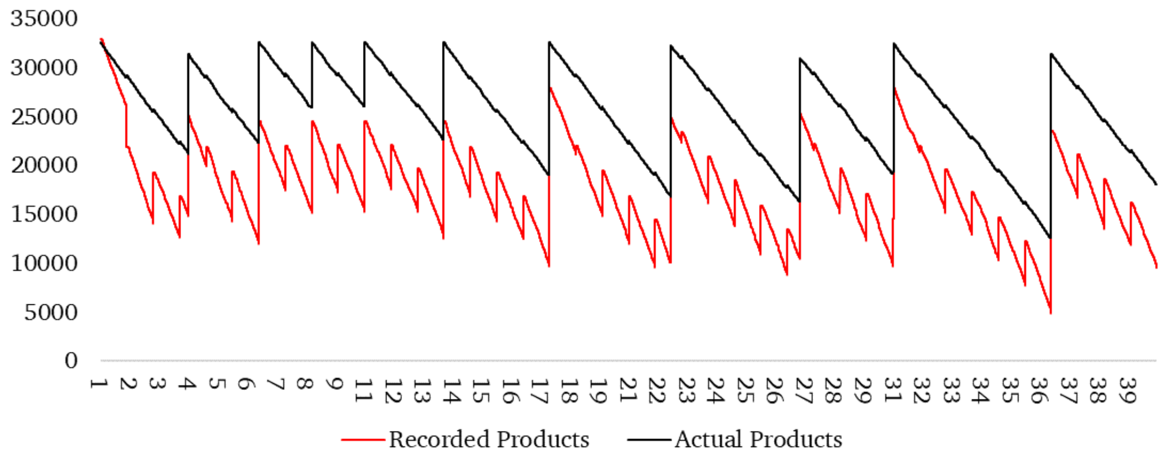


Figure 29: Simulation with both positive and negative discrepancies (own authorship).

Now that the methodology behind how the inventory record inaccuracies were implemented was introduced, it is possible to analyze how the DOE was carried out and, out of it, which conclusion to the research questions were drawn. However, before doing so, it is interesting to present how a simulation carrying both positive and negative discrepancies would look like. For this purpose, figure 28 fills this gap and shows a simulation run with both types of discrepancies.

## 3.2 Design of experiments

The factorial design was the defining tool that drew the conclusions about the research question first presented in section 1.3. This way, section 2.2.3 had the objective of presenting possibilities through which the experiment of this section could be carried out and the current section was responsible for planning how the experiment would be executed. In order to adopt a basic orientation Montgomery's (2009) guidelines introduced in table 4 were folloewd.

The first step then for designing a good experiment was to recognize and state the problem. As already mentioned in the last paragraph, this was done in section 1.3, where the problem was stated as validating the findings of the study of Rekik et al. (2019) regarding the fact that positive inventory record inaccuracies can lead to negative impacts on sales in a typical retail store. A second problem that was likewise stated was to identify the causes of these positive IRIs that most significantly affect the sales of a retailer in a negative way.

Once the problem was properly recognized and stated, the selection of the response variable was needed. This was a quite uncomplicated task, since the response variable could be simply set as something related to sales. This way by setting the price of each SKU as P=$1,00, a measurement for sales was already obtained. To frame the experimental unit, however, it was still necessary to establish a period of time in which the measurements would take place. Then, for this purpose, a time spam of 10 days was chosen (within this period of time, all the most important events such as replenishment of shelf, replenishment of beckroom, customer check-out, stock count, and change of day could already happen).

The next step was the choice of factors, levels, and ranges. The factors, as already presented in the last section, were the same as the ones in table 9. The levels of each factor were chosen in a way that it would be possible to see the changes in the response variable defined in the last paragraph. This way, table 10 summarizes the four parameters from table 9 with their levels.

As already mentioned in section 2.2.4, a 2^k factorial design was chosen because it is already robust enough to present the desired conclusion, but a more complex experiment such as a 3^k was discarded due to its complexity and time constraint of the thesis.

