3 | New Perspectives on Logistics and Supply Chain Management

Analyzing and Managing Consumer Returns

Data-Driven Approaches for Consumer Returns Management

David Karl

UNIVERSITY OF BAMBERG
CHAIR OF OPERATIONS MANAGEMENT & LOGISTICS

3 New Perspectives on Logistics and Supply Chain Management

UNIVERSITY OF BAMBERG
CHAIR OF OPERATIONS MANAGEMENT & LOGISTICS

New Perspectives on Logistics and Supply Chain Management

Volume 3

Edited by
Prof. Dr. Eric Sucky

University of Bamberg
Chair of Operations Management & Logistics
Feldkirchenstr. 21, 96052 Bamberg

The successful management of internal and cross-company value-added processes, value chains, and entire networks depends in particular on the purposeful application of existing and advanced methods and concepts of production and logistics management and operations research, the use of innovative information and communication technologies, and theoretical and practical findings in process management. The series aims to publish new research results in the areas of production and logistics management, supply chain management, and e-commerce. Publications that contribute to scientific progress in these areas are selected.

Analyzing and Managing Consumer Returns

Data-Driven Approaches for Consumer Returns Management

David Karl

UNIVERSITY OF BAMBERG
CHAIR OF OPERATIONS MANAGEMENT & LOGISTICS

The thesis at hand has been submitted to the Faculty of Social Sciences, Economics, and Business Administration of the University of Bamberg as a dissertation.

Reviewers:

Prof. Dr. Eric Sucky,	University of Bamberg
Prof. Dr. Björn Ivens,	University of Bamberg
Prof. Dr. Frank Schiemann,	University of Bamberg
Prof. Dr. Alexander Dobhan,	Technical University of Applied Sciences Würzburg-Schweinfurt

Date of Defense: October 22, 2024

David Karl 0000-0002-0326-5982

Otto-Friedrich-Universität Bamberg, 2025

URN: urn:nbn:de:bvb:473-irb-109981x
DOI: https://doi.org/10.20378/irb-109981
eISSN: 3052-2021

Preface

Electronic commerce between businesses and consumers (B2C e-commerce) has changed the way people purchase goods and services. However, the convenience of shopping from home also has a dark side: product returns. Returns pose a major and complex challenge, especially in the fashion industry.

Clothing, shoes, and fashion accessories are the most popular items among online shoppers. However, the return rate in the German online fashion retail sector averaged between 26 % and 50 % in 2023, with some retailers reporting rates as high as 75 %. The massive negative environmental impact of returns is considerable.

In online fashion retail, returns often cannot be resold due to image concerns or the fast-paced nature of fashion, even if they are in perfect condition. Clearly, processing returns leads to a doubling of logistics emissions from road, sea, or air freight, as well as massive packaging waste. It is estimated that the carbon footprint of returns is 30% greater than that of initial deliveries. In the US, for example, an estimated 24 million tons of CO_2 were generated by product returns in 2022. As mentioned above, it is not uncommon for garments to end up in landfills. In the US alone, returned goods contribute an estimated 4.75 million tons of landfill waste per year.

The success of fashion companies depends significantly on the "tangibility" of their products. Online fashion retailers facilitate this aspect of shopping by offering generous return policies, such as free returns, long return periods, and full refunds. This allows customers to use their homes as dressing rooms. Many online retailers view high return rates as a necessary evil and believe that generous return policies are necessary. While these policies increase sales, they also increase the return rate. This relationship can be described as the trap of lenient return policies. If a lenient return policy is necessary to increase sales, which increases returns, how can online fashion retailers escape this trap?

To resolve this conflict of objectives, the design of returns management must reduce returns without affecting sales or customer loyalty. Returns management is a subdiscipline of supply chain management. It involves cross-institutional planning, implementation, and control of return flows, as well as the associated information and financial flows. The goal is to support profit maximization in the value creation system. To achieve this goal, returns management must fulfill four tasks: (1) processing returns, (2) avoiding returns, (3) preventing returns, and (4) promoting returns. Whether to use processing, avoidance, prevention, and/or promotion measures largely depends on the competitive strategy pursued, loss of value over time, and net return value. Returns management can be divided into preventive and curative components. Preventive returns management involves influencing the volume of returns before they are made. Online retailers can use different data sources (e.g., transaction data) to generate information that supports these tasks.

A data-driven management approach means making decisions based on data analysis. This approach focuses on using numbers, metrics, and insights from various sources to improve strategic and operational decision-making. This approach has many benefits for all organizations, regardless of their industry. However, the success of a data-driven approach hinges on the quality of the data collected and the effectiveness of its analysis and interpretation.

Based on these considerations, David Karl's dissertation aims to provide data-driven insights into consumer returns. David Karl therefore formulated the following main research questions for his dissertation:

1. What is the actual extent of consumer returns in e-commerce? Why is it essential to employ data-driven consumer returns management?
2. What is the current state of research regarding consumer returns forecasting, and how can such forecasting be implemented in business practice?
3. How and to what extent do different return-related factors influence consumers' purchase and return decisions?

This dissertation addresses these questions from a scientific perspective, considering practical applicability. This contributes to a profound understanding of data analysis and data-driven management in the context of returns management. First, an in-depth understanding of the phenomenon is gained by documenting and researching the status quo of consumer returns. Second, data-driven analysis and prediction approaches are reviewed, applied, and evaluated. Third, factors that influence consumer purchasing and return behavior are examined. Influencing variables that predict consumer returns and management decisions, such as return policies, are investigated in particular.

This dissertation impressively reflects David Karl's exceptional technical and methodological expertise. It shows that he is an expert in all methods of data analysis and uses them to thoroughly illuminate the issue at hand.

Overall, David Karl's work significantly contributes to scientific progress in the context of returns forecasting and returns management as a whole. Additionally, his work is highly relevant to business practice.

Univ.-Prof. Dr. Eric Sucky

Acknowledgements

„Rome wasn't built in a day" – John Heywood

The present doctoral thesis is the result of several years of research and teaching at the Chair of Operations Management and Logistics at the University of Bamberg. As with the building of Rome, the completion of this work is the outcome of a prolonged, intricate, and at times demanding process to which several individuals have made invaluable contributions.

First, my special thanks go to my doctoral supervisor and first reviewer, Prof. Dr. Eric Sucky, who consistently provided guidance and motivation. Without his belief in this work, without his ability to repeatedly convey this belief to me, and without his constructive and creative feedback, this work would likely have remained fragmentary.

I would also like to thank Prof. Dr. Björn Ivens, Prof. Dr. Frank Schiemann, and Prof. Dr. Alexander Dobhan for readily accepting their respective reviewer duties, and for ensuring a smooth review process.

Next, I would like to express my gratitude to my former colleague, Dr. Björn Asdecker, with whom I had the privilege of collaborating on the majority of my publications. As my Master's thesis supervisor, Dr. Asdecker was the first to suggest that I should consider undertaking a doctoral thesis. Moreover, he regularly offered direction and advice, indicating the subsequent steps to be taken.

The remaining co-authors, Kilian Vornberger and Dr. Christine Feddersen-Arden, should also be thanked for the excellent collaboration.

Finally, I would like to extend my thanks to all other former colleagues who helped me put together the pieces neccesary for a project like this: Anja Böhnlein, Sabine Ebner, Dr. Vanessa Felch, Sophie Kurbjuhn, Roswitha Laukenmann, Caterina Rauh, Christiane Reinshagen, Luiza Schöffel, Dr. Christian Straubert, Dr. Jonas Wiese, and Dr. Immanuel Zitzmann. Indeed, working with you always felt like family. All of you have made a substantial contribution to the success of this work!

David Karl

Abstract

Online shopping has experienced substantial growth, with the COVID-19 pandemic further accelerating this trend. Since e-commerce customers cannot physically assess products online, returned items are part of the e-commerce business model. Consumer returns are associated with various operational challenges, are costly for retailers and impact the environment due to the additional transport emissions and resource waste they produce. Thus, retailers must actively manage returns in advance to secure profitability and limit their CO_2 footprint. Various data sources (e.g., transaction data) enable retailers to generate helpful information to support these tasks.

Based on these considerations, this dissertation aims to generate data-driven insights into the field of consumer returns. First, an in-depth understanding of the phenomenon is gained through the documentation and exploration of the status quo of consumer returns in Germany. Second, data-driven analytic and predictive approaches are reviewed, applied, and evaluated. Third, influencing factors for consumers' purchase and return behavior are subsequently examined.

With a total of nine papers published in national and international journals or conference proceedings, this dissertation addresses these issues from a scientific perspective, without losing sight of practical applicability, thus contributing to a better understanding of data analysis in the context of consumer returns management.

Table of Contents

List of Publications

Paper I (AK22)	Asdecker, B. & Karl, D. (2022). Shedding Some Light on the Reverse Part of E-commerce: A Systematic Look Into the Black Box of Consumer Returns in Germany. *European Journal of Management, 22*(1), 59–81. DOI: 10.18374/EJM-22-1.4
Paper II (KA21)	Karl, D. & Asdecker, B. (2021). How Does the Covid-19 Pandemic Affect Consumer Returns: An Exploratory Study. In *Proceedings of the 50th European Marketing Academy Conference*, Valencia, Spain. https://proceedings.emac-online.org/pdfs/A2021-94689.pdf
Paper III (K24a)	Karl, D. (2025). Forecasting E-Commerce Consumer Returns – A Systematic Literature Review. *Management Review Quarterly, 75*(1), 1–56. DOI: 10.1007/s11301-024-00436-x
Paper IV (K24b)	Karl, D. (2024). Evaluating Advanced Product Return Forecasting Algorithms – a (Meta-)Review Integrating Consumer Returns Research. *International Journal of Business Forecasting and Marketing Intelligence, 9*(2), 213–241. DOI: 10.1504/IJBFMI.2024.137648
Paper V (K16)	Karl, D. (2016). Data Mining im Retourenmanagement: Evaluation von Retourenmengenprognosen anhand der Transaktionsdaten eines Schuh- und Bekleidungsversandhändlers. In *Mobility in a globalised world 2015* (pp. 190–213), Landsberg a. L., Germany. DOI: 10.20378/irb-40411
Paper VI (AK18)	Asdecker, B. & Karl, D. (2018). Big data analytics in returns management – Are complex techniques necessary to forecast consumer returns properly? In *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics 2018* (pp. 39–46), Valencia, Spain. DOI: 10.4995/CARMA2018.2018.8303
Paper VII (AKS17)	Asdecker, B.; Karl, D. & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In: *Proceedings of the 50th Annual Hawaii International Conference on System Sciences* (pp. 4192–4201), Waikoloa Village, USA. DOI: 10.24251/HICSS.2017.507

Paper VIII (KAF22)	Karl, D.; Asdecker, B. & Feddersen-Arden, C. (2022). The Impact of Displaying Quantity Scarcity and Relative Discounts on Sales and Consumer Returns in Flash Sale E-Commerce. In *Proceedings of the 55th Hawaii International Conference on System Sciences* (pp. 4569–4578), Maui, USA. DOI: 10.24251/HICSS.2022.556
Paper IX (KVA22)	Karl, D.; Vornberger, K. & Asdecker, B. (2022). Return Policy Leniency Impacting Customers' Purchase Intention – A Viable Strategy for E-Tailers? In *Proceedings of the 13th European Marketing Academy Regional Conference*, Kaunas, Lithuania. https://proceedings.emac-online.org/pdfs/R2022-111790.pdf

List of Figures

List of Tables

List of Abbreviations

AdaBoost	Adaptive Boosting
AI	Artificial intelligence
AIC	Akaike information criterion
AISeL	Association for information systems electronic library
AMCIS	American conference on information systems
ANOVA	Analysis of variance
A-P	Aggregate-predict
ARIMA	Autoregressive integrated moving averages
AUC	Area under curve
AVE	Average variance extracted
B2C	Business-to-consumer
BD	Bangladesh
BDA	Big data analytics
BEVH	Bundesverband E-Commerce und Versandhandel e. V.
BIEK	Bundesverband Paket- und Expresslogistik e. V.
BS/BSU	Business Source Ultimate
C2B	Consumer-to-business
CA	Canada
CART	Classification and regression trees
CI	Confidence interval
CLSC	Close loop supply chain
CN	China
CO_2	Carbon dioxide
COVID-19	Coronavirus disease 2019
CRISP-DM	Cross industry standard process for data mining
CRM	Customer relationship management
DE	Germany
DI	Return difficulty
DIY	Do it yourself
DNN	Deep neural network
DP	Discounted product price
DSR	Design science research
DT	Decision trees
EB	EconBiz

ECIS	European conference on information systems
ERP	Enterprise resource planning
ERT	Ensemble of regression trees
EU	European Union
FA	Perceived fairness
FR	France
HDE	Handelsverband Deutschland
HW	Holt-Winters
HWS	Holt-Winters smoothing
ICIS	International conference on information systems
ICT	Information and communications technology
IN	India
IS	Information systems
IT	Information technology
JS	JSTOR
KPI	Key performance indicator
LDA/DA	(Linear) discriminant analysis
Logit	Logistic regression
MA	Moving average
MAD	Mean absolute deviation
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ML	Machine learning
MLP	Multilayer perceptron
MSE	Mean squared error
NN	Neural networks
OLS	Ordinary least squares regression
OPP	Perceived opportunism
OR	Operations research
P-A	Predict-aggregate
PACIS	Pacific asia conference on information systems
PI	Purchase intention
PPD	Post-purchase dissonance
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
RF	Random forest

RL	Reverse logistics
RMSE	Root mean squared error
RP	Return policy
RQ	Research question
RR	Return rate
RRP	Recommended retail price
SCM	Supply chain management
SD	Science Direct
SEM	Structural equation modelling
SES	Simple exponential smoothing
SG	Singapore
SVM	Support vector machines
SW	Sweden
TAE	Total absolute error
TR	Perceived trust
U.S./US/USA	United States of America
WS/WoS	Web of Science
XGBoost	Extreme gradient boosting
α-RR	alpha return rate
β-RR	beta return rate
γ-RR	gamma return rate

Part A – Synopsis

1 Introduction

"*Returns are basically like athlete's foot.*" – Bergt, 2015, taz.

"*US returns alone create 5 billion pounds of landfill waste and 15 million tonnes of carbon emissions annually, equivalent to the amount of trash produced by 5 million people in a year, according to one estimate.*"
– Schiffer, 2019, Vogue Business.

"*Marketers and sellers hate product returns, but smart companies aren't passively accepting them as bitter pills to be swallowed. They're managing product return policies to maximize future profits.*"
– Petersen and Kumar, 2010, MITSloan Management Review, p. 85.

"*Information is the oil of the 21st century, and analytics is the combustion engine.*" – Peter Sondergaard (Gartner Research), 2011.

1.1 Motivation and Relevance

The above introductory quotes underscore the imperative for retailers to address the challenges surrounding consumer returns. As a possible remedy, data-driven approaches provide insights into consumer intentions and behaviors that can be useful in successfully managing consumer returns.

To further substantiate these statements and the relevance of this field, it should be stated that online shopping has experienced substantial growth, reaching 15 % in recent years, surpassing the expansion of traditional brick-and-mortar retailing (Ratchford et al., 2022). The coronavirus disease 2019 (COVID-19) pandemic and the associated restrictions for brick-and-mortar retailers further accelerated this trend (Shaw et al., 2022). Since e-commerce customers cannot physically assess the quality and fit of products online (Hong & Pavlou, 2014; Shulman et al., 2011), returned items are an inevitable part of the e-commerce business model (Asdecker, 2014). Reported return rates or volumes differ widely according to various factors (e.g., country, age, and industry). Fashion retailers have reported return rates higher than 50 % and, thus, are particularly affected (Saarijärvi et al., 2017). Merchandise worth $248 billion is returned to online retailers in the United States (US), equaling 17.6 % of online sales (National Retail Federation / Appriss Retail, 2023). For Europe, a recent study found that almost one in four e-commerce shipments is returned to the retailer at least as a partial return, which results in over half a billion return shipments per year (Asdecker et al., 2022).

Consumer returns are associated with various operational challenges, such as unavoidable processing costs and planning uncertainties regarding logistics capacities, inventory management, procurement decisions, and marketing activities (Chopra, 2019; de Brito & van der Laan, 2009; Duong et al., 2022; Stock & Mulki, 2009). Retailers incur high costs due to operational expenses and possible write-offs (Brohan, 2005; Yan & Pei, 2019). In addition to financial aspects, as depicted by the second introductory quote above (Optoro, 2018; Schiffer, 2019), consumer returns are responsible for high CO_2 emissions during transport, returns processing, and through disposed returns, thus increasing the environmental impact of e-commerce (Pålsson et al., 2017; Tian & Sarkis, 2022). Referring to the first quote above on returns being like “athlete’s foot” for retailers, many e-commerce firms even avoid reporting specific return figures, treating them as trade secrets (Bergt, 2015).

As is unavoidable for retailers, e-commerce customers in the European Union (EU) have the right to withdraw from a purchase contract within at least 14 days for any reason (Directive 2011/83/EU on consumer rights, 2011). Especially in highly developed e-commerce markets all over the world, due to market competition, free return options are the rule rather than the exception. Thus, referring to the third quote above, retailers must actively manage returns in advance to help secure future profits (Petersen & Kumar, 2010), e.g., by using marketing instruments (El Kihal & Shehu, 2022) or by determining a retailer’s return policy (Janakiraman et al., 2016).

Many of these possible measures rely on adequate data to draw the correct conclusions. Referring to the introductory analogy of data as the oil of the 21st century and analytics as the combustion engine (Stach, 2023), e-commerce generates large amounts of data (primarily transaction and clickstream data, but also customer reviews or related social media posts). These valuable resources need to be made usable through adequate analytic approaches to support value creation processes (Akter & Wamba,

2016) by generating helpful information that uncovers hidden relationships and enables a clearer view of consumers. In a review of big data analytics in e-commerce, Akter and Wamba (2016, p. 189) identified this exemplary future research question (RQ): “How can organizations better use insights from big data to achieve operational excellence?”

Linking consumer returns and purposeful data-driven management, several scholars have called for an intensification of research efforts. A solid 20 years ago, Hess and Mayhew (1997, p. 33) were “[...] look[ing] forward to the new knowledge that will be gained as both academic researchers and practitioners share their advances in the modelling of returns.” They suggested the potential of data analytics in this domain as follows: “Larger datasets also would allow the inclusion of more variables [...] adding new insights into the phenomenon of returns.” (Hess & Mayhew, 1997, p. 33) A few years later, Toktay et al. (2004, p. 64) stated the following: “[...T]here is little research on identifying factors that significantly influence return flow characteristics. Developing a good understanding [...] would enable better decision making for influencing return flows.” As returns are a phenomenon relevant for omnichannel and pure e-commerce, Borba et al. (2020, p. 137) recently emphasized the unanswered future RQ in the context of omnichannel retailing: “Can [...] artificial intelligence [...] enhance the efficiency of the return channel?” As an application of data analytics, Ambilkar et al. (2021, p. 17) recommend deepening the research into returns forecasting as follows: “Future studies can focus on product return forecasting [...]. Various methodologies like [...] machine learning [...] can be applied to solve forecasting problems [...].” In the past, “[...] there [was] very limited research conducted on the applications of BDA [big data analytics] in [...] reverse logistics [...]” (Seyedan & Mafakheri, 2020, p. 17). As “[...] the topic of forecasting consumer returns has received little attention in the academic literature [...]” (Shang et al., 2020, p. 342), this dissertation answers this call for research.

Another promising field of research this work refers to concerns the behavioral patterns of consumers in response to data-driven actions at the marketing-operations interface (Duong et al., 2022). For example, as Abdulla et al. (2019, p. 34) stated, “Future research can examine how retailers could design differentiated policies or employ such data-driven return countermeasures to reduce abusive return behavior. Yet, how customers

perceive such countermeasures or differentiated return policies is another open question." Ambilkar et al. (2021, p. 17) also called for empirical studies analyzing the impact of a retailer's return policy, to substantiate managerial decisions with sufficient data as follows: "Deciding the right return policy for products in anticipation of their return plays a crucial role. So, it is recommended to conduct future studies to develop empirical work to understand this important aspect [...] to evaluate the return policy decision-making challenges."
In conclusion, Frei et al. (2020, p. 1613) highlighted the increasing relevance of the research field studied within this thesis in general as follows: "Product returns are a unique and understudied but growing field in academic research [...] with many opportunities to do applied research and achieve considerable impact on society, the environment, the economy whilst contributing to the academic body of knowledge."

1.2 Research Objectives and Structure of the Thesis

Based on the above considerations and calls for research, this cumulative dissertation aims to generate data-driven insights into the field of consumer returns. First, an in-depth understanding of the problem should be gained, documenting and exploring the status quo of consumer returns. Building on this aim, data-driven analytic and predictive approaches should be applied and evaluated in terms of their contribution to the successful management of consumer returns. In particular, influencing variables that act as predictors of consumer returns and managerial decision-making, such as return policies, should be subsequently examined. Thus, the overarching RQs are as follows:

- *RQ I: What is the actual extent of consumer returns in e-commerce, and hence, why is it essential to employ data-driven consumer returns management?*
- *RQ II: What is the current state of research regarding consumer returns forecasting, and how can such forecasting be implemented in business practice?*
- *RQ III: How and to what extent do different return-related factors influence consumers' purchase and return decisions?*

According to these RQs, the main body of research is split into three separate but interdependent research fields. The publications are grouped according to the RQ to which they refer, by the respective methodologic approach, and by the data on which the examination is based.
Figure 1 presents the structure of this cumulative dissertation. After outlining the topic's relevance and presenting the overarching RQs in *Chapter 1*, *Chapter 2* explains the basic concepts relevant to this field of research and classifies the different publications included. *Chapter 3* presents selected results from the publications, draws overall conclusions, discusses limitations, and outlines ideas for future research based on the findings of this work.
To answer RQ I (*Part B*), the first two publications of this dissertation collect, analyze and interpret empirical data from retailer surveys, describing the field of research and shedding light on specific numbers of consumer returns that have not previously been systematically collected. *Part C* includes four essays focusing on RQ II, reviewing the current state of research on returns forecasting and by evaluating various forecasting approaches within case studies based on real-world data from e-commerce retailers. *Part D* refers to the explanatory variables involved in customer decisions in the context of e-commerce (RQ III): two papers examine drivers and factors responsible for consumer returns. Paper VIII alludes to the interdependence of purchase and return decisions, thereby developing a rationale for examining the impact exerted by different return policies on purchase intention during the pre-purchase stage (Paper IX).

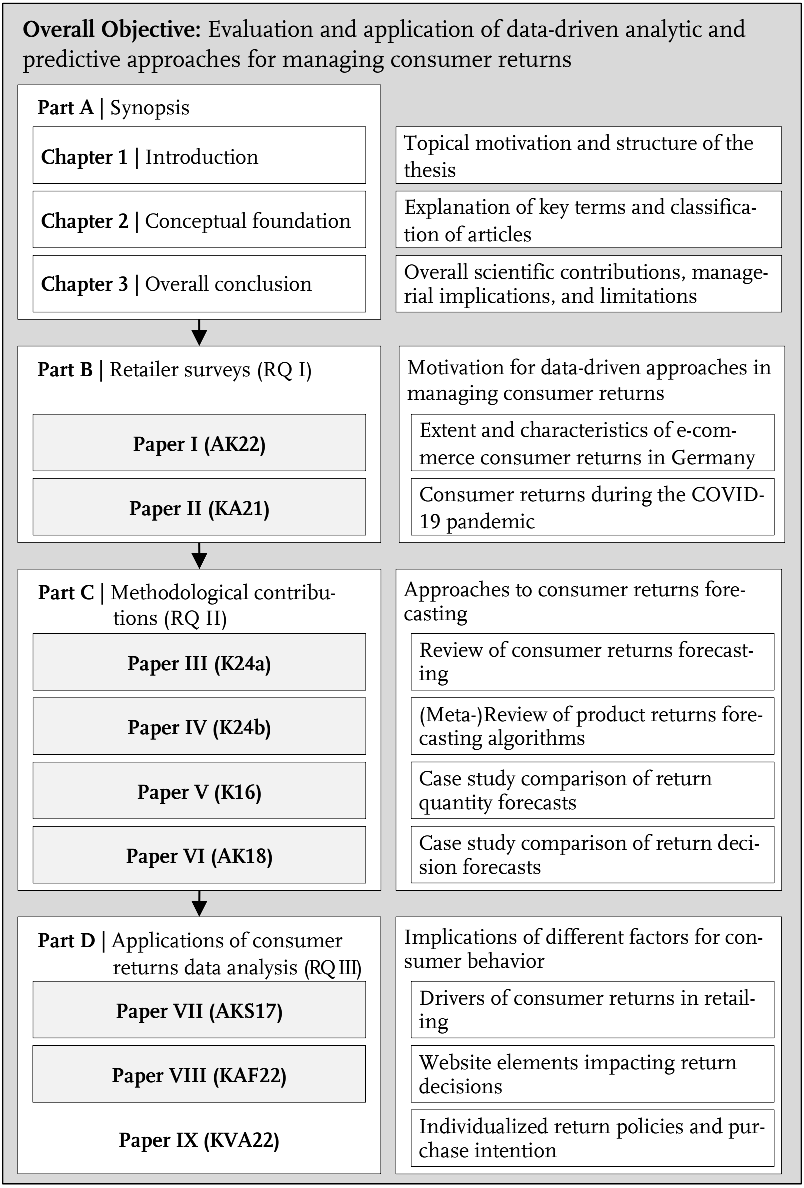

Figure 1. Dissertation structure and objectives.

2 Conceptual Foundation and Publication Overview

This dissertation addresses the application of data-driven approaches to e-commerce returns. The following section presents a conceptual foundation and the link between the topics of consumer returns and data-driven management. Subsequently, the research objectives of the underlying publications are considered from an overall context.

2.1 Consumer Returns and Returns Management

Consumers build the downstream end of a supply chain. Supply chains can be defined as cross-company value creation networks that fulfill customers' requests and demand (Chopra, 2019) by managing goods flows and the associated financial and information flows (Sucky, 2004). Reverse product flows within supply chains to upstream institutions or contracted service providers are called *returns* or *product returns* (Ambilkar et al., 2021). Returns can be categorized into (re)manufacturing returns, distribution returns, and *consumer returns*, sometimes also called customer returns (de Brito & Dekker, 2004; Shaharudin et al., 2015; Tibben-Lembke & Rogers, 2002). Consumer returns describe the product returns from a consumer back to a retailer (Rogers et al., 2002).

In recent years, online consumer returns have accounted for the largest share of returns and thus are highly relevant for e-commerce businesses (Abdulla et al., 2019; Frei et al., 2020). In contrast to returns in closed-loop supply chain management, e.g. end-of-life returns, consumer returns as considered in this thesis are usually sent or given back unused or undamaged shortly after purchase (usually within 14–30 days, while some retailers allow for free returns up to 365 days after purchase). Mostly, this type of return suffers no quality-related defects, should lead to the customer being reimbursed, and is intended to be resold as new ("direct recovery") (de Brito & Dekker, 2004; Melacini et al., 2018). Generally, consumer returns – goods flows from consumers to retailers – arise primarily due to information asymmetries regarding product fit (Hong & Pavlou, 2014) and thus unfulfilled consumer expectations but not due to end-of-life issues, buybacks, or product failure (Shulman et al., 2011).

Returns are an “essential aspect of consumers’ post-purchase decision-making” (Yan & Cao, 2017, p. 211). Consumers’ decision-making process underlying their behavior includes both purchase and returns decisions. Thus, the reasons behind or drivers of returns are an integral part of return-related *consumer behavior research* (Abdulla et al., 2019; Ambilkar et al., 2021). First, the literature differentiates between legitimate and fraudulent or unethical returns (Pei & Paswan, 2018). Pei and Paswan (2018, p. 304) subdivided the reasons for legitimate returns into “lack of needs fulfillment”, “buyers’ remorse”, and “influence of external market”. A more detailed categorization of the reasons for returns was developed by Saarijärvi et al. (2017).

From a retailer’s perspective, consumer returns need to be purposefully managed to satisfy customer needs and meet business objectives. Regarding *returns management*, this dissertation refers to the comprehensive definition by Asdecker (2014, p. 20), reflecting the “[...] planning, execution and control of returned goods flows and the associated information and financial flows, supporting the long-term profit maximization of the value creation system.” Returns management contains *preventive tasks* (prevention, avoidance, and promotion) at the tactical level (Asdecker, 2014), which involve early interventions before a return occurs (Walsh & Möhring, 2017). The literature regarding preventive tasks often refers to a retailer’s *return policy* (Abdulla et al., 2019; Ambilkar et al., 2021). Conversely, *curative tasks* at the operational level aim to achieve the optimal handling of the return process and focus on returns that actually exist or cannot be prevented (Asdecker, 2014; Rogers et al., 2002). Returns management, reverse logistics and closed-loop supply chain management share similarities, especially regarding curative tasks, while preventive returns management tasks are not an inherent part of reverse logistics (Asdecker, 2014). Notably, no clear concept delimitation has prevailed. For instance, in contrast to the conceptualization of this work, Abdulla et al. (2019, p. 562) considered *returns management* to be limited to the “acquisition, processing, and disposition of returns”, separating it from managerial decisions regarding return policies. For a detailed discussion of the delimitation between returns management and reverse logistics, see Asdecker (2011, pp. 427–430).

Returns can be sufficiently planned and controlled only by aggregating and subsequently interpreting data (Röllecke et al., 2018), as illustrated by the following frequently cited manager's proverb: "[I]f you can't measure it, you can't manage it." Accordingly, key performance indicators (KPIs) for managing returns condense operational information to support managerial decisions on a data basis and to reduce the complexity of interpreting business reality (Weber & Wallenburg, 2010). As an example of one of the KPIs used throughout this dissertation, and arguably the most important KPI for returns management, the return rate is calculated as the ratio of returned units to total shipped units. According to Asdecker (2014), the return rate can be differentiated into alpha, beta, and gamma return rates. Each view emphasizes specific business functions. The α-return rate (α-RR, discussed in *Papers I, II, V, VI, and VII*) denotes the ratio of returned to shipped packages (logistics perspective), while the β-return rate (β-RR, included in P*apers I, II, V, VII, and VIII*) expresses this ratio in terms of individual items (marketing perspective). The γ-return rate (γ-RR) describes the financial, value-based ratio of returned to shipped units (finance and controlling perspective).
To calculate such KPIs, the analysis of business data must be carried out. Related streams of research on returns from the retailer's side that include this analytical perspective include the already noted *returns forecasting* and *technology* issues (Ambilkar et al., 2021).

2.2 Data Analytics within the Spheres of E-Commerce Logistics and Consumer Returns

Data analytics for managerial objectives aim to yield commercial advantages and support value creation in the broader sense (Günther et al., 2017). More specifically, value creation pertains to the informational value derived from data and to the transactional value, thus improving operational processes (Elia et al., 2020). Digitized retail platforms present numerous opportunities for data collection, facilitated by digital customer accounts that enable and enhance the efficiency of customer analytics (Akter & Wamba, 2016). Importantly, the challenges and considerations discussed for the e-commerce sphere apply equally to both brick-and-mortar and omnichannel retail (Santoro et al., 2019), both of which benefit from data analytics applications (Hess & Mayhew, 1997). Associated information systems (IS) are necessary to facilitate data-based decision-making (Bernon et al., 2011). The overall goal of management activities within the retail supply chain is to maintain high customer service levels while minimizing the associated costs (Simchi-Levi et al., 2008). Data analytics capabilities can support this objective and, in general, have been shown to positively impact business growth (D. Q. Chen et al., 2015).

While demand forecasting – an activity that makes use of data analytics – is widely acknowledged as a critical supply chain management activity (Carbonneau et al., 2008; Ge et al., 2019), only a limited number of publications link the process of "returns management" with big data analytics (Barbosa et al., 2018). However, retailers deal with massive amounts of data and should effectively exploit available technological advancements and potential (Robertson et al., 2020). Thus, e-commerce platforms are increasingly using machine learning (ML) and artificial intelligence to generate knowledge for both customers and sellers (Micol Policarpo et al., 2021), with the ultimate aim of satisfying customer needs and optimizing retail operations. Consequently, Big Data and Artificial Intelligence are major topics for the future of B2C e-commerce logistics (Straubert et al., 2019).

As mentioned above, an apparent application of (big) data analytics in the context of e-commerce lies in its integration into demand forecasting systems (Hofmann & Rutschmann, 2018). Information sharing between supply chain institutions through approaches such as collaborative forecasting and replenishment (Raghunathan, 1999) can improve prediction accuracy and thus reduce basic supply chain problems such as the well-known bullwhip-effect (Forrester, 1961; Lee et al., 1997). Drawing on current algorithms and emphasizing consumer behavior, demand forecasting falls within the realm of *predictive analytics* (Kumar & Garg, 2018; Shmueli & Koppius, 2011).

Focusing on the opposite direction of goods flows, due to unforseen customer decisions, consumer returns occur stochastically, and thus, also need to be forecasted. Compared to demand forecasting, returns build the "supply" side of the process (Frei et al., 2022). Returns forecasting is more complex than is demand forecasting because customers individually and hardly predictably decide whether or not to keep a product, and these decisions strongly vary among different products (Tibben-Lembke & Rogers, 2002). Interdependencies with sales (forecasts) and promotional activities additionally increase complexity (Govindan & Bouzon, 2018; Tibben-Lembke & Rogers, 2002). However, in all supply chain stages, forecasting accuracy is crucial to the economic, environmental, and social performance of reverse logistics activities (Agrawal & Singh, 2020), justifying a deeper investigation into this domain.

When generating return predictions, there are multiple objectives to consider, such as determining whether a customer will return an item or a package, when he or she will make this return (Hess & Mayhew, 1997), and in what condition the returned items will be. The timing of returns particularly affects resource planning, order quantity planning, and reverse logistics operations in general (Shang et al., 2020) . The condition of the returned items determines their potential for resale and necessitates appropriate returns processing, including recycling, reprocessing, or disposal steps. These measures subsequently impact process costs and supplier order quantities. Evaluating the package-related return probability (α-RR) aligns closely with curative returns management, as this figure and the predicted number of packages support capacity planning for logistics operations (Asdecker, 2014). Predicting the quantity of returned items or articles is relevant to operational and curative decision-making concerning lot size and order quantity planning since resalable returns directly impact inventory levels (e.g., J. Chen & Bell, 2009; Ketzenberg & Zuidwijk, 2009). Accurate returns forecastign can prevent overstocking, which results in increased inventory costs and extended capital lockup. Conversely, understocks lead to lost sales, opportunity costs, and higher procurement expenses and cause customer dissatisfaction, resulting in the loss of goodwill. Furthermore, item-specific return information can prove instrumental in supplier negotiations to secure discounts, improve quality standards, or lower manufacturing tolerance levels.

In addition to forecasting purposes, data analytics form the basis of informed managerial decisions, for example, regarding customer targeting and personalisation (Maheshwari et al., 2021), customer assistance and support, and, most notably, the evaluation of return-related interventions (Röllecke et al., 2018) and return policies (Yu & Wang, 2008). To maintain a high customer lifetime value (Petersen & Kumar, 2015) and to operate profitabely, it is imperative for retailers to analyze all spheres of business, including sales as well as returns (Abbey et al., 2018).

2.3 Overview and Classification of Publications

The papers included in this dissertation contribute to demonstrating the value of data analytics in the context of consumer returns management, and cover various aspects of the research agenda introduced in Chapter 1. Table 1 lists the individual RQs of the publications.

Paper	Research questions
I (AK22)	What is the extent of returns in e-commerce (e.g., return rates) and how are firms performing (e.g., costs)? How do the most relevant key performance indicators change over time?
II (KA21)	What is the extent of consumer returns in e-commerce and how does the Covid-19 pandemic affect it?
III (K24a)	What key research problems (e.g., forecasting purposes, technological approaches) have been addressed in the literature on forecasting consumer returns over time? What are the publication outlets and research disciplines, research types and methodologies, product categories and industries, data sources and characteristics, relevant forecasting predictors, and techniques and algorithms used to address these key problems? How can returns forecasting be integrated into a taxonomy of machine learning in e-commerce? What are promising or emerging future research directions regarding forecasting consumer returns? What key research problems have been addressed in the literature of forecasting consumer returns over time?
IV (K24b)	Which (advanced) algorithms are used in the literature on forecasting product returns in general and consumer returns in particular and how are these approaches evaluated?
V (K16)	How accurately can retailers predict a customer's return decision based on transaction data using data mining methods? Which methods are best suited for predicting return volumes? To what extent do data mining models for return decision-making enable the derivation of an accurate return quantity forecast?
VI (AK18)	How are simple data analysis methods performing compared to complex, more sophisticated ones when predicting consumer returns?

Paper	Research questions
VII (AKS17)	Which drivers influence the flow of returns in e-tailing and to what extent?
VIII (KAF22)	To what extent do the display of quantity scarcity messages and price discounts impact sales and consumer returns in a flash sale e-commerce environment, and how does it influence profitability?
IX (KVA22)	How does return policy leniency influence an e-commerce customer's purchase intention?

Table 1. Overview of the RQs of the respective papers.

The papers in the first part (*Part B*) of this dissertation focus on *exploring and describing the phenomenon* of returns and reveal the need for the holistic and data-driven management of consumer returns. Methodologically, *Papers I and II* collect data by surveying retailers. As demonstrated by the systematic literature reviews within both of these papers, in the past, little was known about the extent of returns and their economic and ecological consequences, as well as their changes over time. Therefore, the first paper presents the results of a long-term trend study conducted among German e-tailers (first survey: n = 143; second survey: n = 68) from various product group clusters. Regarding time-related evolution, in recent years, significant sales shifted from stationary retail to e-tailers not only due to but also further accelerated by the COVID-19 pandemic, which has placed major constraints on brick-and-mortar retailers in most industrialised countries. The second paper systematically reviews the literature on the extent of consumer returns and conducts an exploratory survey (n=103) to capture product category-specific return rates before and during the COVID-19 pandemic, exploring the influences of this crisis on consumer return behavior.

The second part (*Part C*) of this dissertation focuses on *forecasting* as a crucial application of data analytics in consumer returns management. ML techniques have unlocked opportunities for improved forecasting in the last few years. *Papers III and IV* systematically review the current state of research regarding returns forecasting. *Paper III* investigates the research field of consumer returns forecasting in detail, including bibliometric analyses and different factors causing or promoting returns, which

build the exogenous variables of returns forecasting models. *Paper IV* contrasts the former reverse logistics and close-loop supply chain perspective of the returns forecasting field with research on consumer returns forecasting: current algorithms for forecasting returns in e-commerce and the metrics to assess different forecasting approaches are integrated into the product return literature. After synthesizing four literature reviews, a rigorous systematic review is conducted, adding 28 publications related to consumer returns to the forecasting literature on product returns. Based on the results, research propositions are derived within *Papers III and IV*. As the fashion industry is most affected by the phenomenon of consumer returns compared to other industries (see *Paper I*), *Paper V* performs a case study of a German shoe and apparel mail-order company. This paper uses extensive historical transaction data (n = 481.092 ordered items) to generate data mining models. In addition to predicting future return decisions at the time of ordering, from a logistics perspective it is necessary to forecast the expected number of returned parcels, in order to plan the returns management process capacity adequately. Compared with those of time-series forecasting methods, return volume predictions calculated based on ML models allow for conclusions about the benefits of data mining in this context to be drawn.
Although a vast number of ML algorithms can perform such tasks, small and medium-sized e-tailers frequently lack the capabilities and resources to employ complex techniques. Against this background, *Paper VI* evaluates the performance of several techniques for predicting return decisions at the shipment level that differ in application complexity, again using the dataset underlying *Paper V,* but at a different level of aggregation. For model assessment, both *Papers V* and *VI* build on some of the performance indicators that the review in *Paper IV* revealed to be helpful and use ML approaches (e.g., neural networks) that were shown to be common in *Paper III*.

The third part (*Part D*) of this dissertation empirically examines selected factors and return drivers in more detail. Despite increasing interest in this issue, few studies have published empirical results on the drivers of consumer returns. *Paper VII* uses linear and logistic regression models to explore an extensive dataset from an online apparel retailer. Previously untested hypotheses regarding return reasons are formulated and tested, building on established theories and anecdotal information. While *Paper VII* focuses on classical e-commerce, *Paper VIII* is based on sales data from a German flash sale e-tailer. Flash sale e-commerce is a competitive business with low margins. To generate as many orders as possible, techniques that stimulate impulse buying behavior are often used. Two specific instruments that can contribute to impulse buying – the website display of scarcity notifications and the relative discount provided – and their impact on sales and returns are analyzed. Subsequently, a quantitative model synthesizes both perspectives – sales and return decisions. In this way, *Paper IX* ties directly into the results of *Paper VIII*, as the impact of pre-purchase, preventive or promotional measures concerns both purchases and returns. To reduce the return volumes described in *Papers I and II*, retailers can subject high-returning customers to different return policies. In addition to return behavior, purchase intentions can be affected. In an online survey of 197 consumers, this study examines the direct and indirect effects of return policy leniency (including available payment options, as discussed in *Paper VII*) on purchase intentions. Structural equation modeling reveals the influences of other adjacent constructs in consumer research, such as perceived trust, fairness, opportunism, and return difficulty.

To summarize, the publications included in this dissertation can be classified according to their different topical focuses, different research methodologies used or evaluated, and different data used for analysis (see Table 2). Regarding the papers' foci, Part B of this thesis focuses primarily on descriptions about return performance indicators and customers' reactions to various situational changes to illustrate the relevance of the research domain. Part C is dedicated to returns forecasting, including various ML algorithms. Forecasts require an understanding of the reasons and drivers of returns, which are discussed in more detail in Part D, including a discussion of the behavioral consequences of return-related managerial decisions.

From a methodological point of view, the first four papers include systematic literature reviews. In line with their exploratory nature, the first two papers additionally present extensive descriptive statistics. A comparative

quantitative case study approach is the basis for *Paper V and VII*, evaluating different regression and ML models. *Papers VII and VIII* apply regression technique as main research method, in order to verify various hypotheses. *Paper IX* analyzes the data utilizing a structural equation model, again testing different hypotheses.
The dissertation is based on data from three retailer surveys conducted via online questionnaires (*Papers I and II*), three different real-world e-tailer datasets (*Papers V–VIII*), and a consumer survey (*Paper IX*).

		Subject Focus				**Main Research Methods**					**Data**		
Paper		Problem description	Forecasting and ML	Return reasons and drivers	Purchase intention	Systematic literature review	Descriptive statistics	Comparative quantitative case study	Regression models	Structural equation modeling	Retailer survey	Real-world e-tailer data	Consumer survey
I	AK22	✓				✓	✓				✓		
II	KA21	✓				✓	✓				✓		
III	K24a		✓	✓		✓							
IV	K24b		✓			✓							
V	K16		✓	✓				✓				✓	
VI	AK18		✓	✓				✓				✓	
VII	AKS17			✓					✓			✓	
VIII	KAF22			✓	✓				✓			✓	
IX	KVA22			✓	✓					✓			✓

Table 2. Classification of publications.

3 Results and Conclusions

First, the overall research contribution is outlined in Section 3.1. Section 3.2 presents the managerial implications derived from the results. A brief overview of the limitations of the work and an outlook on the need for further research concludes this dissertation in Section 3.3.

3.1 Main Results and Overall Research Contributions

This dissertation contains both *exploratory* and *confirmatory* research. The contribution of *exploratory* research in general is to shed light on phenomena that have not been researched in depth before and to generate ideas and hypotheses that will be verified in the future (Swedberg, 2020). For standard exploratory studies, multimethod approaches and standard statistical procedures are recommended (Swedberg, 2020, p. 41). Exploratory insights are the basis for building and developing theory. As referenced in *Paper I and Paper VIII*, concerning theoretical contributions, Whetten (1989) distinguished (1) factors that explain a phenomenon, (2) the relationship between these factors, (3) logical justifications for altered views, and (4) conditions that limit generalizability. Based on this understanding, some of the publications present theoretical contributions and additions to existing models, especially in contributing to the "what" and "why" questions for various phenomena (e.g., *Papers I and II*, respectively).

In addition to theoretical contibutions, empirical contributions add knowledge to "a phenomenon not previously covered by research or covered to a limited extent" (Ågerfalk & Karlsson, 2021, p. 65). From a *confirmatory* perspective, empirical findings can, but do not necessarily need to fulfill the requirement of generating new theory (Tsang, 2013). Primarily, the empirical papers (*Papers VII, VIII, and IX*) included in this thesis draw on existing models and theories, attempting to falsify them with high-quality empirical evidence, or by strengthening their validity (zur Shapira, 2011) and confirming the current state of research.

Beyond the individual contributions detailed in the respective papers, the dissertation, in its cumulative nature, adds to the consumer returns literature as a whole. One of the key features of this dissertation is the use of various methodologies to generate subsequent findings, the use of various empirical data sources (retailers and consumers) to include multiple perspectives, and the use of diverse secondary data to enhance the generalizability of the case study findings. This diversity also relates to the publication outlets chosen, which reflect the interdisciplinarity of the consumer returns research field, ranging from general management and marketing-related aspects to logicstics aspects, as well as integrating information systems research by including topics such as ML.

The following section answers the RQs formulated in Section 1 by highlighting and merging selected results and conclusions from the publications included in this dissertation. In that sense, the main research contributions elaborated by the respective publications are outlined.

- *RQ I: What is the actual extent of consumer returns in e-commerce, and hence, why is it essential to employ data-driven consumer returns management?*

First, this dissertation underlines the relevance of the research field of consumer returns by systematically collecting extensive empirical descriptive statistics on the number of returns with which retailers have to deal before (*Paper I*) and during the COVID-19 pandemic (*Paper II*). Business-relevant figures such as return rates, return costs, return times, or recovery options elucidate an issue about which retailers are often reluctant to communicate (see the introductory quote regarding “athlete's foot”; Bergt, 2015).

In this way, this dissertation contributes to creating a more accurate, nuanced, and scientifically rigorous characterization of the phenomenon of returns through an explorative approach. Previously, the numbers reported in publications were fragmented or based on single case studies, as documented by a well-structured systematic literature review (*Papers I and II*). Both of these studies substantiate the idea that consumer returns are a massive challenge for e-tailers. The numbers published, e.g. regarding the effects of the COVID-19 pandemic, build a benchmark when continuing and extending the research in this field. The need for further research is especially evident in fashion e-commerce, where *Paper I* confirms the suspected α-return rates of approximately 50 %. However, the

average costs for a returned item (approximately € 2 each for transportation and processing) are significantly lower than the estimates from previous literature. This result is attributable to the large relative share of returned fashion articles, for which below-average costs and faster-than-average return processing are observed. Obviously, handling returns is a core element of the fashion e-commerce business model, prompting retailers to improve process and cost efficency in recent years.
Furthermore, based on the indicators collected, the environmental footprint of return shipments in Germany is systematically quantified to a lower bound of approximately 200,000 tons of CO_2 per year, and the consumption of resources through the disposal of returns can be estimated. In addition, *Paper II* describes a significant decline in return rates observed in the middle of the COVID-19 pandemic for the fashion and interior product groups. Numerous possible causes of this decline are contrasted, which can stimulate further interventions from a retailer's perspective after the pandemic; e.g., avoiding selection orders or attracting new customer groups may be effective in reducing the return numbers.

- *RQ II: What is the current state of research regarding consumer returns forecasting, and how can such forecasting be implemented in business practice?*

Second, this dissertation presents the current state of research regarding forecasting consumer returns with a focus on predictors, advanced algorithms and assessment metrics applied in the field of e-commerce, outlining several paths for research regarding forecasting (*Papers III* and *IV*). Based on real-world retailer data, forecasting models for different purposes are evaluated and compared (*Papers V* and *VI*), adding to the literature by contrasting different prediction models. Furthermore, recommendations are generated as to which forecasting techniques are valuable and appropriate under which circumstances, especially with regard to small and medium-sized e-tailers.
The case study performed within *Paper V* shows that statistical data aggregation from item-level predictions to package-level predictions improves forecasting accuracy. While no prediction algorithm shows superior accuracy under all conditions (*Paper III*), ensemble models prove to be more

accurate and robust than single ML models (*Paper V*), thus confirming the state of research. While *Papers V* and *VI* also confirm the expectation that complex methods outperform simple methods, they demonstrate that correctly specified naïve approaches (e.g., binary logistic models for return decisions) perform almost on par with sophisticated ML algorithms in terms of their technical forecasting accuracy. This finding suggests that even simple forecasting implementation can be helpful for business practice, thus answering the "*why*" question for a somewhat controversial view and making a *minor methodological contribution* according to the typology of Bergh et al. (2022, p. 1837). This insight also calls for future hypothesis testing regarding the interrelationship between forecasting accuracy and business performance.
In addition, time-series forecasting techniques are not suitable for return volumes, while (smoothed) historical order and returns information, even when used as the only predictors, allow for high forecasting precision (*Paper V*).

- *RQ III: How and to what extent do different return-related factors influence consumers' purchase and return decisions?*

Third, to summarize the factors and relationships investigated within this part of the dissertation, Table 3 provides the results of the hypotheses tested in *Papers VII, VIII, and IX*.
This dissertation examines two perspectives on influencing factors to paint a more holistic picture. On the customer side, aspects such as basket composition are shown to interdepend on return behavior (*Paper VII*), whereby, for example, a clear distinction must be made regarding selection orders; there are significant effects for size- and style-related selection orders, while such effects are not evident for color-related selection orders (*Hypothesis VII.H1.2*). In contrast, on the retailer side, the information given on the online store website (*Paper VIII*) and the retailer's return policies, e.g., available payment methods (*Paper VII, Hypothesis VII.H5*), influence consumer behavior, and therefore are effective methods for actively managing returns. Confirming the results from research on forecasting (*Paper V*), historical return patterns are strongly associated with future return behavior (*Hypotheses VII.H6 and VII.H7*).

In conclusion, the studies reveal some of the underlying factors for returns and *how* these variables are related. The methodological approach – using real-world retailer data – distinguishes both studies from other empirical work regarding return drivers and consumer returns in general (e.g., Leeuw et al., 2016), which are often based on survey methods. Furthermore, integrating both the sales and returns perspectives is a unique feature of *Paper VIII*.

As part of the efforts to master the high number of returns, returns management measures must always be considered holistically, as their effects on purchasing behavior, for example, can override the main intended effects. In *Paper IX*, the research on return policy and purchase intention is complemented by examining different return policy manifestations. Uncovering the relationships among different prerequisites (return policy, trust, fairness, opportunism, and return difficulty) for purchasing behavior that have been only partially researched in the past thus addresses the "*how*" question. These findings together contribute to the relevance of data-driven management support systems, as changes in the return policy that might be implemented to influence return rates are also found to directly influence customers' purchase intention.