Table 10: Positive IRI parameters with their simulated levels in the factorial design (own authorship).

| | **Factors** | **-1** | **1** |
|---|---|---|---|
| A | Outbound overcount | 3 | 10 |
| B | Inbound overshipment | +5% | +25% |
| C | Inaccurate stock "undercounting" | +5% | +25% |
| D | Unreported returned products | +5% | +25% |

After choosing the experimental design, the other steps are the experiment itself and the analysis and conclusions of its results. This will be done in the next two sections for the sake of not mixing the methodology with the experimentation.

# 4 RESULTS

The results presented in this section will encompass the steps “Performing the experiment” up until the “Statistical analysis of the data”. The “Conclusions and recommendations” step was left for section 6.

Based on Montgomery's (2009) description of how to carry on with a factorial design, the first step would be to ensure that the three basic design principles were taken into consideration. These principles were first presented in section 2.2.3 and, therefore, they will not be covered in-depth.

The first principle was the randomization principle. In other words, to stick to this principle it was necessary that all experimental observations were done in a random order. The second principle was the replication one, which guaranteed that a standard ANOVA analysis was carried on without resorting to any additional tool that counters the absence of replicates. Finally, the third and last basic principle surrounding the DOE is the blocking principle. As already mentioned in section 2.2.3, blocking itself is a design technique that is often used to improve the precision with which comparisons are made. However, since the experiment of this thesis did not show any defining differences between samples that could justify blocking the treatments, this principle was considered to be not applicable to this case.

Once these first considerations were made, the next step was to conduct the experiment itself. To help with the organization of the experiments, a design matrix such as in table 11 was created. In this matrix, it is possible to see that a total of 3 repetitions per run were made and, to reassure once again, the order in which the data were collected respected the first principle of randomization.

Table 11: Design matrix for the 2^4 experimental design (own authorship).

| Run | A | B | C | D | AB | AC | AD | BC | BD | CD | ABC | ABD | BCD | ABCD | Sample 1 | Sample 2 | Sample 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -1 | -1 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | 1 | -1 | -1 | -1 | 1 | 41826 | 51260 | 41758 |
| 2 | 1 | -1 | -1 | -1 | -1 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | -1 | -1 | 51012 | 51178 | 41707 |
| 3 | -1 | 1 | -1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 | 1 | 1 | 1 | -1 | 41727 | 41697 | 51165 |
| 4 | 1 | 1 | -1 | -1 | 1 | -1 | -1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 | 51042 | 51296 | 48520 |
| 5 | -1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 41977 | 41879 | 41799 |
| 6 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 | 1 | 41772 | 41727 | 41694 |
| 7 | -1 | 1 | 1 | -1 | -1 | -1 | 1 | 1 | -1 | -1 | -1 | 1 | -1 | 1 | 41970 | 51265 | 48103 |
| 8 | 1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 | -1 | 41644 | 41757 | 51161 |
| 9 | -1 | -1 | -1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 | -1 | 1 | 1 | -1 | 58494 | 39803 | 36513 |
| 10 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 | -1 | 1 | 1 | 39894 | 48997 | 49315 |
| 11 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | 1 | -1 | -1 | 1 | 39961 | 39927 | 39813 |
| 12 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 | 30456 | 39848 | 39868 |
| 13 | -1 | -1 | 1 | 1 | 1 | -1 | -1 | -1 | -1 | 1 | 1 | 1 | -1 | 1 | 39790 | 40078 | 49252 |
| 14 | 1 | -1 | 1 | 1 | -1 | 1 | 1 | -1 | -1 | 1 | -1 | -1 | -1 | -1 | 45359 | 39899 | 39894 |
| 15 | -1 | 1 | 1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | -1 | -1 | 1 | -1 | 49155 | 40037 | 45592 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 39880 | 39843 | 39722 |

Before proceeding to the analysis of the effects of each factor and interaction of factors over the response variable, it is fundamental to first assess the quality of the collected data. The so-called residual analysis could be carried out through visual analyses like a control chart and a scatter chart of the residue versos their means. Both of these charts were employed in this thesis, and they are presented in figures 29 and 30, respectively.