Hypothesis		**Results**
VII.H1.1	There is a positive relationship between the number of additional items and the α-/β-returns rate of an order.	α: ✓ β: ✗
VII.H1.2	a) There is a positive relationship between the number of multiple-items regarding size (i.e., same style and same color) and the α-/β-returns rate of an order.	α: ✓ β: ✓
	b) There is a positive relationship between the number of multiple-items regarding color (i.e. same style and same size) and the α-/β-returns rate of an order.	α: ✗ β: ✗
	c) There is a positive relationship between the number of multiple-items regarding style (i.e. color and same size) and the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H2	There is a positive relationship between the mean item price and the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H3.1	The use of coupons codes has a positive effect on the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H3.2	There is a negative relationship between the relative value of a coupon and the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H4	Adding free gifts to an order has a negative effect on the α-/β-returns rate of an order.	α: ✗ β: ✓

Hypothesis		Results
VII.H5	Post-delivery payment options such as invoicing have a positive effect on the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H6	There is a positive relationship between the ordered items' aggregated historical β-returns rate and the α-/β-returns rate of an order.	α: ✓ β: ✓
VII.H7	There is a positive relationship between the customer's historical α-/β-returns rate and the α-/β-returns rate of an order.	α: ✓ β: ✓
VIII.H1	Displaying product scarcity increases sales in a flash sale e-commerce environment.	✓
VIII.H2	Displaying relative price discounts increases sales in a flash sale e-commerce environment.	×
VIII.H3	Displaying product scarcity increases returns in a flash sale e-commerce environment.	✓
VIII.H4	Displaying relative price discounts increases returns in a flash sale e-commerce environment.	×
IX.H1	Customers' purchase intention is positively associated with return policy leniency.	✓
IX.H2a	Customers' perceived fairness is positively associated with return policy leniency.	✓
IX.H2b	Customers' purchase intention is positively associated with perceived fairness.	✓
IX.H3a	Customers' perceived trust is positively associated with return policy leniency.	✓
IX.H3b	Customers' purchase intention is positively associated with perceived trust.	×
IX.H3c	Customers' perceived trust is positively associated with perceived fairness.	✓
IX.H3d	Customers' perceived opportunism is negatively associated with return policy leniency.	✓
IX.H3e	Customers' perceived trust is negatively associated with perceived opportunism.	✓
IX.H4a	Customers' perceived return difficulty is negatively associated with return policy leniency.	✓
IX.H4b	Customers' purchase intention is negatively associated with perceived return difficulty.	×

Legend: ✓ = supported; × = not supported; α: package-related; β: item-related

Table 3. Overview and results of the hypotheses examined.

3.2 Managerial Implications

For businesses, the collected data from *Papers I* and *II* can be used for benchmarking purposes (e.g., regarding return rates, return costs, and processing times) and support the improvement in e-tailers' decision-making systems. Therefore, the data of both retailer surveys included in *Paper I* are available as open data via Figshare (Karl & Asdecker, 2022). As e-tailers dispose of items more frequently than they donate them, new business models should be created to develop solutions for a better coordination of donors and grantees to preserve resources.

Papers III and *IV* can guide managers and data scientists in implementing forecasting solutions in the reverse part of the e-commerce business model by providing an overview of the most discussed algorithms and evaluation metrics used in the scientific literature. Additionally, *Paper III* describes many relevant factors causing or at least predicting consumer returns that can be targeted by retailers when designing preventive measures. *Paper V* shows the benefit of data aggregation after prediction, and derives some practice-oriented recommendations as to which methods are suitable for forecasting the return volume. Conversely, *Paper VI* concludes that the binary logistical regression, as the simplest method analyzed, may already provide satisfactory results when forecasting return decisions. The findings indicate that big data analytics is valuable for managing consumer returns effectively and efficiently – even if not the most sophisticated state-of-the-art method is used.

Paper VII provides suggestions for retailers about how different payment methods can influence consumer returns and provides evidence that measures to avoid size selection orders may be beneficial when trying to reduce the number of returns. Nevertheless, analyzing customer and product history regarding return statistics and the derivation of corresponding measures (e.g., assortment decisions or individual customer measures) may be the most tremendous form of leverage for preventive returns management. *Paper VIII* presents a quantitative model to serve as the basis for a decision support system that enables managers to better address the underlying tradeoff between purchase promotions and return consequences. Thus, decision-makers – beyond the flash sale context – should keep in mind both sides of the coin when designing nudges or other preventive measures for e-commerce websites. *Paper IX* improves companies' understanding of how different return policies affect customer behavior regarding purchase intention under the consideration of various influencing factors. Thus, strategies and opportunities for how to favor or penalize certain customer groups (e.g., low-level returners versus opportunistic returners) by individualizing the return policy to increase profits are outlined.

3.3 Limitations and Directions for Future Research

As with any scientific work, the findings from the publications included in this dissertation must be critically reflected, and their limitations must be noted and discussed. Such limitations also constitute a stimulus for further research.

In general, the research in this dissertation focused primarily on the German e-commerce market. Although Germany is the largest national market within the EU and ranks sixth globally (Lipsman, 2019), recent studies indicate that the behavior of customers and retailers in other countries, both inside and outside the EU, may differ from that in Germany (Asdecker et al., 2022). This situation motivates the call for European and international studies comparing different countries regarding the phenomenon of consumer returns, similar to *Papers I* and *II*. Comprehensive data on returns behavior in Europe have already been collected (Asdecker et al., 2022), but have not yet been published in scientific outlets. Subsequently, this call for internationalization also includes research on return drivers (*Papers VII* and *VIII*), as they may also be composed differently across cultural and legal contexts.

Regarding temporal aspects, the trend study on return performance indicators (*Paper I*) needs to be updated and republished regularly. Furthermore, the results from *Paper II* need to be verified to examine whether COVID-19 merely caused a short-term decline in return rates or whether it triggered long-term structural changes in the retail context and among consumers.

Another limitation concerns the fact that the case studies conducted in *Papers V* and *VI* are each based on a single data set from a fashion e-tailer. A replication with further datasets from other retailers, integrating additional product groups or other periods, can help validate the results. In terms of the product groups evaluated, most articles focus on the apparel sector, which is undoubtedly the most affected by returns (see *Paper I*). However, the results cannot be extrapolated to other product groups where customer behaviors are largely different, as demonstrated in *Papers I* and *II*.

Concerning factors influencing returns, only some of the drivers of returns are analyzed, and no claim is made to completeness (*Paper VII*). This situation calls for further studies based on more extensive data to compare the factors already included with additional return drivers. This kind of study can also be placed within a forecasting context, for which *Paper III* hints at the influencing factors examined in the literature.

Paper VIII investigates selected e-commerce website design options. Future studies should consider further options (e.g., De et al., 2013) to paint a more comprehensive picture of the effect of individual measures on sales and returns. Especially in the context of evolving technologies such as virtual fitting tools, which are based mainly on ML and AI algorithms, a separate research stream has emerged (e.g., Gustafsson et al., 2021; Yang & Xiong, 2019): What is the impact of such tools on returns? Another research stream relevant in this context is (digital) nudging in e-commerce (e.g., Dennis et al., 2020), which should also consider the effect of nudges on returns (e.g., Seewald et al., 2019). Gamification also offers promising avenues that could be applied to preventive returns management by influencing customer behavior (Rauh et al., 2024; Rauh & Asdecker, 2023).

Another limitation of the publications is the transfer of managerial implications back into business practice yet to be shown (*Papers IV* to *XI*). While recommendations and implications are derived, these suggestions have not yet been cross-checked and validated. For example, building on *Paper VIII*, a subsequent study can evaluate whether the targeted hiding of the remaining stock for some customers or products leads to the expected change in purchase and return behavior.

In general, the fundamental focus of this work is on e-commerce returns. In economically challenging times for brick-and-mortar retailers, the omnichannel business model is gaining importance, which also entails consequences for returns (ordered online, tried on or returned in the store, and vice versa) – another viable path for further research insights (e.g., Borba et al., 2020; Nageswaran et al., 2020).

Finally, future research in the context of consumer returns should be strongly motivated from a sustainability perspective (Frei et al., 2020), to manage the ecological footprint of e-commerce. As a form of groundwork, *Paper I* presents a quantification of the ecological impact of the phenomenon of returns, which ought to be researched in more detail in future studies, particularly when it comes to emission reduction efforts. On the one hand, additional topics for scientific discussion regarding operational and curative tasks of returns management include reusable packaging in forward and reverse logistics, which is not discussed within the present work. However, on the other hand, for preventive tasks, measures such as repeatedly discussed return fees (Bower & Maxham, 2012) need a thoughtful and data-driven examination before and during implementation to avoid jeopardizing a retailer's business model and as not to induce any contrary effects. This dissertation provides a solid groundwork for continued research on these topics.

References

Abbey, J. D., Ketzenberg, M. E., & Metters, R. (2018). A More Profitable Approach to Product Returns. *MIT Sloan Management Review, 60*(1), 71–74.

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Ågerfalk, P. J., & Karlsson, F. (2021). Theoretical, Empirical, and Artefactual Contributions in Information Systems Research: Implications Implied In N. R. Hassan & L. P. Willcocks (Eds.), *Technology, Work and Globalization. Advancing Information Systems Theories* (pp. 53–73). Springer International Publishing.

Agrawal, S., & Singh, R. K. (2020). Forecasting product returns and reverse logistics performance: structural equation modelling. *Management of Environmental Quality: An International Journal, 31*(5), 1223–1237.

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194.

Ambilkar, P., Dohale, V., Gunasekaran, A., & Bilolikar, V. (2021). Product returns management: A comprehensive review and future research agenda. *International Journal of Production Research*, 1–25.

Asdecker, B. (2011). Retourenmanagement – Eine Literaturrecherche. In E. Sucky, B. Asdecker, A. Dobhan, S. Haas, & J. Wiese (Eds.), *Logistikmanagement. Herausforderungen, Chancen und Lösungen. Band II: Proceedings of the 7th Conference "Logistikmanagement"* (Vol. 2, pp. 421–461). University of Bamberg Press.

Asdecker, B. (2014). *Retourenmanagement im Versandhandel: Theoretische und empirisch fundierte Gestaltungsalternativen für das Management von Retouren. Logistik und Supply Chain Management: Bd. 10.* Univ. of Bamberg Press.

Asdecker, B., Felch, V., & Karl, D. (2022). *European Return-o-Meter - Results Part 1: Germany vs. Remaining EU*. http://www.retourenforschung.de/info-results-of-the-european-return-o-meter-published.html

Barbosa, M. W., La Vicente, A. d. C., Ladeira, M. B., & Oliveira, M. P. V. de (2018). Managing supply chain resources with Big Data Analytics: a systematic review. *International Journal of Logistics Research and Applications*, *21*(3), 177–200.

Bergh, D. D., Boyd, B. K., Byron, K., Gove, S., & Ketchen, D. J. (2022). What Constitutes a Methodological Contribution? *Journal of Management*, *48*(7), 1835–1848.

Bergt, S. (2015). *Geheimnis um Rücksendequote im Handel: Wie Fußpilz*. Retrieved May 24, 2023, from https://taz.de/Geheimnis-um-Ruecksendequote-im-Handel/!5215913/

Bernon, M., Rossi, S., & Cullen, J. (2011). Retail reverse logistics: a call and grounding framework for research. *International Journal of Physical Distribution & Logistics Management*, *41*(5), 484–510.

Borba, J. L. G. de, Magalhães, M. R. de, Filgueiras, R. S., & Bouzon, M. (2020). Barriers in omnichannel retailing returns: a conceptual framework. *International Journal of Retail & Distribution Management*, *49*(1), 121–143.

Bower, A. B., & Maxham, J. G. (2012). Return Shipping Policies of Online Retailers: Normative Assumptions and the Long-Term Consequences of Fee and Free Returns. *Journal of Marketing*, *76*(5), 110–124.

Brohan, M. (2005). Online retailers learn to live with that persistent problem of returns. *Internet Retailer*, *7*(4), 44–50.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research*, *184*(3), 1140–1154.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, *32*(4), 4–39.

Chen, J., & Bell, P. C. (2009). The impact of customer returns on pricing and order decisions. *European Journal of Operational Research, 195*(1), 280–295.

Chopra, S. (2019). *Supply chain management: Strategy, planning, and operation* (Seventh edition). Pearson.

De, P., Hu, Y., & Rahman, M. S. (2013). Product-Oriented Web Technologies and Product Returns: An Exploratory Study. *Information Systems Research, 24*(4), 998–1010.

de Brito, M. P., & Dekker, R. (2004). A Framework for Reverse Logistics. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 3–27). Springer, Berlin, Heidelberg.

de Brito, M. P., & van der Laan, E. A. (2009). Inventory control with product returns: The impact of imperfect information. *European Journal of Operational Research, 194*(1), 85–101.

Dennis, A. R., Yuan, L., Feng, X., Webb, E., & Hsieh, C. J. (2020). Digital Nudging: Numeric and Semantic Priming in E-Commerce. *Journal of Management Information Systems, 37*(1), 39–65.

Duong, Q. H., Zhou, L., Meng, M., van Nguyen, T., Ieromonachou, P., & Nguyen, D. T. (2022). Understanding product returns: A systematic literature review using machine learning and bibliometric analysis. *International Journal of Production Economics, 243*, 108340.

El Kihal, S., & Shehu, E. (2022). It's not only what they buy, it's also what they keep: Linking marketing instruments to product returns. *Journal of Retailing, 98*(3), 558–571.

Elia, G., Polimeno, G., Solazzo, G., & Passiante, G. (2020). A multi-dimension framework for value creation through Big Data. *Industrial Marketing Management, 90*, 617–632.

Directive 2011/83/EU on consumer rights (2011). https://eur-lex.europa.eu/eli/dir/2011/83/2022-05-28

Forrester, J. W. (1961). *Industrial Dynamics*. Productivity Press; M.I.T. Press Massachusetts Institute of Technology.

Frei, R., Jack, L., & Brown, S. (2020). Product returns: A growing problem for business, society and environment. *International Journal of Operations & Production Management, 40*(10), 1613–1621.

Frei, R., Jack, L., & Krzyzaniak, S.-A. (2022). Mapping Product Returns Processes in Multichannel Retailing: Challenges and Opportunities. *Sustainability, 14*(3), 1382.

Ge, D., Pan, Y., Shen, Z.-J., Di Wu, Yuan, R., & Zhang, C. (2019). Retail supply chain management: a review of theories and practices. *Journal of Data, Information and Management, 1*, 45–64.

Govindan, K., & Bouzon, M. (2018). From a literature review to a multi-perspective framework for reverse logistics barriers and drivers. *Journal of Cleaner Production, 187*, 318–337.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems, 26*(3), 191–209.

Gustafsson, E., Jonsson, P., & Holmström, J. (2021). Reducing retail supply chain costs of product returns using digital product fitting. *International Journal of Physical Distribution & Logistics Management, 51*(8), 877–896.

Hess, J. D., & Mayhew, G. E. (1997). Modeling merchandise returns in direct marketing. *Journal of Direct Marketing, 11*(2), 20–35.

Hofmann, E., & Rutschmann, E. (2018). Big data analytics and demand forecasting in supply chains: a conceptual analysis. *International Journal of Logistics Management, 29*(2), 739–766.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

Janakiraman, N., Syrdal, H. A., & Freling, R. (2016). The Effect of Return Policy Leniency on Consumer Purchase and Return Decisions: A Meta-analytic Review. *Journal of Retailing, 92*(2), 226–235.

Karl, D., & Asdecker, B. (2022). *Retailers' Survey Data on Consumer Returns in Germany 2014 and 2018.* https://doi.org/10.6084/M9.FIGSHARE.14376794

Ketzenberg, M. E., & Zuidwijk, R. A. (2009). Optimal Pricing, Ordering, and Return Policies for Consumer Goods. *Production and Operations Management, 18*(3), 344–360.

Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. *International Journal of Computer Applications, 182*(1), 31-37.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science, 43*(4), 546–558.

Leeuw, S. de, Minguela-Rata, B., Sabet, E., Boter, J., & Sigurðardóttir, R. (2016). Trade-offs in managing commercial consumer returns for online apparel retail. *International Journal of Operations & Production Management, 36*(6), 710–731.

Lipsman, A. (2019). *Global Ecommerce 2019: Ecommerce Continues Strong Gains Amid Global Economic Uncertainty*. Retrieved January 20, 2021, from https://www.emarketer.com/content/global-ecommerce-2019

Maheshwari, S., Gautam, P., & Jaggi, C. K. (2021). Role of Big Data Analytics in supply chain management: current trends and future perspectives. *International Journal of Production Research, 59*(6), 1875–1900.

Melacini, M., Perotti, S., Rasini, M., & Tappia, E. (2018). E-fulfilment and distribution in omni-channel retailing: a systematic literature review. *International Journal of Physical Distribution & Logistics Management, 48*(4), 391–414.

Micol Policarpo, L., da Silveira, D. E., da Rosa Righi, R., Antunes Stoffel, R., da Costa, C. A., Victória Barbosa, J. L., Scorsatto, R., & Arcot, T. (2021). Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review. *Computer Science Review, 41*, 100414.

Nageswaran, L., Cho, S.-H., & Scheller-Wolf, A. (2020). Consumer Return Policies in Omnichannel Operations. *Management Science*. Advance online publication.

National Retail Federation / Appriss Retail. (2023). *2023 Consumer Returns in the Retail Industry*. Retrieved December 30, 2023, from https://cdn.nrf.com/sites/default/files/2024-01/2023%20Consumer%20Returns%20in%20the%20Retail%20Industry.pdf

Optoro. (2018). *2018 Impact Report*. Retrieved November 14, 2023, from https://www.optoro.com/2018-impact-report/

Pålsson, H., Pettersson, F., & Winslott Hiselius, L. (2017). Energy consumption in e-commerce versus conventional trade channels - Insights into packaging, the last mile, unsold products and product returns. *Journal of Cleaner Production, 164*, 765–778.

Pei, Z., & Paswan, A. (2018). Consumers' legitimate and opportunistic product return behaviors in online shopping. *Journal of Electronic Commerce Research, 19*(4), 301–319.

Petersen, J. A., & Kumar, V. (2010). Can Product Returns Make You Money? *MIT Sloan Management Review, 51*(3), 85–89.

Petersen, J. A., & Kumar, V. (2015). Perceived Risk, Product Returns, and Optimal Resource Allocation: Evidence from a Field Experiment. *Journal of Marketing Research, 52*(2), 268–285.

Raghunathan, S. (1999). Interorganizational Collaborative Forecasting and Replenishment Systems and Supply Chain Implications. *Decision Sciences, 30*(4), 1053–1071.

Ratchford, B., Soysal, G., Zentner, A., & Gauri, D. K. (2022). Online and offline retailing: What we know and directions for future research. *Journal of Retailing, 98*(1), 152–177.

Rauh, C., & Asdecker, B. (2023). Can Gamification Help to Unlock the Full Potential of Preventive Returns Management Measures in E-Commerce? – A First Empirical Analysis. In *Proceedings of the 2023 Pacific Asia Conference on Information Systems,* Nanchang, China.

Rauh, C., Straubert, C., & Sucky, E. (2024). Promoting Sustainable Consumer Behavior in E-commerce: An Empirical Study on the Influence of Gamification on Consumers' Return Motivation. In *Proceedings of the 57th Hawaii International Conference on System Sciences,* Oahu, United States.

Robertson, T. S., Hamilton, R., & Jap, S. D. (2020). Many (Un)happy Returns? The Changing Nature of Retail Product Returns and Future Research Directions. *Journal of Retailing, 96*(2), 172–177.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Röllecke, F. J., Huchzermeier, A., & Schröder, D. (2018). Returning Customers: The Hidden Strategic Opportunity of Returns Management. *California Management Review*, *60*(2), 176–203.

Saarijärvi, H., Sutinen, U.-M., & Harris, L. C. (2017). Uncovering consumers' returning behaviour: a study of fashion e-commerce. *International Review of Retail, Distribution & Consumer Research, 27*(3), 284–299.

Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision, 57*(8), 1980–1992.

Schiffer, J. (2019). *The unsustainable cost of free returns | vogue business.* Retrieved May 19, 2023, from https://www.voguebusiness.com/consumers/returns-rising-costs-retail-environmental

Seewald, A., Wernbacher, T., Pfeiffer, A., Denk, N., Platzer, M., Berger, M., & Winter, T. (2019). Towards Minimizing e-Commerce Returns for Clothing. In *Proceedings of the 11th International Conference on Agents and Artificial Intelligence* (pp. 801–808). SCITEPRESS - Science and Technology Publications.

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data, 7*(1).

Shaharudin, M. R., Zailani, S., & Tan, K. C. (2015). Barriers to product returns and recovery management in a developing country: investigation using multiple methods. *Journal of Cleaner Production, 96,* 220–232.

Shang, G., McKie, E. C., Ferguson, M. E., & Galbreth, M. R. (2020). Using transactions data to improve consumer returns forecasting. *Journal of Operations Management, 66*(3), 326–348.

Shaw, N., Eschenbrenner, B., & Baier, D. (2022). Online shopping continuance after COVID-19: A comparison of Canada, Germany and the United States. *Journal of Retailing and Consumer Services, 69,* 103100.

Shmueli, G., & Koppius, O. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly, 35*(3), 553.

Shulman, J. D., Coughlan, A. T., & Savaskan, R. C. (2011). Managing Consumer Returns in a Competitive Environment. *Management Science, 57*(2), 347–362.

Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2008). *Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies* (3rd ed.). McGraw-Hill Irwin.

Stach, C. (2023). Data Is the New Oil–Sort of: A View on Why This Comparison Is Misleading and Its Implications for Modern Data Administration. *Future Internet, 15*(2), 71.

Stock, J. R., & Mulki, J. P. (2009). Product Returns Processing: An Examination of Practices of Manufacturers, Wholesalers/Distributors, and Retailers. *Journal of Business Logistics, 30*(1), 33–62.

Straubert, C., Asdecker, B., & Zitzmann, I. (2019). Current Trends in B2C E-Commerce Logistics – A Content Analysis. In C. Bierwirth, T. Kirschstein, & D. Sackmann (Eds.), *Lecture Notes in Logistics. Logistics Management* (1st ed., pp. 123–140). Springer International Publishing.

Sucky, E. (2004). *Koordination in Supply Chains: Spieltheoretische Ansätze zur Ermittlung integrierter Bestell- und Produktionspolitiken*. Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-322-81691-7

Swedberg, R. (2020). Exploratory research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *Strategies for social inquiry. The production of knowledge: Enhancing progress in social science* (pp. 17–41). Cambridge University Press.

Tian, X., & Sarkis, J. (2022). Emission burden concerns for online shopping returns. *Nature Climate Change, 12*(1), 2–3.

Tibben-Lembke, R. S., & Rogers, D. S. (2002). Differences between forward and reverse logistics in a retail environment. *Supply Chain Management: An International Journal, 7*(5), 271–282.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

Tsang, E. W. (2013). Case study methodology: causal explanation, contextualization, and theorizing. *Journal of International Management, 19*(2), 195–202.

Walsh, G., & Möhring, M. (2017). Effectiveness of product return-prevention instruments: Empirical evidence. *Electronic Markets, 27*(4), 341–350.

Weber, J., & Wallenburg, C. M. (2010). *Logistik- und Supply Chain Controlling* (6. vollständig überarbeitete Auflage). Schäffer-Poeschel Verlag.

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution? *Academy of Management Review, 14*(4), 490–495.

Yan, R., & Cao, Z. (2017). Product returns, asymmetric information, and firm performance. *International Journal of Production Economics, 185*, 211–222.

Yan, R., & Pei, Z. (2019). Return policies and O2O coordination in the e-tailing age. *Journal of Retailing and Consumer Services, 50*, 314–321.

Yang, S., & Xiong, G. (2019). Try It On! Contingency Effects of Virtual Fitting Rooms. *Journal of Management Information Systems, 36*(3), 789–822.

Yu, C.-C., & Wang, C.-S. (2008). A hybrid mining approach for optimizing returns policies in e-retailing. *Expert Systems with Applications, 35*(4), 1575–1582.

zur Shapira (2011). “I've Got a Theory Paper—Do You?”: Conceptual, Empirical, and Theoretical Contributions to Knowledge in the Organizational Sciences. *Organization Science, 22*(5), 1312–1321.

Part B – Contributions on the Status Quo of Consumer Returns

I Shedding Some Light On the Reverse Part of E-Commerce: A Systematic Look into the Black Box of Consumer Returns In Germany

Authors: Björn Asdecker, David Karl

Publication Outlet: European Journal of Management
(ranked as B in VHB-Rating 2024)

DOI: 10.18374/EJM-22-1.4

published as:

Asdecker, B.; Karl, D. (2022): Shedding Some Light on the Reverse Part of E-commerce: A Systematic Look Into the Black Box of Consumer Returns in Germany. In: *European Journal of Management, 22*(1), pp. 59–81. DOI: 10.18374/EJM-22-1.4

Abstract

In recent years, significant sales have shifted from stationary retail to the internet, which has resulted in the enormous growth of e-commerce. Nevertheless, there is one critical aspect that threatens the success of the business model, namely, consumer returns. While the amount of research on this topic has increased noticeably over the last few years, little is known about the extent of returns and their economic and ecological consequences, as well as about changes over time. Against this background, this paper presents the results of a long-term trend study conducted among German e-tailers from various product group clusters. With an annual turnover of more than $80 billion in 2019, Germany is the sixth-largest business-to-consumer e-commerce market. The study shows that consumer returns are a massive challenge for e-tailers that should not be underestimated. This research adds knowledge about the reverse part of the e-commerce business model. For businesses, the collected data can be used for benchmarking purposes and can help to improve e-tailers' decision-making systems. Societally, this study provides a basis for the quantification of the environmental impact of returns.

Keywords: E-commerce, Consumer Returns, Return Rates, Return Costs, Ecological Impact

I.1 Introduction

Despite a promising outlook with double-digit growth rates ahead (Lipsman, 2019, Torry, 2020), the e-commerce business model faces a number of challenges along its growth path. Among them are consumer returns. To create trust and to encourage consumers to order, companies are granting liberal return policies, which in turn lead to more online returns than those experienced by traditional brick-and-mortar retailing (e.g., Tarn et al., 2003, Giménez & Lourenço, 2008, Xia & Zhang, 2010). Since the costs caused by returns do not rise linearly, but instead rise disproportionately with the return rate, returns management is considered a critical success factor (Asdecker, 2015). Similar to that seen in brick-and-mortar retailing (e.g., Cassill, 1998), the return process can generate customer satisfaction and loyalty in online retailing (Vakulenko et al., 2019). Due to the importance of returns, it is not surprising that the number of available publications on this topic has risen steadily over the past years. Despite the increased recent attention among scholars, comparatively little is known about the extent of consumer returns in e-commerce or about important key performance indicators; this is in part because many firms consider such information to be confidential (Stock & Mulki, 2009). However, such data are necessary for many reasons, for example, to substantiate the relevance of the underlying research, to better prepare decision-making, to make realistic assumptions for parameter values in quantitative and simulation models, and to provide some estimates of the ecological impact of returns.

Against this background, one of the most frequently cited academic works in the context of returns management is the study by Rogers and Tibben-Lembke (1998). However, this reference is rather dated and refers to the predecessor of e-commerce, namely, catalog retailing. The more recent studies with a scientific background are either much less extensive (e.g., Stock & Mulki, 2009, Bernon et al., 2016) or are based on data from a single e-tailer (e.g., Walsh & Möhring, 2017, Araújo et al., 2018). The only other data points are practitioner-related publications. However, such publications have not been subject to a peer-review process and their methodological approach is often neither transparent nor replicable. Thus, it is of great importance to shed further light on the black box of consumer returns. In particular, this study aims to contribute to the following two research questions:

- *What is the extent of returns in e-commerce (e.g., return rates) and how are firms performing (e.g., costs)?*
- *How do the most relevant key performance indicators change over time?*

To answer these questions, we use a longitudinal trend study. To contribute to the theoretical background, this research provides a thorough literature review on returns management key performance indicators and collects its own data to provide descriptive statistics on the extent and consequences of consumer returns. These insights add knowledge to the reverse part of the e-commerce business model. In practical application, the collected data can improve e-tailers' decision support systems and can be used as the basis for benchmarking tools. In terms of general knowledge, this study contributes by quantifying the environmental impact of returns. The following paragraph provides a review of the relevant literature, which emphasizes the current research gap.

I.2 Systematic literature review

Unlike other more general literature reviews (e.g., Abdulla et al., 2019), this work specifically focuses on publications that provide insights concerning the extent of consumer returns and important performance indicators to reflect the objectives of this paper. More specifically, this review intends to synthesize available studies about (1) return rates, (2) cycle times, (3) costs, and (4) recovery options.
Methodologically, the review follows the guidelines of Denyer and Tranfield (2009), who structure the research process in the following five steps: (1) question formulation; (2) locating studies; (3) study selection and evaluation; (4) analysis and synthesis; and (5) reporting and using the results. The first step refers to the research questions, which are already derived in the introduction. The second step involves selecting the databases and defining the search terms. In that respect, five scientific databases were selected, namely Business Source Ultimate (BS), Science Direct (SD), JSTOR (JS), Web of Science (WS) and EconBiz (EB). To identify a large number of relevant publications, very general search terms were used,

which combined ‘product returns’, ‘consumer returns’, or ‘customer returns’ with either ‘rate*’, ‘cost*’, ‘time*’, or ‘recovery’. This resulted in 9323 search hits (see Table I.1). In the third step, duplicates and studies that refer to ‘stock returns’ were removed, leaving 846 publications. Thereafter, titles, abstracts, and keywords were scanned to identify potentially relevant studies, which led to a preliminary set of 52 articles that were read in full length. Publications that addressed only brick-and-mortar retailing were excluded. Once confirmed, the references of each relevant article were searched for further publications of interest. In total, the search led to 44 relevant publications. The list of identified references can be found in the digital appendix of this paper.

Search term	**BS**	**SD**	**JS**	**WS**	**EB**
(‘product returns’ OR ‘consumer returns’ OR ‘customer returns’) + ‘rate*’	242	1475	403	119	52
(‘product returns’ OR ‘consumer returns’ OR ‘customer returns’) + ‘time*’	288	1799	622	193	102
(‘product returns’ OR ‘consumer returns’ OR ‘customer returns’) + ‘cost*’	452	1566	507	291	94
(‘product returns’ OR ‘consumer returns’ OR ‘customer returns’) + ‘recovery’	113	689	124	127	65
Combined results without duplicates & without „stock return“	846				
Preliminary set of most relevant articles	52				
Relevant hits + backward search = Final selection	32 + 12 = 44				

Table I.1. Initial search hits and final selection.

In the fourth step, the publications were classified according to their objective. Most notably, only four of the 44 relevant papers have a scientific background and explicitly aim at collecting statistics on the returns management process (Rogers & Tibben-Lembke, 2001, Stock & Mulki, 2009, Bernon et al., 2016, Araújo et al., 2018). Of the remaining 40 papers, 21 are scientific papers with different overall objectives and 19 are reports by practitioners that were published in nonscientific outlets. Furthermore, the results were categorized based on the four relevant fields of interest, that is, return rates, cycle times, costs, and recovery options. The presentation of the results, which is the final fifth step of the review, can be found in the following paragraphs. The presentation distinguishes between peer-reviewed scientific contributions and nonpeer-reviewed practitioner contributions. The works are arranged chronologically.

Return rates. Of the 44 relevant studies, 40 reported about return rates, of which 22 (55 %) were scientific papers and 18 (45 %) were published in nonscientific outlets. It is important to note that retailers can determine their return rates in different ways, as noted in Asdecker et al. (2017) and Hofmann et al. (2020): (1) the shipment-based return rate, which reflects the logistics perspective and is referred to as the alpha return rate (α), (2) the item-based return rate, which represents the sales perspective and is referred to as the beta return rate (β), and (3) the revenue-based return rate, which shows the financial perspective and is referred to as the gamma return rate (γ). The alpha return rate divides the number of returned shipments by the number of outbound shipments, whereas the beta return rate sets the number of returned items in relation to the number of shipped items. The gamma return rate divides the value of returned items by the value of ordered items. A further classification can be made in terms of the product category under investigation. Table I.2 summarizes the published return rates; the appendix contains further details.

	Scientific papers	**Practitioner reports**
Cross-sectional rates	Rogers and Tibben-Lembke (2001): 6 % [n.s.], Hong and Pavlou (2014): 13 % [n.s.], Abbey et al. (2018): 16 % [γ], Araújo et al. (2018): 3–4 % [γ], Wang and Ansell (2020): 4–40.6 % [n.s.]	Del Franco (2000): 5–7 % [n.s.], Thomas (2001): 5.5–7.2 % [n.s.], Stock (2004): 5.6 % [n.s.], Brohan (2005): <10 % [n.s.], Banjo (2013): 33 % [n.s.], Pur et al. (2013): 13 % [β], Ng and Stevens (2015): 10–15 % [n.s.], Ellis (2017): 33 % [n.s.], Woods (2017): 25–40 % [n.s.], Appriss Retail (2019): 9.6 % [n.s.], Nicola (2019): 30–50 % [n.s.], Jack et al. (2019): 1–75 % [n.s.], EHI Retail Institute (2019b): 20 % [β]

	Scientific papers	Practitioner reports
Fashion rates	Rogers et al. (2002): <40 % [n.s.], Mollenkopf et al. (2007): 20–30 % [n.s.], Petersen and Kumar (2010): 16 % [n.s.], Rao et al. (2014): 15 % [β], Bernon et al. (2016): 8.1–38.2 % [γ], Asdecker et al. (2017): 52.1 [β] / 63.3 % [α], Vilar-Zanon et al. (2017): 25 % [n.s.], Walsh and Möhring (2017): 20 % [α], Asdecker and Karl (2018): 59.8 % [α], Difrancesco et al. (2018): 40 % [n.s.], Sahoo et al. (2018): 15–22 % [n.s.], Shang et al. (2019): 7 % [n.s.]	Del Franco (2000): 30 % [n.s.], Thomas (2001): 40 % [n.s.], Catalog Age (2002): 10–40 % [n.s.], Rösch (2011): 50–70 % [n.s.], Pur et al. (2013): 26 % [β], Stevens (2014): 50 % [n.s.], Ng and Stevens (2015): 30 % [n.s.], Pettypiece (2015): 30 % [n.s.], EHI Retail Institute (2019b): 40 % [β]
Entertainment	Mollenkopf et al. (2007): 5–20 % [n.s.], Griffis et al. (2012): 2–11 % [n.s.], Minnema et al. (2016): 8.3 % [n.s.], Bernon et al. (2016): 6.4–10.3 % [γ]	Del Franco (2000): 25–35 % [n.s.], Douthit et al. (2011): 11–20 % [n.s.], EHI Retail Institute (2019b): 10 % [β]
Leisure	Mollenkopf et al. (2007): 5–20 % [n.s.], Cui et al. (2020): 5–14 % [n.s.], Hofmann et al. (2020): 5.1 % [β] / 7.7 % [α]	Del Franco (2000): 2–5 % [n.s.], Catalog Age (2002): 5–10 % [n.s.], EHI Retail Institute (2019b): 10–30 % [β]
Interior	Rabinovich et al. (2011): 2 % [n.s.], Bernon et al. (2016): 5.0–12.7 % [γ], Minnema et al. (2016): 10 % [n.s.], Sahoo et al. (2018): 7 % [n.s.]	EHI Retail Institute (2019b): 20 % [β]
Others	Daily needs: Mollenkopf et al. (2007): <5 % [n.s.]	Daily needs: EHI Retail Institute (2019b): <10 % [β]
	Legend: [α] = Shipment-based return rate; [β] = Item-based return rate; [γ] = Revenue-based return rate; [n.s.] = Return rate not specified	

Table I.2. Return rates reported in the available literature.

Return cycle times. The least amount of data are available in terms of cycle times. Only three peer-reviewed and two nonscientific publications out of the 44 relevant studies refer to the time dimension of the return process.

Rogers and Tibben-Lembke (2001) find that 42.0 % of returns are processed within one week, and that 85.3 % are processed within one month. According to Stock and Mulki (2009), the time used is divided into receiving (17 % of time), processing (31 %), sorting (26 %), and disposing (26 %). According to Bernon et al. (2016), some retailers operate returns within 24–48 hours because "[...] good performance was considered to be 48 hours from receipt at the processing center to being back in to stock [...]" (p. 596). From a customer's perspective, the process still might appear to take six days due to shipping lags. Two nonscientific publications supplement the above results. Because of high customer pressure, some retailers directly reimburse after the items pass the first inspection (Rösch, 2011). More than two-thirds of the surveyed retailers process returns and refund customers within 48 hours (EHI Retail Institute, 2019b).

Return costs. A few more studies refer to costs. However, the number of publications remains limited (13 of the 44 relevant studies). Among those, three papers are scientific and 10 are practitioner-oriented. Rogers and Tibben-Lembke (2001) estimate that 4 % of a firm's total logistics costs can be attributed to managing product returns. Stock et al. (2006) estimate the processing costs for one item to be approximately \$30–\$35. Stock and Mulki (2009) note that 11 out of 16 participating retailers can recover more than 76 % of their original cost, while two retailers are not able to recover more than 25 % of their original cost. On the practitioners' end, Thomas (2001) determines average manual processing costs of \$32.40. Based on a survey, Brohan (2005) reports that the majority (67.4 %) of companies face processing costs of up to \$10. In addition, 20.4 % are in the range of \$11–\$15, and 12 % come out to more than \$15. Douthit et al. (2011) find the handling and disposition costs of returns to account for 2–3 % of sales. The Economist (2013) quantifies the handling costs for returns from \$6–\$18 per item without considering the potential loss of value, cutting a retailer's profit by up to 50 %. According to Pur et al. (2013), fashion returns can be resold more frequently than others. Furthermore, they note that most returns cost less than 15€ per item (cross-sector average 20€). Fashion returns can entail high discounts of up to 80 % when out of season with even lower earnings for items sold to liquidators (Pettypiece, 2015). According to Ellis (2017), the "[...] expense of processing and shipping [...]

can range from 20 percent to 65 percent of an e-tailer's cost of goods sold [...]". Woods (2017) interviews an UPS executive who estimates returns processing costs to be 10–15 % of the goods' value. Nicola (2019) notes that shipping is the dominant part of the total return costs (11€), followed by diminished product value, the inspection of returned items, and collecting/identifying them. An EHI Retail Institute (2019b) study reports average processing costs of 10€ per returned item. In addition, large differences exist between product categories, with the lowest costs in fashion (5€) and the highest costs in consumer electronics and DIY (15€).

Recovery options. Out of the 44 relevant studies, nine refer to recovery options, out of which four have been peer-reviewed. Rogers and Tibben-Lembke (2001) report that 17.6 % of returned goods are resold as is, 15.5 % are remanufactured, 14.7 % are recycled, 13.9 % are disposed of as scrap, 11.0 % are repackaged and sold as new, 9.0 % are sent to a central processing facility, 6.8 % are donated, 5.6 % are sold to a third-party broker, and 5.1 % are sold at an outlet store. Stock and Mulki (2009) find that 88.3 % of survey respondents return products directly to inventory, while 81.8 % do either destroy or sell at least some returns as scrap. Other options used are repackaging and inserting into inventory (61.4 %), donating to charity (37.2 %), selling on third-party/secondary markets (19.0 %), and repairing/refurbishing (4.1 %), among other options (19.9 %). The case study by Araújo et al. (2018) states that 70 % of returned items are in a good-as-new condition and are directly returned to retail inventory. Another 5 % are forwarded to the supplier for direct review and technical assistance. The remaining 25 % are sold at a reduced price during a knockdown sale or disposed of as scrap. Hjort et al. (2019) note that most items are returned to the shelf, but all retailers also dispose some as scrap. Only one in 12 retailers donates any returns to charity. The scientific studies are complemented by five practitioner publications. According to Brohan (2005) approximately 39.5 % of e-tailers use auction platforms to sell returns at a discount. Pur et al. (2013) find that 38 % of returns can be directly reinserted into the sales channels, whereas 24 % need repackaging, 16 % need refurbishment, 12 % can only be sold as B-stock, and 10 % cannot be resold again. A supply chain consultant interviewed by Stevens (2014) estimates that 70 % of returned merchandise is suitable for resale.

Ng and Stevens (2015) describe that many returns are resold in bulk at only 10–20 % of the original value. Reselling online can lift recovery rates to 40–70 %. Besides, according to an interviewee, up to 20 % of returns cannot be resold due to damages. A recent EHI Retail Institute (2019b) study reports that 70 % can be resold as new. For the rest, companies refer to secondary markets or third-party brokers (53 %), recycling or disposal (47 %), outlet stores (36 %), donations (28 %), a return to the supplier (25 %), or staff sales (21 %).

This exhaustive literature review shows that the available data on consumer returns are rather limited. The majority of cited publications refer to data from a single company. In addition, the comparison of the different values per category shows a wide range. The greatest consensus is that the fashion segment has the highest return rates. However, it is not possible to draw conclusions about how high these return rates actually are in the overall market. The same applies to the values for return costs, cycle times, and the employed recovery options. In addition, the available studies only report mean values. Standard deviations and confidence intervals are not provided. A study on developments and trends could not be found. Given this high degree of uncertainty, we conclude that a systematic empirical study is needed to address this knowledge gap. For that purpose, this research draws on Germany, which is the largest national market within the European Union (EU). Globally, Germany ranks sixth, with a turnover of $81.85 billion in 2019, following China ($1,934.78 billion), the United States ($586.92 billion), the United Kingdom ($141.93 billion), Japan ($115.40 billion), and South Korea ($103.48 billion) (Lipsman, 2019). The following section provides a more detailed description of the German e-commerce market.

I.3 E-Commerce in Germany

E-commerce companies in Germany are well organized within associations. These include the leading association of e-commerce companies, the “Bundesverband E-Commerce und Versandhandel e. V.” (BEVH), and the association of delivery companies, the “Bundesverband Paket- und Expresslogistik e. V.” (BIEK). Apart from the thus far neglected consumer returns, the BEVH and the BIEK collect and publish various data on e-commerce in Germany, of which the most relevant are presented in the following in a condensed form.

According to the BEVH, the revenue generated in German e-commerce totaled 72.64 billion € in 2019 up from 65.10 billion € in 2018 (+11.6 %) (BEVH, 2020). At an average exchange rate of $1.120/€ in 2019 (IRS, 2020), this corresponds to $81.36 billion and is very close to the already mentioned estimate by Lipsman (2019). A major share of revenue comes from the product group clusters «Entertainment» and «Fashion» (see Table I.3).

E-commerce in Germany is characterized by many small and medium-sized merchants who sell their goods in their own shops or via marketplace platforms (e.g., Amazon marketplace, eBay). At the same time, a few online shops concentrate a large share of total revenue. According to the EHI Retail Institute (2019a), the 30 e-tailers with the most revenue account for more than one-third (25.3 billion €; 38.9 %) of the total German e-commerce sales.

Cluster	Associated product categories	Revenue 2019 (%)
Entertainment	Books/e-books/audio books, CDs/DVDs, computer/accessories, games/software including downloads, electric goods/telecommunications	25.84 billion € (35.57 %)
Fashion	Clothing, accessories, shoes	18.71 billion € (25.76 %)
Interior	Furniture/lamps/decoration, home textiles, household goods/appliances	10.92 billion € (15.03 %)
Leisure	DIY/flowers, toys, car/motorbike/accessories, hobby/leisure articles	8.66 billion € (11.92 %)
Others	Anything that is be attributed to the other product groups	8.51 billion € (11.72 %)

Table I.3. Detailed origin of the revenue generated in German e-commerce.

The growth in revenue goes hand-in-hand with an increase in the number of parcels transported (BIEK, 2019, BIEK, 2020). In 2019, logistics service providers in Germany transported 3.65 billion courier, express, and parcel shipments (2018: 3.52 billion shipments). Parcels accounted for 84.2 % (2018: 83.9 %), which led to 3.07 billion parcel shipments (2018: 2.95 billion parcels). Of this figure, 65 % (2018: 62 %) was attributed to the business-to-consumer (B2C) segment, including returns. This equaled 1.99 billion shipments in B2C e-commerce. It increased from 1.83 billion in 2018 and corresponded to a growth rate of 8.7 %. Since the BIEK studies draw on transaction data of the member companies, a high validity and reliability can be assumed. With regard to the distribution of the shipment quantities across the respective product categories, a study on behalf of the BEVH concludes that the share of outbound shipments largely corresponds to the revenue share of the product categories (MRU GmbH, 2014), which is also assumed in the remainder of this paper.
The legal basis for returns in member states of the European Union is the consumer rights directive 2011/83/EU, which has been incorporated into national law. Accordingly, consumers in Germany have the right to revoke a purchase on the internet within 14 days after delivery without providing reasons. In principle, consumers bear the direct costs of returning goods to the retailer unless the retailer voluntarily waives this right or fails to inform the customer about the return costs during the order. Despite the legal possibility, in practice, the vast majority refrain from doing so.
Due to the extensive consumer rights and liberal return policies of the major vendors, which are a de facto standard for the entire German e-commerce market, consumers have generally developed an expectation of customer-friendly policies. This study aims to provide information about the extent of consumer returns. For that purpose, data are collected. This collection process is described in the following paragraph.

I.4 Survey

To improve the available data regarding consumer returns, the authors of this paper designed a longitudinal study to observe the same phenomenon over an extended period. More specifically, a trend study was employed, also referred to as a repeated cross-sectional study, which draws different recurring samples over time from a population that answers the same questionnaire (Babbie, 2005). Consequently, trend studies are "[...] a typical instance of a design that is cross-sectional at the level of the sampling units, but longitudinal at the level of research units" (Taris, 2000, p. 6). It is considered suitable to examine change at the aggregate level (Taris, 2000, Babbie, 2005). Moreover, trend studies are unaffected by panel attrition, e.g., due to job and career changes, which would have made the data collection and analysis considerably more difficult if not completely infeasible.

To support data collection, a website was set up in 2013 (http://www.retourenforschung.de) to provide various information and tools for practitioners (e.g., a review of existing literature, a returns management encyclopedia, a tool for calculating the profit-optimal return rate, and study reports). To gain access, interested individuals had to register for a so-called online expert panel and agree to receive invitations to future studies. Only registrations that used a professional e-mail address and could prove their professional experience in regard to e-commerce returns management were accepted. For this last step, the provided data were counterchecked with information that was found on professional social networks (Xing, LinkedIn) to ensure that the registered had access to the information relevant for such studies. The registered experts were invited to participate in the study, which is described in more detail in the following section.

I.4.1 Methodology

For data collection, this research used online questionnaires. Unlike their paper-based counterpart, online surveys allow for a variable survey design and the integration of dynamic elements. Beyond that, the targeted participants, who are responsible for returns management at German e-tailers, most likely prefer the internet as a communication medium. The questionnaire was structured into three parts.

Before the actual questionnaire, a virtual cover letter informed the participants about the background of the survey and assures them of anonymity. To increase the motivation to participate, it was pointed out that only participants would receive a full study report after the collected data were analyzed and that a scientific, publicly accessible report would only be made available after a waiting period of at least 18 months. Next, the first

part queried essential characteristics of the respondents and asked them to identify the product category with the largest share of sales. The purpose of this part was (1) to ensure that only professionals from e-commerce or multichannel retailers participated and (2) to be able to refer to a specific product category in later questions. The second part surveyed performance indicators, for example, the return rates and return costs. The final third part asked for the revenue of each company, which was used to assess the representativeness of the sample. Before the initial field phase in 2014, the survey was pretested by six experts with either an academic or an industry background, of which three used the cognitive pretesting method suggested by Krosnick (1999). Based on their remarks, several changes were made with regard to the relevant key performance indicators, wording, and question sequence.
The study used three sample acquisition channels to address potential participants:

- Members of the expert panel: All returns management experts who successfully registered for the online panel (see previous section) at the time of data collection were invited to participate.
- Members of German e-commerce associations: The two largest German associations (BEVH, Händlerbund) drew attention to the study via a dedicated e-mail to their members or via newsletters. The registered association members represent the top management level.
- Professional social networks. The survey invitation was shared in several e-commerce groups on two of the most popular professional social networks, namely, Xing and LinkedIn.

Empirical studies are subject to various potential biases. Although the influence of such biases can never be completely ruled out (MacKenzie & Podsakoff, 2012), procedural precautions were taken to counteract and minimize such effects and increase the validity of the survey. Special attention was given to four biases, namely, (1) method bias, (2) sampling bias, (3) nonresponse bias, and (4) social-desirability bias. Table I.4 provides an overview of the employed countermeasures.

Bias	Definition	Countermeasure
Method bias	Systematic measurement error due to the inability or unwillingness to provide an accurate answer (MacKenzie & Podsakoff, 2012)	• To minimize the lack of a respondent's ability to provide accurate answers (MacKenzie & Podsakoff, 2012), only professionals with the necessary experience in e-commerce returns management were invited to participate in the study. • A pretest was conducted to increase the understandability and decrease the effect of complex and abstract questions (Doty & Glick, 1998; Krosnick, 1999). Examples complement the survey questions to illustrate more complex issues. A maximum of two questions were asked on a questionnaire page to avoid visual cognitive overload (Schmitt, 1994). • The participants were given instructions about the most relevant information in the cover letter to have the opportunity to collect the data before they start the survey (Krosnick, 1999). In addition, they were asked to answer the questions as accurately as possible and were provided with the possibility to move backward in the survey to change their answers. • To increase the motivation of the respondents and reduce potential satisficing (Krosnick & Alwin, 1987), a virtual cover letter explained the purpose of the study and reminded participants how they and their organizations would benefit from the research. • Contexts that aroused suspicions were mitigated by assuring participants of anonymity and that their data would only be used for research purposes (Schmitt, 1994).
Self-selection bias	Systematic measurement error due to a nonrandom sample of the population (Bethlehem, 2010)	• Participation in a study about returns management might be more appealing to companies with high return rates. Therefore, the study invitation email highlighted the need of participation independent of the respective return rate and volume.

Bias	Definition	Countermeasure
Nonresponse bias	Systematic measurement error because of the underrepresentation of certain respondent groups (Berg, 2005)	• To prevent item nonresponse bias this study referred to complete-case analysis. That is, incomplete responses were removed from data analysis (Little & Rubin, 1989). • With regard to general nonresponse, this study opted for a two-sample validation technique (Berg, 2005) that compares two subsamples from the same population.
Social-desirability bias	Systematic measurement error due to overreporting/underreporting of socially desirable/undesirable outcomes (Krosnick, 1999)	• This study employed an anonymous online questionnaire, which has a lower risk of social-desirability bias than traditional forms of data collection (Grimm, 2010). • Most of the surveyed data does not relate to the respondents' behavior. Instead, the participants reported on the outcomes of consumer behavior, which reduced the susceptibility to social-desirability bias. An exception is the question about recovery options (in particular the disposal/scrap option). To obtain the most valid data possible, the options were queried in an embedded style instead of single questions.

Table I.4. Biases and countermeasures within the study.

To ensure validity, multiple plausibility checks were conducted. For instance, we checked the plausibility of questions 4–6 by comparing the answers. The number of returned shipments (question 5) divided by the number of outbound shipments (question 4) should approximately result in the reported alpha return rate (question 6). In case of inconsistent answers, the participants were excluded. As the questionnaire did not contain any scales, common reliability measures could not be calculated. Nevertheless, the similar results of the two surveys indicated reasonable retest reliability. Furthermore, all results were discussed with and confirmed by industry experts who did not participate in the survey. These experts emphasized that the findings are representative of the German e-commerce market.

After the preparation of the study, data were first collected between September and November 2014. During this period, 270 respondents started the survey of which 143 (53 %) were e-commerce or multichannel merchants who completed it without showing an unusual response behavior. The second data collection period took place between November and December 2018. This time, 143 individuals started the questionnaire, of which 68 (48 %) participants completed the survey. Both samples reflect the diversity of German e-commerce in terms of the most relevant product group clusters (see Table I.5). The total e-commerce revenue that was realized by the participants in Germany in the fiscal year preceding their response amounts to 6.1 billion € (2014) and 1.2 billion € (2018). The companies that participated in the survey thus accounted for 15.6 % (2014) and 2.1 % (2018) of the total German e-commerce sales at the respective times (MRU GmbH, 2014, BEVH, 2018).

	Sample 2014	**Sample 2018**
Product group cluster	Fashion (n=54, 37.8 %), Entertainment (n=16, 11.2 %), Leisure (n=32, 22.4 %), Interior (n=16, 11.2 %), Others (n=25, 17.5 %)	Fashion (n=25, 36.8 %), Entertainment (n=10, 14.7 %), Leisure (n=12, 17.6 %), Interior (n=8, 11.8 %), Others (n=13, 19.1 %)

Table I.5. Sample characteristics.

The sharp decline in the number of participants can be attributed to several factors. These factors include that the second survey period coincided with the busy Christmas business season. Furthermore, the support received from the German e-commerce associations was more reserved, in some part because the survey overlapped with other internal studies. In addition, the 2014 survey benefitted from being one of the first returns management studies in Germany. Furthermore, some participation fatigue can be observed.

To ensure the validity of the second survey despite the smaller sample size, after data collection, the results were presented to selected members of the expert panel who did not participate in the study. They were asked to compare the results with their internal data and industry experience. In this step, retailers with a broad product range proved to be particularly helpful, as they were able to make assessments for several categories. Overall, they attributed a high degree of representativeness and external validity to the presented results. To further account for a potential nonresponse bias, the mean values of the several variables of dropouts were compared with those who finished the survey. All tests were statistically

insignificant at an α=.05 level. Therefore, we are confident that nonresponse is not an issue in this study.
The authors of this study embrace the principles of open science with the overarching goal of increasing the reproducibility and replicability of the study results as well as promoting the collaboration of scholars. For this reason, not only is the questionnaire provided in Appendix 1 of the paper but also the collected data have been made available in an open-access repository. To achieve the desired transparency while respecting the anonymity guaranteed to the survey participants, the companies' revenue figures have been removed, as the presence of this information would allow for an easy identification of some participants. The following section presents the study results.

I.4.2 Results

Because various other studies (Appriss Retail, 2019, EHI Retail Institute, 2019b, Bernon et al., 2016) suggest that the extent of consumer returns depends on the respective product category, this study refers to the previously introduced clusters (entertainment, fashion, leisure, interior, others). Within the clusters, this analysis uses weights to represent the market as realistically as possible. The weights are calculated based on the provided number of outbound packages (for the return rates), received return packages (for the number of articles per returned shipment), and returned articles (for the remaining key figures) to account for company size and the degree to which a company is affected by returns. To avoid sample inflation, the weights are normalized to the respective original sample size as suggested by Maletta (2007). All calculations and statistical tests were performed in SPSS 26.
After calculating the weighted averages (Ø), standard deviations (SD), and confidence intervals (CI) to answer the first research question, the mean values of the two samples from 2014 (S1) and 2018 (S2) are compared using t-tests to identify changes as referred to in the second research question. The following subsections present the results for the return rates, the number of items per returned shipment, processing times, costs, and recovery options. In the case of significant differences (α<.05), the table fields are printed in italics.

I.4.2.1 Return rates and number of items per returned shipment

The «Fashion» cluster has by far the highest return rates (see Table I.6).

Cluster	Alpha return rate	Beta return rate	Number of items per returned shipment
Entertain-ment	S1: 8.39 %, 2.19, [7.22, 9.56] S2: 5.38 %, 2.69, [3.45, 7.31]	S1: 5.27 %, 1.76, [4.33, 6.21] S2: 4.97 %, 2.80, [2.97, 6.98]	S1: 1.31, 0.08, [1.27, 1.36] S2: 1.18, 0.38, [0.91, 1.45]
Fashion	S1: 50.20 %, 7.21, [48.24, 52.17] S2: 46.46 %, 13.49, [40.89, 52.03]	S1: 33.98 %, 5.40, [32.51, 35.46] S2: 20.41 %, 6.71, [17.64, 23.18]	S1: 2.06, 0.37, [1.96, 2.16] S2: 3.20, 1.06, [2.77, 3.64]
Interior	S1: 5.68 %, 2.41, [4.40, 6.97] S2: 5.45 %, 1.35, [4.31, 6.58]	S1: 4.65 %, 1.97, [3.60, 5.70] S2: 4.63 %, 1.04, [3.76, 5.50]	S1: 1.07, 0.07, [1.03, 1.10] S2: 1.03, 0.13, [0.92, 1.14]
Leisure	S1: 12.26 %, 11.12, [8.25, 16.26] S2: 10.30 %, 3.04, [8.37, 12.23]	S1: 8.30 %, 8.95, [5.07, 11.53] S2: 8.47 %, 2.69, [6.76, 10.18]	S1: 1.49, 0.45, [1.33, 1.66] S2: 1.57, 0.37, [1.33, 1.80]
Others	S1: 5.33 %, 8.61, [1.78, 8.88] S2: 5.66 %, 8.12, [0.75, 10.57]	S1: 5.19 %, 8.60, [1.64, 8.74] S2: 4.80 %, 2.44, [3.33, 6.28]	S1: 1.02, 0.10, [0.98, 1.06] S2: 1.25, 0.41, [1.01, 1.50]
Legend: *Sample-ID: Ø, SD, 95 %-CI*			

Table I.6. Return rates statistics and number of items per returned shipment.

With an average alpha return rate of approximately 50 %, half of the packages sent out are later returned. The «Fashion» segment also shows the largest number of items per returned shipment, thereby indicating a high proportion of selection orders where customers order several products for comparison. Despite the often-discussed preventive measures, this study shows that the return rates have hardly changed over time. Only in the «Entertainment» cluster did the alpha return rate decrease significantly ($t(24)=3.124$, $p=.005$). Furthermore, in the «Fashion» cluster, the beta return rate decreased significantly ($t(77)=8.876$, $p<.001$), while the number of items per returned shipment has increased ($t(77)=-5.271$, $p<.001$). These outcomes suggest that fashion e-tailers successfully decreased the return probabilities of individual items by means of preventive measures, e.g., with regard to fit issues (Bertram & Chi, 2018; Saarijärvi et al., 2017). However, since customers have simultaneously formed larger shopping baskets, this decrease is not reflected in the alpha return rate.

I.4.2.2 Cycle time-related performance indicators

In addition to return rates, cycle time-based key figures are an important performance criterion. This research distinguishes between (1) transport time, (2) internal processing time, and (3) reimbursement time (see Table I.7).

Cluster	Transport time	Internal processing time	Reimbursement time
Entertain-ment	S1: 2.08 days, 0.29, [1.92, 2.23] S2: 2.22 days, 0.69, [1.73, 2.72]	S1: 4.48 days, 0.88, [4.01, 4.95] S2: 1.38 days, 0.51, [1.02, 1.75]	S1: 9.15 days, 2.50, [7.81, 10.48] S2: 2.63 days, 3.40, [0.19, 5.06]
Fashion	S1: 2.11 days, 0.40, [2.00, 2.22] S2: 2.24 days, 0.62, [1.98, 2.50]	S1: 2.59 days, 0.76, [2.39, 2.80] S2: 2.50 days, 0.84, [2.16, 2.85]	S1: 5.89 days, 1.86, [5.38, 6.39] S2: 5.14 days, 2.62, [4.06, 6.23]
Interior	S1: 2.49 days, 0.71, [2.12, 2.87] S2: 2.05 days, 0.44, [1.69, 2.42]	S1: 2.13 days, 0.62, [1.81, 2.46] S2: 1.09 days, 0.31, [0.83, 1.34]	S1: 6.89 days, 3.63, [4.96, 8.83] S2: 1.44 days, 1.00, [0.61, 2.28]
Leisure	S1: 2.12 days, 0.74, [1.86, 2.39] S2: 1.95 days, 0.30, [1.76, 2.15]	S1: 4.34 days, 1.61, [3.76, 4.92] S2: 2.70 days, 0.61, [2.31, 3.09]	S1: 4.59 days, 1.67, [3.99, 5.19] S2: 3.96 days, 0.64, [3.56, 4.37]
Others	S1: 2.20 days, 0.82, [1.86, 2.54] S2: 2.57 days, 0.53, [2.25, 2.89]	S1: 1.92 days, 1.59, [1.26, 2.57] S2: 1.88 days, 0.83, [1.38, 2.38]	S1: 8.57 days, 2.15, [7.68, 9.46] S2: 4.99 days, 3.51, [2.86, 7.11]
Legend: *Sample-ID: Ø, SD, 95 %-CI*			

Table I.7. Cycle time-based returns management performance indicators.

In all product categories, transport times are approximately two days and have remained virtually unchanged in the four years under investigation. With regard to internal processing times, the data collected indicate that many e-tailers have sped up their internal processes, particularly with regard to the «Entertainment» ($t(24)=10.074$, $p<.001$), «Interior» (t(22)=4.132, p<.001), and «Leisure» (t(42)=3.425, p=.001) clusters. In addition, the average reimbursement time in the «Entertainment» (t(24)=5.636, p<.001), «Interior» (t(22)=5.604, p<.001), and «Others» (t(36)=3.903, p<.001) clusters has decreased significantly, which benefits customers.

I.4.2.3 Return costs

Consumer returns cause costs for transport and processing. The latter explicitly includes costs for remarketing and depreciation due to deterioration in condition. In 2014, it is shown that online fashion retailers realize very low costs per item (see Table I.8).

Cluster	**Transport costs per returned item**	**Processing costs per returned item**
Entertainment	S1: 4.20 €, 1.64, [3.32, 5.07] S2: 4.44 €, 3.46, [1.97, 6.91]	S1: 10.27 €, 8.16, [5.92, 14.62] S2: 10.46 €, 7.19, [5.31, 15.60]
Fashion	S1: 2.30 €, 0.52, [2.16, 2.44] S2: 1.41 €, 0.66, [1.13, 1.68]	S1: 1.65 €, 0.89, [1.41, 1.89] S2: 1.22 €, 0.52, [1.00, 1.43]
Interior	S1: 13.78 €, 13.32, [6.68, 20.88] S2: 23.04 €, 10.19, [14.52, 31.55]	S1: 10.57 €, 5.34, [7.72, 13.41] S2: 11.10 €, 3.71, [7.99, 14.20]
Leisure	S1: 6.10 €, 1.41, [5.59, 6.61] S2: 5.51 €, 1.34, [4.65, 6.36]	S1: 10.66 €, 2.63, [9.71, 11.61] S2: 4.49 €, 1.66, [3.43, 5.54]
Others	S1: 5.28 €, 2.84, [4.11, 6.46] S2: 3.96 €, 1.47, [3.07, 4.85]	S1: 6.17 €, 5.21, [4.02, 8.32] S2: 3.71 €, 1.98, [2.51, 4.91]

Legend: *Sample-ID: Ø, SD, 95 %-CI*

Table I.8. Costs per returned item.

The costs are particularly high in the «Entertainment» and «Interior» clusters. The data collected in the survey suggest that efficiency gains have significantly reduced costs, particularly in the «Fashion» (transport costs: t(77)=6.503, p<.001; processing costs: t(77)=2.694, p=.009) and «Leisure» (processing costs: t(42)=7.552, p<.001) segments.

I.4.2.4 Recovery options

Part of the processing costs are due to remarketing or depreciation. It is therefore important to determine which share of returned items can be resold as new. This proportion is very high in the «Fashion» sector and particularly low in the «Entertainment» cluster (see Table I.9). Over time, the proportions of the individual recovery options have remained relatively constant.

Despite some significant differences, the overall changes appear limited. The largest absolute shift can be observed in the «Interior» cluster, where the proportion of returned items that were resold as used articles in a secondary market significantly decreased ($t(22)=3.040$, $p=.006$). A similar development is observed in the «Fashion» cluster ($t(77)=3.046$, $p=.003$). Moreover, the relevance of third-party brokers has increased in the cluster «Others» ($t(36)=-2.531$, $p=.020$) and the share of donations has decreased in the category «Leisure» ($t(42)=2.259$, $p=.031$). Other recovery options, such as forwarding the items to suppliers, have lost relevance in the «Entertainment» ($t(24)=4.603$, $p<.001$), «Interior» ($t(22)=3.439$, $p=.004$) and «Leisure» ($t(42)=2.617$, $p=.013$) clusters.

Cluster	Sold as new	Sold as used/returned product	Sold to third party broker
Entertain-ment	S1: 60.84 %, 7.10, [57.05, 64.62] S2: 64.39 %, 20.90, [49.44, 79.34]	S1: 31.09 %, 7.92, [26.87, 35.31] S2: 26.52 %, 12.55, [17.54, 35.49]	S1: 4.72 %, 6.61, [1.20, 8.24] S2: 7.27 %, 10.94, [-0.56, 15.09]
Fashion	S1: 90.89 %, 6.93, [89.00, 92.78] S2: 94.21 %, 13.38, [88.69, 99.74]	S1: 5,34 %, 4.48, [4.11, 6.56] S2: 2.09 %, 4.22, [0.35, 3.84]	S1: 2.68 %, 3.75, [1.65, 3.70] S2: 1.07 %, 4.13, [-0.63, 2.77]
Interior	S1: 60.11 %, 13.50, [52.92, 67.30] S2: 69.54 %, 24.12, [49.37, 89.71]	S1: 28.17 %, 14.77, [20.30, 36.04] S2: 15.49 %, 5.49, [10.90, 20.08]	S1: 1.61 %, 4.69, [-0.90, 4.11] S2: 6.50 %, 10.65, [-2.41, 15.40]
Leisure	S1: 81.66 %, 13.65, [76.74, 86.58] S2: 79.19 %, 11.18, [72.09, 86.29]	S1: 10.67 %, 5.33, [8.75, 12.59] S2: 11.63 %, 6.43, [7.54, 15.71]	S1: 2.12 %, 7.24, [-0.49, 4.73] S2: 4.34 %, 3.46, [2.14, 6.54]
Others	S1: 86.66 %, 25.65, [76.07, 97.24] S2: 81.21 %, 12.93, [73.40, 89.02]	S1: 7.08 %, 21.20, [-1.67, 15.84] S2: 6.82 %, 8.50, [1.68, 11.95]	S1: 0.05 %, 1.11, [-0.41, 0.51] S2: 1.18 %, 1.40, [0.33, 2.03]
Cluster	**Donated**	**Disposal**	**Other recovery option**
Entertain-ment	S1: 0.01 %, 0.08, [-0.03, 0.05] S2: 0.59 %, 1.61, [-0.56, 1.74]	S1: 0,52 %, 2.01, [-0.55, 1.59] S2: 1.20 %, 3.29, [-1.15, 3.55]	S1: 2.82 %, 2.36, [1.57, 4.08] S2: 0.04 %, 0.43, [-0.27, 0.35]
Fashion	S1: 0.01 %, 0.13, [-0.03, 0.04] S2: 1.04 %, 2.75, [-0.10, 2.17]	S1: 1.03 %, 3.81, [-0.01, 2.07] S2: 1.56 %, 2.69, [0.45, 2.67]	S1: 0.06 %, 0.83, [-0.16, 0.29] S2: 0.02 %, 1.25, [-0.49, 0.54]
Interior	S1: 0.01 %, 0.06, [-0.03, 0.04] S2: 0.00 %, 0.00, [0.00, 0.00]	S1: 7.48 %, 11.26, [1.48, 13.48] S2: 8,47 %, 14.09, [-3.31, 20.25]	S1: 2.62 %, 3.04, [1.00, 4.24] S2: 0.00 %, 0.17, [-0.14, 0.14]
Leisure	S1: 0.40 %, 0.92, [0.07, 0.73] S2: 0.02 %, 0.11, [-0.05, 0.09]	S1: 3.32 %, 6.83, [0.86, 5.78] S2: 4.68 %, 4.90, [1.57, 7.80]	S1: 1.83 %, 3.61, [0.53, 3.13] S2: 0.13 %, 0.42, [-0.13, 0.40]
Others	S1: 0.54 %, 2.55, [-0.51, 1.59] S2: 0.29 %, 0.79, [-0.19, 0.77]	S1: 3.12 %, 9.94, [-0.99, 7.22] S2: 3.14 %, 4.21, [0.60, 5.68]	S1: 2.61 %, 3.24, [1.27, 3.94] S2: 7.36 %, 10.42, [1.06, 13.66]
Legend: *Sample-ID: Ø, SD, 95 %-CI*			

Table I.9. Employed recovery options.

I.4.3 Discussion

As previously mentioned (Asdecker, 2015, Hofmann et al., 2020), there are several ways to calculate the return rate. The results of this study show the magnitude of the differences. Future studies should thus always specify and disclose their respective calculation method. It is even possible that the large fluctuations in the figures published to date are partly due to these differences. In addition, this study confirms that large differences exist between product groups and the exceptional relevance of this data for fashion e-tailers (EHI Retail Institute, 2019b). In terms of cycle times, the data show that return transport is considerably slower than the typical next-day delivery that is common practice in outbound distribution, i.e., over two working days on average. Internal processing also takes approximately 1–2 working days. The reimbursement time varies between 1–5 working days, thereby leaving room for some improvement.

The costs per returned article differ significantly depending on the product category. This study further distinguishes between transport and processing costs. A returned article in the «Fashion» cluster costs relatively little, although in the case of transport costs, the low costs can be attributed to the observation that return shipments usually contain multiple items (see Table I.6). In addition, textiles and shoes are processed more quickly compared to other product groups. A major reason for the higher processing costs in the «Interior» and «Entertainment» categories is that the processing is less easy to automate and is characterized by a higher manual share of work.

With regard to the recovery options, this study shows, in line with recent other publications (e.g., Pur et al., 2013, EHI Retail Institute, 2019b), that the overwhelming part of returned goods can be directly restocked and sold as new. The interviews that were conducted to ensure the validity of the second survey in light of the smaller sample size revealed that packaging is the main reason for the comparatively low share of returns that are sold as new in the «Entertainment» cluster. Many brands use packaging seals to guarantee the originality of the goods. If a customer orders an article and opens it, the seal is broken and the goods can no longer be sold as new, even though they work perfectly fine. The significantly lower proportion of donated returns compared to disposal is a source of concern. A member of the expert panel indicated that in many cases, it is cheaper to

dispose of a return than to donate it. Besides, many e-tailers have problems finding trusted donees. Here, information systems, such as designated platforms, could help to better match supply and demand in the future.
Concerning changes over time, despite some significant differences, the overall picture appears rather stable. This is surprising since many e-tailers appear to have invested time and money in preventive measures that help to avoid returns (e.g., Walsh & Möhring, 2017, Wang & Ansell, 2020). However, it seems that these measures only help to keep the return rates from further increasing. At the same time, this finding is an indication that older data are not losing their credibility as quickly as expected or feared and that it is possible to quote such studies for a fairly long time.
Furthermore, the collected data allow for the derivation of market estimates. Weighting the reported 2018 values in each product cluster with the current revenue/shipment shares (see Table I.3) results in an average alpha return rate of 16.59 %. Thus, approximately one in six outbound packages is later returned. With 1.99 billion shipments in German B2C e-commerce in 2019, including returns, this leads to the following estimation:

$$1.99\ \textit{billion shipments} - \frac{1.99\ \textit{billion shipments}}{1+0.1659} = 284\ \textit{million return shipments}.$$

Approximately 72 % of this total can be attributed to the «Fashion» segment (see Table I.10). Using those percentages leads to an average of 2.66 items per returned shipment, which results in a total of 755 million returned articles. Of these, approximately 87 % can be attributed to the «Fashion» cluster. This latter number highlights the importance of reducing the number of fashion returns in particular. In this context, information technologies should be of great value, for example, by reducing size-related returns by means of data analysis and artificial intelligence (e.g., Urbanke et al., 2015, Yang & Xiong, 2019). There is a real lever to make the fashion e-tailers' business model more profitable (Asdecker, 2015).
The estimated share of returned items (see Table 1.10) can also be used to quantify further item-based market figures. For instance, a returned item costs 4.28 € (transport: 2.21 €; processing: 2.07 €) on average. This value is far below the existing estimates. If one were to assume higher

costs, for example 15–20 € (Pur et al., 2013) or $30–$35 (Stock et al., 2006), it is questionable whether the business model could ever be profitable given the high measured return rates.

Cluster	**S1: Revenue/ shipment share**	**S2: Alpha return rate**	**S3: Items per returned shipment**	**Share of returned shipments (=(S1*S2)/ SUM(S1*S2))**	**Share of returned items (=(S1*S2*S3)/ SUM(S1*S2*S3))**
Entertain-ment	35.57 %	5.38 %	1.18	11.53 %	5.11 %
Fashion	25.76 %	46.46 %	3.20	72.13 %	87.32 %
Interior	15.03 %	5.45 %	1.03	4.94 %	1.91 %
Leisure	11.92 %	10.30 %	1.57	7.40 %	4.36 %
Others	11.72 %	5.66 %	1.25	4.00 %	1.88 %

Table I.10. An estimate of return shipments and returned items shares.

Not only the economic dimension of returns can be better evaluated with the collected data but also the ecological effects, which are thus far often neglected. According to Bertram and Chi (2018), the environmental impact of fashion returns depends on the return method. Edwards et al. (2009) estimate the emissions of a consumer return on the last mile to range between 362 g CO_2 and 4455 g CO_2. Despite an extensive search, the authors of this paper could not find any estimate for the total CO_2 footprint of a return. For this reason, a separate approximation is to be made based on the combination of figures that are published in Zalando's annual report. Zalando is one of the largest fashion e-tailers in Europe. Accordingly, Zalando's total carbon footprint in 2019 amounted to 262,511 tons of CO_2 (Zalando, 2020, p. 36). Transport to customers, including returns, accounts for 61 %, or 160,132 tons of CO_2 (Zalando, 2020, p. 36). Given Zalando's documented return rate of 50 % (Gassmann, 2018), of 53,377 tons of CO_2 can be attributed to returns. At the same time Zalando states that in 2019, it processed 144.9 million orders (Zalando, 2020, p. 105). Assuming that each order leads to a package, the above-cited return rate of 50 % results in 72.45 million return shipments. Accordingly, the overall footprint of a return shipment is approximately 737 grams CO_2. It should be emphasized that this estimate is very conservative, because not all companies have Zalando's process efficiency and Zalando is already making great efforts to make its business model CO_2-neutral. Furthermore, some emissions, such as the disposal of unsaleable returns, are not included. If this figure is applied to the German market as a whole, the estimated 284 million return shipments would result in at least 209,308 tons of CO_2.

I.5 Conclusion, implications, and outlook

This paper presents the results of a longitudinal trend study among e-commerce retailers in Germany on consumer returns. It sheds light on the essential reverse part of the e-commerce business model, which companies are usually not particularly willing to provide information about (Stock & Mulki, 2009). To date, there have mainly been data points about individual companies, anecdotal evidence, or practitioner-oriented studies with a methodology that is not fully comprehensive. Furthermore, the literature review has shown that the values published thus far show an enormous range. Valid estimates for the number of returns and the associated costs or ecological effects could not be derived thus far. Given this research gap, the study at hand goes beyond the existing work and presents a systematic approach to gain numerous insights into what is also called the "necessary evil" (Petersen & Kumar, 2010) of e-commerce.

In conclusion, what is the overall contribution of this paper? Whetten (1989) distinguishes the following four categories: factors that explain a phenomenon (*what?*), the relationship between those factors (*how?*), logical justifications for altered views (*why?*), and the conditions that limit generalizability (*who, where, when?*). This research adds to the e-commerce literature by focusing on the often neglected reverse part of the business model and contributing to three of these four categories.

First, we provide a comprehensive focused literature review that summarizes the available descriptive statistics on essential performance indicators *(what?, where?*). This overview allows interested scholars a quick introduction to the topic. As the numbers reported in previous literature are fragmented, the need for a well-structured data collection is evident.

Second, the data obtained for the German market demonstrate the societal, ecological, and economic impact (*what?, where?*) of consumer returns that make active returns management inevitable. Furthermore, the high return rates underline the need for further research specifically regarding technologies such as combining big data analytics with artificial intelligence to provide feasible solutions for preventive measures.

Third, this research enables a better assessment of extant research. In other words, it tries to bring some order to the current chaos and gives justifications for altered views (*why?*). For example, it appears that there have been exaggerations in the past regarding the costs of a returned item (Stock et al., 2006, Pur et al., 2013). Some assumptions concerning return rates must also be critically reflected upon. For instance, Walsh et al. (2016) stated that "[...] computers ‚suffer‘ especially from high product returns". Thus, our results indicate that a generic perspective is neither suitable nor useful – it depends on the context. Return rates for fashion items are roughly five times higher than those for entertainment products such as consumer electronics. In addition, the total cost of processing fashion returns is only one-fifth of the cost of a returned entertainment product. This more nuanced perspective also allows for a more accurate assessment of the representativeness of previously published studies.

Fourth, the data obtained can also stimulate further research in the field of quantitative models and simulation models by providing input for the required parameter values. In addition, it is a source for data triangulation to validate the representativeness of other studies and for the justification of a specific research design, e.g., regarding the product category to focus on (impetus to *why?*).

Finally, fifth, the survey instrument presented in the digital appendix provides the basis for international replication studies (impetus to *where?*). While the current data refer to the largest European domestic market, Germany, which is one of the study's biggest limitations, other scholars can use the questionnaire to investigate the influence of regional differences, as implied by Nicola (2019).

In terms of managerial contribution, the findings of this research can be used to improve decision support systems. For example, the collected data are of particular value to companies that are currently planning to enter the German market and are thus trying to obtain better knowledge for designing return policies, calculating prices, and determining processing capacities. In addition, this study provides the necessary data for benchmarking efforts that help to ensure competitiveness. In this context, reference should again be made to the data provided in the digital appendix,

which enables even finer evaluation options based on 20 product categories, which were aggregated into five clusters for the current analysis.
On the broader scale of the environment, this work helps to draw attention to the ecological consequences of returns in e-commerce. To date, the research community has paid little attention to the intended and unintended consequences of e-commerce. To create a more sustainable future, it is important to clearly address possible shortcomings. Two findings should be mentioned in this respect. First, we present an approach for estimating the total carbon footprint of returns, which needs to be further reduced. Second, this study helps to quantify the consumption of resources through the disposal of returns. Fortunately, the data show that disposal is the exception rather than the rule. Nevertheless, e-tailers obviously dispose of items more frequently than they donate them. On the one hand, these "dark" sides of the business model deserve attention. On the other hand, new business models can arise from them. In particular, solutions are needed to help better coordinate donors and grantees with the goal of obtaining more resources.
In the end, apart from the numbers and figures, consumer returns will remain an essential part of e-commerce. Given the projected high growth rates ahead, it is clear that the problem will become even more prevalent. A prerequisite for change is a thorough review of the status quo. In that regard, this work has made an initial contribution to the field. We hope that other fellow researchers are motivated to build upon the work presented herein.

Data availability statement:
The data that support the findings of this study are openly available in a figshare respository, https://doi.org/10.6084/m9.figshare.14376794.v1

References

Abbey, J. D., Ketzenberg, M. E., & Metters, R. (2018). A More Profitable Approach to Product Returns. *MIT Sloan Management Review, 60*(1), 71–74.

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Appriss Retail. (2019). *Consumer returns in the retail industry*. https://appriss.com/retail/wp-content/uploads/sites/4/2020/01/AR3019-2019-Customer-Returns-in-the-Retail-Industry.pdf

Araújo, A. C. de, Matsuoka, E. M., Ung, J. E., Massote, A., & Sampaio, M. (2018). An exploratory study on the returns management process in an online retailer. *International Journal of Logistics Research and Applications, 21*(3), 345–362.

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Asdecker, B., & Karl, D. (2018). Big data analytics in returns management - Are complex techniques necessary to forecast consumer returns properly? In *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics,* Valencia, Spain.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Babbie, E. R. (2005). *The basics of social research* (3. ed.). Thomson/Wadsworth.

Banjo, S. (2013). Rampant Returns Plague E-Retailers. *Wall Street Journal - Eastern Edition, 22*(148), B1–B2.

Berg, N. (2005). Non-response bias. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement* (pp. 865–873). Elsevier.

Bernon, M., Cullen, J., & Gorst, J. (2016). Online retail returns management. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 584–605.

Bertram, R. F., & Chi, T. (2018). A study of companies' business responses to fashion e-commerce's environmental impact. *International Journal of Fashion Design, Technology and Education, 11*(2), 254–264.

Bethlehem, J. (2010). Selection Bias in Web Surveys. *International Statistical Review, 78*(2), 161–188.

BEVH. (2018). *Zweistelliges Wachstum in 2017 und weiterhin gute Perspektiven im E-Commerce*. Retrieved November 8, 2020, from https://www.bevh.org/presse/pressemitteilungen/details/zweistelliges-wachstum-in-2017-und-weiterhin-gute-perspektiven-im-e-commerce.html

BEVH. (2020). *Vielbesteller treiben E-Commerce-Umsatz in 2019 auf neuen Höchststand*. Bundesverband E-Commerce und Versandhandel e. V. Retrieved November 11, 2020, from https://www.bevh.org/fileadmin/content/05_presse/Pressemitteilungen_2020/2020121_-_PM_Vielbesteller_treiben_E-Commerce-Wachstum.pdf

BIEK. (2019). *KEP-Studie 2019 – Analyse des Marktes in Deutschland*. Bundesverband Paket und Expresslogistik e. V. Retrieved November 11, 2020, from https://www.biek.de/download.html?getfile=2335

BIEK. (2020). *KEP-Studie 2020 – Analyse des Marktes in Deutschland*. Bundesverband Paket und Expresslogistik e. V. Retrieved November 11, 2020, from https://www.biek.de/download.html?getfile=2623

Brohan, M. (2005). Online retailers learn to live with that persistent problem of returns. *Internet Retailer, 7*(4), 44–50.

Cassill, N. L. (1998). Do customer returns enhance product and shopping experience satisfaction? *The International Review of Retail, Distribution and Consumer Research, 8*(1), 1–13.

Catalog Age (2002). Holiday Return Rates Hold Steady for Most Mailers. *Catalog Age, 19*(3), 1.

Cui, H., Rajagopalan, S., & Ward, A. R. (2020). Predicting product return volume using machine learning methods. *European Journal of Operational Research, 281*(3), 612–627.

Del Franco, M. (2000). Good news about returns. *Catalog Age, 17*(4), 41.

Denyer, D., & Tranfield, D. (2009). Producing a Systematic Review. In D. A. Buchanan & A. Bryman (Eds.), *The Sage handbook of organizational research methods* (pp. 671–689). Sage.

Difrancesco, R. M., Huchzermeier, A., & Schröder, D. (2018). Optimizing the return window for online fashion retailers with closed-loop refurbishment. *Omega, 78*, 205–221.

Doty, D. H., & Glick, W. H. (1998). Common Methods Bias: Does Common Methods Variance Really Bias Results? *Organizational Research Methods, 1*(4), 374–406.

Douthit, D., Flach, M., & Agarwal, V. (2011). A returning problem: reducing the quantity and cost of product returns in consumer electronics. *Accenture Communications & High Tech, Accenture*, 3–9.

The Economist. (2013). *Return to Santa*. Retrieved November 18, 2020, from https://www.economist.com/business/2013/12/21/return-to-santa

Edwards, J. B., McKinnon, A. C., & Cullinane, S. (2009). *Carbon auditing the 'last mile': modelling the environmental impacts of conventional and online non-food shopping* (Green Logistics Report). Heriot-Watt University.

EHI Retail Institute. (2019a). *Top 100 umsatzstärkste Onlineshops in Deutschland*. Retrieved November 4, 2020, from https://www.ehi.org/de/top-100-umsatzstaerkste-onlineshops-in-deutschland/

EHI Retail Institute. (2019b). *Versand- und Retourenmanagement im E-Commerce*. Köln.

Ellis, J. E. (2017). Online Shopping. *Bloomberg Businessweek*(4545), 52.

Gassmann, M. (2018). *Retouren gehören ganz klar zum Geschäftsmodell*. Retrieved November 16, 2020, from https://www.welt.de/wirtschaft/plus185244750/Zalando-Vorstand-Schneider-Retouren-gehoeren-ganz-klar-zum-Geschaeftsmodell.html

Giménez, C., & Lourenço, H. R. (2008). e-SCM: internet's impact on supply chain processes. *The International Journal of Logistics Management, 19*(3), 309–343.

Griffis, S. E., Rao, S., Goldsby, T. J., & Niranjan, T. T. (2012). The customer consequences of returns in online retailing: An empirical analysis. *Journal of Operations Management, 30*(4), 282–294.

Grimm, P. (2010). Social Desirability Bias. In J. Sheth & N. Malhotra (Eds.), *Wiley International Encyclopedia of Marketing* (Vol. 50, p. 537). John Wiley & Sons, Ltd.

Hjort, K., Hellstrom, D., Karlsson, S., & Oghazi, P. (2019). Typology of practices for managing consumer returns in internet retailing. *International Journal of Physical Distribution & Logistics Management, 49*(7), 767–790.

Hofmann, A., Gwinner, F., Fuchs, K., & Winkelmann, A. (2020). An Industry-Agnostic Approach for the Prediction of Return Shipments. In *Proceedings of the 26th Americas Conference on Information Systems,* Salt Lake City, United States of America.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

IRS. (2020). *Yearly Average Currency Exchange Rates*. Internal Revenue Service. Retrieved September 3, 2020, from https://www.irs.gov/individuals/international-taxpayers/yearly-average-currency-exchange-rates

Jack, L., Frei, R., & Krzyzaniak, S.-A. (2019). *Buy Online, Return in Store: The Problems & Opportunities of E-Commerce Returns*. Retrieved November 18, 2020, from https://www.ecrloss.com/research/buy-online-return-in-store

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50,* 537–567.

Krosnick, J. A., & Alwin, D. F. (1987). An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly, 51*(2), 201.

Lipsman, A. (2019). *Global Ecommerce 2019: Ecommerce Continues Strong Gains Amid Global Economic Uncertainty*. Retrieved January 20, 2021, from https://www.emarketer.com/content/global-ecommerce-2019

Little, R. J. A., & Rubin, D. B. (1989). The Analysis of Social Science Data with Missing Values. *Sociological Methods & Research, 18*(2-3), 292–326.

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies. *Journal of Retailing, 88*(4), 542–555.

Maletta, H. (2007). *Weighting*. Retrieved November 17, 2020, from https://www.spsstools.net/en/documents/48/WEIGHTING.pdf

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To Keep or Not to Keep: Effects of Online Customer Reviews on Product Returns. *Journal of Retailing, 92*(3), 253–267.

Mollenkopf, D. A., Rabinovich, E., Laseter, T. M., & Boyer, K. K. (2007). Managing Internet Product Returns: A Focus on Effective Service Operations. *Decision Sciences, 38*(2), 215–250.

MRU GmbH. (2014). *E-Commerce und Paketdienste: Lieferdienste und Onlinehandel*. Retrieved November 18, 2020, from https://www.bevh.org/fileadmin/content/05_presse/Pressemitteilungen_2014/140901_E-Commerce_und_Paketdienste.pdf

Ng, S., & Stevens, L. (2015). Where your unwanted christmas gifts get a second life. *Wall Street Journal (December 27)*.

Nicola, S. (2019). *Germany's Returning Customers*. Retrieved November 15, 2020, from https://www.bloomberg.com/news/articles/2019-08-08/germany-s-returning-customers

Petersen, J. A., & Kumar, V. (2010). Can Product Returns Make You Money? *MIT Sloan Management Review, 51*(3), 85–89.

Pettypiece, S. (2015). *Online Surge Means Many Not-So-Happy Returns for Retailers*. Retrieved November 18, 2020, from https://www.bloomberg.com/news/articles/2015-12-29/online-surge-leads-to-many-not-so-happy-returns-for-retailers

Pur, S., Stahl, E., Wittmann, M., Wittmann, G., & Weinfurtner, S. (2013). *Retourenmanagement im Online-Handel - Das Beste daraus machen: Daten, Fakten und Status quo*. Ibi Research.

Rabinovich, E., Sinha, R., & Laseter, T. M. (2011). Unlimited shelf space in Internet supply chains: Treasure trove or wasteland? *Journal of Operations Management, 29*(4), 305–317.

Rao, S., Rabinovich, E., & Raju, D. (2014). The role of physical distribution services as determinants of product returns in Internet retailing. *Journal of Operations Management, 32*(6), 295–312.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Rogers, D. S., & Tibben-Lembke, R. S. (1998). *Going Backwards: Reverse Logistics Trends and Practices.* Reverse Logistics Executive Council.

Rogers, D. S., & Tibben-Lembke, R. S. (2001). An Examination of Reverse Logistics Practices. *Journal of Business Logistics, 22*(2), 129–148.

Rösch, B. (2011). Retouren bringen Textillogistik auf Touren. *Textilwirtschaft*(50), 44.

Saarijärvi, H., Sutinen, U.-M., & Harris, L. C. (2017). Uncovering consumers' returning behaviour: a study of fashion e-commerce. *International Review of Retail, Distribution & Consumer Research, 27*(3), 284–299.

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The Impact of Online Product Reviews on Product Returns. *Information Systems Research, 29*(3), 723–738.

Schmitt, N. (1994). Method bias: The importance of theory and measurement. *Journal of Organizational Behavior, 15*(5), 393–398.

Shang, G., Ferguson, M. E., & Galbreth, M. R. (2019). Where Should I Focus My Return Reduction Efforts? Empirical Guidance for Retailers. *Decision Sciences, 50*(4), 877–909.

Stevens, L. (2014). For Retailers, It's Many Unhappy Returns. *Wall Street Journal - Eastern Edition, 264*(151), B3.

Stock, J. R. (2004). *Product returns/reverse logistics in warehousing: Strategies, policies and programs.* Warehousing Education and Research Council.

Stock, J. R., & Mulki, J. P. (2009). Product Returns Processing: An Examination of Practices of Manufacturers, Wholesalers/Distributors, and Retailers. *Journal of Business Logistics, 30*(1), 33–62.

Stock, J. R., Speh, T., & Shear, H. (2006). Managing Product Returns for Competitive Advantage. *MIT Sloan Management Review, 48*(1), 57–62.

Taris, T. (2000). *A Primer in Longitudinal Data Analysis*. Sage. https://doi.org/10.4135/9781849208512

Tarn, J. M., Razi, M. A., Wen, H. J., & Perez, A. A. (2003). E-fulfillment: the strategy and operational requirements. *Logistics Information Management, 16*(5), 350–362.

Thomas, O. (2001). Gone Today, Here Tomorrow. *Business 2.0, 2*(10), 116.

Torry, H. (2020). *Coronavirus Pandemic Widens Divide Between Online, Traditional Businesses*. Retrieved November 14, 2020, from https://www.wsj.com/articles/coronavirus-pandemic-widens-divide-between-online-traditional-businesses-11585733402

Urbanke, P., Kranz, J., & Kolbe, L. (2015). Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction. In *Proceedings of the 14th International Conference on Computer and Information Science,* Fort Worth, United States of America.

Vakulenko, Y., Shams, P., Hellstrom, D., & Hjort, K. (2019). Online retail experience and customer satisfaction: the mediating role of last mile delivery. *The International Review of Retail, Distribution and Consumer Research, 29*(3), 306–320.

Vilar-Zanon, J. L., Vilar, E., & Heras, A. (2017). Online product returns risk assessment and management. *Top, 25*(3), 445–466.

Walsh, G., Albrecht, A. K., Kunz, W., & Hofacker, C. F. (2016). Relationship between Online Retailers' Reputation and Product Returns. *British Journal of Management, 27*(1), 3–20.

Walsh, G., & Möhring, M. (2017). Effectiveness of product return-prevention instruments: Empirical evidence. *Electronic Markets, 27*(4), 341–350.

Wang, E., & Ansell, J. (2020). When Should E-Retailers Be Lenient? In *Proceedings of the 31st AMA Winter Academic Conference,* San Diego, United States of America.

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution? *Academy of Management Review, 14*(4), 490–495.

Woods, R. (2017). Too Many Happy Returns: The Struggle to Predict Reverse Logistics. *Air Cargo World, 107*(9), 32–34.

Xia, Y., & Zhang, G. P. (2010). The Impact of the Online Channel on Retailers' Performances: An Empirical Evaluation. *Decision Sciences, 41*(3), 517–546.

Yang, S., & Xiong, G. (2019). Try It On! Contingency Effects of Virtual Fitting Rooms. *Journal of Management Information Systems, 36*(3), 789–822.

Zalando. (2020). *Annual report 2019.* Retrieved November 16, 2020, from https://corporate.zalando.com/sites/default/files/media-download/Zalando_SE_SN_Geschäftsbericht%202019_Stand_eng_MASTER_final_Update_sc.pdf

Appendix I.1: Online questionnaire [Excerpt]

Virtual cover letter

Dear expert,

welcome to this study on consumer returns in e-commerce. The primary goal of this research is to gain insights into the often neglected returns management process, which is of great importance. We value your opinion!

Is there a benefit of participation? Yes! Only participants get access to the results directly after data evaluation. A generally accessible scientific publication will be made after a blocking period of 1.5 years. The collected data can improve your decision-making with regard to consumer returns and can be used for benchmarking purposes.
This is an anonymous survey. All information will only be used for scientific purposes.
It takes about 12 to 15 minutes to finish the survey.
The survey requests some mandatory information that is necessary for the data analysis.
Among others, please have an estimate of the following parameters available:
Number of outbound shipments in your most important product category in [year]
Number of returned shipments in your most important product category in [year]
Your company's e-commerce revenue in the last fiscal year
Please answer the questions as accurately as possible. Be assured that there are no right or wrong answers. To achieve meaningful results, the largest possible sample is required. Therefore, please complete the questionnaire.
If you have any questions regarding this study, please contact us by e-mail: xxx.

Questions

Question 1: Please indicate for which kind of company you are participating in this survey. [selection categories: Pure online retailer; Multi-channel or omni-channel retailer (stationary & online); Pure stationary retailer; I do NOT participate for a retailer]

Question 2: Please select the product category in which your company generates the largest share of its e-commerce sales: [selection categories: Clothing/Accessories; Books/E-Books/Audio Books; Car/motorbike/accessories; Computer/Accessories/Games/Software Including Downloads/Electric Goods; DIY/Flowers; Furniture/Lamps/Decoration; Groceries/Food/Beverages; Health/Drugs/Medication; Hobby/Leisure articles;

Home textiles; Household goods/Appliances; Jewelry/Watches; Office supplies; Pet supplies; Shoes/Footwear; Telecommunication/Mobile phone/Accessories; Toiletries/Cosmetics; Toys/Games; CDs/DVDs; Other product category (text field)]

Information: We will again highlight this in the respective questions. Nonetheless, please confirm at this point that most of the following questions refer to the product category [Answer Q2] by clicking on "CONTINUE".

Question 3: For how long can your customers return their ordered articles in product category [Answer Q2]?

Question 4: Please provide your company's total number of outbound shipments that can be accounted for in the product category [Answer Q2] during [year].

Question 5: Please provide the total number of shipments that have been returned to your company (or an appointed third party) that can be attributed to product category [Answer Q2] during [year].

Question 6: How high was your shipment-based return rate in product category [Answer Q2] in [year]? [Example: You have sent out 1000 packages, of which 500 came back --> return rate = 50 %]

Question 7: How high was your item-based return rate in product category [Answer Q2] in [year]?
[Example: You have shipped 3000 articles, of which 1000 came back --> return rate = 33 %]

Question 8: On average, how many articles did a returned shipment in product category [Answer Q2] contain in [year]?

Question 9: What were the average transportation costs per returned item in product category [Answer Q2] in [year]?

Question 10: What were the average processing costs (including possible write-offs and remarketing expenses) per returned item in product category [Answer Q2] in [year]?

Question 11: What is your estimate of the average transport time that occurred after the consumer took the return to the logistics service provider in product category [Answer Q2] in [year] (in working days)?

Question 12: What is your estimate of the internal processing time that occurred after the logistics service provider delivered the return to your facility in product category [Answer Q2] in [year] (in working days)?

Question 13: How long did it take until the customer was reimbursed after taking his/her return to the logistics service provider in product category [Answer Q2] in [year] (in working days)?

Question 14: What was the percentage of returned items that your company could directly resell as new after processing (A-returns) in the product category [Answer Q2] in [year]?

[Only if [Answer Q14] < 100 %] Question 15: Which of the following recovery options did your company use for returned items in product category [Answer Q2] that could not be resold as new in [year]and to what extent?
Sell as used/returned products on secondary markets (outlet, eBay, etc.)
Sell to independent third-party brokers (remarketing specialist)
Donate to charitable organizations
Dispose as scrap
Other recovery option(s)

Question 15/16: Please provide your company's e-commerce net revenue (after returns) during the last fiscal year (in Euro).

Appendix I.2: Literature Review: Return Rates Details

Paper	Peer reviewed	Region	Data type	Year of data collection (if available)	Data source / survey population	Cross-sectional rates (type)	Fashion rates (type)	Entertainment rates (type)	Leisure rates (type)	Interior rates (type)	Other rates (type)
Hofmann et al., 2020	X	DE	Enterprise data	-	Wholesaler	-	-	-	5.1 % (β) 7.7 % (α)	-	-
Cui et al., 2020	X	US	Enterprise data	2012–2015	Retailer	-	-	-	5–14 % (n.s.)	-	-
Wang & Ansell, 2020	X	CN	Enterprise data	2016	Taobao.com stores	4–40,6 % (n.s.)	-	-	-	-	-
Appriss Retail, 2019	-	US	Survey	2019	Retailer	9,6 % (n.s.)	-	-	-	-	-
Nicola, 2019	-	EU	Cited data	2018	Logistics provider	30–50 % (n.s.)	-	-	-	-	-
Shang et al., 2019	X	US	Enterprise data	1998–2004	E-tailer	-	7 % (n.s.)	-	-	-	-
Jack et al., 2019	-	-	Case study	-	Retailer	1–75 % (n.s.)	-	-	-	-	-
EHI Retail Institute, 2019	-	DE	Survey	2018–2019	E-tailer	20 % (β)	40 % (β)	10 % (β)	10–30 % (β)	20 % (β)	< 10 % (β)
Asdecker & Karl, 2018	X	DE	Enterprise data	2012–2013	E-tailer	-	59.8 % (α)	-	-	-	-
Abbey et al., 2018	-	US	Enterprise data	-	Retailer	16 % (γ)	-	-	-	-	-
Araújo et al., 2018	X	BRA	Interviews	2015	Consumer & e-commerce segment manager	3-4 % (γ)	-	-	-	-	-
Difrancesco et al., 2018	X	DE	Enterprise data	-	E-Tailer	-	40 % (n.s.)	-	-	-	-
Sahoo et al., 2018	X	US	Enterprise data	2010–2012	Retailer	-	15–22 % (n.s.)	-	-	7 % (n.s.)	-
Walsh & Möhring, 2017	X	EU	Enterprise data	-	Manufacturer and retailer	-	20 % (α)	-	-	-	-
Vilar-Zanon et al., 2017	X	US	Enterprise data	-	Retailer	-	25 % (n.s.)	-	-	-	-
Ellis, 2017	-	US	-	-	-	33 % (n.s.)	-	-	-	-	-
Woods, 2017	-	US	Cited data	-	Reverse logistics association	25–40 % (n.s.)	-	-	-	-	-
Asdecker et al., 2017	X	DE	Enterprise data	2014–2015	E-tailer	-	52.1 % (β) 63.3 % (α)	-	-	-	-
Bernon et al., 2016	X	UK	Enterprise data	2013–2014	Retailer and returns management 3PL organizations	-	8.1–38.2 % (γ)	6.4–10.3 % (γ)	-	5.0–12.7 % (γ)	-
Minnema et al., 2016	X	EU	Enterprise data	2011–2013	E-tailer	-	-	8.3 % (n.s.)	-	10 % (n.s.)	-
Ng & Stevens, 2015	-	-	Anecdotal	-	Logistics provider (Optoro)	10–25 % (n.s.)	30 % (n.s.)	-	-	-	-
Pettypiece, 2015	-	-	-	-	-	-	30 % (n.s.)	-	-	-	-
Stevens, 2014	-	-	-	-	-	-	50 % (n.s.)	-	-	-	-

Paper	Peer reviewed	Region	Data type	Year of data collection (if available)	Data source / survey population	Cross-sectional rates (type)	Fashion rates (type)	Entertainment rates (type)	Leisure rates (type)	Interior rates (type)	Other rates (type)
Hong & Pavlou, 2014	X	US	Enterprise data, survey	-	Consumers	13 % (n.s.)	-	-	-	-	-
Rao et al., 2014	X	US	Enterprise data	2010	Retailer	-	15 % (β)	-	-	-	-
Banjo, 2013	-	-	Survey	-	-	33 % (n.s.)	-	-	-	-	-
Pur et al., 2013	-	DE	Survey	2012	E-tailer	13 % (β)	26 % (β)	-	-	-	-
Griffis et al., 2012	X	ASIA	Enterprise data	-	E-tailer	-	--	2–11 % (n.s.)		-	-
Douthit et al., 2011	-	-	Survey	-	Manufacturer, retailer	-	-	11–20 % (n.s.)	-	-	-
Rösch, 2011	-	DE	-	-	-	-	50–70 % (n.s.)	-	-	-	-
Rabinovich et al., 2011	X	US	Enterprise data	2005–2006	E-tailer	-	-	-	-	2 % (n.s.)	-
Petersen & Kumar, 2010	-	US	Enterprise data	-	Retailer	-	16 % (n.s.)	-	-	-	-
Mollenkopf et al., 2007	X	US	Interview	2005	Retailer	-	5–20 % (n.s.)	5–20 % (n.s.)	5–20 % (n.s.)	-	< 5 % (n.s.)
Brohan, 2005	-	-	Survey	2005	E-tailer	< 10 % (n.s.)	-	-	-	-	-
Stock, 2004	-	-	Cited data	-	National Retail Federation	5.6 % (n.s.)	-	-	-	-	-
Catalog Age, 2002	-	US	-	-	-	-	10–40 % (n.s.)	-	5–10 % (n.s.)	-	-
Rogers et al., 2002	X	-	-	-	Retailer	-	< 40 % (n.s.)	-	-	-	-
Thomas, 2001	-	-	Anecdotal	-	Consulting firm and research firm	5.5–7.2 % (n.s.)	40 % (n.s.)	-	-	-	-
Del Franco, 2000	-	US	Anecdotal	-	Consulting firm	5–7 % (n.s.)	30 % (n.s.)	25–35 % (n.s.)	2–5 % (n.s.)	-	-
Rogers & Tibben-Lembke, 1998	X	US	Survey	-	Manufacturer, wholesaler, retailer, service provider	6 % (n.s.)	-	-	-	-	-

Legend: - = not available; α = shipment-based return rate; β = item-based return rate; γ = revenue-based return rate; n.s. = return rate not specified

II How Does The Covid-19 Pandemic Affect Consumer Returns: An Exploratory Study

Authors: David Karl, Björn Asdecker

Publication Outlet: Proceedings of the 50th European Marketing Academy Conference (EMAC), 2021
(ranked as D in VHB JOURQUAL 3, MARK)

published as:

Karl, D.; Asdecker, B. (2021): How Does the Covid-19 Pandemic Affect Consumer Returns: An Exploratory Study. In: *Proceedings of the 50th European Marketing Academy Conference*, Madrid, Spain, pp. 1–10. https://proceedings.emac-online.org/pdfs/A2021-94689.pdf

Abstract

In recent years, significant sales shifted from stationary retail to e-commerce. This trend is further accelerated by the Covid-19 pandemic, which has placed major constraints on brick-and-mortar retailers. Yet, there is still one critical aspect that threatens the success of the business model: consumer returns. Against this background, this paper conducts a systematic literature review to provide an overview about the extent of consumer returns. However, the published return rates vary widely and are mostly based on the data of a single e-tailer. Therefore, this study conducts an exploratory survey to capture product category-specific return rates and to explore the influences of Covid-19 on consumer return behavior. Theoretically, this research adds knowledge about the reverse part of the e-commerce business model and the effects of the current pandemic. Managerially, the collected data can be used for benchmarking purposes and can help to improve the e-tailers' returns management.