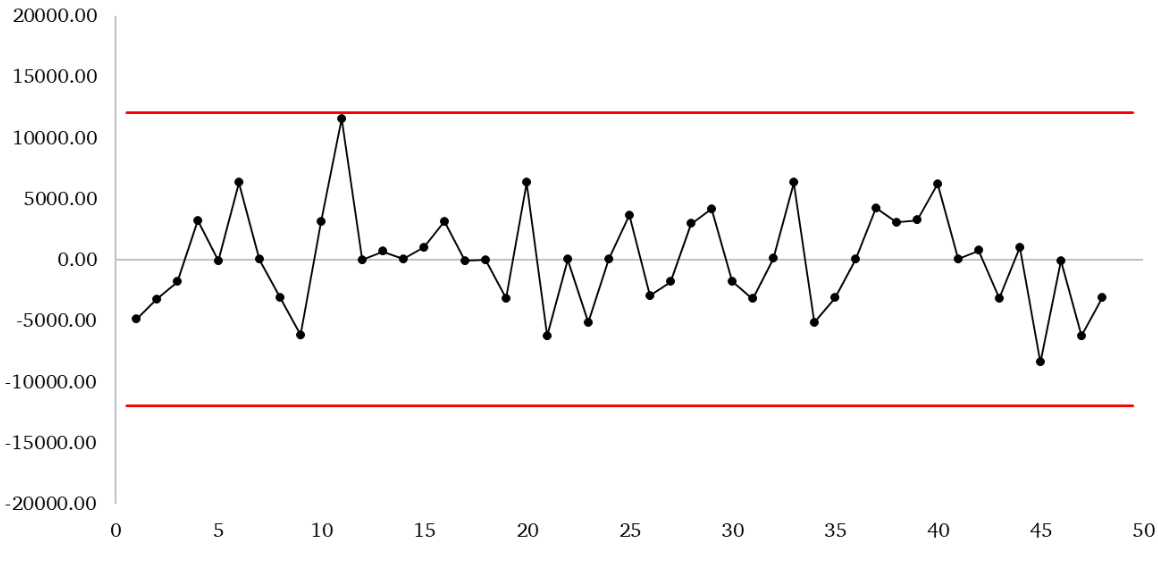


Figure 30: Control chart of the residue from the design matrix in the order they were collected (own authorship).

The two red lines in figure 29 are the control limits defined by plus and minus three standard deviations from the middle line, which in the case of the residual analysis coincides with the y=0 line.

It is important to note that the data in the control chart do not present any kind of runs or trends, which correspond to scenarios in which there are 7 or more points on the same side of the middle line and 7 or more point in ascending or descending order. Additionally to it, there are no points outside the area delimited by the two red lines. These facts visually suggest that the collected data were of good quality and, therefore, suitable to proceed with the experimentation.

However, an additional verification regarding the quality of the collected data can be done in figure 30. There, the scatter chart of the residue of each sample versus the respective mean of the three samples is provided. The analysis here aims at verifying if there are any strong patterns such as a cone shape or a “butterfly knot” shape. Since none of it can be perceived at first glance, there is strong reason to believe that the quality of the collected data is reliable.