Keywords: Consumer Returns, E-Commerce, Covid-19

II.1 Introduction

In the past years, e-commerce experienced enormous growth rates. Recent research shows that this trend is further accelerated by the current Covid-19 pandemic. Faced with government regulations and restrictions, consumers have no alternatives but to shop online. Initial studies suggest that these experiences will lead to a lasting change in the consumer decision journey in favor of e-commerce (e.g., Kulkarni & Barge, 2020; Galhotra & Dewan, 2020). Despite this promising outlook, the e-commerce business model faces several challenges. Among them are consumer returns. To create trust and to encourage consumers to order despite the higher perceived uncertainty, companies are granting liberal return policies, which in turn lead to more returns than in traditional brick-and-mortar retailing (e.g., Giménez & Lourenço, 2008, Xia & Zhang, 2010). While plenty of research is available about e-commerce in general, comparatively little is known about the extent of consumer returns, particularly regarding the current effects of the Covid-19 pandemic. However, such data are necessary for many reasons. For instance, to better prepare pricing decisions or determining the profitability of the e-commerce business model.
Against this background, this exploratory study aims to contribute to the following overarching research question:

- *What is the extent of consumer returns in e-commerce and how does the Covid-19 pandemic affect it?*

To answer this question, this study uses a systematic literature review and conducts an online survey among German business-to-consumer (B2C) e-commerce firms.

II.2 Systematic Literature Review

Unlike other more general literature reviews (e.g., Abdulla et al., 2019), this work specifically focuses on publications that provide insights concerning the extent of consumer returns, which is operationalized through the observed return rates. Methodologically, the review follows the guidelines of Denyer and Tranfield (2009) who structure the research process in five steps: (1) question formulation; (2) locating studies; (3) study selection and evaluation; (4) analysis and synthesis; and (5) reporting and using the results. The first step refers to the research question, which is already derived in the introduction. The second step involves selecting the databases and defining the search terms. In that respect, five scientific databases were selected, namely Business Source Ultimate (BS), Science Direct (SD), JSTOR (JS), Web of Science (WS), and EconBiz (EB). To identify many potentially relevant publications, very general search terms were used, combining 'product returns', 'consumer returns', or 'customer returns' with 'rate*'. This resulted in 2,291 search hits (see Table II.1). In the third step, duplicates and studies that referred to 'stock returns' were removed, leaving 244 publications. Thereafter, titles, abstracts, and keywords were scanned to identify potentially relevant peer-reviewed studies that were read in full length. During this step, publications that only addressed brick-and-mortar retailing were excluded. In addition, the references of each of the 18 relevant articles were searched for further publications of interest. In total, the search led to 22 relevant publications in November 2020.

Search term	**BS**	**SD**	**JS**	**WS**	**EB**
('product returns' OR 'consumer returns' OR 'customer returns') + 'rate*'	242	1475	403	119	52
Combined results without duplicates & without 'stock return'	244				
Relevant hits + backward search = final selection	18 + 4 = 22				

Table II.1. Initial search and final selection.

Concerning the fourth step, it is important to note that retailers can determine their return rates in different ways as noted in Asdecker et al.: (1) the shipment-related return rate, which is referred to as alpha (α), (2) the item-related return rate, which is referred to as beta (β), and (3) the revenue-related return rate, which is referred to as gamma (γ). The α-return rate divides the number of returned shipments by the number of outbound shipments, whereas the β-return rate sets the number of returned items in relation to the number of shipped items. The γ-return rate divides the value of returned items by the value of ordered items. Table II.2 summarizes the relevant publications, which report a wide range of values.

Category	Sources
Cross-sectional rates	Rogers and Tibben-Lembke (2001): 6% [n.s.], Hong and Pavlou (2014): 13 % [n.s.], Abbey et al. (2018): 16 % [γ], Araújo et al. (2018): 3–4 % [γ], Wang and Ansell (2020): 4–40.6 % [n.s.]
Fashion	Rogers et al. (2002): <40 % [n.s.], Mollenkopf et al. (2007): 20–30 % [n.s.], Petersen and Kumar (2010): 16 % [n.s.], Rao et al. (2014): 15 % [β], Bernon et al. (2016): 8.1–38.2 % [γ], Asdecker et al. (2017): 63.3 % [α] / 52.1 [β], Vilar-Zanon et al. (2017): 25 % [n.s.], Walsh and Möhring (2017): 20 % [α], Asdecker and Karl (2018): 59.8 % [α], Difrancesco et al. (2018): 40 % [n.s.], Sahoo et al. (2018): 15–22 % [n.s.], Shang et al. (2020): 7 % [n.s.]
Entertainment	Mollenkopf et al. (2007): 5–20 % [n.s.], Griffis et al. (2012): 2–11 % [n.s.], Minnema et al. (2016): 8.3 % [n.s.], Bernon et al. (2016): 6.4–10.3 % [γ]
Leisure	Mollenkopf et al. (2007): 5–20 % [n.s.], Cui et al. (2020): 5–14 % [n.s.], Hofmann et al. (2020): 7.7 % [α] / 5.1 % [β]
Interior	Rabinovich et al. (2011): 2 % [n.s.], Bernon et al. (2016): 5.0–12.7 % [γ], Minnema et al. (2016): 10 % [n.s.], Sahoo et al. (2018): 7 % [n.s.]
Others	Daily needs: Mollenkopf et al. (2007): <5 % [n.s.]

Legend: [α] = Shipment-based return rate; [β] = Item-based return rate; [γ] = Revenue-based return rate; [n.s.] = Return rate not specified

Table II.2. Return rates reported in the literature.

The greatest consensus is that the fashion segment has the highest return rates. Moreover, only eight of the 22 studies (36 %) clearly define how the reported return rate was calculated, leaving room for speculation. A current study on the effects of the Covid-19 pandemic could not be found. Given this high degree of uncertainty, we conclude that a separate empirical study is needed to address this knowledge gap. For that purpose, this research draws on Germany, which is the largest national market within the European Union (EU). Globally, Germany ranks sixth with a turnover of $81.85 billion in 2019, following China ($1,934.78 billion), the United States ($586.92 billion), the United Kingdom ($141.93 billion), Japan ($115.40 billion), and South Korea ($103.48 billion) (Lipsman, 2019). Consumers in Germany have the right to revoke a purchase on the Internet within 14 days after delivery without providing reasons. In principle, consumers bear the direct costs of returning the goods to the retailer, unless the retailer voluntarily waives this right or the retailer fails to inform the customer about the return costs during the order. Despite the legal possibility, in practice, almost all major vendors refrain from doing so, which is why consumers have generally developed an expectation of customer-friendly policies and make frequent use of it. The following paragraph further describes the used methodology.

II.3 Methodology

For data collection, this research uses an online questionnaire. Unlike the paper-based counterpart, online surveys allow for a variable survey design and the integration of dynamic elements. Beyond that, the targeted participants that work in e-commerce most likely prefer the Internet as communication medium. The questionnaire can be structured in three parts.

Followed by a virtual cover letter, the first part queries essential characteristics of the respondents and asks for the respective product category with the largest share of sales. The purpose of this part is (1) to ensure that only professionals from e-commerce or multi-/omni-channel retailers participate and (2) to be able to refer to a specific product category in later questions. In the second part, changes in the outbound shipment quantities and the α-/β-return rate before and after the beginning of the Covid-19 pandemic in March 2020 are surveyed. The final third part asks for potential causes and other observed changes in the return behavior. Before the field phase, the survey was pretested by five experts with either an academic or an industry background. Based on their remarks, several minor changes were made concerning the wording of the questions. Sample acquisition was supported by the three leading German e-Commerce Associations (BEVH e.V., BVOH e.V., Händlerbund Management AG) who drew attention to the study via dedicated e-mail to their members, newsletters, or other communication channels such as XING or LinkedIn. The data collection took place between September and October 2020. During this period, 215 respondents started the survey of which 103 (48 %) e-commerce or multi-channel-merchants completed it without showing an unusual response behavior. The sample reflects the diversity of German e-commerce in terms of the most relevant product group clusters and company size (see Table II.3). The total e-commerce revenue that the participants realized in Germany in 2019 amounts to 12.1

billion €, which accounts for 16.6 % of total German e-commerce sales (BEVH, 2018).

Product cluster	Fashion (n=28, 27.2 %), leisure (n=23, 22.3 %), entertainment (n=13, 12.6 %), interior (n=12, 11.7 %), others (n=27, 22.3 %)
E-commerce revenue	Small companies with revenue < 10 mio. € (n=64, 62.1 %), medium-sized companies with revenue 10–100 mio. € (n=21, 20.4 %), large companies with revenue > 100 mio. € (n=18, 17.5 %)
Type of e-tailer	Pure e-tailer (n=63, 61.2 %), multi-/omni-channel retailer (n=40, 38.8 %)

Table II.3. Sample characteristics.

II.4 Results

Concerning the presentation of the results, this study refers to the most important product clusters ('entertainment', 'fashion', 'interior', 'leisure', 'others'), which are also used by the BEVH to categorize the German e-commerce market (BEVH, 2018). The first survey question referred to the relative changes in the number of outbound shipments. It shows an increase across all product categories, with the clusters 'leisure', 'entertainment' and 'interior' benefitting the most (see Table II.4).

Product cluster	**Entertainment (n=13)**	**Fashion (n=28)**	**Interior (n=12)**	**Leisure (n=23)**	**Others (n=27)**
Number of outbound shipments	+20.13 %	+13.76 %	+19.10 %	+22.56 %	+10.07 %
Items per shipment	+0.28 %	-6.37 %	+2.84 %	-0.84 %	-3.78 %

Table II.4. Average percentage changes in outbound e-commerce logistics (n=103).

This observation for the German e-commerce market resembles those that have been reported in other markets. These assessments are shared by several study participants in a field for open comments, as aptly summarized by the following statement: "*As bitter as it is, Corona was a stroke of luck for my [...] company.*" Since the study only considers the period from March to August with a lockdown of stationary retail from mid-March to early May, the figures for the entire year are very likely higher, because the strong Christmas business with a further lockdown of stationary retail in Germany is not taken into account. The next questions referred to the return rates. After calculating the means (Ø), standard deviations (SD) and confidence intervals (CI) of the two surveyed periods (P1: March–August 2019; P2: March–August 2020), the values are compared using t-tests to identify changes. Calculations were performed in SPSS 26. In case of significant differences ($\alpha<.05$), the fields in Table II.5 are shaded in gray.

Product cluster	**α-return rate** **Sample-ID: Ø, SD, 95 %-CI**	**β-return rate** **Sample-ID: Ø, SD, 95 %-CI**
Entertain-ment (n=13)	P1: 6.55 %, 2.75, 4.89 to 8.22 P2: 6.63 %, 1.40, 5.78 to 7.48	P1: 6.07 %, 2.87, 4.33 to 7.80 P2: 6.12 %, 1.19, 5.40 to 6.84
Fashion (n=28)	P1: 47.50 %, 13.96, 42.09 to 52.91 P2: 41.18 %, 11.66, 36.66 to 45.71	P1: 28.00 %, 10.05, 24.10 to 31.89 P2: 23.53 %, 9.83, 19.72 to 27.34
Interior (n=12)	P1: 7.48 %, 1.76, 6.36 to 8.60 P2: 5.56 %, 1.72, 4.46 to 6.65	P1: 7.30 %, 1.40, 6.41 to 8.20 P2: 5.48 %, 1.52, 4.52 to 6.45
Leisure (n=23)	P1: 14.11 %, 8.48, 10.44 to 17.78 P2: 13.73 %, 7.45, 10.51 to 16.95	P1: 11.74 %, 4.99, 9.59 to 13.90 P2: 11.50 %, 5.31, 9.20 to 13.79
Others (n=27)	P1: 3.88 %, 2.61, 2.85 to 4.91 P2: 3.77 %, 2.97, 2.60 to 4.95	P1: 3.77 %, 2.66, 2.71 to 4.82 P2: 3.68 %, 3.07, 2.47 to 4.89

Table II.5. Average return rates before (P1) and after (P2) the beginning of the Covid-19 pandemic (n=103).

The analysis shows the differences between the two kinds of return rates as well as the large variations among the respective product clusters. Concerning the effect of the Covid-19 pandemic, it shows that both returns rates were significantly lower in the 'fashion' (α: $t(27)=-10.739$, $p<.000$; β: $t(27)=-15.844$, $p<.000$) and 'interior' (α: $t(11)=-12.488$, $p<.000$; β: $t(11)=-8.554$, $p<.000$) product cluster. Therefore, in these clusters, e-tailers did not only benefit from more orders but also from lower return rates. The combination of these two effects acts as a lever on the profitability of the e-commerce business model (Asdecker, 2015). Participants who indicated a reduction in either the α- or β-return rate were asked for possible reasons. On the one hand, predefined categories on a Likert scale ranging from 1–5 (1=not relevant, 5=very relevant) were queried (see Table II.6).

Potential reasons	Mean value
During the pandemic, customers placed less selection orders.	+3.16
During the pandemic, our company attracted new customers with lower return rates than existing customers.	+3.09
During the pandemic, customers informed themselves more thoroughly about the items they ordered.	+3.06
During the pandemic, customers increasingly ordered product assortments with lower return rates.	+2.88
During the pandemic, our company reduced promotional activities that previously encouraged impulse purchases.	+1.59

Table II.6. Provided reasons by companies that reported a reduction in the return rate (n=32).

The assessments made provide evidence that customers changed their ordering behavior during the Covid-19 pandemic. Orders were placed more in line with actual requirements. In addition, e-commerce appears to have won over customer groups with a more moderate return behavior who previously shopped mainly in brick-and-mortar stores. On the other hand, additional reasons could be given by means of a free text field. Several study participants noted that they benefited from an internationalization, which the following statement summarizes: "*The customer segment has changed completely and therefore the return rate has halved. In 2019, I had 80 % German customers. Since the pandemic, it is only 40 % German customers, which means the demand in Italy, France, and Spain has increased dramatically. [...] Only in Germany are returns considered a legitimate behavior [...]. In all other EU countries, it is considered inappropriate.*" In addition, several retailers suggest that customers shied away from going to the post office for fear of catching the disease: "*[...] They returned less because they didn't want to go to the post office to return the package. They preferred to keep the goods.*" Other counteracting factors may have prevented even lower return rates. These include longer delivery times due to capacity problems at logistics service providers: "*[...] At the same time, there was a noticeable increase in returns for orders with longer delivery times, especially at the beginning of the pandemic. DHL did not meet the SLAs in some cases.*" Furthermore, some respondents observed opportunistic ordering behavior: "*Multiple orders from different companies. Only the goods that were delivered the fastest [...] were kept.*"

The final survey section dealt with further changes in the return behavior as observed by e-tailers. The collected data indicates (1) that the proportion of returns sent back in the last third of the return period has increased, (2) that the proportion of returns with obvious signs of use has increased and (3) that the proportion of returns that can only be disposed of due to their condition has slightly increased (see Table II.7).

	Decreased	**No change**	**Increased**
Share of returns sent back in the last third of the return period	3 % (n=3)	71 % (n=73)	26 % (n=27)
Share of returns with obvious signs of use	2 % (n=2)	74 % (n=76)	24 % (n=25)
Share of returns that was disposed due to their condition	4 % (n=4)	79 % (n=81)	17 % (n=18)

Table II.7. Evaluation of other relevant return behaviors.

Consequently, during the first months of the Covid-19 pandemic customers return later and use the ordered items more intensively before ultimately deciding whether to keep the ordered items, which – from an economic perspective – may offset some of the e-tailers' benefits from lower return rates.

II.5 Summary and Conclusion

This paper conducts a systematic literature review and an exploratory survey. The literature review summarizes the available descriptive statistics on return rates published in peer-reviewed papers to date and demonstrates the relevance of consumer returns to interested scholars. It also shows the need to clearly define the calculation method of the return rate in future studies to avoid comparing apples with oranges. The wide range of published values justifies an exploratory survey, which takes into account current influences from the Covid-19 pandemic. It shows that e-commerce in Germany is massively benefiting from the pandemic. In addition to double-digit growth rates, e-tailers with fashion and interior product assortments are also benefiting from significantly lower return rates. The study also identifies reasons for this change in behavior and points to further changes in the timing and the condition of returns. Thus, this study contributes to an emerging strand of literature analyzing the impact of the Covid-19 pandemic. While there are already several studies on the forward distribution process (e.g., Kulkarni & Barge, 2020; Galhotra & Dewan, 2020), this is the first to focus on the often neglected reverse part of the business model. Accordingly, it appears that the current Covid-19 pandemic not only affects where and what consumers purchase but also how they return. Managerially, e-tailers should consider those changes in their returns management processes. Besides, the collected data can be used for benchmarking purposes.

One limitation of this research is that it only considers German e-tailers. As the remarks of the study respondents point out, Germany is a country with comparatively high return rates. Other regions could therefore experience different changes. Furthermore, the big question is whether the observed changes are merely a temporary phenomenon. Further studies will be needed to investigate whether the observed and described changes in the consumers' decision journeys have long-term effects or not.

References

Abbey, J. D., Ketzenberg, M. E., & Metters, R. (2018). A More Profitable Approach to Product Returns. *MIT Sloan Management Review, 60*(1), 71–74.

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Araújo, A. C. de, Matsuoka, E. M., Ung, J. E., Massote, A., & Sampaio, M. (2018). An exploratory study on the returns management process in an online retailer. *International Journal of Logistics Research and Applications, 21*(3), 345–362.

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Asdecker, B., & Karl, D. (2018). Big data analytics in returns management - Are complex techniques necessary to forecast consumer returns properly? In *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics,* Valencia, Spain.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Bernon, M., Cullen, J., & Gorst, J. (2016). Online retail returns management. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 584–605.

BEVH. (2018). *Zweistelliges Wachstum in 2017 und weiterhin gute Perspektiven im E-Commerce*. Retrieved November 8, 2020, from https://www.bevh.org/presse/pressemitteilungen/details/zweistelliges-wachstum-in-2017-und-weiterhin-gute-perspektiven-im-e-commerce.html

Cui, H., Rajagopalan, S., & Ward, A. R. (2020). Predicting product return volume using machine learning methods. *European Journal of Operational Research, 281*(3), 612–627.

Denyer, D., & Tranfield, D. (2009). Producing a Systematic Review. In D. A. Buchanan & A. Bryman (Eds.), *The Sage handbook of organizational research methods* (pp. 671–689). Sage.

Difrancesco, R. M., Huchzermeier, A., & Schröder, D. (2018). Optimizing the return window for online fashion retailers with closed-loop refurbishment. *Omega, 78,* 205–221.

Galhotra, B., & Dewan, A. (2020). Impact of COVID-19 on digital platforms and change in E-commerce shopping trends. In*2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC),* Palladam, India, pp. 861–866.

Giménez, C., & Lourenço, H. R. (2008). e-SCM: internet's impact on supply chain processes. *The International Journal of Logistics Management, 19*(3), 309–343.

Griffis, S. E., Rao, S., Goldsby, T. J., & Niranjan, T. T. (2012). The customer consequences of returns in online retailing: An empirical analysis. *Journal of Operations Management, 30*(4), 282–294.

Hofmann, A., Gwinner, F., Fuchs, K., & Winkelmann, A. (2020). An Industry-Agnostic Approach for the Prediction of Return Shipments. In *Proceedings of the 26th Americas Conference on Information Systems,* Salt Lake City, United States of America.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

Kulkarni, S. R., & Barge, P. (2020). Effect of COVID-19 on the Shift in Consumer Preferences with Respect to Shopping Modes (Offline/Online) for Groceries: An Exploratory Study. *International Journal of Management, 11*(10), 581–590.

Lipsman, A. (2019). *Global Ecommerce 2019: Ecommerce Continues Strong Gains Amid Global Economic Uncertainty*. Retrieved January 20, 2021, from https://www.emarketer.com/content/global-ecommerce-2019

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To Keep or Not to Keep: Effects of Online Customer Reviews on Product Returns. *Journal of Retailing, 92*(3), 253–267.

Mollenkopf, D. A., Rabinovich, E., Laseter, T. M., & Boyer, K. K. (2007). Managing Internet Product Returns: A Focus on Effective Service Operations. *Decision Sciences, 38*(2), 215–250.

Petersen, J. A., & Kumar, V. (2010). Can Product Returns Make You Money? *MIT Sloan Management Review, 51*(3), 85–89.

Rabinovich, E., Sinha, R., & Laseter, T. M. (2011). Unlimited shelf space in Internet supply chains: Treasure trove or wasteland? *Journal of Operations Management, 29*(4), 305–317.

Rao, S., Rabinovich, E., & Raju, D. (2014). The role of physical distribution services as determinants of product returns in Internet retailing. *Journal of Operations Management, 32*(6), 295–312.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Rogers, D. S., & Tibben-Lembke, R. S. (2001). An Examination of Reverse Logistics Practices. *Journal of Business Logistics, 22*(2), 129–148.

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The Impact of Online Product Reviews on Product Returns. *Information Systems Research, 29*(3), 723–738.

Shang, G., McKie, E. C., Ferguson, M. E., & Galbreth, M. R. (2020). Using transactions data to improve consumer returns forecasting. *Journal of Operations Management, 66*(3), 326–348.

Vilar-Zanon, J. L., Vilar, E., & Heras, A. (2017). Online product returns risk assessment and management. *Top, 25*(3), 445–466.

Walsh, G., & Möhring, M. (2017). Effectiveness of product return-prevention instruments: Empirical evidence. *Electronic Markets, 27*(4), 341–350.

Wang, E., & Ansell, J. (2020). When Should E-Retailers Be Lenient? In *Proceedings of the 31st AMA Winter Academic Conference,* San Diego, United States of America.

Xia, Y., & Zhang, G. P. (2010). The Impact of the Online Channel on Retailers' Performances: An Empirical Evaluation. *Decision Sciences, 41*(3), 517–546.

Part C – Contributions on Returns Forecasting

III Forecasting E-Commerce Consumer Returns – A Systematic Literature Review

Author: David Karl

Publication Outlet: Management Review Quarterly
(ranked as B in VHB-Rating 2024)

DOI: 10.1007/s11301-024-00436-x

submitted as:

Karl, D. (2025): Forecasting e-commerce consumer returns – a systematic literature review. In: *Management Review Quarterly*, *75*(1), 1–56. DOI: 10.1007/s11301-024-00436-x

Abstract

The substantial growth of e-commerce during the last years has led to a surge in consumer returns. Recently, research interest in consumer returns has grown steadily. The availability of vast customer data and advancements in machine learning opened up new avenues for returns forecasting. However, existing reviews predominantly took a broader perspective, focussing on reverse logistics and closed-loop supply chain management aspects. This paper addresses this gap by reviewing the state of research on returns forecasting in the realms of e-commerce.
Methodologically, a systematic literature review was conducted, analyzing 25 relevant publications regarding methodology, required or employed data, significant predictors, and forecasting techniques, classifying them into several publication streams according to the papers' main scope. Besides extending a taxonomy for machine learning in e-commerce, this review outlines avenues for future research. This comprehensive literature review contributes to several disciplines, from information systems to operations management and marketing research, and is the first to explore returns forecasting issues specifically from the e-commerce perspective.

Keywords: Consumer returns; product returns; forecasting; prediction; literature review; e-commerce.

III.1 Introduction

E-commerce has witnessed substantial growth rates in recent years and continues growing by double-digit margins (National Retail Federation / Appriss Retail, 2023). However, lenient consumer return policies have resulted in $212 Billion worth of merchandise being returned to online retailers in the U.S. in 2022, accounting for 16.5 % of online sales (National Retail Federation / Appriss Retail, 2023). While high rates of consumer returns mainly concern specific sectors and product categories, online fashion retailing is particularly affected (Diggins et al., 2016). Recent studies report average shipment-related return rates for fashion retailers in the 40-50 % range (Difrancesco et al., 2018; Karl & Asdecker, 2021). In addition to missed sales and reduced profits (Zhao et al., 2020), consumer returns pose operational challenges (Stock & Mulki, 2009), including unavoidable processing costs (Asdecker, 2015) and uncertainties regarding logistics capacities, inventory management, procurement decisions, and marketing activities. Hence, effectively managing consumer returns is an essential part of the e-commerce business model (Urbanke et al., 2015).

Similar to the research conducted by Abdulla et al. (2019), this work focuses on consumer returns in online retailing (e-commerce), excluding the larger body of closed-loop supply chain (CLSC) management, which encompasses product returns related to end-of-life and end-of-use scenarios involving raw material recycling or remanufacturing. In contrast to CLSC returns, retail consumer returns are typically sent or given back unused or undamaged shortly after purchase, without any quality-related defects. These returns should be reimbursed to the consumer and are intended to be resold "as new" (de Brito et al., 2005; Melacini et al., 2018; Shang et al., 2020).

Regarding forecasting aspects, demand forecasting is a crucial activity for successful retail management (Ge et al., 2019). In contrast to demand and sales, returns constitute the "supply" side of the return process (Frei et al., 2022). Consequently, forecasting becomes a complex task and a significant challenge in managing returns due to the inherently uncertain nature of customer decisions regarding product retention (Frei et al., 2022). Moreover, return forecasts are interconnected with sales forecasts and promotional activities (Govindan & Bouzon, 2018; Tibben-Lembke & Rogers, 2002). Hence, forecasting objectives may vary, encompassing return quantities, timing (Hachimi et al., 2018), and even individual return probabilities. Minimizing return forecast errors is critical to reduce and minimize reactive planning (Hess & Mayhew, 1997). Accurate forecasts rely on (1) comprehensive data collection, e.g., regarding consumer behavior, and (2) information and communications technology (ICT) for data processing, such as big data analytics. Despite extensive research in supply chain management (SCM), Barbosa et al. (2018) noted a lack of relevant publications exploring the "returns management" process of SCM in conjunction with big data analytics. Specifically, “the topic of forecasting consumer returns has received little attention in the academic literature” (Shang et al., 2020, p. 342). Nonetheless, precise return forecasts positively impact reverse logistics activities’ economic, environmental, and social performance, primarily concerning quantity, quality, and timing predictions (Agrawal & Singh, 2020). Hence, forecasting returns holds significant relevance across various supply chain stages.

III.1.1 Previous Meta-Research

Hess and Mayhew (1997) emphasized the need for extensive data analysis concerning reverse flows, which forms the basis for returns forecasting. Subsequently, research on consumer returns and reverse logistics has proliferated. Thus, before collecting data and reviewing the topic of consumer returns forecasting, we first examined existing reviews and meta-studies relevant to the subject matter. To accomplish this, we referred to Web of Science, Business Source Ultimate via EBSCOhost, JSTOR and the AIS Electronic Library as primary sources of knowledge (search term: "literature review" AND "return*" AND "forecast*”). As a secondary source, we appended the results of Google Scholar[1], for which a different search term was used (intitle:"literature review" ("product return" OR "consumer return" OR "retail return" OR "e-commerce return") forecast) due to unavailable truncations and to reduce the vast amount of literature with financial focus the search term “return” would lead to. Table III.1

[1] The use of Google Scholar for systematic scientific information search is controversely discussed (e.g., Halevi et al. (2017)) due to the missing quality control and indexing guidelines, as well as limited advanced search options. But as an additional database for an initial search, the wide coverage of this search system can enrich the results.

presents the most pertinent literature reviews related to the scope of this paper.

Authors	Title	Period	Sources
Agrawal et al., 2015	A literature review and perspectives in reverse logistics	1986-2015	242
Hachimi et al., 2018	The optimization of reverse logistics activities: A literature review and future directions	1986-2018	45
Abdulla et al., 2019	Taking stock of consumer returns: A review and classification of the literature	2012-2018	100
Ambilkar et al., 2021	Product returns management: A comprehensive review and future research agenda	1986-2020	518
Micol Policarpo et al., 2021	Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review	2015-2020	70
Duong et al., 2022	Understanding product returns: A systematic literature review using machine learning and bibliometric analysis	until 2021	167

Table III.1. Relevant literature reviews related to consumer returns or returns forecasting, sorted by year of publication.

Agrawal et al. (2015) identified research gaps within the realm of reverse logistics, finding "forecasting product returns" as a crucial future research path. However, among 21 papers focusing on "forecasting models for product returns", the emphasis was predominantly on CLSC, reuse, remanufacturing, and recycling, which do not align with the aim of this review. Agrawal et al. also noted a lack of comprehensive analysis of underlying factors in returns forecasting, such as demographics or consumer behavior.

Similarly, Hachimi et al. (2018) addressed forecasting challenges within the broader context of reverse logistics. They classified their literature using various forecasting approaches: time series and machine learning, operations research methods, and simulation programs. The research gaps they identified included a limited number of influencing factors taken into account, the absence of established performance indicators, and methodological issues related to dynamic lot-sizing with returns. Although this review focused on reverse logistics, the call for research into predictors of future returns is equally applicable to consumer returns in e-commerce.

The review of Abdulla et al. (2019) centers on consumer returns within the retail context, particularly in relation to return policies. While they discuss consumer behavior and planning and execution of returns, they do not present any sources explicitly focused on forecasting issues.
Micol Policarpo et al. (2021) reviewed the literature on the use of machine learning (ML) in e-commerce, encompassing common goals of e-commerce studies (e.g., purchase prediction, repurchase prediction, and product return prediction) and the ML techniques suitable for supporting these goals. Their primary contribution is a novel taxonomy of machine learning in e-commerce, covering most of the identified goals. However, within the taxonomy developed, the aspect of return predictions is disregarded.
The most exhaustive literature review to date regarding product returns, conducted by Ambilkar et al. (2021), analyzed 518 papers and adopted a holistic reverse logistics approach encompassing all supply chain stages. The authors categorized the papers into six categories, including "forecasting product returns", for which they found and concisely described 13 papers. Due to the broader research scope, none of the analyzed papers focused on consumer returns within the retail context.
The review by Duong et al. (2022) employed a hybrid approach combining machine learning and bibliometric analysis. Regarding forecasts of product returns, they identified three relevant papers (Clottey & Benton, 2014; Cui et al., 2020; Shang et al., 2020) within the "operations management" category. They explicitly call for further research on predicting customer returns behavior in the pre-purchase stage, highlighting the importance of a better understanding of online product reviews and customers' online interactions.

III.1.2 Research Gap and Research Questions

Why is a systematic literature review necessary for investigating consumer returns and forecasting? On the one hand, there are empirical and conceptual papers that touch upon this topic, including brief literature reviews that align with the subject's focus (e.g., Hofmann et al., 2020). However, narrative reviews lack transparency and replicability (Tranfield et al., 2003) and often induce selection bias (Srivastava & Srivastava, 2006) as they tend to approach a field from a specific perspective. In contrast, systematic reviews strive to present a holistic, differentiated, and more detailed picture, incorporating the complete available literature (Uman, 2011). On the other hand, existing systematic reviews provide structured yet relatively superficial overviews of literature on end-of-use and end-of-life forecasting (Shang et al., 2020), but they do not specifically address consumer returns. Furthermore, we contend that a review dedicated to general reverse logistics forecasting would not adequately capture the distinctive context and requirements inherent in the consumer-retailer relationship within the realm of e-commerce (Abdulla et al., 2019).

Consequently, based on existing reviews and papers, we have identified research gaps worth examining more in detail: (1) Returns forecasting techniques and relevant predictors for the respective underlying purposes, especially in the context of e-commerce (RQ1 and RQ2); (2) the integration of return forecasts into an existing but incomplete taxonomy of machine learning in e-commerce (Micol Policarpo et al., 2021; RQ3); and (3) future research directions pertaining to e-commerce returns forecasting (RQ4). Therefore, this review aims to shed more light on consumer returns forecasting in the retail context. The following research questions outline the primary objectives:

- *RQ1: What key research problems (e.g., forecasting purposes, technological approaches) have been addressed in the literature on forecasting consumer returns over time?*
- *RQ2: What are the ...*
    - *publication outlets and research disciplines,*
    - *research types and methodologies,*
    - *product categories and industries,*
    - *data sources and characteristics,*
    - *relevant forecasting predictors,*
    - *techniques and algorithms*

  ... used to address these key problems?
- *RQ3: How can returns forecasting be integrated into a taxonomy of machine learning in e-commerce?*
- *RQ4: What are promising or emerging future research directions regarding forecasting consumer returns?*

The paper is organized as follows: Section 2 describes selected fundamental concepts and the delimitation of the research field on consumer returns forecasting. Section 3 contains the methodology for the review, drawing on the PRISMA guideline (Page et al., 2021) while integrating the approaches of Denyer and Tranfield (2009) and Webster and Watson (2002). Section 4 presents the review's main results, answering RQ1 (Section 4.1), RQ2 (Section 4.2-4.5), and RQ3 (Section 4.6). A research framework developed in Section 5 structures the discussion regarding future research directions (RQ4). Section 6 subsumes the overall contribution of this review.

III.2 Consumer Returns and Forecasting

III.2.1 Consumer Returns and Return Reasons

Reverse product flows, commonly referred to as product returns, can be classified into three categories: manufacturing returns, distribution returns, and consumer returns (Shaharudin et al., 2015; Tibben-Lembke & Rogers, 2002). Among these, *consumer returns* are further differentiated between returns in brick-and-mortar retail or mail-order / e-commerce returns (Tibben-Lembke & Rogers, 2002) and are also known as *commercial returns* (de Brito et al., 2005) or *retail (product) returns* (Bernon et al., 2016). With sky-rocketing e-commerce sales, online consumer returns have emerged as the dominant segment, making them a highly relevant field of research (Abdulla et al., 2019; Frei et al., 2020). Additionally, the digitization of retail provides numerous opportunities for data collection, as digital customer accounts facilitate more efficient analytical monitoring of customer behavior (Akter & Wamba, 2016). Simultaneously, as competitive pressures intensify in e-commerce due to increased price transparency and substitution possibilites, retailers aiming to stimulate impulse purchases face hightened return rates (Cook & Yurchisin, 2017; Karl et al., 2022).

The spatial decoupling of supply and demand introduces a higher level of uncertainty for e-commerce customers regarding various product attributes compared to bricks-and-mortar retailing (Hong & Pavlou, 2014). As consumers are unable to physically assess the products they order, this translates into returns being essential part of the e-commerce business model. Besides fit uncertainty, other reasons for returns exist. Stöcker et al. (2021) classify the drivers triggering consumer returns into *consumer behavior* related reasons (e.g., impulsive purchases, showrooming), *fulfillment/service* related reasons (e.g., wrong/delayed delivery) and *information gap* related reasons (product fit, insufficient visualization). By mitigating customers' return reasons, retailers try to reduce the return likelihood ("return avoidance") (Rogers et al., 2002). Another, but less promising way of reducing returns, is preventing customers who intend to return from actually doing so (e.g., by incurring additional effort or by rejecting returns) (Rogers et al., 2002).

Adapted from Abdulla et al. (2019) and Vakulenko et al. (2019), a simplified parallel process of a return transaction from the consumer's and retailer's perspective is visualized in Figure III.1. Retailers can use forecasting in all transaction phases (Hess & Mayhew, 1997). Targeting customer interventions pre-purchase (real-time forecasting) could be implemented by using dynamically generated (Dalecke & Karlsen, 2020) digital nudging elements (Kaiser, 2018; Thaler & Sunstein, 2009; Zahn et al.) in case of a predicted high return propensity. In the post-purchase phase, forecasting could stimulate different interventions (e.g., customer support) or can be helpful for logistics and inventory planning activities (Hess & Mayhew, 1997). In the phase after the return decision, data analysis, including segmentation on different levels, e.g., for customers, products, or brands (Shang et al., 2020), can support managerial decision-making regarding assortment or (individualized) return policies for future orders (Abdulla et al., 2019). In other words, forecasting (or modeling) of returns in later phases of the process can substantiate interventions in earlier phases of the process (e.g., a temporary return policy change, or the suspension of product promotions due to particular forecasts). However, such data-driven interventions itself also represent an influencing factor to be taken into account in future forecasts; thus, different forecasting purposes can be linked, at least when it comes to the data required. All these interdependencies hint at the circularity of the returns process, with an adequate management of returns representing an opportunity for generating customer satisfaction and retention (Ahsan & Rahman, 2016; Röllecke et al., 2018).

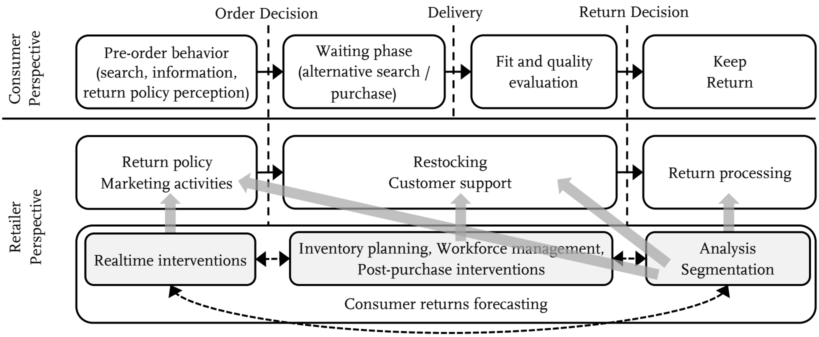

Figure III.1. Purchase and return process concerning forecasting issues (adapted from Abdulla et al., 2019; Vakulenko et al., 2019)

III.2.2 Forecasting Purposes and Corresponding Techniques

Accurate forecasting holds significant importance in the realm of e-commerce. Precise demand forecasts ("predictions") play a pivotal role in inventory planning, pricing, and promotions and ultimately impact the commercial success of retailers (Ren et al., 2020). Forecasting consumer returns affects similar business aspects and resorts to comparable existing technical procedures. The data science and statistics literature offers diverse methods and algorithms for forecasting consumer returns. The choice of approach depends on the specific objective, with the outcome variable being scaled accordingly. For instance, when forecasting whether a single product will be returned, the dependent variable is either binary or expressed as a propensity value ranging form 0 to 1. On the other hand, forecasting the quantitay or timing of returns entails continuous outcome variables. As a result, various techniques, from time-series forecasting to machine learning approaches can be applied, which will be briefly outlined in the subsequent sections.

III.2.2.1 Return Classifications and Propensities

A naïve method for determining the propensity or return decision forecast is using lagged (historical) return information (return rates), either for a given product, a given customer, or any other reference, to calculate a historical return probability (Hess & Mayhew, 1997). Return rate forecasts are a reference-specific variant of forecasting return propensities.

Simple causal models based on statistical regression methods utilize one or more independent exogenous variables. The *logistic regression (logit model)* is employed when the dependent variable is binary or contains more nominal outcomes (multinomial logistic regression). For each observation, the binary logistic regression assesses the probability that the dependent variable takes the value “1” (Hastie et al., 2017). Consequently, this approach finds application for return decisions and return propensities. Comparatively, *linear discriminant analysis* (Fisher, 1936) bears a resemblance to logistic regression by generating a linear combination of independent variables to best classify available data. This classification process involves determining a score for each observation, subsequently compared to a critical discriminant score threshold, and distinguishing between return and keep.

More sophisticated machine learning (ML) techniques such as neural networks, decision tree-based methods, ensemble learning, and boosting methods are highly suitable for this forecasting purpose. For a general exposition of ML techniques in the domain of e-commerce, we refer to Micol Policarpo et al. (2021). Additionally, for a comparative study of several state-of-the-art ML classification techniques, see Fernández-Delgado et al. (2014). Artificial *Neural Networks* (NN) consist of interconnected nodes (“neurons”) organized in layers, exchanging signals to ascertain a function that accurately assigns input data to corresponding outputs. Typically, supervised learning techniques such as backpropagation compare the network outputs with known actual values (Hastie et al., 2017). Notably, neural networks are the most popular machine learning algorithm in last years’ e-commerce research (Micol Policarpo et al., 2021), and deep learning extensions like Long Short-Term Memory (Bandara et al., 2019) are gaining attention. *Decision Trees* (DT) manifest as hierarchical structures of branches representing conjunctions of specific characteristics and leaf nodes denoting class labels. This approach endeavors to construct an optimal decision tree for classifying available observations. Many decision tree algorithms have been introduced to serve this purpose (e.g., Breiman et al., 1984; Pandya & Pandya, 2015). *Ensemble learning methods* adopt a voting mechanism involving multiple algorithms to enhance predictive performance (Polikar, 2006). Analogously, boosting and bagging techniques are incorporated in algorithms like *AdaBoost* or the tree-based

Random Forest (RF) to augment the input data, aiming at more generalizable forecasting models less prone to overfitting issues (Hastie et al., 2017). *Support Vector Machines* (SVM) stand as another example of a supervised ML algorithm, having demonstrated efficacy in tackling classification problems within e-commerce (Micol Policarpo et al., 2021).

III.2.2.2 Return Timing and Volume Forecasts

For product returns, timing is crucial in forecasting end-of-life, end-of-use, or remanufacturing returns that can occur years after the initial purchase (Petropoulos et al., 2022). In contrast, for consumer returns, the possible time window in which products are regularly returned in new condition with the aim of a refund is much shorter (usually less than 100 days and mostly less than 30 days), and priorities are more on forecasting return volumes. Forecasting return volumes can be multi-faceted, ranging from forecasting the total return volume a retailer has to process within its logistics department through forecasting product-specific return numbers up to forecasting costly return shares, e.g., return fraud volume. Because returns depend on fluctuating sales, time-series forecasting of return volumes performs only well with constant sales volumes or under risk-pooling (Petropoulos et al., 2022). Thus, for a naïve return volume forecast, sales forecasts for a given timeframe are multiplied by the lagged return rate (historical data of products/consumers or any other reference). Possible algorithms for estimating historical return rates include time series forecasting to causal predictions comprising ML approaches (Hachimi et al., 2018).

Time-series techniques, e.g., single exponential smoothing (SES) or Holt-Winters-approaches (HW), are based on the assumption that the future development of an outcome variable (e.g., return volume) is dependent on its past numbers, while time acts as the only predictor. Most of these models can be generalized as autoregressive moving averages (ARIMA) models, for which numerous extensions are available. These models can approximate more complex temporal relationships. Similarly, time-series regression models use univariate linear regression with time as a single exogenous variable.

The mentioned multivariate regression models are essential statistical tools and can predict metric variables such as return volume or time.

The logic is to fit a linear function of a given set of input variables ("features") to the outcome variable with the criteria of minimizing the residual sum of squares (Hastie et al., 2017). Many variants of regression models are derived from this logic (e.g., generalized linear models), and various extensions are built upon this base (e.g., LASSO for variable selection, Tibshirani, 1996).
Emerging from more complex statistical methods and using the possibilities of continuously increasing computing power, IT-based machine learning (ML) approaches were developed. Some of these approaches have already been presented in Section 2.2.1, being suitable for predicting metric variables in addition to classification tasks, e.g., neural networks, decision tree algorithms, and especially ensemble techniques like random forests.

III.3 Methodology

Methodologically, the research process of this review follows the PRISMA guideline (Page et al., 2021) where applicable and is structured in five steps (Denyer & Tranfield, 2009; Webster & Watson, 2002): (1) question formulation; (2) locating studies; (3) study selection and evaluation; (4) (concept-centric) analysis and synthesis; and (5) reporting and using the results for defining an agenda for future research.
The first step refers to the research questions already formulated in the introduction. The second step involves selecting the databases and defining the search terms. In that respect, five scientific databases were selected, aiming at journal as well as conference publications: AIS Electronic Library (AISeL), Business Source Ultimate (BS) via EbscoHost, JSTOR (JS), Science Direct (SD), and Web of Science (WoS). To ensure inclusivity and to account for potential variations in spelling or phrasing, the final search strings incorporate truncations where applicable. The search query utilized in this review comprises two key components. Firstly, it pertains to consumer returns, encompassing products returned by consumers, primarily in the context of e-commerce, to the retailer. While it is recommended to use reasonably general search terms, the term "return" alone would yield results for various stages of reverse logistics and a vast amount of financial literature. Therefore, we conducted a more specific search using the phrase "consumer return*" and the related

terms “e-commerce return*”, “product return*”, “return* product”, “customer return*”, and “retail return*”. Secondly, this paper specifically focuses on forecasting (“forecast*”), which can be alternately referred to as “predict*” or “prognos*”. The combination of these terms was searched for in the Title, Abstract and Keywords fields.
The search includes results up to the middle of 2022 and resulted in 725 initial search hits (see Figure III.2). As this review aims to identify papers dealing with consumer returns and forecasting, the inclusion criteria for eligibility were:

- The title or keywords referred to consumer returns or forecasting (in a broader sense, including data preparation). A connection to the respective subject area and applicability to the retail domain should at least be plausible.
- Manuscript in English: No important study would be written and published in a language different than English.
- The paper has undergone a single- or double-blind peer-review process, either as a journal publication or as a publication in peer-reviewed conference proceedings.

In the third step, duplicates were removed, resulting in a set of 650 unique records. Subsequently, the papers underwent screening based on title, keywords, and language to determine whether they warranted further examination. This preliminary screening phase reduced the number of papers to 85. These papers’ abstracts and full texts were thoroughly reviewed to assess their relevance. This step encompasses all papers pertaining to returns forecasting for retailers or direct-selling manufacturers while excluding those focused on closed-loop supply chain management or remanufacturing, recycling, and end-of-life returns. Ultimately, a final sample of 20 publications was identified, serving as a foundation for identifying additional relevant papers (vom Brocke et al., 2009; Webster & Watson, 2002) through a forward search using Google Scholar and snowballing via backward search. This process yielded an additional five papers, resulting in a total of 25 papers included for review (Table III.2).

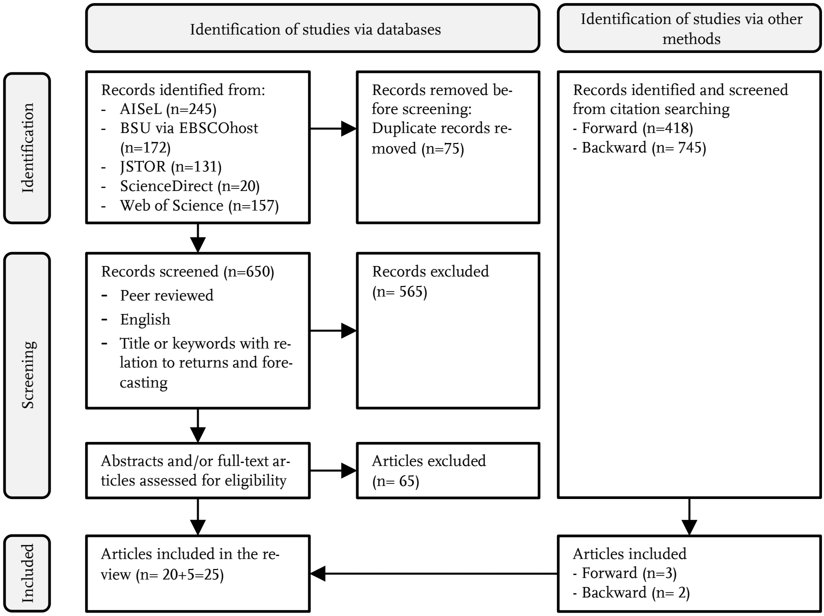

Figure III.2. Research process flow diagram.

The fourth step comprises the analysis and synthesis of the relevant papers. Data, including bibliographic statistics, were collected in accordance with the research questions. A two-way concept-centric analysis, as described by Webster and Watson (2002), was conducted, encompassing confirmatory aspects based on the fundamentals outlined in Section 2 of this paper, as well as exploratory elements aimed at enriching existing categories and concepts. The objective was to comprehensively describe the relevant concepts, approaches, and dimensions discussed in the literature.

Moving on to the fifth and final step (Denyer & Tranfield, 2009), the results are presented. Initially, the main scope of the papers included in the analysis is presented. Next, bibliographic data pertaining to the included papers are provided to offer a concise overview of the research area and its recent developments, followed by a content analysis and synthesis of the relevant literature to delve into the current state of research and highlight key findings. Finally, Section 5 outlines a research agenda for the domain (vom Brocke et al., 2009).