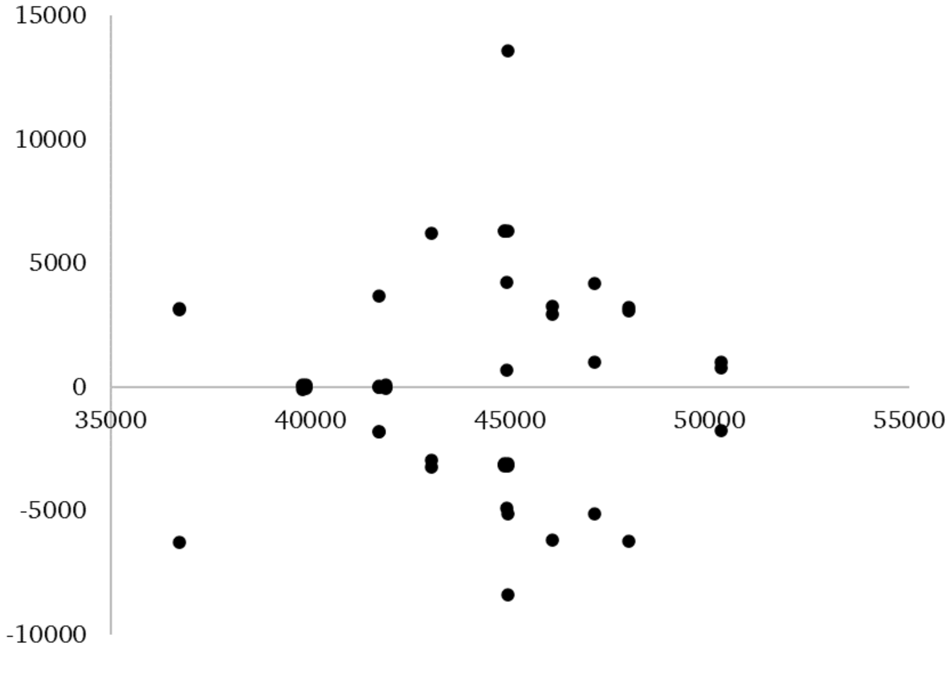


Figure 31: Scatter chart of each residuum versus its respective associated mean (own authorship).

Once the quality of the collected data was assured, it was possible proceed and calculate the effects of each factor and interaction of factors over the response variable. These results are presented in table 12.

Table 12: Effects and F-values of each factor and interaction of factors over the response variable (own authorship).

| | A | B | C | D | AB | AC | AD | BC | BD | CD | ABC | ABD | BCD | ABCD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Effect** | -306.5 | -476.167 | -1326.17 | -3314.42 | -974.75 | -1905.58 | -1813.5 | 2560.25 | -3122.67 | 1793.833 | -499 | -1049.92 | 1031.417 | 628.5 |
| **$SS_X$** | 187884.5 | 453469.4 | 3517436 | 21970716 | 1900275 | 7262496 | 6577565 | 13109760 | 19502094 | 6435676 | 498002 | 2204650 | 2127641 | 790024.5 |
| **DOF** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **$MS_X$** | 187884.5 | 453469.4 | 3517436 | 21970716 | 1900275 | 7262496 | 6577565 | 13109760 | 19502094 | 6435676 | 498002 | 2204650 | 2127641 | 790024.5 |
| **$DOF_T$** | 31 | | | | | | | | | | | | | |
| **$SS_T$** | 1331595870 | | | | | | | | | | | | | |
| **$DOF_E$** | 17 | | | | | | | | | | | | | |
| **$SS_E$** | 1245058181 | | | | | | | | | | | | | |
| **$MS_E$** | 73238716.55 | | | | | | | | | | | | | |
| **Fcalc** | 0.002565 | 0.006192 | 0.048027 | 0.299988 | 0.025946 | 0.099162 | 0.08981 | 0.179 | 0.266281 | 0.087873 | 0.0068 | 0.030102 | 0.029051 | 0.010787 |

The first important conclusion to extract from table 12 is that, apart from the interactions BC, CD, BCD, and ABCD, all other factors and interaction of factors resulted in negative effects on the response variable. Since the response variable was set as being the total sales of the simulated retail store over the course of 10 day, these results mean that there is in fact a negative impact on the retailer’s sales caused by positive discrepancies between the recorded inventory and the actual one.

This finding is in line with the paper of Rekik et al. (2019) and, more than that, it answers the first research question presented in section 1.3, which aimed at validating the findings of the researchers regarding the fact that positive inventory record inaccuracies indeed lead to negative impacts on sales in a typical retail store.

Once this first question was answered, it was possible to move further and try to identify the main factors that influence the response variable out of the ones considered in table 10. To this end a Pareto chart such as the one presented in figure 31 provided a valuable tool to visually draw conclusions.

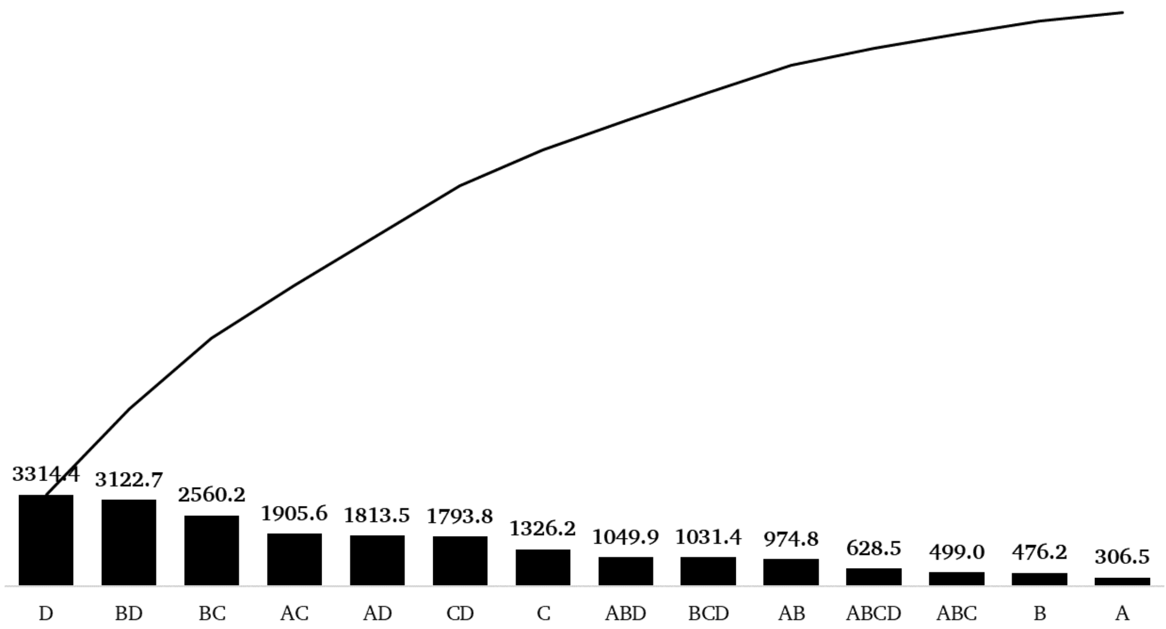


Figure 32: Pareto chart of the main factors and interactions of factors over the response variable (own authorship).

Note that, in order to build the chart in figure 31, the absolute values of each factor and interaction of factors were taken into consideration, instead of accounting for their individual signs.

Another important conclusion that can be taken from the chart is that, for instance, factor D is clearly more influent than factors A and C. In other words, unreported returns, in an apparel store for instance, plays a greater role in generating IRIs than outbound overcount (A) and inbound overshipment (C).

These conclusions will be once again presented in the next and last section, together with some limitations of this thesis and some next steps that could be taken.

# 5 CONCLUSION

As already stated in the last section, the main conclusions that are worth mentioning here are that the findings of Rekik et al. (2019) were verified and the simulation model indeed showed that positive discrepancies do lead to negative impact on sales. Another key conclusion that came out of the results exposed in the last section was that unreported returns overshadow other causes of IRI such as outbound overcount and inbound overshipment.

Although these conclusions are very important to support not only the research of Rekik et al. (2019) but also the literature around IRI as a whole, some limitations of this thesis must be mentioned for future development.

The first limitation to be mentioned is the fact that a single-item simulation model was adopted. As already mentioned by DeHoratius & Raman (2008), one important generator of IRI is the high level of product variety in a store. When simulating under the assumption of a single-item model, such parameters as this one cannot be accounted for. Another downside is also the impossibility of analyzing how this parameter would interact with the other parameters in the simulation model presented in this work.

Another limitation of this work is that some setup parameters for the simulation model, such as the opening hours of the retail store and average time spent in the store per customer were simply guessed out of common sense. Although these parameters played a minor role in this study, it is not possible to know if they would significantly influence the outcomes of the simulation and the DOE if they were estimated with greater accuracy (e.g., by carrying out a survey).

The next limitation is associated with the assumptions made to develop the simulation model. In other words, the parameters such as for example theft as a percentage of turnover were primarily obtained from the studies of Bamfield (2004), Moraca et al. (2015), and Hollinger & Davis (2002, p. 4). Since these researchers limited their observations to Europe and the US, the conclusions of this thesis might be geographically limited. That means, the same precision of the results obtained here might not be met in other places like East Asia or Latin America.