Author(s), Year	Country	J/C	Rank	Outlet Title	Paper Title
Hess & Mayhew, 1997	US	J	Q2	Journal of Direct Marketing	Modeling merchandise returns in direct marketing
Potdar & Rogers, 2012	US	J	Q2	Foresight	Reason-code based model to forecast product returns
Drechsler & Lasch, 2015	DE	C	/	Logistics Management	Forecasting Misused E-Commerce Consumer Returns
Urbanke et al., 2015	DE	C	A*	International Conference on Information Systems (ICIS)	Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction
Fu et al., 2016	US/CN	J	Q1	Decision Support Systems	Fused latent models for assessing product return propensity in online commerce
Ahmed et al., 2016	CA/US	C	B	IEEE Big Data	Advantage of Integration in Big Data: Feature Generation in Multi-Relational Databases for Imbalanced Learning
Heilig et al., 2016	DE	C	A	European Conference on Information Systems (ECIS)	Data-Driven Product Returns Prediction: A Cloud-Based Ensemble Selection Approach
Ding et al., 2016	SG	C	A	Pacific Asia Conference on Information Systems (PACIS)	Predicting Product Return Rate with "Tweets"
Samorani et al., 2016	US/CA	C	B	IEEE Big Data	Automatic Generation of Relational Attributes: An Application to Product Returns
Urbanke et al., 2017	DE	C	A*	International Conference on Information Systems (ICIS)	A Customized and Interpretable Deep Neural Network for High-Dimensional Business Data - Evidence from an E-Commerce Application

Author(s), Year	Country	J/C	Rank	Outlet Title	Paper Title
J. Li et al., 2018	US	C	A*	International Conference on Knowledge Discovery and Data Mining (ACM)	E-tail Product Return Prediction via Hypergraph-based Local Graph Cut
Zhu et al., 2018	US	C	A*	International Joint Conference on Artificial Intelligence (IJCAI)	A Local Algorithm for Product Return Prediction in E-Commerce
Asdecker & Karl, 2018	DE	C	/	International Conference on Advanced Research Methods and Analytics (CARMA)	Big data analytics in returns management - Are complex techniques necessary to forecast consumer returns properly?
Joshi et al., 2018	IN	C	/	International Conference on Advances in Social Networks Analysis and Mining (IEEE/ACM)	One Size Does Not Fit All: Predicting Product Returns in E-Commerce Platforms
X. Li et al., 2019	CN	J	Q1	Science China Information Sciences	A trust-aware random walk model for return propensity estimation and consumer anomaly scoring in online shopping
Cui et al., 2020	US	J	Q1	European Journal of Operational Research	Predicting product return volume using machine learning methods
Shang et al., 2020	US	J	Q1	Journal of Operations Management	Using transactions data to improve consumer returns forecasting
Imran & Amin, 2020	BD/FR	C	B	International Conference on Knowledge-Based and Intelligent Information & Engineering Systems Systems	Predicting the Return of Orders in the E-Tail Industry Accompanying with Model Interpretation
Ketzenberg et al., 2020	US	J	Q1	Journal of Operations Management	Assessing customer return behaviors through data analytics
Hofmann et al., 2020	DE	C	A	Americas Conference on Information Systems (AMCIS)	An Industry-Agnostic Approach for the Prediction of Return Shipments
John et al., 2020	IN	J	Q3	Journal of Business Analytics	Refund fraud analytics for online retail purchases

Author(s), Year	Country	J/C	Rank	Outlet Title	Paper Title
Rezaei et al., 2021	CA/US	J	Q1	Annals of Operations Research	A clustering-based feature selection method for automatically generated relational attributes
Rajasekaran & Priyadarshini, 2021	IN	C	/	Advances in Computing and Data Sciences (ICACDS)	An E-Commerce Prototype for Predicting the Product Return Phenomenon Using Optimization and Regression Techniques
Sweidan et al., 2020	SW	C	C	International FLINS Conference (FLINS)	Predicting returns in men's fashion
Fuchs & Lutz, 2021	DE	C	A	European Conference on Information Systems (ECIS)	A stitch in time saves nine – A meta-model for real-time prediction of product returns in ERP systems

Table III.2. Summary of the reviewed papers, sorted by year of publication (J=Journal, C=Conference, Journal Ranking according to Scimago JR, Conference Ranking according to CORE if available).

III.4 Results of the Systematic Review

After outlining the main scope of the relevant publications (4.1), a short bibliographic characterization (4.2) is given. Next, this section presents the results of the systematic review, focussing on the methodology and datasets used (4.3), predictors used for returns forecasting (4.4), and forecasting techniques employed (4.5). The integration of consumer returns forecasting into an existing taxonomy for e-commerce and machine learning (Micol Policarpo et al., 2021) summarizes and concludes the presentation of the results.

III.4.1 Overview and Main Scope of the Relevant Publications

Table III.3 provides an overview of the forecasting purpose of the papers, the data source for the forecasting, the algorithms employed, and the predictors used in the forecasting models. The contributions of the respective papers regarding forecasting issues are summarized in the Appendix.

Paper	Forecasting Purpose	Data Source	Predictors
Journal papers			
Hess & Mayhew, 1997	Return volume over time	Fashion direct marketer, and simulated data	Product attributes (price, fit importance), lagged timestamps
Potdar & Rogers, 2012	Return volume	Consumer electronics, and simulated data	Return reason codes, behavior data
Fu et al., 2016	Return propensity on customer-product-level	Taobao, cosmetics	Product features, user profiles
X. Li et al., 2019	Return propensity, customer clustering	Taobao, cosmetics	Transaction level data, customer profile, product profile
Shang et al., 2020	Return volume and timing	Offline electronics, jewelry	Transaction level data (timestamps, lagged returns)
Cui et al., 2020	Return volume (C-R and C-OEM)	Automotive accessories	Sales volume, product return rate, sales channel, time, product attributes
Ketzenberg et al., 2020	Return abuser detection	Department store retailer (fashion, jewelry, etc.)	Transactional data, customer attributes
Rezaei et al., 2021	Return decision (feature selection)	Simulated data, and consumer electronics	Price, customer attributes
John et al., 2020	Return fraud detection	E-commerce generalist	Agent attributes, order attributes, return details
Conference Papers			
Urbanke et al., 2015	Return decision (feature selection)	Fashion e-tailer	Product attributes, user profiles, basket attributes
Asdecker & Karl, 2018	Return decision	Fashion e-tailer	Product attributes, customer attributes, shipment attributes
Drechsler & Lasch, 2015	Misused return volume	Simulated data	Lot-based sales volume, timing information, lagged return rates
Sweidan et al., 2020	Return decision (order-based)	Fashion e-tailer	Order information, customer attributes
Ahmed et al., 2016	Return decision	Electronics retailer	Automatically aggregated and selected
Samorani et al., 2016	Return decision	Electronics retailer	Automatically generated, e.g., price, brand perception

Paper	Forecasting Purpose	Data Source	Predictors
Heilig et al., 2016	Return decision	Fashion e-tailer	Order information, product attributes, customer attributes
Ding et al., 2016	Overall daily return rate	E-commerce Website	Sentiment analysis of tweets regarding the company
Urbanke et al., 2017	Return decision	Fashion and sports e-tailer	Order information, product features
Zhu et al., 2018	Return decision	Fashion e-tailer	Customer attributes, product attributes
Joshi et al., 2018	Return propensity and decision	E-commerce / apparel	Customer purchase / return history
J. Li et al., 2018	Return decision (order- and product-based)	Omnichannel fashion	Product attributes, historical purchase/return records
Imran & Amin, 2020	Return decision	E-commerce Bangladesh	Order attributes
Hofmann et al., 2020	Return decision (order-based)	Technical wholesaler	Order / return numbers
Fuchs & Lutz, 2021	Real-time basket return detection	Slow-moving products	Basket composition
Rajasekaran & Priyadarshini, 2021	Product-based return probability	E-commerce dataset	Product feedbacks, product attributes, time

Table III.3. Overview of relevant publications and their forecasting models.

For identifying research streams, the publications are analyzed regarding the intention and main scope, as described in the abstract, the respective research questions, and the remainder of the papers. Most papers were assigned to an unequivocal research scope, while some contributed to two key topics (see Figure III.3).

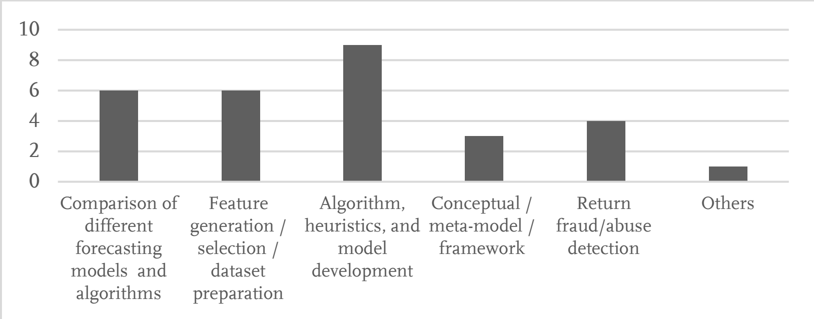

Figure III.3. Classification of main scopes (n=25; not mutually exclusive).

At first, we identified a stream of literature regarding the *comparison of different forecasting models and algorithms* (Asdecker & Karl, 2018; Cui et al., 2020; Drechsler & Lasch, 2015; Heilig et al., 2016; Hess & Mayhew, 1997; Hofmann et al., 2020; Imran & Amin, 2020). These papers use existing approaches, adapt them for individual forecasting purposes, apply models to one or more datasets, and compare and evaluate the resulting forecasting performance. One paper claims that the difference in forecasting accuracy of easily interpretable algorithms is relatively small compared to more sophisticated ML algorithms (Asdecker & Karl, 2018). This statement is partially confirmed (Cui et al., 2020), as the ML algorithms show advantages over simpler models in the training data set but have lower prediction quality due to overfitting issues in the test data. Nevertheless, fine-tuned ML approaches (e.g., deep learning with TabNet) outperform simpler models and gain accuracy when correcting class imbalances during the data preparation phase (Imran & Amin, 2020). When confronted with large class imbalances (e.g., low return rates), boosting algorithms like Gradient Boosting work well without oversampling (Hofmann et al., 2020). Fundamentally, ensemble models incorporating different techniques show the maximum possible accuracy (Asdecker & Karl, 2018; Heilig et al., 2016). Forecasting of return timing is more erroneous than return decisions, and split-hazard-models outperform simple OLS approaches (Hess & Mayhew, 1997). Time series prediction only works reliably when return rates do not fluctuate heavily (Drechsler & Lasch, 2015).

The second stream we identified focuses on *feature generation or selection and dataset preparation* (Ahmed et al., 2016; Ding et al., 2016; Hofmann et al., 2020; Rezaei et al., 2021; Samorani et al., 2016; Urbanke et al., 2015; Urbanke et al., 2017). Besides this central topic, some papers also compare different forecasting algorithms (Ahmed et al., 2016; Hofmann et al., 2020; Rezaei et al., 2021; Urbanke et al., 2015; Urbanke et al., 2017). For example, random oversampling of data with large class imbalances can improve the performance of different forecasting algorithms, while models based only on sales/return history perform worse than models with more features (Hofmann et al., 2020). Two similar approaches are based on product, basket, and clickstream data, using different algorithms for feature extraction (Urbanke et al., 2015; Urbanke et al., 2017). The first developed a Mahalanobis Feature Extraction algorithm, proving superior to other algorithms like principal component analysis or non-negative matrix factorization (Urbanke et al., 2015). The second develops a NeuralNet algorithm to extract interpretable features from a high-dimensional dataset, showing superior performance and giving reasonable interpretability of the most important factors (Urbanke et al., 2017). For the automated integration of different data sources into single flat tables and the generation of discriminating features, a rolling-path algorithm is developed, improving performance when data is imbalanced (Ahmed et al., 2016). Similarly, the software "Dataconda" can automatically generate and integrate relational attributes from different sources into a flat table, which is often the required prerequisite for forecasting algorithms (Samorani et al., 2016). A different selection approach clusters the features into groups and applies selection algorithms to the groups, aiming to select a smaller set of attributes (Rezaei et al., 2021). As quite an offshoot, one paper predicts a seller's overall daily return volume dependent on his current "reputation" measured by tweets (Ding et al., 2016), which needs sentiment analysis to be integrated into the forecast.

A quite heterogenous research stream belongs to the *development of algorithms, heuristics, and models* that go beyond a straightforward adaption of existing approaches (Fu et al., 2016; Joshi et al., 2018; J. Li et al., 2018; Potdar & Rogers, 2012; Rajasekaran & Priyadarshini, 2021; Shang et al., 2020; Sweidan et al., 2020; Zhu et al., 2018). Potdar and Rogers (2012) developed a methodology for forecasting product returns based on reason codes and consumer behavior data. Fu et al. (2016) developed a conditional probability-based statistical model for predicting return propensities while revealing return reasons and outperforming some baseline benchmark models. Li et al. (2018) describe their "HyperGo" approach as a 'framework' and develop an algorithm for forecasting return intention after basket composition. Zhu et al. (2018) describe a "LoGraph" random walk algorithm for predicting returned customer/product combinations within their framework. Although Joshi et al. (2018) label their approach as a "framework", they describe a specific two-stage algorithm for forecasting return decisions based on network science and ML. Rajasekaran and Priyadarshini (2021) developed a hybrid metaheuristic-based regression approach to predict return propensities.

Seven papers deal with *concepts, meta-models, or substantial frameworks* for returns forecasting (Fu et al., 2016; Fuchs & Lutz, 2021; Heilig et al., 2016; Hofmann et al., 2020; J. Li et al., 2018; Shang et al., 2020; Zhu et al., 2018). A generic framework for a scalable cloud-based platform, which enables a vertical and horizontal adjustment of resources, could enable the practical real-time use of computationally intensive ML algorithms for forecasting returns in an e-commerce platform (Heilig et al., 2016). Two papers (Fuchs & Lutz, 2021; Hofmann et al., 2020) are based on design science research (DSR, Hevner et al., 2004) for developing artifacts like meta models and frameworks. The first also refers to CRISP-DM, the "Cross Industry Standard Process for Data Mining" (Wirth & Hipp, 2000), and develops a shopping-basket-based general forecasting approach suitable across different industries without domain knowledge and attributes needed (Hofmann et al., 2020). In a similar approach, based on the basket composition and user interactions, a generic model for real-time return prediction and intervention is developed (Fuchs & Lutz, 2021) and prepared for integration into an ERP system. Fu et al. (2016) present a generalized return propensity latent model framework by decomposing returns into different inconsistencies (unmet product expectations, shipping issues, and both factors combined) and enriching the derived propensities with product features and customer profiles. Li et al. (2018) developed a "HyperGo" framework for forecasting the return intention in real-time after basket composition, including a hypergraph representation of historical purchase and return information. Similarly, Zhu et al. (2018) developed a

"HyGraph" representation of historical customer behavior and customer/product similarity, combined with a "LoGraph" random-walk-based algorithm for predicting customer/product combinations that will be returned. Shang et al. (2020) discuss two opposing forecasting concepts, demonstrating that their predict-aggregate framework is superior to common and more naïve aggregate-predict approaches.

The last stream covers the *detection and forecasting of return fraud and abuse* (Drechsler & Lasch, 2015; John et al., 2020; Ketzenberg et al., 2020; X. Li et al., 2019). On the employees' side, one paper tries to automatically predict fraudulent return behavior of agents (employees), e.g., regarding unjustified refunds, by a penalized logit model, enabling a lift in detection (John et al., 2020). On the customers' side, misused returns as a cost-incurring problem are the forecasting purpose of different time series prediction models (Drechsler & Lasch, 2015). Instead of focussing on fraudulent transactions, a trust-aware random walk model identifies consumer anomalies, enabling retailers to apply targeted measures to specific customer groups (selfish, honest, fraud, and irrelevant customers) (X. Li et al., 2019). Similarly, returning customers can be categorized into abusive, legitimate, and nonreturners (Ketzenberg et al., 2020). Based on the characterization of abusive return behavior, a neural network classifier recaptures almost 50 % of lost profits due to return abuse (Ketzenberg et al., 2020).

One paper (Sweidan et al., 2020) could not be assigned to the other scopes. It applies a single algorithm (RF) to a given dataset, and it contributes to the idea that only forecasted return decisions with high confidence should be used for targeted interventions due to their overproportional reliability.

III.4.2 Bibliographic Literature Analysis

Forecasting consumer returns has gained more research attention since 2016 (Figure III.4). The majority of the sample are conference publications, a couple of years ahead of the rise in journal publications. Compared to the publications on returns forecasting in the broader context of reverse logistics, which emerged in 2006 (Agrawal et al., 2015), the research on consumer returns moved into the spotlight about ten years later. This development is linked to a massive increase in e-commerce sales pre- and in-pandemic (Alfonso et al., 2021).

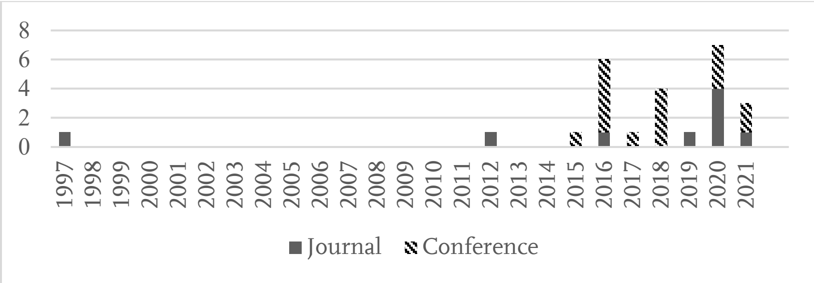

Figure III.4. Publication trend, by publication outlet.

Out of 9 journal publications in the final sample, only two are published in the same journal (Journal of Operations Management). Out of 16 conference papers, 6 are published at conferences of the Association for Information Systems. In total, 16 of the 25 papers found are published in Information Systems (IS) and related outlets. Others can be assigned to the Management Science / Operations Research discipline (3), Strategy & Management in a broader sense (4), Marketing (1), and Research Methods (1) (Figure III.5).

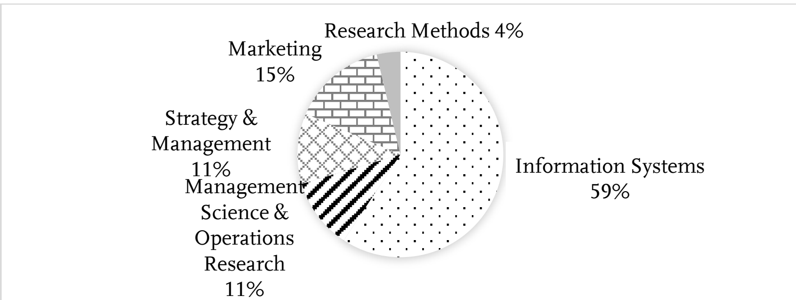

Figure III.5. Distribution of publication disciplines.

Regarding the researchers' geographical perspective, one paper was jointly published by authors from the US and China, 10 of 25 papers were authored from North America, followed by authors from Germany (7), India (3), China (1), and one paper each from Bangladesh, Singapore, and Sweden.

The most cited paper (200 external citations[2]) from Hess and Mayhew (1997) could be thought of as the root of this research field (Table III.4). However, only 10 out of 24 papers reference this work. Although Urbanke et al. (2015) received only 15 citations in total, within the sample, it is the second most cited paper (8 citations) and could eventually be classified as a research strand and origin of returns forecasting in the IS domain. Concerning the remaining papers, no unique strands of literature are recognizable based on citation analysis.

Author(s) and Year	External Citations	Internal Citations
Hess & Mayhew, 1997	200	10
Cui et al., 2020	39	3
Potdar & Rogers, 2012	33	3
Fu et al., 2016	22	0
J. Li et al., 2018	16	3
Shang et al., 2020	16	0
Urbanke et al., 2015	15	8
Imran & Amin, 2020	15	0
Ahmed et al., 2016	12	0
Ketzenberg et al., 2020	11	1

Table III.4. Top 10 papers within the sample (minimum of 10 external citations).

III.4.3 Methodology and Data Characterization

Regarding methodology, most of the papers start with a short narrative literature review regarding their respective focus. Not a single paper was based on interviews, surveys, questionnaires, or field experiments. 3 out of 25 papers formulated and tested conventional hypotheses. All of the publications use quantitative data for analysis and forecasting in a “case study” style, including numerical experiments based on real or simulated data.

[2] External citations according to Google Scholar, which is preferable for citation tracking over controlled databases (Halevi et al. 2017).

Paper	Business Type			Data Type		Industry							
	Brick & Mortar Store	Direct Marketer	E-Commerce	Simulated Data	Real Data	Fashion	(Consumer) Electronics	General Retailer / Wholesaler	Cosmetics	Building material, hardware store articles	Jewelry, personal items, home accessories	Car accessories	Not named
Hess & Mayhew, 1997		X		X	X	X							
Potdar & Rogers, 2012	X			X			X						
Urbanke et al., 2015			X		X	X							
Ahmed et al., 2016	X		X		X		X						
Heilig et al., 2016			X		X	X							
Ding et al., 2016			X		X								X
Fu et al., 2016					X				X				
Samorani et al., 2016	X		X		X		X						
Drechsler & Lasch, 2015				X							X		
Urbanke et al., 2017			X		X	X							
J. Li et al., 2018			X		X	X							
Zhu et al., 2018			X		X	X							
Asdecker & Karl, 2018			X		X	X							
Joshi et al., 2018			X		X	X							
X. Li et al., 2019			X		X				X				
Cui et al., 2020	X	X			X							X	
Shang et al., 2020	X		X		X	X	X						
Imran & Amin, 2020			X		X								X
Ketzenberg et al., 2020	X				X	X			X		X		
Hofmann et al., 2020			X		X					X			
John et al., 2020			X		X			X					
Rezaei et al., 2021	X		X	X	X		X						
Rajasekaran & Priyadarshini, 2021			X		X								X
Sweidan et al., 2020			X		X	X							
Fuchs & Lutz, 2021			X		X					X			
Σ	7	2	19	4	23	11	5	1	3	2	2	1	3

Table III.5. Dataset characteristics.

Table III.5 lists further details about the data used in the publications. 4 out of 25 papers rely on simulated data, and 23 out of 25 integrate actual data gained from a retailer. Two papers use both data types. 5 papers use more than one dataset (Ahmed et al., 2016; Cui et al., 2020; Rezaei et al., 2021; Samorani et al., 2016; Shang et al., 2020). The most frequently studied industry is fashion/apparel (10 papers), followed by five consumer electronics datasets. Two publications are based on data from a Taobao cosmetics retailer, and two datasets originate from general and wide assortment retailers. Two datasets incorporate building material and hardware store articles, and the detailed products are not named for three publications. Based on the previous studies, it is evident that consumer returns forecasting is most relevant for e-commerce, as 19 of the 25 publications refer to e-tailers. Nevertheless, 7 publications refer to brick-and-mortar retailing. Direct selling/marketing is represented in 2 data sets.

III.4.4 Predictors for Consumer Returns

There is an individual stream of research into factors that influence or help avoid consumer returns (e.g., Asdecker et al., 2017; De et al., 2013; Walsh & Möhring, 2017), which is not part of this review. Nevertheless, the forecasting literature gives insights into return drivers, as the input variables (features, predictors, exogenous variables) for forecasting models represent some of these factors. Table III.6 presents the most used predictors and tries to map these to the return driver categorization from Section 2.2 (Stöcker et al., 2021).
Although only a part of the publications interprets the predictors, some insights can be extracted. For *total return volume*, sales volume is the most critical predictor (Cui et al., 2020; Shang et al., 2020). Historical return volume trends can include behavioral aspects (e.g., impulse purchases) in a given timeframe (Cui et al., 2020; Shang et al., 2020). The product type significantly impacts the volume of returns (Cui et al., 2020), confirmed by widely varying return rates between different industries/sectors. Adding transaction-, customer-, or product-level predictors led to a surprisingly small forecasting accuracy gain (4 % reduction of RMSE, Shang et al., 2020). The latter input variables may be more critical in forecasting return decisions and propensities.

Paper	Product/Order Price / Discounts	Return Attributes (e.g., Reason Codes, Fit Importance)	Product Attributes (Category, Brand, Size ...)	Product Order History	Product Return History	Basket Composition	Sales Channel	Sales Volume	Customer Return History	Customer Order History	Other Customer Attributes	Employee Attributes	Order attributes (e.g., Payment)	Order/Return timing	Device / Browser Information	Network Information (other Customers)	Logistics attributes (e.g., delivery time)	Qualitative Information (Reviews, Tweets)	Automated Feature Selection / Generation
Hess & Mayhew, 1997	X	X																	
Potdar & Rogers, 2012	X	X	X																
Urbanke et al., 2015			X			X	X		X	X			X	X	X		X		X
Ahmed et al., 2016	X								X										X
Heilig et al., 2016	X		X			X					X								
Ding et al., 2016																		X	
Fu et al., 2016	X	X	X	X	X				X	X	X								
Samorani et al., 2016	X		X						X	X						X			X
Drechsler & Lasch, 2015					X			X						X					
Urbanke et al., 2017	X		X			X	X		X	X			X	X	X				X
J. Li et al., 2018			X			X			X	X				X		X			
Zhu et al., 2018			X						X	X						X			
Asdecker & Karl, 2018	X										X			X			X		
Joshi et al., 2018			X						X	X									
X. Li et al., 2019	X		X	X	X				X	X	X					X			
Cui et al., 2020			X	X	X	X	X	X						X					
Shang et al., 2020								X						X					
Imran & Amin, 2020						X							X	X	X				
Ketzenberg et al., 2020	X		X				X		X	X	X		X						
Hofmann et al., 2020				X	X	X													
John et al., 2020		X										X	X						
Rezaei et al., 2021	X								X	X	X								X
Rajasekaran & Priyadarshini, 2021				X	X									X				X	
Sweidan et al., 2020	X					X			X	X	X			X					
Fuchs & Lutz, 2021	X					X													
Return Reason Category*	I	I/B/F	I	I	B	I/B	-	-	B	B	B	F	-	B/F	I/B	I	F	I	-

* Legend: I = Information gap-related reasons,
B = (Consumer) Behavior-related reasons, F = Fulfillment-related reasons

Table III.6. Predictors used in return forecasting models.

Regarding *product attributes*, product or order price is one of the most common predictors, while some papers also include price discounts. In most models, price is hypothesized to increase returns (e.g., Asdecker & Karl, 2018; Hess & Mayhew, 1997). Promotional (discounted) orders also seem to result in more returns (Imran & Amin, 2020), which could be explained by the stimulation of impulse purchases.[3] Brand perception influences return decisions (positive brands, lower returns) (Samorani et al., 2016). The order and return history of products are also relevant for predicting future orders and returns (Hofmann et al., 2020). Fit importance as a product attribute does not significantly change return propensities (Hess & Mayhew, 1997).

Concerning *customer attributes*, gender seems essential, as female customers return significantly more items than men (Asdecker & Karl, 2018; Fu et al., 2016). Younger customers show a slightly lower propensity to return (Asdecker & Karl, 2018), but age played a more prominent role in predicting return fraud among employees than in customers (John et al., 2020 observed more fraud among younger employees). Customers with low credit scores returned more (Fu et al., 2016). The return history of a customer is possibly the most important predictor of future return behavior (Samorani et al., 2016). Some papers argue that consumer attributes, including purchase and return history (e.g., number and value of orders), are more relevant predictors than product or transaction profiles, reflecting more or less stable consumer preferences (X. Li et al., 2019).

Basket interactions are significant (Urbanke et al., 2017) in returns prediction. E.g., the larger the basket, the higher the return propensity will be (Asdecker & Karl, 2018). Selection orders (same product in different sizes or colors) increase the return propensity (J. Li et al., 2018). Logistics attributes like delivery times only show minor effects (Asdecker & Karl, 2018). Regarding the payment method, prepaid products are sent back less frequently than those with post-delivery payment options (Imran & Amin, 2020), confirming other research results (Asdecker et al., 2017).

One literature stream focuses on the automated *generation of features*, as different and large-scale data sources need to be integrated and prepared for forecasting algorithms. Thus, possible interrelationships are complex to find manually, and ML approaches might outperform human analysts (Rezaei et al., 2021). While some approaches generate a large number of

[3] Other literature also describes a counteracting effect of a reduced price due to lowered quality expectations or a higher perceived value of the “deal” itself (e.g., Sahoo et al. 2018).

features that are hard to make sense of (Ahmed et al., 2016), the approach of Urbanke et al. (2017) aims to maintain the interpretability of automatically generated input variables. Some unexpected but meaningful interrelations might be found by automatic feature generation, e.g., the price of the last returned orders (Samorani et al., 2016). Nevertheless, automatic feature generation might be computation-intensive; thus, a parallel integration of feature selection could be advantageous for large data sets (Rezaei et al., 2021).

A remarkable research path based on artificial intelligence is integrating qualitative information like product reviews as predictors, going beyond numerical feedback (Rajasekaran & Priyadarshini, 2021) or tweets. These data can be processed and made accessible for forecasting with ML-based sentiment analysis techniques (Ding et al., 2016).

III.4.5 Forecasting Techniques and Algorithms

To describe the techniques and algorithms employed, we sorted the papers by forecasting purpose as described in Section 2, then assigned them to different algorithms, either from time series forecasting, statistical techniques, or ML algorithms. Table III.7 lists all papers for which an assignment was possible, and the respective techniques used. If a comparison was possible, the best-performing algorithm is marked in this table.

	Time Series				Statistical							Machine Learning														
Paper	MA	SES	HWS	ARIMA	OLS	Logit	LDA	Hazard Model	Probit	Vector Autoregression	LASSO / Elastic Net	NN	RF	SVM	AdaBoost	CART	ERT	XGBoost	GradientBoost	JRIP	C4.5 / C5.0	BayesNetwork	Ensemble	k-NearestNeighbor	CatBoost	Others (e.g., RandomWalk)
Return decision																										
Hess & Mayhew, 1997					X	X		X																		
Urbanke et al., 2015						X	X					X	X	X	X	X	X									
Ahmed et al., 2016												X	X	X						X	X	X				
Heilig et al., 2016						X						X	X	X	X	X							X			
Samorani et al., 2016													X													
Urbanke et al., 2017							X						X	X	X	X	X									
J. Li et al., 2018																								X		X

Paper	Time Series				Statistical							Machine Learning														Others (e.g., RandomWalk)
	MA	SES	HWS	ARIMA	OLS	Logit	LDA	Hazard Model	Probit	Vector Autoregression	LASSO / Elastic Net	NN	RF	SVM	AdaBoost	CART	ERT	XGBoost	GradientBoost	JRIP	C4.5 / C5.0	BayesNetwork	Ensemble	k-NearestNeighbor	CatBoost	
Zhu et al., 2018																										X
Asdecker & Karl, 2018						X	X					X									X		X			
Joshi et al., 2018														X												
Imran & Amin, 2020												X						X	X							
Hofmann et al., 2020														X					X						X	
Rezaei et al., 2021											X		X	X												X
Sweidan et al., 2020													X													
Fuchs & Lutz, 2021																			X							
Return propensity																										
X. Li et al., 2019																										X
Rajasekaran & Priyadarshini, 2021					X								X						X							
Fu et al., 2016																										X
Return fraud / abuse																										
Ketzenberg et al., 2020						X						X	X	X												
John et al., 2020						X																				
Return volume (and timing)																										
Potdar & Rogers, 2012	X				X																					
Drechsler & Lasch, 2015	X	X	X																							
Cui et al., 2020					X						X		X						X							
Shang et al., 2020				X					X																	
Ding et al., 2016										X																
Σ	2	1	1	1	4	3	3	1	1	1	2	6	10	8	3	3	2	1	5	1	2	1	2	1	1	5

Table III.7. Employed algorithms for return forecasting (for comparisons: best-performing algorithm with shaded background, if named).

The approaches listed in Table III.7 are overlap-free, but some papers use more than one version of an approach, i.e., more than one algorithm from a category. E.g., TabNet is a DeepLearning version of neural networks (NN), and different variants of GradientBoosting are compared in one paper (CatBoost/LightGBM, not differentiated in the table below) (Imran & Amin, 2020).

The algorithm used most frequently (Figure III.6) is the Random Forest algorithm (RF, 10 papers), followed by Support Vector Machines (SVM, 8 papers), Neural Networks (NN, 6 papers), logistic regression (Logit, 6 papers), GradientBoosting (5 papers), Ordinary Least Squares regression (OLS, 4 papers), Adaptive Boosting (AdaBoost), Linear Discriminant Analysis (LDA), and CART (Classification and Regression Trees, 3 papers each).

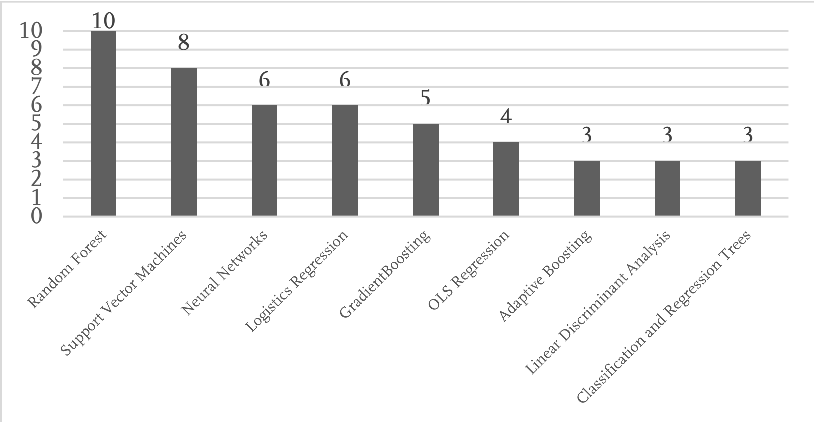

Figure III.6. Most frequently used algorithms (used in at least three papers).

The papers focusing on return volume use time series forecasts like (AutoRegressive) Moving Averages (MA), Single Exponential Smoothing (SES), and Holt-Winters Smoothing (HWS) more frequently than ML algorithms. Nevertheless, when considering a predict-aggregate approach as proposed by Shang et al. (2020), these ML techniques could be helpful in forecasting return decisions first and cumulating the propensity results for the volume prediction in the second step.

In forecasting binary return decisions, Random Forests (RF) (Ahmed et al., 2016; Heilig et al., 2016; Ketzenberg et al., 2020), Neural Networks (NN) (Imran & Amin, 2020; Ketzenberg et al., 2020), as well as Adaptive Boosting (AdaBoost) (Urbanke et al., 2015; Urbanke et al., 2017) showed high prediction performance. The performance of different algorithms varies depending on the data set, the implementation, and the parameterization used. For this reason, it is hardly possible to make a generally valid statement regarding performance levels. Combining several algorithms in ensembles (Asdecker & Karl, 2018; Heilig et al., 2016) seems advantageous, at least for retrospective analytical purposes, when the required computing resources are less relevant.

When evaluating different forecasting algorithms for return decisions, imbalanced classes (especially evident for low return shares in non-fashion datasets) seem to be handled differently depending on the algorithms. Class imbalances might distort comparison results in some publications. Random oversampling as a measure of data preparation can solve this problem (Hofmann et al., 2020).

High-performance algorithms are needed for real-time predictions, e.g., graph and random-walk-based (J. Li et al., 2018; Zhu et al., 2018). According to J. Li et al. (2018), the proposed algorithm "HyperGo" performs best for most performance metrics.

III.4.7 E-Commerce and Machine Learning Taxonomy Extension

In their literature review regarding the use of ML techniques in e-commerce, Micol Policarpo et al. (2021, p. 13) propose a taxonomy to visualize specific ML algorithms in the context of e-commerce platforms. This novel kind of taxonomy is based on direct acyclic graphs, i.e., all input variables need to be fulfilled to reach the target. The first level of the taxonomy represents different target goals for the use of ML in e-commerce. While returns forecasting ("product return prediction") is identified as an essential goal among others (purchase prediction, repurchase prediction, customer relationship management, discovering relationships between data, fraud detection, and recommendation systems), it was excluded from the taxonomy they developed, possibly because the review comprised only two relevant papers on this topic (Micol Policarpo et al., 2021). The review at hand proposes an extension of Micol Policarpo's taxonomy, renaming the goal to "consumer returns forecasting". This extension reflects and synthesizes the consumer returns forecasting studies reviewed. The *middle level* of the taxonomy represents properties and features that support this superordinate goal. On this level, our extension does not include return fraud detection, which we propose to be integrated into the existing category of "fraud detection", separated into transaction analysis and consumer analysis (Micol Policarpo et al., 2021). *Circles* represent the necessary data to execute the analysis, referring to categories introduced in (Micol Policarpo et al., 2021), with an additional "return history" category. The *bottom level* presents the algorithms described frequently, while some streamlining is required regarding the tools and approaches that seem the most common or most appropriate.

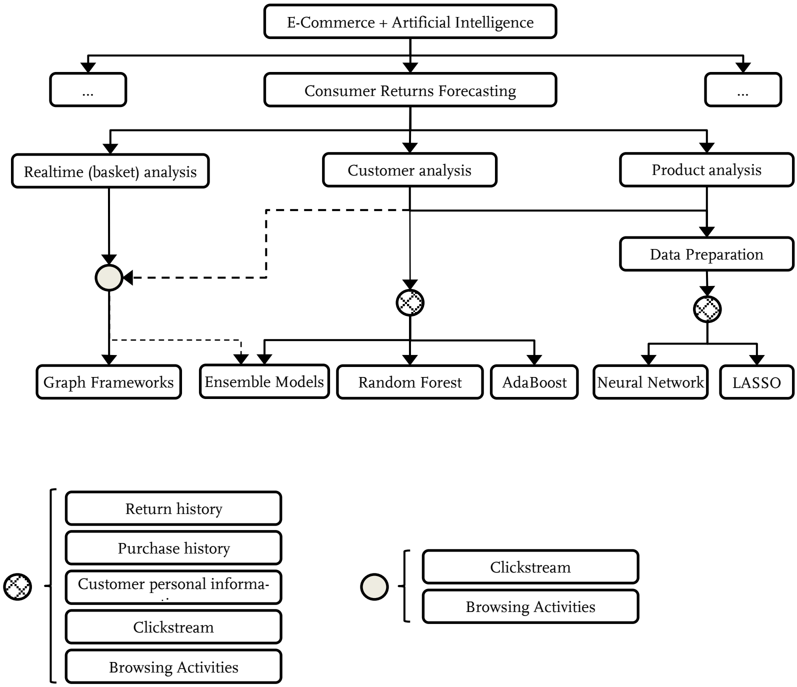

Figure III.7. Proposed consumer returns forecasting extension to the E-commerce and Machine Learning techniques taxonomy of Micol Policarpo et al. (2021, p. 13).

The schematic above (Figure III.7) is to be read as follows: In the context of *E-Commerce + Artificial Intelligence* (Layer 1), *Consumer Return Forecasting* (Layer 2) is an essential goal among six other goals. Layer 3 presents different purposes of analysis, which are the base for return forecasting. *Realtime Basket Analysis* is based on clickstream data and basket composition (browsing activities) to target interventions. Basket analysis benefits from customer and product information (dotted line). Graph-based approaches (J. Li et al., 2018; Zhu et al., 2018) are promising for real-time analysis due to their lower computing requirements, although cloud-based implementation of more complex algorithms or ensemble models might be feasible (Fuchs & Lutz, 2021; Heilig et al., 2016; Hofmann et al., 2020). *Customer Analysis* and *Product Analysis* (e.g., Potdar & Rogers, 2012) require adequate *Data Preparation* in the sense of input variable generation, extraction, and selection (Urbanke et al., 2015; Urbanke et al., 2017). For these purposes, data regarding return history (e.g., Hofmann et al., 2020; Ketzenberg et al., 2020), purchase history (e.g., Cui et al., 2020; Fu et al., 2016), customer personal information (e.g., Heilig et al., 2016; Ketzenberg et al., 2020), clickstream data, and browsing activities are required as input (shown by cross-hatched circles). For each purpose, one or more possible algorithms are shown.

Compared to predicting purchase intention, return predictions seem to require more levels of data. Nevertheless, even simple rule-based interventions can promise benefits, e.g., selection orders that inevitably lead to a return shipment can be easily recognized (Hofmann et al., 2020; Sweidan et al., 2020). Different ML techniques are helpful for data preparation and input variable (feature) extraction and generation when considering more complex interrelations. NeuralNet is one example of an automatic selection of relevant features (Urbanke et al., 2017). These approaches are not only able to enhance forecasting accuracy (Rezaei et al., 2021) but can also render the many possible variables interpretable about their content.

III.5 Discussion

The analysis of the papers above revealed that research in this discipline seems heterogeneous and partly fragmented, and clear-cut research strands are still hard to identify. Thus, the existing literature calls for further publications to render this research field more comprehensive. Below, research opportunities are derived and embedded in a *conceptual research framework* derived from the results of the existing literature, also integrating the extension of the E-Commerce and Machine Learning taxonomy (Figure III.7). A conceptual framework improves the understanding of a complex topic by naming and explaining key concepts and their relationships important to a specific field (Jabareen, 2009; Miles et al., 2020). Thus, this framework aims to organize problems and solutions discussed in the consumer returns forecasting literature and to embed and classify potential future research topics in the existing knowledge base (Ravitch & Riggan, 2017). The subsections following the framework outline some potential research avenues (P1 – P6) that have been touched on in the past but still leave considerable opportunities for further insights. These proposals should not be seen as comprehensive due to numerous other research opportunities in this field but rather as prioritization based on the current literature.

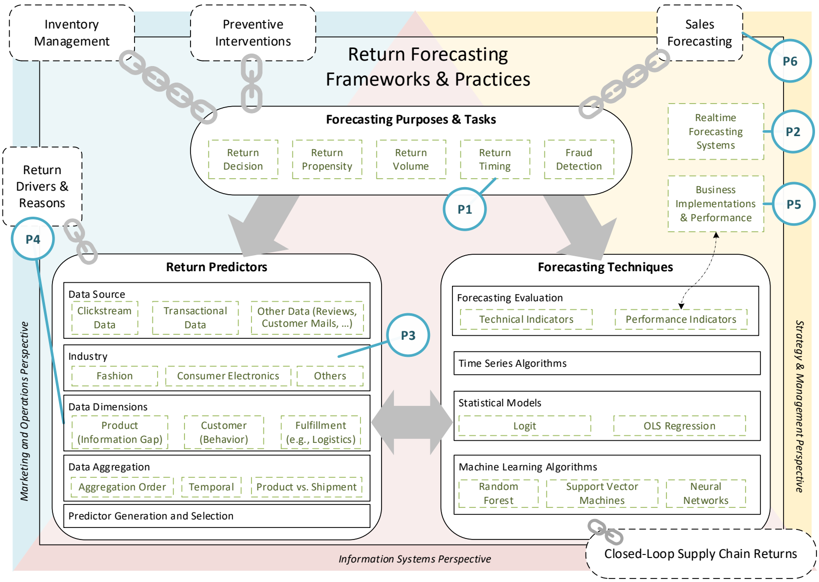

Figure III.8 Conceptual consumer returns forecasting framework.

The framework derived (Figure III.8) underlines the interdisciplinary nature of this research field, integrating different perspectives (information systems research, marketing and operations perspective, and strategy and management perspective). From a managerial point of view, the literature included in this review is biased towards the information systems perspective. Thus, in contrast to the framework developed by Cirqueira et al. (2020) for purchase prediction, we do not take a process perspective but instead emphasize the interdependencies and interactions between research topics and highlight the managerial need to take a strategical perspective similar to the framework developed by Winklhofer et al. (1996). Consequently, a meta-layer on forecasting frameworks and practices includes the mainly technical development frameworks in this review but also accentuates the need for further research regarding actual organizational forecasting practices (e.g., P2, P5, P6). Around this meta-layer, some related research strands are linked in order to embed the topic of returns forecasting in the research landscape. E.g., in general, forecasting purchases and returns could be linked (P6), also effecting inventory decisions.

The center of the framework consists of three dimensions, namely purposes and tasks, predictors, and techniques. Depending on the strategical purpose, tasks are derived that determine (1) the data (predictors) needed and (2) the usable techniques to execute the forecasting. Different forecasting techniques require an individual set of predictors, whereas the availability of specific data allows and determines the use of more or less sophisticated algorithms.

In the literature, some forecasting purposes were more pronounced (return decisions or propensities), while others have gained less attention (return timing, P1). Regarding the data necessary for accurate forecasting, the return predictors discussed often were hardly comparable, as they originated from different data sources, different industries, were related to different dimensions, or were aggregated in another way. Systematically linking forecasting predictors and research on return drivers and reasons could contribute significant insights (P4) that, from a marketing perspective, may support the development of effective preventive instruments. Furthermore, the literature mainly refers to the fashion or consumer electronics industry, leaving room to validate the findings in the context of other industries (P3).

When (automatically) selecting or creating predictors, the boundaries between predictors and prediction techniques are blurred as machine learning algorithms prepare the input data before executing a forecasting model. Regarding forecasting techniques, time series forecasting was seldom used in recent publications. Machine learning algorithms were the most popular subject of investigation, with random forests, support vector machines, and neural networks as the most popular implementations. Classical statistical models like logit models for return decisions or OLS regression gained less research attention. Literature on end-of-life return forecasting could complement the research on techniques and their accuracy. Most publications used technical indicators for assessing the accuracy of forecasting models, which is the information systems perspective. From a managerial position, evaluating (monetary) performance outcomes (e.g., Ketzenberg et al., 2020) of forecasting systems should be more relevant.

III.5.1 Research Proposal P1: Return Timing for Consumer Returns

Toktay et al. (2004) encouraged the integrated forecasting of the return rate and the return time lag. In line with this, Shang et al. (2020) criticize the missing focus on the timing of return forecasts. The reviewed literature confirms that forecasting return propensities and decisions are more prominent than timing and volume forecasts. While the knowledge of when a return is expected is vital in managing end-of-life returns that occur over the years, for retail consumer returns, return periods are mostly 14-30 days. Thus, the variability of return timing seems limited compared to end-of-life returns in this context, which makes this forecasting purpose less critical. Nevertheless, some retailers offer up to 100 days of free returns (e.g., Zalando). Consequently, more studies about the importance of return timing forecasts in the e-commerce context from a business and planning perspective and their interdependence with return processing or warehousing issues could shed light on this topic and complement the current literature (Shang et al., 2020; Toktay et al., 2004).

III.5.2 Research Proposal P2: Realtime Forecasting Systems

Another research gap became apparent regarding the real-time use of forecasting systems and the associated activities and interventions, building on the initial research and the frameworks already published (e.g., Heilig et al., 2016; Urbanke et al., 2015). The generic framework developed by Fuchs and Lutz (2021) could serve as a launching pad for this stream of research.

- The paper from Ketzenberg et al. (2020) could act as a stimulus and inspiration for a similar approach, not only focusing on return abuse as already examined but on return forecasting in general, the possible associated interventions for various consumer groups, and the resulting consequences for the retailer's profit. Even the methodology of customer classification could be helpful for many retailers in targeting interventions.
- Before real-time return forecasting is implemented, associated preventive return management instruments need to be designed and evaluated. Many of these measures are discussed (e.g., Urbanke et al., 2015; Walsh et al., 2014), but an overview of which preventive measures (for some examples, see Walsh & Möhring, 2017) are effective in general (1) and how forecasting accuracy interdepends with their usefulness (2) is still missing, to substantially link the topics of forecasting and interventions. No answers could be found to the call by Urbanke et al. (2015) for field experiments to investigate such a link.
- Thanks to cloud and parallelization technologies and the associated scalability of computing power (Bekkerman et al., 2011), algorithm runtimes are becoming less relevant. However, especially for real-time use, it should be evaluated which algorithms and underlying datasets exhibit an appropriate relationship between the targeted forecasting accuracy, the expected benefit, and the required computing power.
- Recommendations concerning the algorithms and techniques can be derived (Urbanke et al., 2015, p. 15), and a generic implementation framework was developed (Fuchs & Lutz, 2021). However, from a business perspective, no contributions could be found regarding the actual implementation of real-time forecasting systems, the interventions involved, and their impact on consumer behavior or profit (also see proposal P5). In addition, the implementations of such systems need to be analyzed concerning the cost-effectiveness of the required investments.

III.5.3 Research Proposal P3: Cross-Industry and Multiple Dataset Studies

Many publications rely on a single data set from a specific industry or retailer. Only a few compare several retailers (e.g., Cui et al., 2020). Studies including and comparing different countries are missing, which is especially interesting since legal regulations for returns vary. For example, in contrast to the U.S., citizens within the EU are granted a 14-day right of withdrawal for distance selling purchases.[4] Although in most developed countries, liberal and broadly comparable returns policies are standard in practice due to competitive pressure, the generalizability of the results is frequently limited. One remedy for this problem is to use multiple data sets from different retailers (e.g., electronics vs. jewelry, Shang et al., 2020). Admittedly, it is challenging to simultaneously collaborate with several retailers and to combine different data sets, due to reasons of preserving corporate privacy and synchronizing various data sources. Nevertheless, research needs to draw conclusions from single data points, as well as logically replicate or falsify those results by integrating more data points to find patterns of similarities and differences, either within or cross-study (Hamermesh, 2007). Therefore, we suggest that future studies acquire industry-related datasets from several retailers at once or replicate existing studies, which aligns with the aim and scope of Management Review Quarterly (Block & Kuckertz, 2018). Cross-industry or cross-country manuscripts, which go beyond the mere assertion of an industry-agnostic approach (Hofmann et al., 2020) and jointly investigate data from several sectors, would promise an additional gain in knowledge and could be less challenging from a privacy perspective.

III.5.4 Research Proposal P4: Extended Study of Relevant Predictors in Forecasting Applications

Although not the main focus of this review, predictors of consumer returns are especially interesting for marketing and e-commerce research,

[4] It should be noted that the relevance of the forecasting topic depends on the maturity of the e-commerce sector. In most developing countries, B2C e-commerce is comparatively young and returns are not yet a common phenomenon, which is why research on return forecasts is relatively insignificant for these countries.

for example, regarding preventive measures for avoiding returns. In the past, many consumer return papers highlighted single aspects or a limited selection of return drivers or preventive measures employed but rarely attempted to model return behavior as comprehensively as possible. However, the latter is the very objective of returns forecasting, which is why the findings on influencing factors in articles with a forecasting focus tend to be more holistic, although not sufficiently complete (Hachimi et al., 2018). Some return reasons named in the literature (e.g., Stöcker et al., 2021) have not yet been included in forecasting approaches, and vice versa, only a part of the influencing factors investigated could be mapped to a return reason categorization. The reason categories assigned (Section 4.4, Table III.6) still contain some uncertainty. For example, a customer's product return history may reflect the general returning behavior of a customer to some extent, while it can not be ruled out that repeated logistical problems caused the returns. Product attributes may reflect information gaps that consumers can only assess after physically inspecting the product, whereas product price – a frequently cited and influential product attribute – is only related to information gaps when considering the price-performance ratio (Stöcker et al., 2021). Technical information about the web browser or device used by the customer is difficult to categorize, as it may reflect behavioral (impulse-driven mobile shopping) as well as informational (small display with few visible information) aspects. The payment method chosen by a customer, for example, could not be linked to one of the reason categories.