The final limitation to be stated is concerning the factorial design. Due to time constraints and lack of reliable data regarding positive discrepancies in the IRI context, a 2^k factorial design with three replications was carried out. The problem with this approach is that it assumes that the relationship between the two analyzed factor levels is linear. If this is not the case, the conclusions might suffer some precision problems. The other limitation within this topic is that only three repetitions were analyzed. More repetitions could potentially reduce the error within samples and, by doing so, a better-defined Pareto chart could have been obtained in comparison to the one presented in Figure 32. The last limitation to be mentioned targets the chosen factor levels: if a test to the statistical significance of the results obtained was to be made (comparing the F_calc of each factor and interaction of factors with the F_crit for the conditions of the experiment), then a study to better determine more accurate levels -1 and +1 should be first carried out.

In contrast to the just mentioned limitations, the current work also presents some strengths in comparison to some of the reviewed literature. For instance, since stock counts are prone to human error as mentioned by Iglehart & Morey (1972) in section 2.1.2., this phenomenon was considered in the model.

Another strength of this thesis can be seen when compared to Condea et al. (2012). In the current work, some checkout error, as well as inbound mistakes (e.g., overshipment), are considered and implemented into the simulation model to get a model that more closely resembles the real-world system. Condea et al. (2012) assumed that the detection rate of incoming and outgoing products was 100% because this process was supported by RFID technology.

Finally, the current thesis aimed at studying a very important sector of the economy, the retail industry. Within this industry, one of the great points of pain is the problem of inventory record inaccuracies. These inaccuracies, both positive and negative ones, have the potential to lead to severe issues regarding loss of sales. Although the relationship between negative IRI and sales loss might be direct and easy to see, the same cannot be said about positive discrepancies. However, through simulation modeling and experimental design, it was possible to verify in this thesis that positive IRI does lead to sales loss and that some parameters play a bigger role than others in this matter. Due to some constraints such as time and expertise, the conclusions of this thesis have some limitations as mentioned previously in this section. However, it also tackles some gaps in peer literatures that were identified throughout the literature review. This leaves us with the next steps can be taken in this research field as already identified previously in this section.

**REFERENCES**

ALEXANDER, B., BLAZQUEZ Cano, M. (2020). Store of the future: Towards a (re)invention and (re)imagination of physical store space in an omnichannel context. Journal of Retailing and Consumer Services, 55, Article 101913. https://doi.org/10.1016/j.jretcon-ser.2019.101913

AMAZON.COM. (2019a). Amazon Q1 2019 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2019b). Amazon Q2 2019 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2019c). Amazon Q3 2019 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2020). Amazon Q4 2019 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2021a). Amazon Q1 2021 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2021b). Amazon Q2 2021 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2021c). Amazon Q3 2021 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2022a). Amazon Q1 2022 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMAZON.COM. (2022b). Amazon Q4 2021 Earnings Release. Amazon.com, Inc. https://ir.aboutamazon.com/quarterly-results/default.aspx

AMIT, P. R., KAMESHVARI, B. J. (2012). A Study on Consumer Behaviour of Organized and Unorganized Retail Outlets in Vadodara City. International Journal of Engineering and Management Sciences(3), 466–474.

BAMFIELD, J. (2004). Shrinkage, shoplifting and the cost of retail crime in Europe: a cross-sectional analysis of major retailers in 16 European countries. International Journal of Retail & Distribution Management, 32(5), 235–241. https://doi.org/10.1108/09590550410699233

BANKS, J., CARSON, J. S., NELSON, B. L. (Eds.). (1996). Discrete-Event System Simulation (2nd ed.). Prentice-Hall.

BOSTON CONSULTING GROUP. (2022). The Future of Sales and Marketing Is Here (Executive Perspectives). Boston Consulting Group.

BRADLEY, C., KOHLI, S., Kuijpers, D., SMITH, T. R. (2021). Why retail outperformers are pulling ahead (Consumer & Retail Practices). McKinsey & Company.

CENTERS FOR DISEASE CONTROL AND PREVENTION. (2022). COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/#trends_dailycases

CHUANG, H. H.-C., OLIVA, R. (2015). Inventory record inaccuracy: Causes and labor effects. Journal of Operations Management, 39-40(1), 63–78. https://doi.org/10.1016/j.jom.2015.07.006

CONDEA, C., THIESSE, F., FLEISCH, E. (2012). RFID-enabled shelf replenishment with backroom monitoring in retail stores. Decision Support Systems, 52(4), 839–849. https://doi.org/10.1016/j.dss.2011.11.018

DEHORATIUS, N., RAMAN, A. (2008). Inventory Record Inaccuracy: An Empirical Analysis. Management Science(54), Article 4, 627–641.