This reasoning should serve as a basis for linking forecasting predictors and return reasons more closely in the future. For example, the respective relative weighting of return drivers is more likely to be obtained considering as many factors involved as possible, minimizing the unexplained variation. From the reviewed literature, we extracted 18 different return predictor categories. For instance, seven papers (Cui et al., 2020; Fu et al., 2016; Ketzenberg et al., 2020; J. Li et al., 2018; X. Li et al., 2019; Urbanke et al., 2015; Urbanke et al., 2017) integrated more than five predictor categories. But even though some papers integrate more than 5,000 features for automated feature selection (Ketzenberg et al., 2020), there are still combinations of input variable categories that have not been investigated and, more importantly, interpreted yet. Therefore, we call for more comprehensive research on return predictors and their interpretation, including associated preventive return measures, in the context of return forecasting.

III.5.5 Research Proposal P5: Descriptive Case Studies and Business Implementations

This review identified a lack of publications regarding the actual benefit and the diffusion of consumer returns forecasting systems in different scopes and industries, building on the papers presenting return forecasting frameworks. In 2013, less than half of German retailers analyzed the likelihood of returns (Pur et al., 2013). Most of those who did were using naïve approaches that might be outperformed by the models presented in this review. Still, we do not know the status quo regarding the degree of adoption and implementation of forecasting systems for consumer returns in e-commerce firms (e.g., see Mentzer & Kahn, 1995 for sales forecasting systems), country-specific and internationally.

Furthermore, the impact of return forecasting practices on company performance should be examined not only based on modeling, but on retrospective data (e.g., see Zotteri & Kalchschmidt, 2007 for a similar study on demand forecasting practices in manufacturing). A possible hypothesis to examine might be that accuracy measures like RMSE or precision/recall and subsequently even the choice of the most accurate machine learning algorithm (e.g., see Asdecker & Karl, 2018) are less relevant from a business perspective: (1) No algorithm clearly outperforms all other algorithms, and (2) the correlation between technical indicators and business value is unstable (Leitch & Tanner, 1991). Methodologically, implementations of consumer returns forecasting in e-commerce should thus be surveyed and analyzed with multivariate statistical methods to examine critical factors and circumstances of return forecasting systems – similar to publications on reverse logistics performance (Agrawal & Singh, 2020).

III.5.6 Research Proposal P6: Holistic Forward and Backward Forecasting Framework for E-Tailers

Some publications present frameworks for forecasting returns (Fuchs & Lutz, 2021). Nevertheless, in the past, forecasting in retail and especially e-commerce commonly focused more on demand (Micol Policarpo et al., 2021) than returns. Current approaches for demand forecasting try to predict individual purchase intentions based on click-stream data, online session attributes, and customer history (e.g., Esmeli et al., 2021). Our systematic approach could not identify any paper that connects and integrates both directions in e-commerce forecasting, neither conceptual (frameworks) nor with a quantitative or case-study-like approach. Nevertheless, first implementations of return predictions in inventory management are presented (e.g., Goedhart et al., 2023). Subsequently, similar to Goltsos et al. (2019), we call for research addressing both demand and return uncertainties by providing a holistic forecasting framework in the context of e-commerce.

III.6 Conclusion

To date, no systematic literature review has undertaken an in-depth exploration of the topic of forecasting consumer returns in the e-commerce context. Previous reviews have primarily focused on product returns forecasting within the broader context of reverse logistics or closed-loop supply chain management (Agrawal et al., 2015; Ambilkar et al., 2021; Hachimi et al., 2018). Regrettably, the interdisciplinary nature of this subject has often been overlooked, also neglecting the inclusion of results from information systems research.

The review first aims to provide an overview of the existing literature (Kraus et al., 2022) on forecasting consumer returns. The findings confirm that this once novel topic has significantly evolved in recent years. Consequently, this review is timely in examining current gaps and establishing a robust foundation for future research, which forms a second goal of systematic reviews (Kraus et al., 2022). The current body of work encompasses various aspects from different domains, including marketing, operations management/research, and information systems research, highlighting the interdisciplinary nature of e-commerce analytics and research. As a result, future studies can find suitable publication outlets in domain-specific as well as methodologically oriented journals and conferences.

Scientifically, the algorithms and predictors investigated in previous research serve as a foundational reference for subsequent publications and informed decisions regarding research design, ensuring that specific predictors and techniques are not overlooked. Researchers can utilize this review and the research framework developed as a structuring guide, e.g., regarding relevant publications on already examined algorithms or predictors.

Managerially, the extended taxonomy for machine learning in e-commerce (Micol Policarpo et al., 2021) can serve as a guideline for implementing forecasting systems for consumer returns. This review classifies possible prediction purposes, allowing businesses to apply them based on their respective challenges. Exploring the most frequently used predictors reveals the data that must be collected for the respective purposes. This review also offers valuable insights into data (pre-)processing and highlights popular algorithms. Furthermore, frameworks are outlined that support the design and implementation phase of such forecasting systems, supporting analytical purposes or enabling direct interventions during the online shopping process flow. As an exemplary and promising application, return policies could be personalized (Abbey et al., 2018) by identifying opportunistic or fraudulent basket compositions or high-returning customers, thereby reducing unwanted returns (Lantz & Hjort, 2013).

Finally, a limitation of this review is the exclusion of forecasting algorithms for end-of-use returns, which could potentially be applicable to forecasting shorter-term retail consumer returns. However, the closed-loop supply chain and reverse logistics literature has been systematically excluded. Hence, future reviews could synthesize previous reviews on reverse logistics forecasting with the more detailed findings presented in this paper.

Appendix III.1 Journal publications

Hess and Mayhew (1997) describe a forecasting approach, taking the example of a direct marketer for apparel with a lenient consumer return policy (free returns anytime). The analysis can plausibly be applied to a general retailer, although return time windows are somewhat different. A regression approach and a hazard model are compared. The regression approach itself is split into an OLS estimation of return timing (with poor fit) and a logit model of return propensities, which is in turn used for the split function of the box-cox-hazard approach for estimating the probability of a return over time. The accuracy was measured by fit statistics regarding the absolute deviation from the actual cumulative return proportion, with the split-hazard model outperforming the regression model. Besides price, the importance of fit of the respective product is used as a predictor.

Potdar and Rogers (2012) propose a method using reason codes combined with consumer behavior data for forecasting returns volume in the consumer electronics industry, aiming at the retailer stage as well as the preceding supply chain stages. The subject of their study is an offline retailer, which allows generalization for e-tailers due to a similar return policy (14 days free returns with no questions asked). In a multi-step approach, the authors are using essential statistical methods (moving averages, correlations, and linear regression), but use sophisticated domain and product knowledge like product features or price in relation to past return numbers, aiming to rank different competing products regarding their quality, and to predict the volume of returns for a given product for each given period of time.

Fu et al. (2016) derive a framework for the forecasting of product- and consumer-specific return propensities, i.e., the return propensity for individual purchases. Their study is directed at online shopping and is evaluated using the data from an online cosmetic retailer selling via Taobao.com. The predictors are categorized into inconsistencies in the buying and in the shipping phase of a transaction. A latent factor model is introduced for return propensities capturing differences between expectations and performance. This model is extended by product (e.g., warranty) and customer information (e.g., gender, credit score). The model is based on conditional probabilities, and an iterative expectation-maximization approach derives its parameters. MAE and RMSE, precision/recall, and AUC metrics assess the forecast accuracy. As benchmark models, two matrix factorization models and two memory-based models (historical consumer or product return rates) are compared, while the proposed model outperforms the references. Furthermore, this model allows identifying various return reasons, e.g., return abuse and fraud.

Building on the work of Fu et al. (Fu et al., 2016), Li et al. (2019) investigate underlying reasons for consumer returns, taking the example and data of an online cosmetic retailer via Taobao.com. They examine the customers' return propensity for product types, aiming at detecting abnormal returns suspecting abuse. Different from purchase decisions, they find customer profile data to be more important predictors for return decisions than product information or transaction details. The authors detect "selfish" or "fraud" consumers based on this rationale. For estimating return propensities for a given consumer and product, they calculate the return behavior depending on the return decision of similar consumers ("trust network") and the amount of trust in these other consumers. MAE and precision-recall-measures are used to assess the prediction of different random walk models. The employed trust-based random walk model outperforms the other models on most indicators, building the basis for anomaly detection of consumers to cluster them into groups (honest/selfish/fraud) and individually address the return issues of these groups.

Although the paper from Cui et al. (2020) aims at product return forecasts from the perspective of the manufacturer, their case can be generalized for classic e-tailers, as the manufacturer is responsible for the return handling in their scenario – a task often performed by the retailer. They used a comprehensive data set from an automotive accessories manufacturer aiming to forecast return volume for sales channels and different products. The observed return rates lower than 1 % are uncommonly low, and therefore the results must be interpreted with caution. First, a hierarchical OLS regression step-by-step incorporates up to 40 predictors regarding sales, time, product type, sales channel, and product details, including return history. The full model shows a significantly increased performance measured by a more than 50 % decrease of MSE, which was used as the primary performance measure. Interestingly, relatively small differences in model quality (R^2) led to overproportional changes in the MSE. Using a machine-learning approach for predictor selection (“LASSO”), another MSE reduction of about 10 % was achieved. Data Mining approaches (random forest, gradient boosting) could not outperform the LASSO approach. Forecasting performance was strongly dependent on the variation of the data. The two best predictors for return volume were past sales volume and lagged return statistics. The authors were wondering about the importance of lagged return information, failing to acknowledge that this predictor includes the consumer reaction to detailed product information, which has not been a significant predictor.

Ketzenberg et al. (2020) segment customers and target detecting the small number of abusive returners, as these are unprofitable for the retailer and generate significant losses over a long time. In general, high-returning customers are usually more profitable. The data used for this study is from a department store retailer with various product groups in the assortment. Predictors are transactional data and customer attributes. For classification, different algorithms like logit, Support Vector Machines (SVM), Random Forests (RF), Neural Networks (NN) are used in combination with different shrinkage methods like LASSO, ridge regression, and elastic net. Random Forests and especially Neural Networks outperform the other algorithms, assessed by sensitivity, precision, and AUC. In conclusion, a low rate of false positives could assure retailers of using abuse detection systems.

Shang et al. (Shang et al., 2020) developed a predict-aggregate (P-A) model adaptable both for retailers and manufacturers for forecasting return volume in a continuous timeframe, in contrast to commonly used aggregate-predict (A-P) models. Instead of aggregating data first (i.e., sales volume and returns volume), they first aggregate product-specific return probabilities and then aggregate the purchases by addition of the individual probabilities. As predictors, they only use timestamps and lagged return information. They tune and assess their models on two datasets from an offline electronics and an online jewelry retailer. ARIMA and lagged return models known from end-of-life forecasting (de Brito et al., 2005) are used as benchmarks, using RMSE as an assessment criterion. The authors show that even a basic version of their approach outperforms the benchmark models in almost all observed cases by up to 19 %, though using only lagged returns and timestamps as input. Different extensions, e.g., including more predictor variables, can easily be integrated and are shown to further improve the forecasting performance.

John et al. (2020) try to predict the rare event of return fraud from customer representatives that make use of exactly knowing the e-commerce company's return policy framework and buying and returning items fraudulently. Therefore, predictors range from transaction details to customer service agent attributes. A penalized likelihood logit model was chosen by the authors and was evaluated by precision and recall, focussing on maximizing recall and minimizing false negatives. The most important predictors were communication type and reason for interaction.

The paper by Rezaei et al. (2021) introduces a new algorithm to automatically select attributes from high-dimensional databases for forecasting purposes. As a demonstration sample, they use simulated data as well as the publicly available ISMS Durable Goods dataset (Ni et al., 2012) for consumer electronics. The results are assessed by AUC, precision, recall, and f1-score. They compare different configurations. For the simulated data, LASSO as shrinkage method generally works best, outperforming RF and BaggedTrees. For real-world data, based on a forecast with a logit model, they show that the proposed selection algorithm performs similar or better compared to LASSO, SVM, and RF, while the complexity of the chosen variables is lower.

Appendix III.2 Conference Publications

Urbanke et al. (2015) describe a decision support system to better direct return-reducing interventions at e-commerce purchases with highly likely returns. They compare different approaches for extracting input variables for return propensity forecasting. They use a large dataset from a fashion e-tailer, aiming to reduce the input variables regarding consumer profile, product profile, and basket information from over 5,000 binary variables to 10 numeric variables by different algorithms (e.g., principal component analysis, non-negative matrix factorization, etc.). The results are then used to predict return propensities with a wide variety of state-of-the-art algorithms (AdaBoost, CART, ERT, GB, LDA, LR, RF, SVM), thus also revealing both feature selection and prediction performance. The proposed Ma-

halanobis feature extraction algorithm used as input for AdaBoost outperforms all other combinations presented, while interestingly, a logit model with all original inputs delivers relatively precise forecasts.
Building on some parts of this study, the paper of Urbanke et al. (2017) presents a return decision forecasting approach and aims at two targets, (1) high predictive accuracy and (2) interpretability of the model. Based on real-world data of a fashion and sports e-tailer, they first hand-craft 18 input variables and then use NN to extract more features and compare this approach to other feature extraction algorithms based on different forecasting algorithms. For assessment, they measure correlations between out-of-sample-predictions and class labels and AUC. The best performing classifier was AdaBoost, while the contribution of NN-based feature extraction shows interpretability as well as superior predictive performance.
Ahmed et al. (2016) focus on the automatic aggregation and integration of different data sources to generate input variables (features). They use return forecasting just as an exemplary classification problem for their data preparation approach, using various ML algorithms, e.g., RF, NN, DT-based algorithms, to detect returned purchases of an electronics retailer. Based on AUC measure, the results of their GARP-approach are superior to not using aggregations while generating an extensive amount of features with no pruning approach. In general, SVM and RF work best in combination with the proposed GARP approach. The data is based on the publicly available ISMS durable goods data sets (Ni et al., 2012).
A similar group of authors published another paper (Samorani et al. (2016)), again using the aforementioned ISMS dataset as an example for data preparation and automatic attribute generation. Besides forecasting performance, in this paper, they want to generate knowledge about important return predictors; e.g., a higher price is associated with more returns, but only as long price levels are below a 1,500$ threshold. AUC is used to assess different levels of data integration, confirming that overfitting might happen when too many attributes are used.
Heilig et al. (2016) describe a Forecasting Support System (FSS) to predict return decisions in a real environment. First, they compare different forecasting approaches for data from a fashion e-tailer, assessed by AUC and accuracy metrics. The ensemble selection approach outperforms all other classifiers, with RF being the closest competitor. Computational times

grow exponentially when using more data. Based on these results, they secondly describe a cloud framework for implementing such ensemble models for live use in a real shop environment.

Ding et al. (2016) present an approach to predict the daily return rate of an e-commerce company based on sentiment analysis of tweets regarding this company in the categories of news, experience, products, and service. Therefore, they use sophisticated text mining technologies, while the forecasting approach of an econometric vector autoregression is more or less common. The emotion of posts regarding different variables (news, product, service) impacts the returns rate negatively, while the emotion of purchasing experience impacts it positively, showing that the prediction accuracy enhances through classifying social network posts.

Drechsler and Lasch (2015) aim at forecasting the volume of fraudulent returns in e-commerce over several periods of time. They present different approaches multiplying the sales volume and the relative return rate, the first referring to Potdar and Rogers (2012), estimating the rate of misused returns directly based on time-lag-specific return rates. In a second approach referring to Toktay et al. (2000), they estimate the overall returns rate and multiply it by the time-specific ratio of fraudulent returns. The return rates were forecasted by moving averages and exponential smoothing techniques. Assessment criteria for performance comparison based on simulated data were MAE, MAPE, and TIC, showing the first approach to be superior, but both methods are not sufficiently robust. Therefore, the authors include further time-specific information (like promotions or special events, which could foster fraudulent returns) in a model using a Holt-Winters approach, showing superior performance. All of the models are highly dependent on low fluctuation in return rates, showing a shortcoming of these more or less naive forecasting techniques.

Asdecker and Karl (2018) compare the performance of different algorithms for forecasting binary return decisions: logit, linear discriminant analysis, neuronal networks, and a decision-tree-based algorithm (C5.0). Their analysis is based on the data of a fashion e-tailer, including price, consumer information, and shipment information (number of articles in shipment, delivery time). For the assessment of different algorithms, they use the total absolut error (TAE) and relative error. An ensemble learning approach performs best and similar to the C5.0 algorithm. Though, differences in performance are relatively small, while only about 68 % of return decisions are forecasted correctly.

Li et al. (2018) propose a hypergraph representation of historical purchase and return information combined with a random-walk-based local graph cut algorithm to forecast return decisions on order (basket) level as well as on product level. By this, they aim to detect the underlying return causes. They use data from two omnichannel fashion e-tailers from the US and Europe to assess the performance of their approach, using precision/recall/$F_{0.5}$/AUC metrics while arguing that precision is the most important indicator for targeted interventions. Three similarity-based approaches (e.g., a k-Nearest Neighbor model) are used as reference. The proposed approach performs best regarding AUC, precision, and $F_{0.5}$ metrics.

Zhu et al. (2018) developed a weighted hybrid graph algorithm representing historical customer behavior and customer/product similarity, combined with a random-walk-based algorithm for predicting customer/product combinations that will be returned. They report an experiment based on data from a European fashion e-tailer suffering from return rates as high as 50 %. For assessment, they use precision, recall, and $F_{0.5}$ metrics. Their approach is superior to two reference competitors (similarity-based and a bipartite graph algorithm). As predictors, they use product similarities and historical return information, while their approach can be enriched with detailed customer attributes.

Joshi et al. (2018) model the return decisions based on the data of an Indian e-commerce company, especially dealing with returns for apparel due to fit issues. In a two-step approach, they first model return probabilities using concepts from network science based on a customer's historical purchase and return decisions, and secondly use a SVM implementation with return probabilities as a single input to classify for the return decision. Assessed by F_1/precision/recall scores, their approach is superior to a reference random-walk baseline model.

Imran and Amin (2020) compare different forecasting algorithms (XGBoost, CatBoost, LightGBM, TabNet) for return classification based on the data of a general e-commerce retailer from Bangladesh. As input variables, only order attributes, including payment method and order medium, are used. For evaluation, they use metrics like true negative rate, false-positive rate, false-negative rate, true positive rate, AUC, F2-score, precision, and accuracy. In the end, they chose TPR, AUC, and F2-score, claiming that misclassifying high return probability objects were the first thing to avoid. According to these metrics, TabNet as a deep learning algorithm outperforms the other models. The most important predictors were payment method, order location, and promotional orders.

As returns are most prominent in fashion e-commerce, most of the forecasting papers take this industry as an example, as forecasting models are more precise when returns are more frequent. Hofmann et al. (2020) develop a more generalized order-based return decision forecasting approach, appropriate for different industries and suitable also for low return rates. For their analysis, they use a dataset from a german technical wholesaler with a return rate as low as 5 %. Input variables were just basket composition and return information. For assessment, they used precision and recall metrics. RF did not perform superior to a statistical baseline approach, nor with oversampling as data preparation, to deal with the group imbalance. The DART algorithm makes use of the group imbalance correction by random oversampling. In general, gradient boosting performs best with imbalanced groups, also without oversampling, but forecasting quality is lower than with more specialized forecasting approaches as described for fashion. Furthermore, results were more accurate on basket level than on single-item level.

Fuchs and Lutz (2021) use Design Science Research (DSR) principles to design a meta-model for the real-time prediction of returns. The goal is to influence consumer decisions by triggering a feedback system based on the basket composition and its return probability. For forecasting, which is not the primary focus of their paper, they build upon a gradient boosting model taken from existing research (Hofmann et al., 2020) and describe possible implementations into an ERP system regarding asynchronous communication requirements and possible architecture.

The paper by Sweidan et al. (2020) evaluates the forecasting performance of a random forest model for a shipment-based return decision, using real-world data of a fashion e-tailer. For their model, they use customer (e.g., lagged return rate) and order information as inputs. They find that predictions with high confidence are very precise (i.e., low false-positive rate). Thus, interventions can be targeted at such orders already when the items are in the consumers' basket without risk of a misdirected intervention. For assessment, accuracy, AUC, precision, recall and specificity are used. Regarding the predictors, they note that selection orders (a product in different sizes) are the best predictor for order-based returns.

Rajasekaran and Priyadarshini (2021) develop a metaheuristic for forecasting the product-based return probabilities. In the first step, they determine return probabilities based on product feedback, time, and product attributes regarding manufacturer return statistics. Secondly, they compare different algorithms (OLS, RF, Gradient Boosting) by MAE, MSE, and RMSE metrics. Interestingly, linear regression performs best in all metrics, but no explanation and a misinterpretation regarding the best algorithm are given.

References

Abbey, J. D., Ketzenberg, M. E., & Metters, R. (2018). A More Profitable Approach to Product Returns. *MIT Sloan Management Review, 60*(1), 71–74.

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Agrawal, S., & Singh, R. K. (2020). Forecasting product returns and reverse logistics performance: structural equation modelling. *Management of Environmental Quality: An International Journal, 31*(5), 1223–1237.

Agrawal, S., Singh, R. K., & Murtaza, Q. (2015). A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling, 97*, 76–92.

Ahmed, F., Samorani, M., Bellinger, C., & Zaiane, O. R. (2016). Advantage of Integration in Big Data: Feature Generation in Multi-Relational Databases for Imbalanced Learning. In *Proceedings of the 4th IEEE International Conference on Big Data,* Washington DC, United States of America.

Ahsan, K., & Rahman, S. (2016). An investigation into critical service determinants of customer to business (C2B) type product returns in retail firms. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 606–633.

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194.

Alfonso, V., Boar, C., Frost, J., Gambacorta, L., & Liu, J. (2021). E-commerce in the pandemic and beyond. *BIS Bulletin, 36*(9).

Ambilkar, P., Dohale, V., Gunasekaran, A., & Bilolikar, V. (2021). Product returns management: A comprehensive review and future research agenda. *International Journal of Production Research*, 1–25.

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Asdecker, B., & Karl, D. (2018). Big data analytics in returns management - Are complex techniques necessary to forecast consumer returns properly? In *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics,* Valencia, Spain.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q., & Seaman, B. (2019). Sales Demand Forecast in E-commerce Using a Long Short-Term Memory Neural Network Methodology. In T. Gedeon, K. W. Wong, & M. Lee (Chairs), *Neural Information Processing: Proceedings of the 26th International Conference on Neural Information Processing,* Sydney, Australia.

Barbosa, M. W., La Vicente, A. d. C., Ladeira, M. B., & Oliveira, M. P. V. de (2018). Managing supply chain resources with Big Data Analytics: a systematic review. *International Journal of Logistics Research and Applications, 21*(3), 177–200.

Bekkerman, R., Bilenko, M., & Langford, J. (2011). Scaling up machine learning. In *Proceedings of the 17th ACM SIGKDD International Conference Tutorials,* San Diego, United States of America.

Bernon, M., Cullen, J., & Gorst, J. (2016). Online retail returns management. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 584–605.

Block, J., & Kuckertz, A. (2018). Seven principles of effective replication studies: strengthening the evidence base of management research. *Management Review Quarterly, 68*(4), 355–359.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Wadsworth & Brooks/Cole Advanced Books & Software.

Cirqueira, D., Hofer, M., Nedbal, D., Helfert, M., & Bezbradica, M. (2020). Customer Purchase Behavior Prediction in E-commerce: A Conceptual Framework and Research Agenda. In M. Ceci, C. Loglisci, G. Manco, E. Masciari, & Z. Raś (Eds.), *Lecture Notes in Artificial Intelligence: Vol. 11948. New Frontiers in Mining Complex Patterns* (2020th ed., Vol. 11948, pp. 119–136). Springer International Publishing; Imprint: Springer.

Clottey, T., & Benton, W. C. (2014). Determining core acquisition quantities when products have long return lags. *IIE Transactions, 46*(9), 880–893.

Cook, S. C., & Yurchisin, J. (2017). Fast fashion environments: consumer's heaven or retailer's nightmare? *International Journal of Retail & Distribution Management, 45*(2), 143–157.

Cui, H., Rajagopalan, S., & Ward, A. R. (2020). Predicting product return volume using machine learning methods. *European Journal of Operational Research, 281*(3), 612–627.

Dalecke, S., & Karlsen, R. (2020). Designing Dynamic and Personalized Nudges. In R. Chbeir, Y. Manolopoulos, R. Akerkar, & J. Mizera-Pietraszko (Eds.), *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics* (pp. 139–148). ACM.

De, P., Hu, Y., & Rahman, M. S. (2013). Product-Oriented Web Technologies and Product Returns: An Exploratory Study. *Information Systems Research, 24*(4), 998–1010.

de Brito, M. P., Dekker, R., & Flapper, S. D. P. (2005). Reverse Logistics: A Review of Case Studies. In A. Klose & B. Fleischmann (Eds.), *Lecture Notes in Economics and Mathematical Systems. Distribution Logistics* (Vol. 544, pp. 243–281). Springer Berlin Heidelberg.

Denyer, D., & Tranfield, D. (2009). Producing a Systematic Review. In D. A. Buchanan & A. Bryman (Eds.), *The Sage handbook of organizational research methods* (pp. 671–689). Sage.

Difrancesco, R. M., Huchzermeier, A., & Schröder, D. (2018). Optimizing the return window for online fashion retailers with closed-loop refurbishment. *Omega, 78*, 205–221.

Diggins, M. A., Chen, C., & Chen, J. (2016). A Review: Customer Returns in Fashion Retailing. In T.-M. Choi (Ed.), *Springer Series in Fashion Business. Analytical Modeling Research in Fashion Business* (pp. 31–48). Springer Singapore.

Ding, Y., Xu, H., & Tan, B. C. Y. (2016). Predicting Product Return Rate with "Tweets". In *Proceedings of the 20th Pacific Asia Conference on Information Systems,* Chiayi, Taiwan.

Drechsler, S., & Lasch, R. (2015). Forecasting Misused E-Commerce Consumer Returns. In *Logistics Management: Proceedings of the 9th Conference "Logistikmanagement",* Braunschweig, Germany.

Duong, Q. H., Zhou, L., Meng, M., van Nguyen, T., Ieromonachou, P., & Nguyen, D. T. (2022). Understanding product returns: A systematic literature review using machine learning and bibliometric analysis. *International Journal of Production Economics, 243,* 108340.

Esmeli, R., Bader-El-Den, M., & Abdullahi, H. (2021). Towards early purchase intention prediction in online session based retailing systems. *Electronic Markets, 31*(3), 697–715.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research, 15*(1), 3133–3181.

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics, 7*(2), 179–188.

Frei, R., Jack, L., & Brown, S. (2020). Product returns: A growing problem for business, society and environment. *International Journal of Operations & Production Management, 40*(10), 1613–1621.

Frei, R., Jack, L., & Krzyzaniak, S.-A. (2022). Mapping Product Returns Processes in Multichannel Retailing: Challenges and Opportunities. *Sustainability, 14*(3), 1382.

Fu, Y., Liu, G., Papadimitriou, S., Xiong, H., Li, X., & Chen, G. (2016). Fused latent models for assessing product return propensity in online commerce. *Decision Support Systems, 91,* 77–88.

Fuchs, K., & Lutz, O. (2021). A stitch in time saves nine – A meta-model for real-time prediction of product returns in ERP systems. In *Proceedings of the 29th European Conference on Information Systems,* Marrakech, Morocco.

Ge, D., Pan, Y., Shen, Z.-J., Di Wu, Yuan, R., & Zhang, C. (2019). Retail supply chain management: a review of theories and practices. *Journal of Data, Information and Management, 1,* 45–64.

Goedhart, J., Haijema, R., & Akkerman, R. (2023). Modelling the influence of returns for an omni-channel retailer. *European Journal of Operational Research, 306*(3), 1248–1263.

Goltsos, T. E., Ponte, B., Wang, S. X., Liu, Y., Naim, M. M., & Syntetos, A. A. (2019). The boomerang returns? Accounting for the impact of uncertainties on the dynamics of remanufacturing systems. *International Journal of Production Research, 57*(23), 7361–7394.

Govindan, K., & Bouzon, M. (2018). From a literature review to a multi-perspective framework for reverse logistics barriers and drivers. *Journal of Cleaner Production, 187,* 318–337.

Hachimi, H. E. L., Oubrich, M., & Souissi, O. (2018). The optimization of reverse logistics activities: A literature review and future directions. In *Proceedings of the 5th IEEE International Conference on Technology Management, Operations and Decisions,* Marrakech, Morocco.

Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. *Journal of Informetrics, 11*(3), 823–834.

Hamermesh, D. S. (2007). Viewpoint: Replication in economics. *Canadian Journal of Economics/Revue Canadienne D'économique, 40*(3), 715–733.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (Second edition). Springer.

Heilig, L., Hofer, J., Lessmann, S., & Voß, S. (2016). Data-Driven Product Returns Prediction: A Cloud-Based Ensemble Selection Approach. In *Proceedings of the 24th European Conference on Information Systems,* Istanbul, Turkey.

Hess, J. D., & Mayhew, G. E. (1997). Modeling merchandise returns in direct marketing. *Journal of Direct Marketing, 11*(2), 20–35.

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75.

Hofmann, A., Gwinner, F., Fuchs, K., & Winkelmann, A. (2020). An Industry-Agnostic Approach for the Prediction of Return Shipments. In *Proceedings of the 26th Americas Conference on Information Systems,* Salt Lake City, United States of America.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

Imran, A. A., & Amin, M. N. (2020). Predicting the Return of Orders in the E-Tail Industry Accompanying with Model Interpretation. *Procedia Computer Science, 176,* 1170–1179.

Jabareen, Y. (2009). Building a Conceptual Framework: Philosophy, Definitions, and Procedure. *International Journal of Qualitative Methods, 8*(4), 49–62.

John, S., Shah, B. J., & Kartha, P. (2020). Refund fraud analytics for an online retail purchases. *Journal of Business Analytics, 3*(1), 56–66.

Joshi, T., Mukherjee, A., & Ippadi, G. (2018). One Size Does Not Fit All: Predicting Product Returns in E-Commerce Platforms. In *Proceedings of the 10th IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining,* Barcelona, Spain.

Kaiser, D. (2018). *Individualized Choices and Digital Nudging: Multiple Studies in Digital Retail Channels.* Karlsruher Institut für Technologie (KIT). https://doi.org/10.5445/IR/1000088341

Karl, D., & Asdecker, B. (2021). How Does the Covid-19 Pandemic Affect Consumer Returns: An Exploratory Study. In *Proceedings of the 50th European Marketing Academy Conference,* Madrid, Spain.

Karl, D., Asdecker, B., & Feddersen-Arden, C. (2022). The Impact of Displaying Quantity Scarcity and Relative Discounts on Sales and Consumer Returns in Flash Sale E-Commerce. In *Proceedings of the 55th Hawaii International Conference on System Sciences,* Maui, United States of America.

Ketzenberg, M. E., Abbey, J. D., Heim, G. R., & Kumar, S [Subodha] (2020). Assessing customer return behaviors through data analytics. *Journal of Operations Management*, *66*(6), 622–645.

Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S [Satish], Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E., Palacios-Marqués, D., Schiavone, F., Ferraris, A., Fernandes, C., & Ferreira, J. J. (2022). Literature reviews as independent studies: guidelines for academic practice. *Review of Managerial Science, 16*(8), 2577–2595.

Lantz, B., & Hjort, K. (2013). Real e-customer behavioural responses to free delivery and free returns. *Electronic Commerce Research, 13*(2), 183–198.

Leitch, G., & Tanner, J. E. (1991). Economic Forecast Evaluation: Profits Versus the Conventional Error Measures. *The American Economic Review, 81*(3), 580–590.

Li, J., He, J., & Zhu, Y. (2018). E-tail Product Return Prediction via Hypergraph-based Local Graph Cut. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,* London, United Kingdom.

Li, X., Zhuang, Y., Fu, Y., & He, X. (2019). A trust-aware random walk model for return propensity estimation and consumer anomaly scoring in online shopping. *Science China Information Sciences*, *62*(5).

Melacini, M., Perotti, S., Rasini, M., & Tappia, E. (2018). E-fulfilment and distribution in omni-channel retailing: a systematic literature review. *International Journal of Physical Distribution & Logistics Management, 48*(4), 391–414.

Mentzer, J. T., & Kahn, K. B. (1995). Forecasting technique familiarity, satisfaction, usage, and application. *Journal of Forecasting, 14*(5), 465–476.

Micol Policarpo, L., da Silveira, D. E., da Rosa Righi, R., Antunes Stoffel, R., da Costa, C. A., Victória Barbosa, J. L., Scorsatto, R., & Arcot, T. (2021). Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review. *Computer Science Review, 41*, 100414.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (Fourth edition). Sage.

National Retail Federation / Appriss Retail. (2023, May 23). *Consumer Returns in the Retail Industry 2022*. Retrieved May 23, 2023, from https://nrf.com/research/2022-consumer-returns-retail-industry

Ni, J., Neslin, S. A., & Sun, B. (2012). Database Submission The ISMS Durable Goods Data Sets. *Marketing Science, 31*(6), 1008–1013.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews, 10*, 89.

Pandya, R., & Pandya, J. (2015). C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning. *International Journal of Computer Applications, 117*(16), 18–21.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., . . . Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting, 38*(3), 705–871.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine, 6*(3), 21–45.

Potdar, A., & Rogers, J. (2012). Reason-code based model to forecast product returns. *Foresight, 14*(2), 105–120.

Pur, S., Stahl, E., Wittmann, M., Wittmann, G., & Weinfurtner, S. (2013). *Retourenmanagement im Online-Handel - Das Beste daraus machen: Daten, Fakten und Status quo*. Ibi Research.

Rajasekaran, V., & Priyadarshini, R. (2021). An E-Commerce Prototype for Predicting the Product Return Phenomenon Using Optimization and Regression Techniques. In M. Singh, V. Tyagi, P. K. Gupta, J. Flusser, T. Ören, & V. R. Sonawane (Chairs), *Advances in Computing and Data Sciences: Proceedings of the 5th International Conference on Advances in Computing and Data Sciences,* Nashik, India.

Ravitch, S. M., & Riggan, M. (2017). *Reason & Rigor: How conceptual frameworks guide research* (Second edition). Sage.

Ren, S., Chan, H.-L., & Siqin, T. (2020). Demand forecasting in retail operations for fashionable products: methods, practices, and real case study. *Annals of Operations Research, 291*(1-2), 761–777.

Rezaei, M., Cribben, I., & Samorani, M. (2021). A clustering-based feature selection method for automatically generated relational attributes. *Annals of Operations Research, 303*(1-2), 233–263.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Röllecke, F. J., Huchzermeier, A., & Schröder, D. (2018). Returning Customers: The Hidden Strategic Opportunity of Returns Management. *California Management Review, 60*(2), 176–203.

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The Impact of Online Product Reviews on Product Returns. *Information Systems Research, 29*(3), 723–738.

Samorani, M., Ahmed, F., & Zaiane, O. R. (2016). Automatic Generation of Relational Attributes: An Application to Product Returns. In *Proceedings of the 4th IEEE International Conference on Big Data,* Washington DC, United States of America.

Shaharudin, M. R., Zailani, S., & Tan, K. C. (2015). Barriers to product returns and recovery management in a developing country: investigation using multiple methods. *Journal of Cleaner Production, 96,* 220–232.

Shang, G., McKie, E. C., Ferguson, M. E., & Galbreth, M. R. (2020). Using transactions data to improve consumer returns forecasting. *Journal of Operations Management, 66*(3), 326–348.

Srivastava, S. K., & Srivastava, R. K. (2006). Managing product returns for reverse logistics. *International Journal of Physical Distribution & Logistics Management, 36*(7), 524–546.

Stock, J. R., & Mulki, J. P. (2009). Product Returns Processing: An Examination of Practices of Manufacturers, Wholesalers/Distributors, and Retailers. *Journal of Business Logistics, 30*(1), 33–62.

Stöcker, B., Baier, D., & Brand, B. M. (2021). New insights in online fashion retail returns from a customers' perspective and their dynamics. *Journal of Business Economics, 91*(8), 1149–1187.

Sweidan, D., Johansson, U., & Gidenstam, A. (2020). Predicting returns in men's fashion. In *Proceedings of the 14th International Fuzzy Logic and Intelligent Technologies in Nuclear Science Conference,* Cologne, Germany.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness.* Penguin.

Tibben-Lembke, R. S., & Rogers, D. S. (2002). Differences between forward and reverse logistics in a retail environment. *Supply Chain Management: An International Journal, 7*(5), 271–282.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological), 58*(1), 267–288.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

Toktay, L. B., Wein, L. M., & Zenios, S. A. (2000). Inventory Management of Remanufacturable Products. *Management Science, 46*(11), 1412–1426.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management, 14*(3), 207–222.

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 20*(1), 57–59.

Urbanke, P., Kranz, J., & Kolbe, L. (2015). Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction. In *Proceedings of the 14th International Conference on Computer and Information Science,* Fort Worth, United States of America.

Urbanke, P., Uhlig, A., & Kranz, J. (2017). A Customized and Interpretable Deep Neural Network for High-Dimensional Business Data - Evidence from an E-Commerce Application. In *Proceedings of the 38th International Conference on Information Systems,* Seoul, South Korea.

Vakulenko, Y., Shams, P., Hellstrom, D., & Hjort, K. (2019). Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research, 101,* 461–468.

vom Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *Proceedings of the 17th European Conference on Information Systems,* Verona, Italy.

Walsh, G., & Möhring, M. (2017). Effectiveness of product return-prevention instruments: Empirical evidence. *Electronic Markets, 27*(4), 341–350.

Walsh, G., Möhring, M., Koot, C., & Schaarschmidt, M. (2014). Preventive Product Returns Management Systems - A Review and Model. In *Proceedings of the 22nd European Conference on Information Systems,* Tel Aviv, Israel.

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii–xxiii.

Winklhofer, H., Diamantopoulos, A., & Witt, S. F. (1996). Forecasting practice: A review of the empirical literature and an agenda for future research. *International Journal of Forecasting, 12*(2), 193–221.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining,* Manchester, United Kingdom.

Zahn, M. von, Bauer, K., Mihale-Wilson, C., Jagow, J., Speicher, M., & Hinz, O. *The Smart Green Nudge: Reducing Product Returns through Enriched Digital Footprints & Causal Machine Learning.* https://doi.org/10.2139/ssrn.4262656

Zhao, X., Hu, S., & Meng, X. (2020). Who should pay for return freight in the online retailing? Retailers or consumers. *Electronic Commerce Research, 20*(2), 427–452.

Zhu, Y., Li, J., He, J., Quanz, B. L., & Deshpande, A. (2018). A Local Algorithm for Product Return Prediction in E-Commerce. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence,* Stockholm, Sweden.

Zotteri, G., & Kalchschmidt, M. (2007). Forecasting practices: Empirical evidence and a framework for research. *International Journal of Production Economics, 108*(1-2), 84–99.

IV Evaluating Advanced Product Return Forecasting Algorithms – a (Meta-)Review Integrating Consumer Returns Research

Author: David Karl

Publication Outlet: International Journal of Business Forecasting and Marketing Intelligence

DOI: 10.1504/IJBFMI.2023.10057211

published as:

Karl, D. (2024): Evaluating Advanced Product Return Forecasting Algorithms – a (Meta-)Review Integrating Consumer Returns Research. In: *International Journal of Business Forecasting and Marketing Intelligence, 9*(2), 213–241. DOI: 10.1504/IJBFMI.2024.137648

Abstract

While e-commerce has recently experienced substantial growth rates, retailers face increasing consumer returns. Machine learning techniques opened up opportunities for improved consumer returns forecasting. In the past, returns forecasting was analyzed predominantly from a broader reverse logistics and closed-loop supply chain perspective. This paper extends this view by reviewing the state of research on current algorithms for forecasting returns in e-commerce in particular and integrating it into the body of knowledge regarding forecasting product returns extracted from previous reviews. Methodologically, four reviews were synthesized first. Subsequently, a systematic literature review was conducted, analyzing 28 additional publications related to consumer returns and enriching the literature on product returns. Thus, this comprehensive review is the first to analyze current forecasting issues while integrating the e-commerce perspective and emphasizing relevant developments regarding advanced algorithms and metrics for their assessment in returns forecasting.

Keywords: Forecasting, product returns, consumer returns, literature review, performance indicators, machine learning, evaluation

IV.1 Introduction

Online shopping experienced substantial growth rates of up to 15 % during the last few years and is growing way faster than traditional brick-and-mortar retailing (Ratchford et al., 2022). Since customers cannot physically assess the quality and fit of products online, returns are an inevitable part of the business model (Hong & Pavlou, 2014). Merchants of fashion products experience return rates of 50 % and are particularly affected (Asdecker et al., 2017). Merchandise worth $218 Billion was returned to online retailers in the U.S., equalling 20.8 % of online sales (National Retail Federation / Appriss Retail, 2021). Consumer returns are associated with operational challenges, such as unavoidable processing costs and planning uncertainties regarding logistics capacities, inventory management, procurement decisions, and marketing activities (Chopra, 2019; Stock & Mulki, 2009). Nevertheless, returns do not just happen but can be actively influenced and managed, e.g., by marketing instruments (El Kihal & Shehu, 2022). Upstream in the supply chain, end-of-use or end-of-life returns are linked to similar issues while trying to recover as much value as possible by recycling or remanufacturing (Guide, 2000).
Returns belong to the related term *reverse logistics*. Reverse product flows, also called *product returns*, can be categorized into (re-)manufacturing returns, distribution returns, and *consumer returns* (Shaharudin et al., 2015; Tibben-Lembke & Rogers, 2002). Consumer returns (see Table IV.1) describe returns from a consumer back to a retailer. Over the past years, online consumer returns have accounted for the largest share of returns and thus are a highly relevant field of research (Abdulla et al., 2019; Frei et al., 2020). In contrast to returns in closed-loop supply chain management (CLSC), consumer returns are usually sent or given back unused or undamaged shortly after purchase (usually within 30 days, while some retailers allow free returns up to 365 days after purchase). Generally, they suffer no quality-related defects, should be reimbursed to the consumer, and are intended to be resold "as new" (de Brito et al., 2003; Melacini et al., 2018). In summary, consumer returns describe returns from consumers to retailers primarily due to information asymmetries regarding the product fit to their expectations, but not due to end-of-life issues, durable goods buybacks, or product failure (Shulman et al., 2011).

Dimension	Consumer returns	CLSC returns in general ("product returns")
Time of return	Shortly after purchase (14-365 days)	Indefinite, including end-of-life returns
Return source	Retail customer ("consumer")	Any supply chain actor
Return destination	Retailer	Any supply chain actor except the consumer
Purpose	Resale as new	Resale, Repair, Refurbish, Remanufacture, Recycle, Dispose
Financial compensation	Full refund	Full/partial/no refund
Return reason	Mainly pre-purchase lack of information	Various, including defective products

Table IV.1. Delimitation between different kinds of returns.

In addition, digitized retail offers numerous opportunities for data collection, as digital customer accounts make customer analytics more efficient (Akter & Wamba, 2016). Nevertheless, the problem presented equally applies to brick-and-mortar retail (Santoro et al., 2019), which can benefit from data analytics for forecasting purposes (Hess & Mayhew, 1997). While demand forecasting is widely recognized as critical supply chain management activity (Ge et al., 2019), only a few publications exist combining the "returns management" process with big data analytics (Barbosa et al., 2018). However, retailers managing returns deal with massive amounts of data and use available technological advances (Robertson et al., 2020). Compared to demand forecasting, returns build the "supply" side of the process (Frei et al., 2022). Thus, forecasting is more complex due to customers ultimately and only hardly predictable deciding about their intention whether to keep a product or not. Additionally, interdependencies with sales forecasts and promotional activities increase complexity (Govindan & Bouzon, 2018; Tibben-Lembke & Rogers, 2002), justifying a deeper investigation of this domain.
In all supply chain stages, the accuracy of return forecasting positively impacts the economic, environmental, and social performance of reverse logistics activities (Agrawal & Singh, 2020). In general, "there is a very limited research conducted on the applications of BDA [Big Data Analytics] in [...] reverse logistics" (Seyedan & Mafakheri, 2020, p. 17). Similar, but even more than forecasting in reverse logistics in general (Petropoulos et

al., 2022), "the topic of forecasting consumer returns has received little attention in the academic literature" (Shang et al., 2020, p. 342) in the past. For example, a recent review on retail forecasting completely ignored the prediction of product returns (Fildes et al., 2022). However, e-commerce platforms increasingly use machine learning (ML) and artificial intelligence to generate beneficial knowledge for customers and sellers (Micol Policarpo et al., 2021), e.g., by better addressing customer needs and optimizing retail operations.

IV.1.1 Previous literature reviews

About twenty years ago, de Brito et al. (2003) called for analyzing reverse flows, forming a basis for forecasting returns. Since then, online commerce has skyrocketed, and research into reverse logistics in general and consumer returns in particular has multiplied. Thus, we first examined existing literature reviews and meta-studies regarding the subject of this study, referring to Google Scholar and Web of Science (see Table IV.2), to identify apparent research gaps.

Review	Sources	Scope
Duong et al., 2022	167	Operations management, retailer and (re)manufacturer issues, customer's psychology, experience, and perception
Ambilkar et al., 2021	518	Reverse logistics: product recovery, forecasting product returns, consumer behaviour, return policy, uncertainty, technology
Micol Policarpo et al., 2021	70	E-Commerce and machine learning: goals, techniques, a taxonomy of machine learning in e-commerce
Abdulla et al., 2019	100	Consumer returns: consumer behaviour, return policy, planning, and execution, return management
Goltsos et al., 2019	25	Closed-loop supply chain dynamics, uncertainties for remanufacturing systems: forecasting, collection, inventory, and production control
Hachimi et al., 2018	45	Reverse logistics forecasting techniques
Agrawal et al., 2015	242	Adoption and implementation, forecasting product returns, outsourcing, RL networks, disposition decisions

Table IV.2. Relevant reviews discussing consumer/product returns or returns forecasting.

Duong et al. (2022) analyze 167 papers. Regarding forecasts of product returns, they find three papers (Clottey & Benton, 2014; Cui et al., 2020; Shang et al., 2020) and called for research regarding predicting customer returns behaviour in the pre-purchase stage. Ambilkar et al. (2021) examine 518 papers and consider product returns in a holistic reverse logistic approach. Although not analyzed in detail, 13 papers are categorized under "forecasting product returns" in the digital appendix. None of these papers focuses on consumer returns.

Micol Policarpo et al. (2021) review 70 papers on machine learning in e-commerce regarding common goals of e-commerce studies (e.g., (re-) purchase prediction, product return prediction) and techniques to support these goals. A taxonomy of ML in e-commerce includes most of the identified goals but ignores the aspect of return predictions, apparently due to a lack of sufficient studies on this topic.

In contrast to other reviews, Abdulla et al. (2019) focus on consumer returns. Although their review discusses consumer behaviour, planning, and execution of returns, no sources focussing on forecasting issues are included.

Goltsos et al. (2019) investigate the impact of different uncertainties on the dynamics of remanufacturing systems. While their systematic review focuses on closed-loop supply chain dynamics, in a non-systematic review, they elaborate on one of the most relevant "pillars" of uncertainty, namely forecasting in closed-loop supply chains.

Hachimi et al. (2018) consider reverse logistics forecasting issues and categorize 45 relevant works regarding various forecasting approaches: time series and ML, operations research (OR) methods, and simulation programs. As a research gap, missing performance indicators for assessment are identified. Agrawal et al. (Agrawal et al., 2015) point out research gaps in the reverse logistics domain. Reviewing 242 papers, they identify "forecasting product returns" as a crucial opportunity for future research. 21 papers regarding "forecasting models for product returns" mainly focus on CLSC, reuse, remanufacturing, and recycling. No literature on performance indicators for assessing return forecast model accuracy was found.

IV.1.2 Research objectives

Generally, "returns forecasting has not received sufficient attention in the academic literature" (Petropoulos et al., 2022, p. 67). More specifically, existing reviews (e.g., Ambilkar et al., 2021) include a structured but rather superficial overview of the literature regarding end-of-use and end-of-life forecasting (J. Li et al., 2018; Shang et al., 2020) while almost omitting the topic of consumer returns. Some empirical or conceptual papers cover this field with brief literature sections (e.g., Hofmann et al., 2020), but these short narrative reviews are neither transparent nor replicable (Tranfield et al., 2003). Unsystematic reviews often induce selection bias, while systematic reviews attempt to draw a holistic, differentiated, and more detailed picture, incorporating the complete available literature when possible (Uman, 2011). Additionally, only Hachimi et al. focus on ML algorithms for forecasting returns, an application partly ignored by Micol Policarpo et al. Several reviews (Agrawal et al., 2015; Hachimi et al., 2018) name the lack of established assessment indicators for return forecasts. Accordingly, this (meta-)review aims to close gaps in existing reviews by providing a comprehensive overview of available knowledge regarding forecasting approaches for product returns management, including consumer returns, and to elaborate on areas for further scholarly investigation. Against this background, we address the following research question:

- *Which (advanced) algorithms are used in the literature on forecasting product returns in general and consumer returns in particular, and how are these approaches evaluated?*

To answer this question, we derive knowledge from existing reviews concerning product returns forecasting (Agrawal et al., 2015; Ambilkar et al., 2021; Hachimi et al., 2018) and conduct a systematic literature review on consumer returns forecasting, in particular, to integrate this field into the existing body of literature. This study contributes to multiple disciplines, including the literature around decision support and machine learning for reverse logistics as well as the e-commerce literature. We examine algorithms currently researched, performance indicators used to assess and

compare these approaches, and extend the previous research by integrating the consumer returns perspective.
The remainder of this paper is organized as follows: Section 2 presents backgrounds and concepts essential for assessing the relevant literature. The methodology used for the meta-review on product return forecasting and the systematic review on consumer return forecasting is presented in Section 3. Section 4 contains the data analysis and discusses the results by outlining opportunities for future research. A brief conclusion forms Section 5.

IV.2 Background and relevant concepts

In general, accurate forecasting is crucial in managing the supply chain. E.g., reliable demand forecasts are essential for inventory planning, pricing, promotions and affect a retailer's success. Forecasting returns affects similar business aspects but also resorts to similar technical procedures. Depending on the purpose, the outcome variable of a return forecast is scaled differently. When forecasting whether a single product will be returned, the dependent variable is either binary in nature or a continuous propensity value between 0 and 1. When forecasting the number or the timing of returns, outcomes are continuous variables. The timing of product returns is crucial in forecasting end-of-life, end-of-use, or remanufacturing returns that usually occur years after the initial purchase. Forecasting return volumes can be multi-faceted, ranging from forecasting the total return volume a retailer has to process within its logistics department through forecasting product-specific return numbers up to forecasting costly return shares, e.g., return fraud volume.

IV.2.1 Fundamental forecasting techniques

Because returns depend on fluctuating sales, sales forecasts are often naively combined with lagged return rates. Nevertheless, various other forecasting techniques can be applied. For the most part, we follow the categorization suggested by Hachimi et al. (2018) while emphasizing the machine learning category (Carbonneau et al., 2008).
In a narrower sense, **Operations Research (OR) methods** aim at optimizing a given decision problem utilizing mathematical modelling of the

structure of a real-world situation (e.g., Meisel & Mattfeld, 2010). For a more detailed classification of OR methods, we draw on Brandenburg et al. (2014), differentiating between mathematical programming methods, analytical models, and (meta-)heuristic methods (Hachimi et al., 2018). **Simulation programs** sometimes enable what-if-analyses to compare operational alternatives before making real-world decisions (Chang & Makatsoris, 2001). **Time-series techniques** (e.g., exponential smoothing) predict the future development of an outcome on its past numbers and time as the only predictor. Most time-series models can be generalized as autoregressive moving averages (ARIMA) models. Similarly, time-series regression models use univariate linear regression with time as an exogenous variable.

Traditional regression models comprise other independent exogenous variables. Multivariate models are essential statistical tools (Hastie et al., 2017). Many variants are derived from this logic (e.g., generalized linear models), and various extensions are built upon this base. The logistic regression and the linear discriminant analysis extend the linear regression, where the dependent variable is binary (Hastie et al., 2017). These approaches can model return decisions as well as non-binary return propensities.

Expanding on these traditional approaches, an almost unlimited computing power facilitated the establishment of **advanced algorithms** (Carbonneau et al., 2008) based on more complex statistical methods. Some popular algorithms are shortly outlined: Artificial Neural Networks (NN) are the most popular ML algorithm in last years' e-commerce research (Micol Policarpo et al., 2021), and deep learning extensions like Long Short-Term Memory (Bandara et al., 2019) are gaining attention. Many Decision Tree (DT) algorithms have been presented for classification, e.g., the C5.0 algorithm (Pandya & Pandya, 2015). Ensemble learning methods use several algorithms to improve predictive performance (Polikar, 2006; Tsoumakas et al., 2008). Similar to the ensemble principle, boosting and bagging techniques are used in algorithms like AdaBoost or Random Forest (RF) (tree-based) to expand the input data artificially, aiming at more generalizable forecasting models that are less prone to overfitting, a problem of algorithms not able to generalize rules to other datasets apart from those they are trained with (Hastie et al., 2017). Support Vector Machines

(SVM) proved effective for classification problems in e-commerce (Micol Policarpo et al., 2021). For a comparison of state-of-the-art ML classification techniques, we refer to Fernández-Delgado et al. (2014). Besides classification (return decision and propensities), almost all supervised ML approaches are suitable for different forecasting purposes, e.g., continuous forecasts (return quantities and timing).

IV.2.2 Forecasting evaluation

Forecasts must be assessed regarding their quality to select the suitable forecasting algorithm, its settings, and the necessary input variables for the individual forecasting purpose. Often, after a prediction model has been fitted using a subset of data (training data), a comparison of predicted values with known values for another subset of the data set (test or check data) is performed (Efendigil et al., 2009). For assessing the prediction quality, performance indicators represent the error between the forecasted and the actual value. E.g. for classification, the accuracy of an algorithm applied to a given dataset can be measured by true and false positives, prediction accuracy, sensitivity, specificity, or graphically by receiver operating characteristics (Baldi et al., 2000). When predicting continuous variables, accuracy measures widespread in literature are mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), or the mean absolute percentage error (MAPE) (Galdi & Tagliaferri, 2018).

IV.3 Methodology

The methodology of this paper is split into a (meta-)review of product returns forecasting, which is complemented by a systematic review of consumer returns forecasting. For both parts, we draw on and combine established suggestions for reviewing the literature (Denyer & Tranfield, 2009; vom Brocke et al., 2009; Webster & Watson, 2002). The process is structured in several steps: (1) question formulation, (2) identifying research gaps and defining the scope, (3) conceptualizing the topic; (4) locating studies; (5) study selection and evaluation; (6) analysis and synthesis; and (7) reporting and using the results for defining a research agenda.

The first two steps are described in Section 1, while Section 2 describes the research topic's conceptualization (step 3).

IV.3.1 Meta-review on forecasting product returns

For the (meta-)review, in step 4 (locating studies) and 5a (study selection), we refer to the literature already located and selected by previous reviews concerning forecasting product returns. As already described in the introduction, four reviews included the topic of forecasting product returns at least as a relevant section. We are confident that the sample derived from these reviews (see Table IV.3) can be considered saturated and relatively up-to-date, as it covers works from 1986 until 2020, found via different databases and search strategies.

Review	Years	Relevant sources	Sum of relevant sources
Ambilkar et al. (2021)	1986-2020	13	78
Goltsos et al. (2019)	1994-2017	8	
Hachimi et al. (2018)	1986-2018	43	
Agrawal et al. (2015)	1986-2015	21	

Table IV.3. Relevant reviews for product returns (n=4 reviews).

One paper (Efendigil et al., 2009) has been miscategorized in recent reviews (Agrawal et al., 2015; Hachimi et al., 2018) while exclusively covering demand forecasting. After removing this paper and after removing duplicates, the sample consists of 56 works. In sum, 15 papers were added to the most extensive review by Hachimi et al. (2018). The complete list is provided in Table IV.4.

Paper	Agrawal et al., 2015	Hachimi et al., 2018	Goltsos et al., 2019	Ambilkar et al., 2021
Goh & Varaprasad, 1986	✓	✓	✓	
Kelle & Silver, 1989	✓		✓	
Hess & Mayhew, 1997		✓		
Murayama et al., 2000		✓		
Toktay et al., 2000	✓	✓	✓	
Golany et al., 2001		✓		
Kiesmüller & van der Laan, 2001			✓	
Richter & Weber, 2001		✓		

Paper	Agrawal et al., 2015	Hachimi et al., 2018	Goltsos et al., 2019	Ambilkar et al., 2021
Linton et al., 2002		✓		
Marx-Gómez et al., 2002	✓	✓		
Terzi & Cavalieri, 2004		✓		
Toktay et al., 2004		✓		
Kleyner & Sandborn, 2005		✓		
Y. Li et al., 2006		✓		
Peralta & Fontanos, 2006	✓			
Srivastava & Srivastava, 2006	✓	✓		
Teunter et al., 2006		✓		
Hanafi et al., 2007	✓	✓	✓	
Serrato et al., 2007		✓		
Wen-Jie & Zhi-Geng, 2007	✓	✓		
Carrasco-Gallego & Ponce-Cueto, 2009		✓	✓	
Xiaofeng & Tijun, 2009	✓			
Chen & He, 2010	✓			
Yu et al., 2010	✓			
Barker & Zabinsky, 2011		✓		
Masmoudi, 2011		✓		
Schulz, 2011		✓		
Clottey et al., 2012		✓	✓	
Plewa & Jodejko-Pietruczuk, 2012		✓		
Potdar & Rogers, 2012	✓			✓
Shih et al., 2012	✓	✓		
Benedito & Corominas, 2013	✓	✓		
Krapp et al., 2013a	✓	✓		✓
Krapp et al., 2013b	✓	✓		
Agrawal et al., 2014	✓	✓		
Ayvaz et al., 2014		✓		
Baki et al., 2014		✓		✓
Clottey & Benton, 2014	✓	✓	✓	✓
Kumar, Palaniappan, et al., 2014	✓			
Kumar, Soleimani, & Kannan, 2014	✓	✓		✓
Liang et al., 2014		✓		✓
Tekin Temur et al., 2014	✓	✓		✓
Canda et al., 2015		✓		
Mota et al., 2015		✓		

Paper	Agrawal et al., 2015	Hachimi et al., 2018	Goltsos et al., 2019	Ambilkar et al., 2021
Parsopoulos et al., 2015		✓		
Sifaleras et al., 2015		✓		
Ma & Kim, 2016		✓		
Zhou et al., 2016		✓		✓
Ene & Öztürk, 2017		✓		✓
Sifaleras & Konstantaras, 2017		✓		
Wu et al., 2017		✓		
Banasik et al., 2018		✓		
Mota et al., 2018		✓		
Tsiliyannis, 2018				✓
Agrawal & Singh, 2020				✓
Chou et al., 2020				✓

Table IV.4. Sample derived from reviews, sorted by year of publication (n=56 papers).

IV.3.2 Review on forecasting consumer returns

For the location of studies regarding consumer returns forecasting in particular, five scientific databases were selected, aiming at journal and conference publications: (1) AIS Electronic Library (AISeL), (2) Business Source Ultimate (BSU) via EBSCOhost, (3) JSTOR, (4) Science Direct, and (5) Web of Science.

Consumer return related keywords	Forecasting related keywords
consumer return* product return* return* product retail return* customer return* e-commerce return*	forecast* predict* prognos*

Table IV.5. Search string combinations for Title/Abstract/Keywords.

The search string consists of two parts. On the one hand, we specifically searched for "consumer returns" and related terms like "e-commerce returns", "customer returns", and "retail returns" but also included the term "product returns", as the delimitation of these concepts might not be used coherently in the existing literature. Secondly, “forecasting” is paraphrased with “prediction” or “prognosis”. The final search strings include truncations, where applicable, to capture similar spellings or wordings (see Table IV.5).

The search performed in early 2023 resulted in 802 hits (see Figure IV.1).

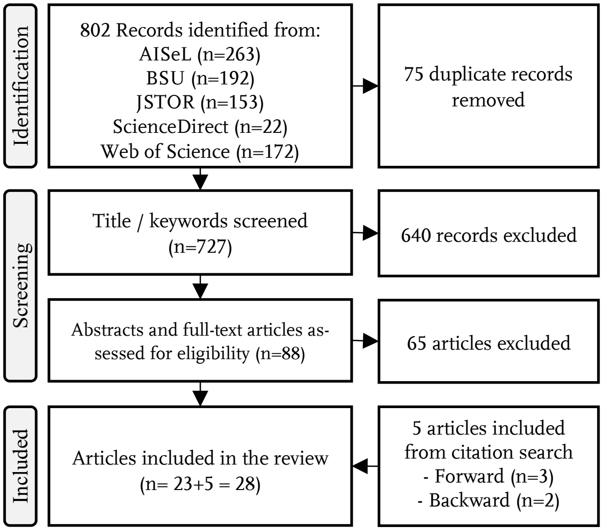

Figure IV.1. Research process of the systematic review.

Inclusion criteria were the English language, the title or keywords referring to returns or forecasting, and the paper being peer-reviewed and published as a journal or conference publication. After determining the papers’ relevance regarding consumer returns (in the sense of the delimitation to other return types outlined in the introduction) based on abstract and full text, the preliminary set consisted of 23 publications. Conducting a forward and backward search via Google Scholar, among 1264 references, we found five more relevant papers, leading to a total of 28 papers included for review. Wherever possible, we classified the results according to their publication outlet, the forecasting outcome, the focus of the paper, and the industry the paper refers to (see Table IV.6).

Paper	Outlet		Outcome				Focus						Industry			
	Journal	Conference	Quantity	Decision	Propensity	Fraud/abuse detection	Algorithm comparison	Data preparation	Development	Framework	Fraud/abuse	others	Fashion	Consumer electronics	Cosmetics	Others
Hess and Mayhew (1997)	✓		✓				✓						✓			
Potdar and Rogers (2012)	✓		✓						✓					✓		
Fu et al. (2016)	✓				✓				✓	✓					✓	
X. Li et al. (2019)	✓				✓						✓				✓	
Cui et al. (2020)	✓		✓				✓									✓
John et al. (2020)	✓					✓					✓					✓
Ketzenberg et al. (2020)	✓					✓					✓		✓			✓
Shang et al. (2020)	✓		✓						✓	✓			✓	✓		
Rezaei et al. (2021)	✓			✓				✓						✓		
Xu et al. (2023)	✓			✓					✓	✓						✓
Urbanke et al. (2015)		✓		✓				✓					✓			
Ahmed et al. (2016)		✓		✓				✓						✓		
Ding et al. (2016)		✓			✓			✓								✓
Drechsler and Lasch (2015)		✓	✓				✓				✓					✓
Heilig et al. (2016)		✓		✓			✓			✓			✓			
Samorani et al. (2016)		✓		✓				✓						✓		
Urbanke et al. (2017)		✓		✓				✓					✓			
Asdecker and Karl (2018)		✓		✓			✓						✓			
Joshi et al. (2018)		✓		✓	✓				✓				✓			
J. Li et al. (2018)		✓		✓					✓	✓			✓			
Zhu et al. (2018)		✓		✓					✓	✓			✓			
Hofmann et al. (2020)		✓		✓			✓	✓		✓						✓
Imran and Amin (2020)		✓		✓			✓									✓
Sweidan et al. (2020)		✓		✓					✓			✓	✓			
Fuchs and Lutz (2021)		✓		✓						✓						✓
Rajasekaran and Priyadarshini (2021)		✓			✓				✓							✓
Meißner et al. (2022)		✓	✓							✓		✓				✓
Zhang et al. (2022)		✓				✓						✓				✓

Table IV.6. Relevant works from the consumer returns forecasting database search (n=28 papers), sorted by outlet and year of publication.

IV.3.3 Combined sample

After joining the results from Section 3.2 with the dataset derived from product return reviews (Section 3.1), only two duplicates existed (Hess & Mayhew, 1997; Potdar & Rogers, 2012). This small overlap between consumer and product returns domains additionally justifies the review conducted under Section 3.2 as an extension for existing product return reviews. Furthermore, these two works are fundamental for the whole area of research. The final sample consists of 82 works (see Table IV.7). Of note, the sample also includes the three works referenced in the most recent review by Duong et al. (2022).

Paper	Simulation	OR: Mathematical	OR: Analytical	OR: Heuristic	Time series & regresion	Machine learning
Goh & Varaprasad, 1986					✓	
Kelle & Silver, 1989	✓					
Hess & Mayhew, 1997					✓	
Murayama et al., 2000	✓					
Toktay et al., 2000					✓	
Golany et al., 2001		✓				
Kiesmüller & van der Laan, 2001		✓				
Richter & Weber, 2001		✓				
Linton et al., 2002	✓					
Marx-Gómez et al., 2002						✓
Terzi & Cavalieri, 2004	✓					
Toktay et al., 2004					✓	
Kleyner & Sandborn, 2005					✓	
Y. Li et al., 2006		✓				
Peralta & Fontanos, 2006		✓				
Srivastava & Srivastava, 2006	✓					
Teunter et al., 2006				✓		
Hanafi et al., 2007						✓
Serrato et al., 2007		✓				
Wen-Jie & Zhi-Geng, 2007					✓	
Carrasco-Gallego & Ponce-Cueto, 2009					✓	

Paper	**Simulation**	**OR: Mathematical**	**OR: Analytical**	**OR: Heuristic**	**Time series & regresion**	**Machine learning**
Xiaofeng & Tijun, 2009		✓				
Chen & He, 2010					✓	
Yu et al., 2010					✓	
Barker & Zabinsky, 2011			✓			
Masmoudi, 2011					✓	
Schulz, 2011				✓		
Clottey et al., 2012					✓	
Plewa & Jodejko-Pietruczuk, 2012	✓					
Potdar & Rogers, 2012					✓	
Shih et al., 2012			✓			
Benedito & Corominas, 2013		✓				
Krapp et al., 2013a					✓	
Krapp et al., 2013b					✓	
Agrawal et al., 2014			✓			
Ayvaz et al., 2014					✓	
Baki et al., 2014				✓		
Clottey & Benton, 2014					✓	
Kumar, Palaniappan, et al., 2014			✓			
Kumar, Soleimani, & Kannan, 2014						✓
Liang et al., 2014					✓	
Tekin Temur et al., 2014				✓		✓
Canda et al., 2015					✓	
Mota et al., 2015		✓				
Parsopoulos et al., 2015				✓		
Sifaleras et al., 2015				✓		
Urbanke et al., 2015						✓
Ahmed et al., 2016						✓
Ding et al., 2016					✓	
Drechsler & Lasch, 2015					✓	
Fu et al., 2016			✓			
Heilig et al., 2016						✓
Ma & Kim, 2016					✓	
Samorani et al., 2016						✓
Zhou et al., 2016			✓			
Ene & Öztürk, 2017					✓	

Paper	Simulation	OR: Mathematical	OR: Analytical	OR: Heuristic	Time series & regresion	Machine learning
Sifaleras & Konstantaras, 2017				✓		
Urbanke et al., 2017						✓
Wu et al., 2017				✓		
Asdecker & Karl, 2018						✓
Banasik et al., 2018			✓			
Joshi et al., 2018						✓
J. Li et al., 2018						✓
Mota et al., 2018		✓				
Tsiliyannis, 2018	✓					
Zhu et al., 2018			✓			
X. Li et al., 2019						✓
Agrawal & Singh, 2020						
Chou et al., 2020		✓				
Cui et al., 2020						✓
Hofmann et al., 2020						✓
Imran & Amin, 2020						✓
John et al., 2020					✓	
Ketzenberg et al., 2020						✓
Shang et al., 2020					✓	
Sweidan et al., 2020						✓
Fuchs & Lutz, 2021						✓
Rajasekaran & Priyadarshini, 2021						✓
Rezaei et al., 2021						✓
Meißner et al. (2022)					✓	
Zhang et al. (2022)	✓					
Xu et al. (2023)						✓

Table IV.7. Final sample (n=82 papers) according to the categories of Hachimi et al., 2018, sorted by year of publication.

IV.4 Results

The presentation of the results is split into (standard) forecasting algorithms in general (Section 4.1), presenting an extension for the classification provided by Hachimi et al. (2018) (Simulation, Operations Research: mathematical, Operations Research: analytical, Operations Research: heuristic, time series and regression) with some additional sources. Section 4.2 focuses on advanced (machine learning) algorithms. Section 4.3 covers indicators for evaluating time series, regression, and advanced algorithms. Section 4.4 details the recent research discussion in academic journals, and Section 4.5 discusses some future research opportunities.

IV.4.1 Forecasting algorithms in general

Regarding **simulation programs**, an often-cited paper (Kelle & Silver, 1989) uses simulation to compare different forecasting procedures. Tsiliyannis (2018) uses a Markov-chain Monte Carlo simulation to assess return forecasting. Additional **OR models** divide into mathematical programming (Chou et al., 2020; Kiesmüller & van der Laan, 2001; Peralta & Fontanos, 2006; Xiaofeng & Tijun, 2009) and analytical models (Kumar, Palaniappan, et al., 2014). Some papers contribute to **time series and regression approaches**: In a multi-step method, Potdar and Rogers (2012) use essential statistical methods (moving averages, correlations, regression). Logistic regression was used by Yu et al. (2010). For grey forecasting, an additional source was found (Chen & He, 2010).
Only two papers from the field of consumer returns fall into the category of OR models. Fu et al. (2016) publish an analytical model for product- and customer-individual return propensities. Zhu et al. (2018) examine a graph-clustering-based analytical model for return decisions. For forecasting consumer return volumes, time series techniques are widely used (Drechsler & Lasch, 2015; Potdar & Rogers, 2012; Shang et al., 2020). Within this strain of literature, regression-based models are used for various purposes, but only two works exclusively rely on these procedures (Ding et al., 2016; John et al., 2020).

IV.4.2 Advanced forecasting algorithms

Twenty-two works contain advanced algorithms from the machine learning domain (see Table IV.8).

Algorithm	Forecasting purpose and references
RF	Decision (Ahmed et al., 2016; Heilig et al., 2016; Rezaei et al., 2021; Samorani et al., 2016; Sweidan et al., 2020; Urbanke et al., 2015; Urbanke et al., 2017), propensity (Rajasekaran & Priyadarshini, 2021), abuse (Ketzenberg et al., 2020), quantity (Cui et al., 2020)
SVM	Decision (Ahmed et al., 2016; Heilig et al., 2016; Hofmann et al., 2020; Joshi et al., 2018; Rezaei et al., 2021; Urbanke et al., 2015; Urbanke et al., 2017), abuse (Ketzenberg et al., 2020)
NN	Decision (Ahmed et al., 2016; Asdecker & Karl, 2018; Heilig et al., 2016; Imran & Amin, 2020; Urbanke et al., 2015; Xu et al., 2023), abuse (Ketzenberg et al., 2020)
GradientBoost	Decision (Fuchs & Lutz, 2021; Hofmann et al., 2020; Imran & Amin, 2020), propensity (Rajasekaran & Priyadarshini, 2021), quantity (Cui et al., 2020)
Fuzzy	Quantity (Hanafi et al., 2007; Kumar, Soleimani, & Kannan, 2014; Marx-Gómez et al., 2002; Tekin Temur et al., 2014)
AdaBoost	Decision (Heilig et al., 2016; Urbanke et al., 2015; Urbanke et al., 2017)
CART	Decision (Heilig et al., 2016; Urbanke et al., 2015; Urbanke et al., 2017)
DT	Decision (Ahmed et al., 2016; Asdecker & Karl, 2018)
ERT	Decision (Urbanke et al., 2015; Urbanke et al., 2017)
Ensemble	Decision (Asdecker & Karl, 2018; Heilig et al., 2016)
XGBoost	Decision (Imran & Amin, 2020)
CatBoost	Decision (Hofmann et al., 2020)
JRIP	Decision (Ahmed et al., 2016)
BayesNet	Decision (Ahmed et al., 2016)
kNN	Decision (J. Li et al., 2018)
Random Walk	Propensity (X. Li et al., 2019)
Clustering	Feature generation (Rezaei et al., 2021)
HyperGo	Decision (J. Li et al., 2018)

Table IV.8. Machine learning algorithms & forecasting purpose (n=22 papers).

Most advanced algorithms were found within consumer returns papers; only four were selected via product returns reviews. Consequently, the most frequent forecasting purpose was determining return decisions or propensities in an e-commerce context. Forecasting return quantities via machine learning has not been studied frequently, while classical time series procedures seem to fit this purpose better.

The most popular algorithms for various purposes were RF (10), SVM (8), NN (7), GradientBoost (5), and fuzzy algorithms. Interestingly, RF and SVM algorithms were used more frequently than NN, which was the most popular algorithm in general e-commerce applications, according to Micol Policarpo et al. (2021).

Some papers focus on the comparison of algorithms (Asdecker & Karl, 2018; Cui et al., 2020; Heilig et al., 2016; Hofmann et al., 2020; Imran & Amin, 2020) based on performance indicators (see Section 4.3), but no definitive ranking can be derived. In forecasting a binary return decision, Random Forests (RF) (Ahmed et al., 2016; Heilig et al., 2016; Ketzenberg et al., 2020), Neural Networks (NN) (Imran & Amin, 2020; Ketzenberg et al., 2020), as well as Adaptive Boosting (AdaBoost) (Urbanke et al., 2015; Urbanke et al., 2017) achieve high performance. Ensemble or ensemble-like algorithms (e.g., RF) are robust and highly accurate (Asdecker & Karl, 2018; Heilig et al., 2016).

A few works cover forecasting algorithms more or less incidentally, concentrating on data preparation (e.g., automated feature generation or selection via ML algorithms) for forecasting (Ahmed et al., 2016; Rezaei et al., 2021; Samorani et al., 2016; Urbanke et al., 2015; Urbanke et al., 2017).

IV.4.3 Forecasting evaluation

Table IV.9 categorizes assessment indicators extracted from 36 papers regarding time series, regression, or ML product return forecasts by their respective outcome, integrating the consumer returns forecasting literature into the literature from existing reviews. Not all works used a numerical assessment; e.g., Marx-Gómez et al. (2002) evaluated the forecasting quality based on graphical representations.

Some works from time series forecasting use expected-cost measures in the context of inventory management (Toktay et al., 2000; Toktay et al., 2004). However, most papers from this domain use classical indicators like MAPE or MAE, supplemented by a coefficient of determination for regression procedures. Interestingly, papers predicting return propensities often assess the results using classical error metrics

(MAE/MSE/RMSE) and indicators often applied in forecasting dichotomous return decisions (precision/recall/F-score/AUC). The F-score individually weighs precision and recall; e.g., $F_{0.5}$ weights precision higher than recall, as in an e-commerce scenario, the share of true positives among all positives is more important for targeting possible interventions (J. Li et al., 2018) to reduce the risk of unjustified limiting the shopping experience.

	Indicator	Purpose	References
Indicators for dichotomous outcomes	Precision	Decision	Hofmann et al., 2020; Imran & Amin, 2020; Joshi et al., 2018; J. Li et al., 2018; Rezaei et al., 2021; Sweidan et al., 2020; Urbanke et al., 2015; Xu et al., 2023; Zhu et al., 2018
		Propensity	Fu et al., 2016; X. Li et al., 2019
		Fraud/abuse	John et al., 2020; Ketzenberg et al., 2020
	Recall	Decision	Hofmann et al., 2020; Imran & Amin, 2020; Joshi et al., 2018; J. Li et al., 2018; Rezaei et al., 2021; Sweidan et al., 2020; Urbanke et al., 2015; Xu et al., 2023; Zhu et al., 2018
		Propensity	Fu et al., 2016; X. Li et al., 2019
		Fraud/abuse	John et al., 2020; Ketzenberg et al., 2020
	F-Score	Decision	Imran & Amin, 2020; Joshi et al., 2018; J. Li et al., 2018; Rezaei et al., 2021; Xu et al., 2023; Zhu et al., 2018
		Propensity	X. Li et al., 2019
	AUC	Decision	Ahmed et al., 2016; Heilig et al., 2016; Imran & Amin, 2020; J. Li et al., 2018; Rezaei et al., 2021; Samorani et al., 2016; Sweidan et al., 2020; Urbanke et al., 2017; Xu et al., 2023
		Propensity	Fu et al., 2016
		Abuse	Ketzenberg et al., 2020
	Accuracy	Decision	Asdecker & Karl, 2018; Heilig et al., 2016; Imran & Amin, 2020; Sweidan et al., 2020; Xu et al., 2023
	Specificity	Decision	Sweidan et al., 2020

	Indicator	Purpose	References
Indicators for continuous outcomes	MAPE	Quantity	Canda et al., 2015; Chen & He, 2010; Clottey et al., 2012; Clottey & Benton, 2014; Drechsler & Lasch, 2015; Ene & Öztürk, 2017; Hanafi et al., 2007; Krapp et al., 2013a, 2013b; Kumar, Soleimani, & Kannan, 2014
	MAE	Propensity	Fu et al., 2016; X. Li et al., 2019; Rajasekaran & Priyadarshini, 2021
		Quantity	Canda et al., 2015; Drechsler & Lasch, 2015; Krapp et al., 2013a, 2013b; Kumar, Soleimani, & Kannan, 2014
	MSE	Propensity	Rajasekaran & Priyadarshini, 2021
		Quantity	Ayvaz et al., 2014; Cui et al., 2020; Krapp et al., 2013a, 2013b
	RMSE	Propensity	Fu et al., 2016; Rajasekaran & Priyadarshini, 2021
		Quantity	Canda et al., 2015; Kumar, Soleimani, & Kannan, 2014; Meißner et al., 2022; Shang et al., 2020
	TAE	Decision	Asdecker & Karl, 2018
		Return Rate	Masmoudi, 2011
	TIC	Quantity	Drechsler & Lasch, 2015; Krapp et al., 2013a
	Relative Error	Quantity	Ene & Öztürk, 2017
Regression metrics & others	R / R^2	Decision	Tekin Temur et al., 2014; Urbanke et al., 2017
		Quantity	Cui et al., 2020; Kumar, Soleimani, & Kannan, 2014; Potdar & Rogers, 2012; Tekin Temur et al., 2014
	AIC	Quantity	Ding et al., 2016
	Tracking Signal	Quantity	Ayvaz et al., 2014

Table IV.9. Usage of performance indicators (n=36 papers).

IV.4.4 Recent discussion in academic journals

This section presents an overview of the latest discussion published in academic journals found via the database search and published in well-renowned journals since 2020.

Cui et al. (2020) use a comprehensive data set from an automotive accessories manufacturer aiming to forecast return volume for sales channels and different products. A hierarchical OLS regression incorporates up to 40 predictors. Relatively small differences in model quality (R^2) led to overproportional changes in the MSE. A machine-learning approach for predictor selection ("LASSO") resulted in a significant MSE reduction, while other data mining approaches (random forest, gradient boosting) could not outperform the LASSO approach.

The paper from Chou et al. (2020) models inventory policies integrating return forecasts for used products that are intended to be remanufactured ("buyback cores"). Past demand and sales serve as predictors for returns. When comparing two different models, they only find minor differences, which is confirmed by numerical examples based on data from a remanufacturing company.

Ketzenberg et al. (2020) segment customers and target detecting the small number of abusive returners, as these are unprofitable for the retailer and generate significant losses over a long time. In general, high-returning customers are usually more profitable. Different algorithms like logit, SVM, RF, and NN are combined with shrinkage methods like LASSO, ridge regression, and elastic net for classification. RF and especially NN outperform the other algorithms, assessed by sensitivity, precision, and AUC.

Shang et al. (2020) developed a predict-aggregate (P-A) model adaptable both for retailers and manufacturers for forecasting return volume in a continuous timeframe, in contrast to commonly used aggregate-predict (A-P) models. ARIMA and lagged return models known from end-of-life forecasting (de Brito et al., 2003) are used as benchmarks, using RMSE as an assessment criterion. The authors show that even the basic version of their approach outperforms the benchmark models, though using only lagged returns and timestamps as input.

John et al. (2020) predict the rare event of return fraud from customer representatives buying and returning items fraudulently. They chose a penalized likelihood logit model and evaluated it by precision and recall, focusing on maximizing recall and minimizing false negatives.

The paper by Rezaei et al. (2021) introduces an algorithm to select attributes from high-dimensional databases for forecasting purposes automatically. They use simulated data and the publicly available ISMS Durable Goods dataset (Ni et al., 2012) for consumer electronics as demonstration sample. AUC, precision, recall, and f1-score assess the results. LASSO as shrinkage method generally works best for the simulated data, outperforming RF and BaggedTrees. For real-world data, based on a forecast with a logit model, they show that the proposed selection algorithm with a lower variable complexity performs similarly or better compared to LASSO, SVM, and RF.

The framework of Xu et al. (2023) predicts problematic products with high return rates (> 30 %) in live-streaming e-commerce, which is prone to impulse purchases and suffers from high return rates. Based on multimodal features, including visuals, texts, and audio sources, their model is tested with real-world data from Taobao.com. For classification, a deep neural network (DNN) is complemented by a multihead attention mechanism for combining the different predictor modalities. Evaluated by AUC, precision, recall, F1-score, and accuracy, the framework proved suitable for live return predictions with high accuracy (> 90 %).

IV.4.5 Research gaps and paths

This review identified a research gap regarding ML algorithms for forecasting (e.g., end-of-life) returns in closed-loop supply chain management. Previous reviews presented ML approaches based only on different fuzzy algorithms, while many different ML algorithms were used in forecasting consumer returns. The likely explanation is that consumer return research predominantly models return decisions (classification problems), while the CLSC domain predominantly intends to forecast return quantities. Nevertheless, the paper by Cui et al. (2020) could serve as a role model when adjusting the input parameters for the respective reverse logistics objectives. Furthermore, when considering a predict-aggregate approach proposed by Shang et al. (2020), classification techniques could be helpful in forecasting return decisions before cumulating the propensity results for the volume prediction in the second step. Vice-versa, fuzzy algorithms have not been considered in forecasting consumer return quantities thus far.

Another significant gap could be recognized regarding the implementation and real-world application of the proposed algorithms. Indeed, some recent papers focus on the practical use of forecasting, e.g., by incorporating a management dashboard in the framework of Meißner et al. (2022) or developing a management tool for retailers (Zhang et al., 2022). However, most papers present models and their evaluation based on single case studies. We neither know which algorithms are used in general business practice nor if and how the model assessment with the proposed performance metrics takes place apart from research purposes. Generally, forecasting accuracy in RL should increase operational performance (Agrawal & Singh, 2020). We call for more empirical studies focussing on the business value to substantiate this relation. For advanced forecasting algorithms and their application in e-commerce, we recommend evaluating their use based on multiple case studies and surveys on operational performance (Agrawal & Singh, 2020) or based on cost metrics or profit (Ketzenberg et al., 2020) instead of more technical performance indicators.

To further elaborate on this, one could argue that analyzing the assessment metrics used in the papers can only deliver limited insights, as many of these indicators are used in interchangeable forecasting domains like finance. Nevertheless, this leads to a long-standing discussion of whether these indicators correlate to the business value, which should not be taken for granted (e.g., Leitch & Tanner, 1991). Thus, besides the scientific discussion about algorithms and assessment indicators, it is due to study actual consumer returns forecasting implementations for ERP or CRM systems in business practice, in the short and in the long term. Frameworks for these systems exist (Fuchs & Lutz, 2021; Heilig et al., 2016; Hofmann et al., 2020; Xu et al., 2023), but we did not find any papers assessing the real-time use of such systems and the respective algorithms.

Some studies compare forecasting algorithms (e.g., Asdecker & Karl, 2018; Urbanke et al., 2015). These comparisons rarely consider that each algorithm, algorithm implementation or parametrization may require tailored data preprocessing, affecting the performance (Alshdaifat et al., 2021). Moreover, the specific parameterization is rarely made transparent. We also call for sensitivity analyses and different input variable sets for the forecasting algorithms examined.

Regarding the predictors for forecasting models, it is well known that product returns from the customer of unused products are due to the customer ordering the same product in different sizes to be sure that one fits their needs (Asdecker et al., 2017). Fashion retailers heavily work on solutions for customers to find the perfect fit in the pre-purchase phase and refrain from a practice that guarantees return flows. This emphasizes the need to adequately model such initiatives to anticipate their effectiveness, similar to the approach of Zhang et al. (2022) for fraudulent returns. Thus, research on consumer return forecasts should, in addition to discussing algorithms and techniques, focus on the contextual aspects (i.e., relevant predictors): When and why do returns occur, and to what extent can forecasts help prevent them?

IV.5 Conclusion

The research area of product returns forecasting is highly interdisciplinary, referring to information systems research, operations research, operations & logistics, and marketing. Motivated by skyrocketing e-commerce sales and the associated returns, ML methods that have become more and more common in management and for decision support systems, and encouraged by research gaps resulting in calls for research regarding these issues in previous literature reviews, we examined the scientific literature on the use of different return forecasting algorithms and metrics for their evaluation. Based on existing reviews and an additional database search on consumer returns forecasting, we deliver an overview of advanced forecasting techniques and the indicators to assess their performance. Based on this, we outline several research propositions.

Of note, this review entails some limitations: While previous reviews mostly excluded conference papers from their data collection, we integrated them into our database search on consumer returns, making the search less comparable. However, e.g., the AIS eLibrary delivers insightful, up-to-date, and relevant conference works on this topic. Furthermore, no overarching meta-comparison of algorithm performance was possible given the studies at hand, as controlling for the crucial model parameters and the data preprocessing is hardly possible. This fact adds to our call for more empirical multiple case study research, either qualitative or quantitative, which could reveal the algorithm's performance in business use.

This work reviewed the state of research on return forecasting and outlined some future research paths. We are confident that this paper is helpful for other scholars researching this topic, e.g., when determining which evaluation metrics to use. For practitioners, we provide an overview and kickoff for their enquiry when implementing return forecasting algorithms, especially in e-commerce.

References

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Agrawal, S., K. Singh, R., & Murtaza, Q. (2014). Forecasting product returns for recycling in Indian electronics industry. *Journal of Advances in Management Research, 11*(1), 102–114.

Agrawal, S., & Singh, R. K. (2020). Forecasting product returns and reverse logistics performance: structural equation modelling. *Management of Environmental Quality: An International Journal, 31*(5), 1223–1237.

Agrawal, S., Singh, R. K., & Murtaza, Q. (2015). A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling, 97*, 76–92.

Ahmed, F., Samorani, M., Bellinger, C., & Zaiane, O. R. (2016). Advantage of Integration in Big Data: Feature Generation in Multi-Relational Databases for Imbalanced Learning. In *Proceedings of the 4th IEEE International Conference on Big Data,* Washington DC, United States of America.

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194.

Alshdaifat, E., Alshdaifat, D., Alsarhan, A., Hussein, F., & El-Salhi, S. M. F. S. (2021). The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms' Performance. *Data, 6*(2), 11.

Ambilkar, P., Dohale, V., Gunasekaran, A., & Bilolikar, V. (2021). Product returns management: A comprehensive review and future research agenda. *International Journal of Production Research,* 1–25.

Asdecker, B., & Karl, D. (2018). Big data analytics in returns management - Are complex techniques necessary to forecast consumer returns properly? In *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics,* Valencia, Spain.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Ayvaz, B., Bolturk, E., & Kaçtıoğlu, S. (2014). A Grey System for the Forecasting of Return Product Quantity in Recycling Network. *International Journal of Supply Chain Management, 3*(3), 105–112.

Baki, M. F., Chaouch, B. A., & Abdul-Kader, W. (2014). A heuristic solution procedure for the dynamic lot sizing problem with remanufacturing and product recovery. *Computers & Operations Research, 43*, 225–236.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics, 16*(5), 412–424.

Banasik, A., Bloemhof-Ruwaard, J. M., Kanellopoulos, A., Claassen, G. D. H., & van der Vorst, J. G. A. J. (2018). Multi-criteria decision making approaches for green supply chains: a review. *Flexible Services and Manufacturing Journal, 30*(3), 366–396.

Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q., & Seaman, B. (2019). Sales Demand Forecast in E-commerce Using a Long Short-Term Memory Neural Network Methodology. In T. Gedeon, K. W. Wong, & M. Lee (Chairs), *Neural Information Processing: Proceedings of the 26th International Conference on Neural Information Processing,* Sydney, Australia.

Barbosa, M. W., La Vicente, A. d. C., Ladeira, M. B., & Oliveira, M. P. V. de (2018). Managing supply chain resources with Big Data Analytics: a systematic review. *International Journal of Logistics Research and Applications, 21*(3), 177–200.

Barker, T. J., & Zabinsky, Z. B. (2011). A multicriteria decision making model for reverse logistics using analytical hierarchy process. *Omega, 39*(5), 558–573.

Benedito, E., & Corominas, A. (2013). Optimal manufacturing policy in a reverse logistic system with dependent stochastic returns and limited capacities. *International Journal of Production Research, 51*(1), 189–201.

Brandenburg, M., Govindan, K., Sarkis, J., & Seuring, S. (2014). Quantitative models for sustainable supply chain management: Developments and directions. *European Journal of Operational Research, 233*(2), 299–312.

Canda, A., Yuan, X.-M., & Wang, F. (2015). Modeling and forecasting product returns: An industry case study. In *Proceedings of the 2015 IEEE International Conference on Industrial Engineering and Engineering Management,* Singapore.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*(3), 1140–1154.

Carrasco-Gallego, R., & Ponce-Cueto, E. (2009). Forecasting the returns in reusable containers' closed-loop supply chains. A case in the LPG industry. In *Proceedings of the 3rd International Conference on Industrial Engineering and Industrial Management,* Barcelona, Spain.

Chang, Y., & Makatsoris, H. (2001). Supply chain modelling using simulation. *International Journal of Simulation, 2*(1), 24–30.

Chen, H., & He, H. (2010). Reverse Logistics Demand Forecasting under Demand Uncertainty. In *Proceedings of the 2010 International Conference of Logistics Engineering and Management,* Chengdu, China.

Chopra, S. (2019). *Supply chain management: Strategy, planning, and operation* (Seventh edition). Pearson.

Chou, M. C., Sim, C.-K., & Yuan, X.-M. (2020). Policies for inventory models with product returns forecast from past demands and past sales. *Annals of Operations Research, 288*(1), 137–180.

Clottey, T., & Benton, W. C. (2014). Determining core acquisition quantities when products have long return lags. *IIE Transactions, 46*(9), 880–893.

Clottey, T., Benton, W. C., & Srivastava, R. (2012). Forecasting Product Returns for Remanufacturing Operations. *Decision Sciences, 43*(4), 589–614.

Cui, H., Rajagopalan, S., & Ward, A. R. (2020). Predicting product return volume using machine learning methods. *European Journal of Operational Research, 281*(3), 612–627.

de Brito, M. P., Dekker, R., & Flapper, S. D. P. (2003). *Reverse Logistics - a Review of Case Studies*. Retrieved May 28, 2023, from https://papers.ssrn.com/Sol3/Delivery.cfm/277.pdf?abstractid=411649

Denyer, D., & Tranfield, D. (2009). Producing a Systematic Review. In D. A. Buchanan & A. Bryman (Eds.), *The Sage handbook of organizational research methods* (pp. 671–689). Sage.

Ding, Y., Xu, H., & Tan, B. C. Y. (2016). Predicting Product Return Rate with "Tweets". In *Proceedings of the 20th Pacific Asia Conference on Information Systems,* Chiayi, Taiwan.

Drechsler, S., & Lasch, R. (2015). Forecasting Misused E-Commerce Consumer Returns. In *Logistics Management: Proceedings of the 9th Conference "Logistikmanagement",* Braunschweig, Germany.

Duong, Q. H., Zhou, L., Meng, M., van Nguyen, T., Ieromonachou, P., & Nguyen, D. T. (2022). Understanding product returns: A systematic literature review using machine learning and bibliometric analysis. *International Journal of Production Economics, 243*, 108340.

Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications, 36*(3), 6697–6707.

El Kihal, S., & Shehu, E. (2022). It's not only what they buy, it's also what they keep: Linking marketing instruments to product returns. *Journal of Retailing, 98*(3), 558–571.

Ene, S., & Öztürk, N. (2017). Grey modelling based forecasting system for return flow of end-of-life vehicles. *Technological Forecasting and Social Change, 115,* 155–166.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research, 15*(1), 3133–3181.

Fildes, R., Ma, S., & Kolassa, S. (2022). Retail forecasting: Research and practice. *International Journal of Forecasting, 38*(4), 1283–1318.

Frei, R., Jack, L., & Brown, S. (2020). Product returns: A growing problem for business, society and environment. *International Journal of Operations & Production Management, 40*(10), 1613–1621.

Frei, R., Jack, L., & Krzyzaniak, S.-A. (2022). Mapping Product Returns Processes in Multichannel Retailing: Challenges and Opportunities. *Sustainability, 14*(3), 1382.

Fu, Y., Liu, G., Papadimitriou, S., Xiong, H., Li, X., & Chen, G. (2016). Fused latent models for assessing product return propensity in online commerce. *Decision Support Systems, 91*, 77–88.

Fuchs, K., & Lutz, O. (2021). A stitch in time saves nine – A meta-model for real-time prediction of product returns in ERP systems. In *Proceedings of the 29th European Conference on Information Systems,* Marrakech, Morocco.

Galdi, P., & Tagliaferri, R. (2018). Data mining: accuracy and error measures for classification and prediction. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 431–436).

Ge, D., Pan, Y., Shen, Z.-J., Di Wu, Yuan, R., & Zhang, C. (2019). Retail supply chain management: a review of theories and practices. *Journal of Data, Information and Management, 1*, 45–64.

Goh, T. N., & Varaprasad, N. (1986). A Statistical Methodology for the Analysis of the Life-Cycle of Reusable Containers. *IIE Transactions, 18*(1), 42–47.

Golany, B., Yang, J., & Yu, G. (2001). Economic lot-sizing with remanufacturing options. *IIE Transactions, 33*(11), 995–1003.

Goltsos, T. E., Ponte, B., Wang, S. X., Liu, Y., Naim, M. M., & Syntetos, A. A. (2019). The boomerang returns? Accounting for the impact of uncertainties on the dynamics of remanufacturing systems. *International Journal of Production Research, 57*(23), 7361–7394.

Govindan, K., & Bouzon, M. (2018). From a literature review to a multi-perspective framework for reverse logistics barriers and drivers. *Journal of Cleaner Production, 187*, 318–337.

Guide, V. D. R. (2000). Production planning and control for remanufacturing: industry practice and research needs. *Journal of Operations Management, 18*(4), 467–483.

Hachimi, H. E. L., Oubrich, M., & Souissi, O. (2018). The optimization of reverse logistics activities: A literature review and future directions. In *Proceedings of the 5th IEEE International Conference on Technology Management, Operations and Decisions,* Marrakech, Morocco.

Hanafi, J., Kara, S., & Kaebernick, H. (2007). Generating Fuzzy Coloured Petri Net Forecasting Model to Predict the Return of Products. In *Proceedings of the 15th IEEE International Symposium on Electronics & the Environment,* Orlando, United States of America.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (Second edition). Springer.

Heilig, L., Hofer, J., Lessmann, S., & Voß, S. (2016). Data-Driven Product Returns Prediction: A Cloud-Based Ensemble Selection Approach. In *Proceedings of the 24th European Conference on Information Systems,* Istanbul, Turkey.

Hess, J. D., & Mayhew, G. E. (1997). Modeling merchandise returns in direct marketing. *Journal of Direct Marketing, 11*(2), 20–35.

Hofmann, A., Gwinner, F., Fuchs, K., & Winkelmann, A. (2020). An Industry-Agnostic Approach for the Prediction of Return Shipments. In *Proceedings of the 26th Americas Conference on Information Systems,* Salt Lake City, United States of America.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

Imran, A. A., & Amin, M. N. (2020). Predicting the Return of Orders in the E-Tail Industry Accompanying with Model Interpretation. *Procedia Computer Science, 176,* 1170–1179.

John, S., Shah, B. J., & Kartha, P. (2020). Refund fraud analytics for an online retail purchases. *Journal of Business Analytics, 3*(1), 56–66.

Joshi, T., Mukherjee, A., & Ippadi, G. (2018). One Size Does Not Fit All: Predicting Product Returns in E-Commerce Platforms. In *Proceedings of the 10th IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining,* Barcelona, Spain.

Kelle, P., & Silver, E. A. (1989). Forecasting the returns of reusable containers. *Journal of Operations Management, 8*(1), 17–35.

Ketzenberg, M. E., Abbey, J. D., Heim, G. R., & Kumar, S. (2020). Assessing customer return behaviors through data analytics. *Journal of Operations Management, 66*(6), 622–645.

Kiesmüller, G. P., & van der Laan, E. A. (2001). An inventory model with dependent product demands and returns. *International Journal of Production Economics, 72*(1), 73–87.

Kleyner, A., & Sandborn, P. (2005). A warranty forecasting model based on piecewise statistical distributions and stochastic simulation. *Reliability Engineering & System Safety, 88*(3), 207–214.

Krapp, M., Nebel, J., & Sahamie, R. (2013a). Forecasting product returns in closed-loop supply chains. *International Journal of Physical Distribution & Logistics Management, 43*(8), 614–637.

Krapp, M., Nebel, J., & Sahamie, R. (2013b). Using forecasts and managerial accounting information to enhance closed-loop supply chain management. *OR Spectrum, 35*(4), 975–1007.

Kumar, D. T., Palaniappan, M., Kannan, D., & Shankar, K. M. (2014). Analyzing the CSR issues behind the supplier selection process using ISM approach. *Resources, Conservation and Recycling, 92*, 268–278.

Kumar, D. T., Soleimani, H., & Kannan, G. (2014). Forecasting return products in an integrated forward/reverse supply chain utilizing an ANFIS. *International Journal of Applied Mathematics and Computer Science, 24*(3), 669–682.

Leitch, G., & Tanner, J. E. (1991). Economic Forecast Evaluation: Profits Versus the Conventional Error Measures. *The American Economic Review, 81*(3), 580–590.

Li, J., He, J., & Zhu, Y. (2018). E-tail Product Return Prediction via Hypergraph-based Local Graph Cut. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, United Kingdom.

Li, X., Zhuang, Y., Fu, Y., & He, X. (2019). A trust-aware random walk model for return propensity estimation and consumer anomaly scoring in online shopping. *Science China Information Sciences, 62*(5).

Li, Y., Chen, J., & Cai, X. (2006). Uncapacitated production planning with multiple product types, returned product remanufacturing, and demand substitution. *OR Spectrum, 28*(1), 101–125.

Liang, X., Jin, X., & Ni, J [Jun] (2014). Forecasting product returns for remanufacturing systems. *Journal of Remanufacturing, 4*(1).

Linton, J. D., Yeomans, J. S., & Yoogalingam, R. (2002). Supply planning for industrial ecology and remanufacturing under uncertainty: a numerical study of leaded-waste recovery from television disposal. *Journal of the Operational Research Society, 53*(11), 1185–1196.

Ma, J., & Kim, H. M. (2016). Predictive Model Selection for Forecasting Product Returns. *Journal of Mechanical Design, 138*(5), Article 054501.

Marx-Gómez, J., Rautenstrauch, C., Nürnberger, A., & Kruse, R. (2002). Neuro-fuzzy approach to forecast returns of scrapped products to recycling and remanufacturing. *Knowledge-Based Systems, 15*(1-2), 119–128.

Masmoudi, M. (2011). Forecasting returns in reverse logistics: Application to catalog and mail-order retailing. In *Proceedings of the 4th International Conference on Industrial Engineering and Systems Management,* Metz, France.

Meisel, S., & Mattfeld, D. (2010). Synergies of Operations Research and Data Mining. *European Journal of Operational Research, 206*(1), 1–10.

Meißner, K., Stevenson, A. B., & Rieck, J. (2022). Identifying Non-intuitive Relationships Within Returns Data of a Furniture Online-Shop Using Temporal Data Mining. In E. Szczerbicki, K. Wojtkiewicz, S. van Nguyen, M. Pietranik, & M. Krótkiewicz (Eds.), *Communications in Computer and Information Science, Recent Challenges in Intelligent Information and Database Systems: Proceedings of the 14th Asian Conference on Intelligent Information and Database Systems* (pp. 299–312). Springer Nature Singapore.

Melacini, M., Perotti, S., Rasini, M., & Tappia, E. (2018). E-fulfilment and distribution in omni-channel retailing: a systematic literature review. *International Journal of Physical Distribution & Logistics Management, 48*(4), 391–414.

Micol Policarpo, L., da Silveira, D. E., da Rosa Righi, R., Antunes Stoffel, R., da Costa, C. A., Victória Barbosa, J. L., Scorsatto, R., & Arcot, T. (2021). Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review. *Computer Science Review, 41*, 100414.