ERNST, R., GUERRERO, J.-L., ROSHWALB, A. (1993). A Quality Control Approach for Monitoring Inventory Stock Levels. The Journal of the Operational Research Society, 44, Article 11, 1115–1127.

EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL. (2022). Data on the daily number of new reported COVID-19 cases and deaths by EU/EEA country. https://www.ecdc.europa.eu/en/publications-data/data-daily-new-cases-covid-19-eueea-country

GRUEN, T. W., CORSTEN, D. S., SUNDAR, B. (Eds.). (2002). Retail Out-of-Stocks: A Worldwide Examination of Extent, Causes and Consumer Response [Special issue], 76. Grocery Manufacturers of America.

HÄNNINEN, M., SMEDLUND, A., MITRONEN, L. (2018). Digitalization in retailing: multi-sided platforms as drivers of industry transformation. Baltic Journal of Management, 13(2), 152–168. https://doi.org/10.1108/BJM-04-2017-0109

HAUSRUCKINGER, G., HASSE, F., THIEDE, H. (Eds.). (2003). Optimal Shelf Availability: Increasing Shopper Satisfaction at the Moment of Truth [Special issue]. Efficient Consumer Response, 65.

HILLIER, F. S., LIEBERMAN, G. J. (Eds.). (2015). Introduction to Operations Research (10th ed.). McGraw-Hill Education.

HOLLINGER, R. C., DAVIS, J. L. (Eds.). (2002). The 2002 National Retail Security Survey [Special issue]. National Retail Security Survey, 28. Gainesville, FL 32611-7330. University of Florida. www.soc.ufl.edu/srp.htm

IGLEHART, D. L., MOREY, R. C. (1972). Inventory Systems with Imperfect Asset Information. Management Science, 18(8), 388-394. https://doi.org/10.1287/mnsc.18.8.B388

KANG, Y., GERSHWIN, S. B. (2005). Information inaccuracy in inventory systems: stock loss and stockout. IIE Transactions, 37(9), 843–859. https://doi.org/10.1080/07408170590969861

KOTLER, P., ARMSTRONG, G. (Eds.). (2012). Principles of Marketing (14th ed.). Pearson Prentice Hall. https://doi.org/10.1017/CBO9780511761751.003

LAW, A. M. (Ed.). (2007). Simulation Modeling & Analysis (4th ed.). McGraw-Hill Education.

MONTGOMERY, D. C. (Ed.). (2009). Design and Analysis of Experiments (7th ed.). John Wiley & Sons.

MORACA, B., HOLLINGER, R., CANTRELL, V. (Eds.). (2015). The 2015 National Retail Security Survey [Special issue]. National Retail Security Survey, 22. National Retail Federation.

MOSTAGHEL, R., OGHAZI, P., PARIDA, V., SOHRABPOUR, V. (2022). Digitalization driven retail business model innovation: Evaluation of past and avenues for future research trends. Journal of Business Research, 146, 134–145. https://doi.org/10.1016/j.jbusres.2022.03.072

PAPAKIRIAKOPOULOS, D. (2012). Predict on-shelf product availability in grocery retailing with classification methods. Expert Systems with Applications, 39(4), 4473–4482. https://doi.org/10.1016/j.eswa.2011.09.141

REKIK, Y., SYNTETOS, A. A., GLOCK, C. H. (Eds.). (2019). Inventory Inaccuracy in Retailing: Does it Matter? [Special issue]. Efficient Consumer Response, 39. Efficient Consumer Response.

RINEHART, R. F. (1960). Effects and Causes of Discrepancies in Supply Operations. Operations Research, 8(4), 543–564.

SABANOGLU, T. (2022). Retail market worldwide - Statistics & Facts. Statista. https://www.statista.com/topics/5922/retail-market-worldwide/#dossierKeyfigures

YAW WONG, C., MCFARLANE, D. (2007). Radio frequency identification data capture and its impact on shelf replenishment. International Journal of Logistics Research and Applications, 10(1), 71–93. https://doi.org/10.1080/13675560600813117