Mota, B., Gomes, M. I., Carvalho, A., & Barbosa-Povoa, A. P. (2015). Towards supply chain sustainability: economic, environmental and social design and planning. *Journal of Cleaner Production, 105*, 14–27.

Mota, B., Gomes, M. I., Carvalho, A., & Barbosa-Povoa, A. P. (2018). Sustainable supply chains: An integrated modeling approach under uncertainty. *Omega, 77*, 32–57.

Murayama, T., Obata, H., Narutaki, N., Oba, F., & Shu, L. H. (2000). Production Planning Based on Reliability Models for Part Reuse. In *Proceedings of 7th CIRP International Seminar on Life Cycle Engineering*, Tokyo, Japan.

National Retail Federation / Appriss Retail. (2021). *Customer Returns in the Retail Industry 2021*. https://cdn.nrf.com/

Ni, J [Jian], Neslin, S. A., & Sun, B. (2012). Database Submission The ISMS Durable Goods Data Sets. *Marketing Science, 31*(6), 1008–1013.

Pandya, R., & Pandya, J. (2015). C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning. *International Journal of Computer Applications, 117*(16), 18–21.

Parsopoulos, K. E., Konstantaras, I., & Skouri, K. (2015). Metaheuristic optimization for the Single-Item Dynamic Lot Sizing problem with returns and remanufacturing. *Computers & Industrial Engineering, 83*, 307–315.

Peralta, G. L., & Fontanos, P. M. (2006). E-waste issues and measures in the Philippines. *Journal of Material Cycles and Waste Management, 8*(1), 34–39.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., . . . Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting, 38*(3), 705–871.

Plewa, M., & Jodejko-Pietruczuk, A. (2012). The reverse logistics forecasting model with whole product recovery. *Reliability: Theory & Applications, 7*(1 (24)), 98–113.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine, 6*(3), 21–45.

Potdar, A., & Rogers, J. (2012). Reason-code based model to forecast product returns. *Foresight, 14*(2), 105–120.

Rajasekaran, V., & Priyadarshini, R. (2021). An E-Commerce Prototype for Predicting the Product Return Phenomenon Using Optimization and Regression Techniques. In M. Singh, V. Tyagi, P. K. Gupta, J. Flusser, T. Ören, & V. R. Sonawane (Chairs), *Advances in Computing and Data Sciences: Proceedings of the 5th International Conference on Advances in Computing and Data Sciences,* Nashik, India.

Ratchford, B., Soysal, G., Zentner, A., & Gauri, D. K. (2022). Online and offline retailing: What we know and directions for future research. *Journal of Retailing, 98*(1), 152–177.

Rezaei, M., Cribben, I., & Samorani, M. (2021). A clustering-based feature selection method for automatically generated relational attributes. *Annals of Operations Research, 303*(1-2), 233–263.

Richter, K., & Weber, J. (2001). The reverse Wagner/Whitin model with variable manufacturing and remanufacturing cost. *International Journal of Production Economics, 71*(1-3), 447–456.

Robertson, T. S., Hamilton, R., & Jap, S. D. (2020). Many (Un)happy Returns? The Changing Nature of Retail Product Returns and Future Research Directions. *Journal of Retailing, 96*(2), 172–177.

Samorani, M., Ahmed, F., & Zaiane, O. R. (2016). Automatic Generation of Relational Attributes: An Application to Product Returns. In *Proceedings of the 4th IEEE International Conference on Big Data,* Washington DC, United States of America.

Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision, 57*(8), 1980–1992.

Schulz, T. (2011). A new Silver–Meal based heuristic for the single-item dynamic lot sizing problem with returns and remanufacturing. *International Journal of Production Research, 49*(9), 2519–2533.

Serrato, M. A., Ryan, S. M., & Gaytán, J. (2007). A Markov decision model to evaluate outsourcing in reverse logistics. *International Journal of Production Research, 45*(18-19), 4289–4315.

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data, 7*(1).

Shaharudin, M. R., Zailani, S., & Tan, K. C. (2015). Barriers to product returns and recovery management in a developing country: investigation using multiple methods. *Journal of Cleaner Production, 96*, 220–232.

Shang, G., McKie, E. C., Ferguson, M. E., & Galbreth, M. R. (2020). Using transactions data to improve consumer returns forecasting. *Journal of Operations Management, 66*(3), 326–348.

Shih, H.-S., Stanley Lee, E., Chuang, S.-H., & Chen, C.-C. (2012). A forecasting decision on the sales volume of printers in Taiwan: An exploitation of the Analytic Network Process. *Computers & Mathematics with Applications, 64*(6), 1545–1556.

Shulman, J. D., Coughlan, A. T., & Savaskan, R. C. (2011). Managing Consumer Returns in a Competitive Environment. *Management Science, 57*(2), 347–362.

Sifaleras, A., & Konstantaras, I. (2017). Variable neighborhood descent heuristic for solving reverse logistics multi-item dynamic lot-sizing problems. *Computers & Operations Research, 78*, 385–392.

Sifaleras, A., Konstantaras, I., & Mladenović, N. (2015). Variable neighborhood search for the economic lot sizing problem with product returns and recovery. *International Journal of Production Economics, 160*, 133–143.

Srivastava, S. K., & Srivastava, R. K. (2006). Managing product returns for reverse logistics. *International Journal of Physical Distribution & Logistics Management, 36*(7), 524–546.

Stock, J. R., & Mulki, J. P. (2009). Product Returns Processing: An Examination of Practices of Manufacturers, Wholesalers/Distributors, and Retailers. *Journal of Business Logistics, 30*(1), 33–62.

Sweidan, D., Johansson, U., & Gidenstam, A. (2020). Predicting returns in men's fashion. In *Proceedings of the 14th International Fuzzy Logic and Intelligent Technologies in Nuclear Science Conference,* Cologne, Germany.

Tekin Temur, G., Balcilar, M., & Bolat, B. (2014). A fuzzy expert system design for forecasting return quantity in reverse logistics network. *Journal of Enterprise Information Management, 27*(3), 316–328.

Terzi, S., & Cavalieri, S. (2004). Simulation in the supply chain context: a survey. *Computers in Industry, 53*(1), 3–16.

Teunter, R. H., Bayindir, Z. P., & van Heuvel, W. den (2006). Dynamic lot sizing with product returns and remanufacturing. *International Journal of Production Research, 44*(20), 4377–4400.

Tibben-Lembke, R. S., & Rogers, D. S. (2002). Differences between forward and reverse logistics in a retail environment. *Supply Chain Management: An International Journal, 7*(5), 271–282.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

Toktay, L. B., Wein, L. M., & Zenios, S. A. (2000). Inventory Management of Remanufacturable Products. *Management Science, 46*(11), 1412–1426.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management, 14*(3), 207–222.

Tsiliyannis, C. A. (2018). Markov chain modeling and forecasting of product returns in remanufacturing based on stock mean-age. *European Journal of Operational Research, 271*(2), 474–489.

Tsoumakas, G., Partalas, I., & Vlahavas, I. (2008). A taxonomy and short review of ensemble selection. In *Proceedings of the 18th European Conference on Artificial Intelligence,* Patras, Greece.

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 20*(1), 57–59.

Urbanke, P., Kranz, J., & Kolbe, L. (2015). Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction. In *Proceedings of the 14th International Conference on Computer and Information Science,* Fort Worth, United States of America.

Urbanke, P., Uhlig, A., & Kranz, J. (2017). A Customized and Interpretable Deep Neural Network for High-Dimensional Business Data - Evidence from an E-Commerce Application. In *Proceedings of the 38th International Conference on Information Systems,* Seoul, South Korea.

vom Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *Proceedings of the 17th European Conference on Information Systems,* Verona, Italy.

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii–xxiii.

Wen-Jie, L., & Zhi-Geng, F. (2007). A conceptual framework for reverse logistics. In *Proceedings of the 2007 IEEE International Conference on Grey Systems and Intelligent Services,* Nanjing, China.

Wu, Z., Kwong, C. K., Aydin, R., & Tang, J. (2017). A cooperative negotiation embedded NSGA-II for solving an integrated product family and supply chain design problem with remanufacturing consideration. *Applied Soft Computing, 57,* 19–34.

Xiaofeng, X., & Tijun, F. (2009). Forecast for the Amount of Returned Products Based on Wave Function. In *Proceedings of the 2009 International Conference on Information Management, Innovation Management and Industrial Engineering,* Xi'an, China.

Xu, W., Zhang, X [Xiaohan], Chen, R., & Yang, Z. (2023). How do you say it matters? A multimodal analytics framework for product return prediction in live streaming e-commerce. *Decision Support Systems,* 113984.

Yu, J., Williams, E., Ju, M., & Yang, Y. (2010). Forecasting global generation of obsolete personal computers. *Environmental Science & Technology, 44*(9), 3232–3237.

Zhang, D., Bayer, S., Willis, G., Frei, R., Gerding, E., & Senyo, P. (2022). Using Big Data Analytics to Combat Retail Fraud. In *Proceedings of the 4th International Conference on Finance, Economics, Management and IT Business* (pp. 85–92). SCITEPRESS.

Zhou, L., Xie, J., Gu, X., Lin, Y., Ieromonachou, P., & Zhang, X [Xiaole] (2016). Forecasting return of used products for remanufacturing using Graphical Evaluation and Review Technique (GERT). *International Journal of Production Economics, 181*, 315–324.

Zhu, Y., Li, J., He, J., Quanz, B. L., & Deshpande, A. (2018). A Local Algorithm for Product Return Prediction in E-Commerce. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence,* Stockholm, Sweden.

V Data Mining in Returns Management: Evaluation of Returns Volume Forecasts Based on Transaction Data of a Shoe and Fashion Mail-Order Company

Author: David Karl

Publication Outlet: Proceedings of the 5th conference on Mobility in a globalised world 2015

DOI: 10.20378/irb-40411

published as and **translated from:**

Karl, D. (2016): Data Mining im Retourenmanagement: Evaluation von Retourenmengenprognosen anhand der Transaktionsdaten eines Schuh- und Bekleidungsversandhändlers. In: *Proceedings of the 5th Conference on Mobility in a globalised world*, Landsberg am Lech, Germany, pp. 190–213. DOI: 10.20378/irb-40411

Abstract

The digitization of all aspects of life is causing online retailing to become increasingly important: Growing mail-order sales and increasing independence from brick-and-mortar stores lead to higher logistics volumes and, consequently, to more mobility of goods. Due to legal requirements and liberal return policies, this applies in particular to the reverse logistics and gives rise to operational challenges in the area of inventory and returns management. The fashion industry is particularly affected by the returns phenomenon, as customers often return every second package.

This paper first derives models from extensive historical transaction data of a German shoe and apparel mail-order company using data mining methods. Subsequently, these models are used to predict future returns at the time of ordering. By evaluating the volume-related prediction based on these models in comparison to established forecasting methods, conclusions can be drawn about the benefits of data mining in returns management. In addition, practical recommendations can be generated as to which methods are most suitable for returns volume forecasting.

V.1 Introduction

"Retailers have long understood that the digitization of all our business processes will completely revolutionize the century-old retail business model. Better put, digitization will turn retail into an experimental field for unleashed data analytics of all kinds in the coming years." (translated from Wenzel & Dziemba, 2014, p. 95)

This quote illustrates that data analysis is and will be a critical success factor for retail. Demand forecasts – often estimated intuitively without information technology support (Crone, 2010, p. 92) – have long been essential for brick-and-mortar retailers to be able to anticipate and satisfy customer needs. The ongoing digital transformation of processes and entire business models regarding information and communication technologies expands the potential data base and analysis options for retailers: Operational information systems such as enterprise resource planning (ERP) systems hold large data sets that are either a by-product of business processes or are deliberately collected (e.g., expressed customer preferences). However, the phrase "experimental field" (Wenzel & Dziemba, 2014, p. 95) from the above-mentioned quotation indicates that there is considerable uncertainty as to how retailers can exploit and use this data in a meaningful and beneficial way.

Mail-order retailers in particular, whose business model is characterized by information flows to and from the customer in a special manner due to the spatial discrepancy between supply and demand, can benefit from the purposeful use of data. Compared to stationary retail, mail-order business is continuously gaining in importance (Handelsverband Deutschland [HDE], 2023b; Jahn, 2013, p. 3). In the footwear and apparel industry, online retailing already accounts for almost 20 % of sales (HDE, 2023a). Since there is no inspection of the item of interest in the store, customers have the right under EU Consumer Rights Directive 2011/83/EU to inspect ordered goods at home and return them (Directive 2011/83/EU on consumer rights, 2011). Due to the associated expenses and possible loss of revenue for the retailer, the planning, management, and control of these returns is an elementary part of the distance-selling business model (Asdecker, 2014, back cover). Due to an increasing fusion of brick-and-mortar and online retailing, traditional retailers cannot ignore this issue and must consider their return policies (Jahn, 2013, pp. 8, 13; Wenzel & Dziemba, 2014, p. 112).

A survey of returns management practitioners by Asdecker (2014, p. 144) regarding the research gaps and the associated practical significance supports the relevance of data analysis in returns management: Many decision-makers perceive a need for research on “estimating [the] quantity of returns” using “forecasting models” as part of tactical planning, which is a key component of returns management (translated from Asdecker, 2014, p. 160).

Based on real-world transaction data from a shoe and apparel retailer over a period of one year, the following research questions shall be answered:

- *How accurately can retailers predict a customer’s return decision based on transaction data using data mining techniques?*
- *Which techniques are best suited for forecasting return volumes?*
- *To what extent do data mining models used for return decision-making enable the derivation of accurate return quantity forecasts?*

For this purpose, Chapter V.2 briefly introduces the conceptual and methodological basics of returns management and data mining. Chapter V.3 presents the data set used by the retailer and the associated data mining models for forecasting the return decision. Chapter V.4 uses these models to generate predictions of the volume of returned items and packages, and compares these with naïve estimates of return volume based on sales figures and historical return rates. Finally, Chapter V.5 briefly summarizes the results obtained, outlines the limitations of the study, and identifies future research needs.

V.2 Conceptual and methodological basics

Section V.2.1 provides a basic introduction into the field of returns management and explains the key performance indicators that are relevant for the remainder of the paper. Section V.2.2 indicates potential benefits of data analysis in this area. Section V.2.3 outlines the conceptual basics of data mining as well as methods relevant for this paper, while Section V.2.4 briefly summarizes simple forecasting methods commonly used. Section

V.2.5 provides an overview of frequently used measures for evaluating forecasting models, which will be employed in the remainder of this paper.

V.2.1 Returns management and relevant key performance indicators

The term *returns* generally encompasses all returns to upstream institutions in the value creation process or to contracted service providers (Asdecker, 2011, p. 258; Rogers et al., 2002, pp. 3–4). Returns management refers to the "cross-institutional planning, control, and monitoring of return flows as well as associated information and financial flows, with the aim of supporting the profit maximization of the value creation system" (translated from Asdecker, 2012, p. 495).

Returns management tasks are divided into two areas: on the one hand, preventive tasks at the tactical level (prevention, avoidance, promotion) that involve early interventions before a return occurs and on the other hand, predominantly curative tasks at the operational level for optimal handling of the return process, focusing on actual or unavoidable returns (Asdecker, 2014, p. 23; Rogers et al., 2002, p. 5). This article primarily covers curative, internal returns management and is also limited to returns from end customers to retailers, in other words consumer-to-business (C2B) transactions ("consumer product returns").

Key performance indicators generally serve as condensed operational information to support management decisions with a data basis and to reduce complexity in interpreting business reality (Weber & Wallenburg, 2010, pp. 332–333). Returns can only be planned, controlled, and monitored meaningfully with aggregated data, as this frequently cited management wisdom illustrates: "[I]f you can't measure it, you can't manage it." (Garvin, 1993, p. 89)

Probably the most important key figure for returns management (Asdecker, 2014, pp. 229–230) is calculated from the ratio of returned units to the total number of units shipped:

$$RR_{unspecific} = \frac{units\ returned}{units\ shipped}$$

The return rate (RR) can be differentiated into α-, β-, and γ-return rates, particularly for the purposes of distance selling, with each perspective highlighting certain entrepreneurial functional areas (Asdecker, 2014, pp. 42–44): The α-return rate (α-RR) refers to the ratio of returned to shipped packages and, consequently, focuses on the logistical perspective, while the β-return rate (β-RR) expresses this ratio in relation to individual items and takes the marketing perspective. The γ-return rate (γ-RR) describes the monetary, value-based ratio of returned to shipped units, which corresponds to the financial and controlling perspective. Due to the operational focus of this article, the remainder of the analysis considers α- and β-return rates.

A more detailed classification of the return rate can be conducted with regard to the perspective or aggregation level. Common reference units are customers, items, item variants (e.g., colors), manufacturers, or an overall view from the company's perspective. To avoid confusion, this reference must always be specified, leading to two-dimensional manifestations of this key figure.

In addition to the return rate, practitioners use other key figures such as the absolute number of returns, lead times, return reasons, or error rates (Asdecker, 2014, pp. 227–231). This study focuses on operational sales key figures derived from historical transaction data and examines their influence on returns management.

V.2.2 Context-related benefits of data analytics

Data analytics for management purposes aim to achieve business benefits, which, in a broader sense, means value creation. The following section outlines potential ways of using returns-related data analytics in the e-commerce sector.

When predicting a return, the following factors are of interest: whether a customer returns an item or package, the timing of the return, and the condition of the returned items. The timing of the return is particularly relevant to resource allocation and order quantity planning. The condition of the item affects its resale value and the returns processing, including necessary steps for reuse, refurbishment, or disposal. These, in turn, affect process costs and order quantities from suppliers. No conclusions can be drawn about the condition and timing of a return based on the available

transaction data. The subsequent discussions focus solely on the decision to return and its probability.
Analyzing the probability of package-related returns or the α return rate is more applicable to curative returns management, as this metric and the resulting number of packages support capacity planning for specific logistical operations, such as workforce scheduling (Asdecker, 2014, p. 261). Predicting the quantity of returns related to specific items is operationally/curatively relevant for batch size and order quantity planning because returns directly affect inventory levels assuming the items are reusable. Managing returns is one of the main challenges in e-commerce. According to Loebich (2014), the central question for retailers is how to control their inventory and the number of returns. Accurate return forecasts can prevent overstocks, which lead to increased warehousing costs and longer capital lockup (Feindt & Sinn, 2014, p. 24). Understocks result in sales losses, opportunity costs, higher procurement costs, and customer dissatisfaction or loss of goodwill (Jarke, 2014, p. 330; Werner, 2013, p. 229). Item-specific return information can also be useful in supplier negotiations to obtain higher discounts, improved quality standards, or lower manufacturing tolerances.

V.2.3 Data mining fundamentals

According to Ratner (2012, pp. 13–14), data mining combines three components: classical statistical methods of exploratory data analysis, big data, and machine learning using information technology (IT). Therefore, it is not possible to make a clear distinction between data mining and the concept of big data. "Mining" is more about the process of extracting valuable information from complex datasets using classical statistical or modern IT algorithms, as summarized by Brethenoux of the Gartner Group:
"Data mining is the process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories. Data mining employs pattern recognition technologies, as well as statistical and mathematical techniques." (Gartner, 2023; for a German translation, see Kuß et al., 2014)
Data mining methods can be broadly categorized into four categories based on their application. Nevertheless, the classifications are not mutu-

ally exclusive, as identical algorithms can be used in multiple areas: numerical methods, clustering methods, association analyses, and classification methods, with only the latter being relevant for the further course of this study (Cleve & Lämmel, 2014, p. 57; Petersohn, 2009, p. 25).

Classification methods, including discriminant analysis, classification and decision trees, and neural networks, determine the class membership of a nominal-scaled (or dichotomous) dependent variable based on a set of predictors (Küsters, 2001, p. 107). In addition to class membership, the probability or prediction confidence is of interest (Cleve & Lämmel, 2014, p. 61). Unlike cluster analyses, the classes are known a priori, although both approaches may use the same algorithms, such as the aforementioned neural networks.

Combining different data mining methods into hybrid models can compensate for weaknesses in a single model and improve the overall model's efficiency, especially since data mining software packages enable automated model linking (Fidan & Ozkok, 2013, p. 121). Cross-validation of the model through subsample evaluation and aggregating multiple submodels can further stabilize or enhance predictions (Domingos, 2012, pp. 85–86; Fromm, 2005, p. 23).

For the present study, the software packages IBM SPSS Statistics 23 (classical statistical package), IBM SPSS Modeler 17 (data mining tool), and Microsoft Excel 2013 (spreadsheet for all numerical calculations) are used.

V.2.4 Common forecasting techniques

Forecasts or predictions refer to statements about future events and developments based on qualitative or quantitative foundations that promote rational decision-making (Hansmann, 1983, pp. 7; 12; Rosentreter, 1977, p. 5). Qualitative forecasting methods, which aim to estimate long-term trends, include expert surveys (e.g., Delphi method) or scenario techniques, which are not relevant to the current topic (Hansmann, 1983, pp. 18–26). Quantitative forecasting techniques, oriented toward the short to medium term, can be classified into two subcategories: time series extrapolation and causal analysis. A time series extrapolation calculates future values solely based on past values of the same variable and does not consider any influences other than time, which is also a major criticism of such methods (Hansmann, 1983, p. 45). By contrast, causal analysis estimate the target variable based on explanatory variables, considering their relationships and enabling the interpretation of directions and strengths of effects.

The first class of forecasting techniques includes simple time series forecasts such as moving averages or simple exponential smoothing. These two methods assume a constant model, meaning the forecasted value is assumed to have no underlying trend or seasonality. The forecast value t $\hat{y}_{T,T+1}$ for period $T+1$, based on the moving average (MA) method, is calculated as the arithmetic mean of the last N periods (Schröder, 2012, pp. 22–23).

$$\hat{y}_{T,T+1} = \frac{1}{N} \cdot \sum_{t=T+1-N}^{T} y_t$$

The principle of simple exponential smoothing (SES) considers all past values for the forecast but with exponentially decreasing weights: the farther back a value is in the past, the less it influences the next forecasted value. The smoothing parameter α, set in the interval between 0 and 1, determines how quickly the weight of older values decreases (Schröder, 2012, pp. 25–27). The forecasted value using simple exponential smoothing is defined as

$$\hat{y}_{T,T+1} = \alpha \cdot y_T + (1 - \alpha) \cdot \hat{y}_{T-1,T}.$$

However, the quantity of returns is directly dependent on the quantity of units shipped because a return to the retailer can occur only if the customer has previously received the item or package. In this regard, causal forecasting methods seem more appropriate, as they consider such functional relationships, unlike time series forecasting. Formally, bivariate linear regression can be described as follows:

$$\hat{y}_{T,T+1}(x) = a + b \cdot x_T$$

x_T represents the independent variable (i.e., the number of orders in period T) that explains the dependent variable y (number of returns). The parameters a and b are estimated by minimizing the squared deviation of the predicted value from the actual value in the historical data using the least squares method (Dreger et al., 2014, pp. 24–25). The regression coefficients can be recalculated at fixed intervals (e.g., annually), but it is also possible to perform a continuous, quasirolling calculation based on the information from the last N periods.

V.2.5 Evaluation of forecasting models

The coefficient of determination (R^2) of a linear regression model measures the goodness of fit of the model to given (training) data. However, it is less suitable for the evaluation of forecasting problems, especially since there is no comparable measure for time series forecasts. The goal of a forecasting model is generally to have good generalizability for new data (Domingos, 2012, 80). Only the ex-post comparison of the forecasted values with the actual values from an evaluation or test data set can objectively assess the forecasting quality, assuming the model was specified with training data (Hansmann, 1983, p. 14).
Such an ex-post forecast error can be calculated in several ways, with the metrics differing in terms of calculation, assessment of the error measure, and interpretation possibilities. When comparing different models with minor differences, it may be useful to consider multiple error measures to incorporate heterogeneous perspectives (e.g., different weighting of outliers). Table V.1 provides an overview of some common evaluation metrics suitable for assessing numerical forecasting models, but it does not claim to be exhaustive.

Formula	**Description**
$AE = \sum_{i=1}^{N} \lvert \hat{x}_i - x_i \rvert$	Absolute Error
$MAE = \frac{1}{N} \sum_{i=1}^{N} \lvert \hat{x}_i - x_i \rvert$	Mean Absolute Error
$MSE = \frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2$	Mean Squared Error
$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2}$	Root Mean Squared Error
$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left\lvert \frac{\hat{x}_i - x_i}{x_i} \right\rvert$	Mean Absolute Percentage Error

Table V.1. Selected assessment measures for forecasting (Hansmann, 1983, pp. 15–16; Küsters, 2012, p. 434; Petersohn, 2009, pp. 171–172).

The absolute error (AE) is generally applicable only when all forecasts are based on the same number of observations. Since it is not standardized, it is not relatively interpretable (Barrot, 2009, pp. 548–550; Petersohn, 2009, pp. 171–172). When forecasting a binary dependent variable (0/1), there is no difference between the mean absolute error (MAE) and the mean squared error (MSE), while the mean absolute percentage error (MAPE) is not applicable due to a possible division by zero.

The advantages of the MAE and the root mean squared error (RMSE) for numerical values lie in their simple interpretability when compared to the predicted value. Conversely, the MSE, due to the squared result that includes the unit, can hardly be meaningfully related to the original value. The MAPE is the easiest to interpret due to its relative perspective, but, similar to the MAE, its sensitivity to outliers is lower compared to the RMSE (Küsters, 2012, pp. 433–434).

V.3 Data mining models for the prediction of return decisions

Chapter V.3 covers the generation and evaluation of various data mining models that aim to predict whether a return will occur for each individual item or package shipped or if the item or package will remain with the customer. Section V.3.1 introduces the data set used as the basis for building the models. Section V.3.2 explains the modeling approach, while Sections V.3.3 and V.3.4 assess the achievable prediction accuracy for the item and package levels, respectively.

V.3.1 Dataset details

The data used for the study consist of historical transaction data over one year (n = 481,092) from a real fashion and footwear online store. The observed retailer is a medium-sized business (Commission Recommendation of 6 May 2003 concerning the definition of micro, small, and medium-sized enterprises, 2003) with an annual turnover of €13.2 million during the observation period, with returns amounting to €17.8 million during the same period. The return rates during this time are above 50% for all metrics (α-RR: 59.85 %; β-RR: 52.55 %; γ-RR: 57.38 %). There is a clear relationship between orders and the number of returns (Figure V.1). Over the course of the year, a relatively constant number of orders and returns can be observed with a weak trend. The monthly β-return rate fluctuates only slightly between 50 % and 55 % (right scale).

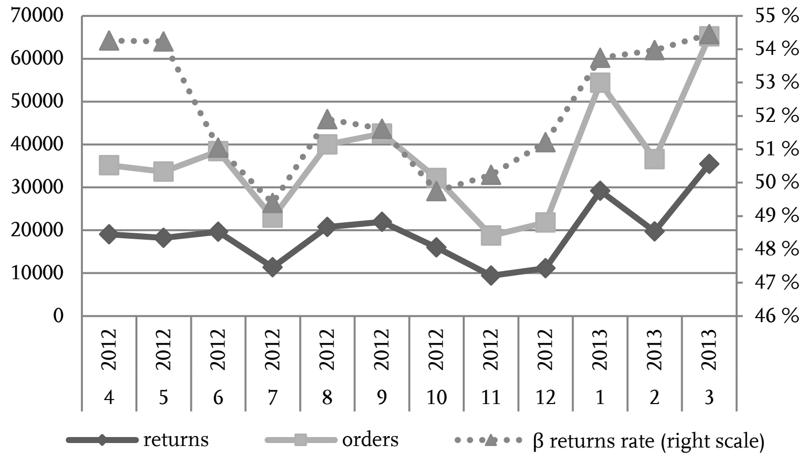

Figure V.1. History of orders, returns, and beta return rate in the transaction data.

The data set contains information about the transaction itself (sales and shipping date, ID), the sold product (ID, size, color, manufacturer, price), customer details (ID, gender/title, date of birth, state, customer account creation date), and whether the customer returned the respective product. In addition to the mentioned data, the transactions of the immediately following month (04/2013; n = 45,810) are used as *test data*, where the return decision is initially assumed to be unknown, which serves the evaluation of the different models. The preceding one-year transaction data are used for training the models and is henceforth referred to as *training data*. A total of 93.5 % of the customers in the analyzed data set are female. Table V.2 provides further characterization of the customer base and the items shipped by the retailer.

Variable	N	Mean	SD	Min.	Max.
Age [years] (without missings)	53,549	47.80	9.94	0	112
Lead time [days]	482,283	11.25	18.343	0	175
Product price [Euro]	531,170	70.51	44.82	0.00	999.00
Account age [days]	531,170	476.46	273.661	1	805

Table V.2. Measures of central tendency and of spread for various relevant variables for test and training data set together (13 months).

The original data contain missing, inconsistent, or implausible values. For example, an order placed by a 112-year-old customer is highly unlikely. Therefore, comprehensive data preprocessing is necessary. Furthermore, the existing data have been enhanced through data aggregation (e.g., identification of selection orders) and integration of data from external sources (such as population density of the customers' originating states) to enrich and expand the available information.

V.3.2 Approach to model specification for return decision forecast

In the first step, it is necessary to filter out all transactions with missing shipping or delivery dates. These items cannot be returned because they never reached the customer.

The approach for model specification follows these systematics at the article and package levels (Figure V.2): Initially, the *base model* includes only the continuous (metric) variables contained in the data set. Together with additional explanatory or control variables, they form an *extended model*. The *history model* exclusively uses historical return rates of different reference variables (article, manufacturer, color, customer, state) but ignores information about prices, selection orders, et cetera. A comprehensive model ("*full model*") combines all influencing factors available in the data, including additional factors generated during data preprocessing.

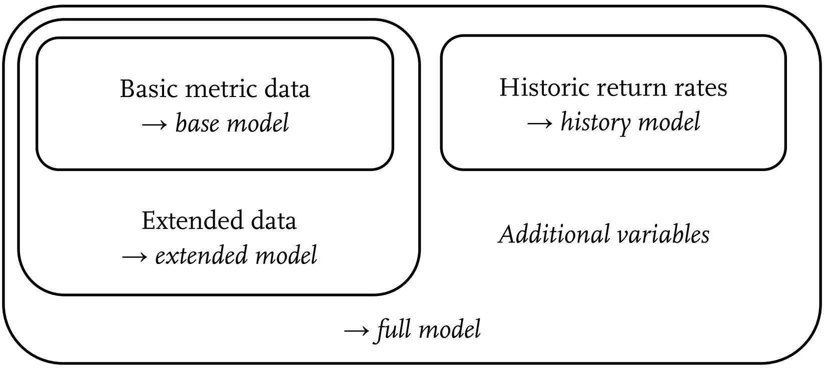

Figure V.2. Schematic assignment of different variables to the models.

Different algorithms such as logistic regression, neural networks, C5.0 decision trees, discriminant analysis, and ensemble models, as well as various preprocessing steps are used to create multiple models for each of the described data foundations. The models' precision is evaluated using the test data.

V.3.3 Models for the return decision forecast on article level

At the article level, the comprehensive full model encompassing all available input variables enables the best prediction: It accurately forecasts two-thirds of all returns (Table V.3). The elimination of weak or implausible variables can further enhance this overall model, resulting in a relatively high accuracy with a forecast error of 14,987 misclassified transactions (out of 45,810).
The greatest improvement is observed between the base model and the extended model, indicating the significant role played by the identification of selection orders within the extended model. The performance of the full model stands out distinctly from the history model or the extended model. Therefore, in pursuit of high prediction accuracy, it is worthwhile to generate and incorporate additional variables.
A careful iterative examination of a variable's relative influence is also essential, as demonstrated by the improvements of a refined full model through variable elimination. Although the refined full model exhibits poorer fit to the training data, it can better predict unknown data, indicating higher generalization capability.
The performance of the history model highlights that minimal effort and no consideration of order-specific information still allow an acceptable forecast by averaging over a large data set.

Data basis	Model	Absolute (relative) error
Base model	C5.0	17,869 (39.01 %)
Extended model	NN	16,682 (36.42 %)
History model	NN	16,080 (35.10 %)
Full model	NN	15,274 (33.34 %)
Refined full model	DA, NN	14,987 (32.71 %)

Table V.3. Comparison of model precision of article-based classification models (n=45,810 articles).

V.3.4 Models for return decision forecasting at the package level

In addition to article-level predictions, forecasting at the package level is of particular interest for logistical purposes. Package-specific information such as delivery time, package size, and package value are used as exogenous variables. Article-specific return rates incorporate additional variable details about the articles contained in the package into the package models. Similar to the article level, the package models employ a progressively expanded set of predictors to compare their influence in the overall evaluation and to assess the models based on an appropriate balance between complexity and goodness of fit.
An additional alternative article based model aggregates the information from the best article model for the package level and utilizes its probabilities for the respective decision on an article basis (return or not) to determine the return probability of a package:

$$p_j(package) = 1 - \prod_{i=1}^{N} (1 - p_i(article))$$

Subsequently, the model classifies packages with a calculated return probability greater than 50% as expected returns. In comparison (Table V.4), this article-based alternative model performs the best and also enables a slightly better prediction than at the article level. Generally, the prediction is more precise at the package level, which is related to the typically higher α-return rate: For example, a naïve model that classifies all units as returns would make an error of approximately 40 % for packages, while it would misassign 48 % of the articles.

Data basis	**Algorithm**	**Absolute (relative) error**
Base model	Ensemble	7,215 (32.69 %)
Extended model	Ensemble	7,004 (31.74 %)
History model	C5.0	6,897 (31.25 %)
Full model	Ensemble	6,672 (30.23 %)
Refined full model	Ensemble	6,544 (29.65 %)
Article-based model	Alternative model	6,195 (28.07 %)

Table V.4. Model precision comparison of packet-based classification models (n=22,069 packets).

The ensemble models, which consist of combined models from different method classes, exhibit the best and most stable forecasts. The differences between the various databases are relatively small at the package level. Therefore, from a practical perspective, it is advisable to either rely on less resource-intensive history models or use the detailed modeling performed at the article level directly for package prediction using the alternative model.

V.4 Evaluation of return volume forecasts

Section V.4.1 explains an additional naïve forecasting model for the total volume of returned articles that should be evaluated afterward. Subsequently, Sections V.4.2 and V.4.3 evaluate the weekly forecasting accuracy of different techniques for both article and package total volume.
To verify the usefulness of the created data mining models for logistical planning processes, Section V.4.4 compares the forecast for the expected returns volume over a one-month period based on the most accurate data mining models from Section V.3.2 with the techniques used in Sections V.4.2 and V.4.3. A subsequent analysis at the individual product level in Section V.4.4 completes the analysis of return volume forecasts.

V.4.1 Naïve return volume forecast

In addition to the techniques explained in Section V.2.4, a third calculation methodology is employed. This model takes into account the mentioned causal relationship – in other words, the dependence of returns on current sales figures – and, therefore, exhibits similarities to the regression model mentioned above. The multiplication of the volume ordered by customers within a reference period by the historical return rate can be formally represented as a naïve estimation (Toktay, 2001, p. 4) as follows:

$$N_{units\ ordered} \cdot RR_{historic} = N_{units\ returned}$$

In a sense, it can be interpreted as regression model without a constant, in which the historical return rate serves as the slope parameter. The temporal reference of this return rate is variable and can, for example, refer to the most recent period, an exponentially smoothed value of past periods, or a constant time frame, such as the return rate of the previous fiscal year.

V.4.2 Article-based return volume forecast

The investigation now focuses on the number of returned articles, with one week corresponding to one period. The explained quantitative methods are applied to the transaction data of the shoe and fashion retailer and are evaluated based on the error metrics MAPE, MAE, and RMSE. The evaluation (Table V.5) implicitly assumes that forecasts are generated at the end of the period and that returns occur in the period following the order.

Model	Parameter	MAPE	MAE	RMSE
Moving Avereage (MA)	N = 1 period	29.87 %	1,397	2,040
	N = 4 periods	36.36 %	1,575	2,260
	N = 8 periods	33.81 %	1,403	2,079
Single Exponential Smoothing (SES)	$\alpha = 0.2$	32.76 %	1,366	2,062
	$\alpha = 0.5$	31.36 %	1,378	2,032
	$\alpha = 0.8$	30.44 %	1,385	2,031
Holt	$\alpha = 0.5;\ \beta = 0.2$	35.07 %	1,548	2,209
	$\alpha = 0.5;\ \beta = 0.5$	40.15 %	1,781	2,499
	$\alpha = 0.5;\ \beta = 0.8$	45.94 %	2,092	2,810
Bivariate linear regression	periods 1–8	6.04 %	213	260
	rolling, N = 4 periods	3.69 %	143	187
	rolling, N = 8 periods	3.37 %	123	182
Naïve model	$\alpha = 0.2$	2.60 %	118	186
	$\alpha = 0.5$	2.36 %	101	152
	$\alpha = 0.8$	2.41 %	102	155
	$\alpha = 1$ (previous RR value)	2.50 %	106	162

Table V.5. Comparison of returned article volume forecasts. Quality measures calculated for 49 periods starting from period 9 (total number of returns in this period: 221,968, average: 4,530 returns/period).

Since the forecast accuracy of simple exponential smoothing improves with increasing α and achieves the best value for α = 1, the assumption of a constant model needs to be questioned: Such behavior suggests the existence of a trend or seasonality (Schröder, 2012, p. 30). For this reason, the double exponential smoothing method, known as Holt's method, is

additionally employed, which takes into account a possible trend (Holt, 2004, pp. 5–10; Schröder, 2012, pp. 41–42). However, according to the results, no such trend is present, since the forecasts deteriorate for double exponential smoothing, and the parameter $\beta = 0$ yields the best forecasts for this method, which aligns with simple exponential smoothing. Seasonal factors could be calculated only with a multi-year database.
The moving average over one period or exponential smoothing with a smoothing parameter $\alpha = 1$ corresponds to the value of the previous period as the forecast value, which still represents the best forecast within the time series forecasts but exhibits an average error (MAPE) of nearly 30 %. Generally, due to the lack of consideration for current sales figures, time series forecasts are not well suited for forecasting the return volume.
The causal methods perform much better, with the naïve method of linear regression being preferable due to its simplicity and better forecast performance. Exponential smoothing of the historic return rate with a moderate to high parameter α can yield the best possible results depending on the volatility of the underlying data. However, in the data used for this study, using the return rate from the previous period ($\alpha = 1$) already provides a very low forecast deviation of only 2.60 %.

V.4.3 Package-related return volume forecast

The primary benefit of volume forecasting at the enterprise level lies in predicting the number of packages to be processed by the company, as these data, for example, supports operational workforce planning. The results in Table V.6 confirm the findings from the article volume forecast. Once again, the naïve model estimates the volume most precisely, which holds for the package level. Bivariate linear regression yields usable results, but due to the higher complexity, the naïve model is preferable without smoothing or with simple exponential smoothing using a moderate to high smoothing parameter α. Time series methods perform approximately an order of magnitude worse.

Model	Parameter	MAPE	MAE	RMSE
MA	N = 1 period	30.70 %	714	1,040
	N = 4 periods	37.39 %	814	1,174
	N = 8 periods	36.18 %	759	1,093
SES	$\alpha = 0.2$	35.06 %	728	1,074
	$\alpha = 0.5$	32.73 %	713	1,050
	$\alpha = 0.8$	31.37 %	706	1,041
Holt	$\alpha = 0.5;\ \beta = 0.2$	35.80 %	801	1,145
	$\alpha = 0.5;\ \beta = 0.5$	41.59 %	917	1,293
	$\alpha = 0.5;\ \beta = 0.8$	47.89 %	1,077	1,447
Bivariate linear regression	periods 1–8	5.31 %	113	140
	rolling, N = 4 periods	4.50 %	86	117
	rolling, N = 8 periods	3.48 %	71	95
Naïve model	$\alpha = 0.2$	2.70 %	56	72
	$\alpha = 0.5$	2.48 %	50	65
	$\alpha = 0.8$	2.39 %	48	65
	$\alpha = 1$ (previous RR value)	2.45 %	51	67

Table V.6. Comparison of returned package volume forecasts. Quality measures calculated for 49 periods starting from period 9 (total number of returns in this period: 114,921, average: 2,345 returns/period).

V.4.4 Comparison of the naïve volume forecast with the volume forecast of the data mining model

Since the data mining model is based on the complete training data (April 2012 – March 2013), an objective evaluation with unknown data is possible only for the month of the test data set (April 2013). The comparison is conducted for the total number of returns in that month, and the previously described naïve model serves as reference model.

	25,029 articles returned		*13,166 packages returned*	
Model	**Forecast**	**Percentage error**	**Forecast**	**Percentage error**
Data mining model	28,671	15.08 %	14,799	12.40 %
Naïve model	24,072	3.82 %	13,208	0.32 %

Table V.7. Forecast comparison of the data mining model and the naïve model for the total number of returns in one month (test data: April 2013).

The naïve model for articles in Table V.7 is based on the historical return rate averaged over the previous fiscal year. By using the return rate of the previous month instead of the entire year, the deviation is only 0.34 %,

but it is more susceptible to short-term fluctuations in this rate. Exponential smoothing of this return rate with a smoothing parameter $\alpha = 0.5$ results in a prediction error of 1.55 % and should represent a satisfactory compromise in terms of responsiveness to changes in return behavior.
At the package level, the forecast is slightly more precise than at the article level: With 13,166 actually returned packages in the test data, the error is only 0.32 %, while using the α-return rate of the previous month increases it to 0.61 %. When using a monthly exponentially smoothed return rate with a smoothing parameter of $\alpha = 0.5$, the error is only 0.71 %, and for $\alpha = 0.25$, it is 0.07 %.
Compared to the best data mining model, the naïve model deviates significantly less from the true value and, therefore, is better suited for the overall quantity forecast (see Table V.7).

V.4.5 Forecasting the returns volume at the product level

At the product level, the forecast of the expected return quantity holds the highest practical value, as these numbers can be directly used for order quantity and inventory planning. Forecasting returns for items with low sales and return volumes is likely to be difficult, which is why the analysis covers the entire month of the test data to include an adequate number of units. Nevertheless, this section also demonstrates the extent to which low-selling and low-return items can be predictable.
The naïve volume forecast is calculated by multiplying the product-specific β-return rate aggregated over one year (corresponding to the training data) with the sales figures of the following month (i.e., from the test data). The data mining-based forecast is derived from the most accurate article model, in other words, the refined full model, for which the sum of predicted returns was calculated for each product.
Figure V.3 illustrates the higher forecast accuracy of the naïve model and the relationship between the forecast accuracy measured by the MAPE and the number of returned items. The return volume of rarely purchased items is more difficult to predict and, therefore, can be forecasted only over longer periods compared to top sellers. In this unweighted analysis of the articles, multiple iterations progressively limit the number of considered products based on their order frequency: For example, during the analysis period, 575 different products were returned more than five

times, while only 256 products caused more than 25 returns. Items without returns are not considered, as their deviation from the true value (zero) cannot be determined as a percentage.

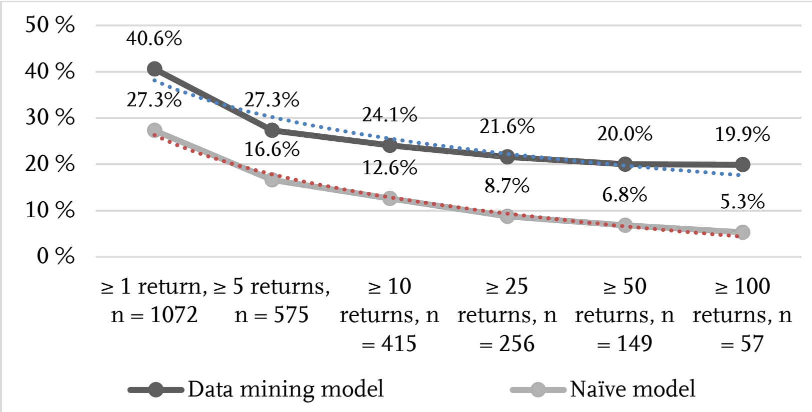

Figure V.3. Mean absolute percentage error (MAPE) of the product-specific volume forecast as a function of the minimum number of returned articles. Data basis: 45,810 transactions in one month (45,378 transactions for products with at least one return).

With each restriction on the number of examined products, gradually filtering out the rarely returned items, the forecast of the returned quantity improves as expected. The naïve model predicts the return volume more accurate than the data mining model for products with more than five returns. The average deviation of the naïve model is 16.6 %, compared to the data mining model restricted to the products with more than 100 returns (19.9 % average forecast deviation).

Finally, weighting all products with at least one return by their respective sales volume and subsequently calculating the normalized relative deviation confirms this result. An average percentage error of 13.0 % (SD = 0.31) for the naïve forecast compares to 26.6 % (SD = 0.34) for the most accurate data mining model.

V.5 Conclusion

A fundamental limitation of this study is the fact that only data from a single online shoe and fashion retailer with a rather specific customer base (93.5 % female customers of higher age) were considered. Therefore, the findings can only be partially extrapolated to other online retailers or even other industries. Consequently, the results should be critically examined in the future using additional real-world data sets or through simulation. Additionally, such a study could also consider return timings, which could not be closely examined in the present analysis due to data limitations.
Referring to the research questions, the results can be summarized as follows:

- *How accurately can retailers predict a customer's return decision based on transaction data using data mining techniques?*

The best prediction models misclassify 14,987 (32.71 %) out of 45,810 transactions. Taking into account the previously filtered-out 4,268 canceled items, the overall precision in the test data set is approximately 70 %. A package-level model based on these item-level models can correctly assign 15,874 (71.93 %) out of 22,069 packages. Thus, aggregating data from the item level to the package level slightly improves the forecasting accuracy.

- *Which techniques are best suited for forecasting return volumes?*

The choice of the appropriate forecasting method depends on usually complementary criteria (particularly "accuracy, cost, and complexity"; Hansmann, 1983, p. 141), requiring careful prioritization (Hansmann, 1983, pp. 141–142). This study demonstrates that this trade-off does not always hold true, as the naïve model for quantity forecasting based on historical return rates and current order quantity outperforms more complex regression models. Time series methods, in general, are less suitable for forecasting return quantities.

- *To what extent do data mining models used for return decision-making enable the derivation of accurate return quantity forecasts?*

In the overall analysis, the data mining models used to forecast return volume, which were actually created to predict return decisions, are not well-suited for estimating the total return quantity of the merchant or product-specific return quantities. However, in a future study, alternative numerical data mining models specifically designed for these purposes could be generated, which may achieve a level of precision similar to that of the naïve volume forecast.

The naïve forecasting tool presented in this paper is recommended for practical use in quantity forecasting due to its low deviations and intuitive handling. In the next step, for early return estimation, it is conceivable to link it with product-specific demand forecasts instead of order quantities.

References

Asdecker, B. (2011). Retourenmanagement – Eine Literaturrecherche. In E. Sucky, B. Asdecker, A. Dobhan, S. Haas, & J. Wiese (Eds.), *Logistikmanagement. Herausforderungen, Chancen und Lösungen. Band II: Proceedings of the 7th Conference "Logistikmanagement"* (Vol. 2, pp. 421–461). University of Bamberg Press.

Asdecker, B. (2012). Retourenmanagement. In P. Klaus, W. Krieger, & M. Krupp (Eds.), *Gabler Lexikon Logistik* (5th ed., p. 495). Gabler.

Asdecker, B. (2014). *Retourenmanagement im Versandhandel: Theoretische und empirisch fundierte Gestaltungsalternativen für das Management von Retouren. Logistik und Supply Chain Management: Bd. 10.* Univ. of Bamberg Press.

Barrot, C. (2009). Prognosegütemaße. In S. Albers, D. Klapper, U. Konradt, A. Walter, & J. Wolf (Eds.), *Methodik der empirischen Forschung* (3rd ed., pp. 547–560). Springer.

Cleve, J., & Lämmel, U. (2014). *Data Mining. De Gruyter eBook-Paket Informatik.* De Gruyter. https://doi.org/10.1524/9783486720341

Crone, S. F. (2010). *Neuronale Netze zur Prognose und Disposition im Handel* (1. Aufl.). *Gabler Research : Betriebliche Forschung zur Unternehmensführung.* Gabler.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78–87.

Dreger, C., Eckey, H. F., & Kosfeld, R. (2014). *Ökonometrie: Grundlagen - Methoden - Beispiele* (5th ed.). *Lehrbuch.* Gabler Verlag.

Directive 2011/83/EU on consumer rights (2011). https://eur-lex.europa.eu/eli/dir/2011/83/2022-05-28

Commission Recommendation of 6 May 2003 concerning the definition of micro, small, and medium-sized enterprises, May 20, 2003. https://eur-lex.europa.eu/eli/reco/2003/361

Feindt, M., & Sinn, M. (2014). *Blue Yonder Big Data Analytics / Big Data Analytics – wie automatisierte Entscheidungen OTTO entlang des gesamten Produktlebenszyklus unterstützen*. Retrieved March 3, 2015, from http://www.bitkom-bigdata.de/sites/default/files/BDS2014_1400%20FeindtProf%20BlueYonder_Sinn%20Otto.pdf

Fidan, N., & Ozkok, B. A. (2013). A Review on Applied Data Mining Techniques to Stock Market Prediction. In P. J. Papajorgji (Ed.), *Gale virtual reference library. Enterprise business modeling, optimization techniques, and flexible information systems* (pp. 108–126). Business Science Reference.

Fromm, S. (2005). *Binäre logistische Regressionsanalyse : eine Einführung für Sozialwissenschaftler mit SPSS für Windows. Bamberger Beiträge zur empirischen Sozialforschung: Vol. 11.*

Gartner. (2023). *Information Technology Glossary: Data Mining*. Retrieved June 12, 2023, from https://www.gartner.com/en/information-technology/glossary/data-mining

Garvin, D. A. (1993). Building a learning organization. *Harvard Business Review, 71*(4), 78–91.

Handelsverband Deutschland. (2023a). *Anteile der Branchen am Onlineumsatz*. Retrieved June 9, 2023, from https://einzelhandel.de/index.php?option=com_content&view=article&id=4133

Handelsverband Deutschland. (2023b). *E-Commerce-Umsätze*. Retrieved June 9, 2023, from https://einzelhandel.de/index.php?option=com_content&view=article&id=1889

Hansmann, K.-W. (1983). *Kurzlehrbuch Prognoseverfahren: Mit Aufgaben und Lösungen. Gabler Lehrbuch*. Gabler.

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting, 20*(1), 5–10.

Jahn, M. (2013). *Online versus stationär? Lieber komplementär*. Retrieved February 5, 2015, from http://www.gfk.com/de/loesungen/geomarketing/Seiten/Landingpage/White-Paper-Handel.aspx

Jarke, M. (2014). Interview mit Michael Feindt zum Thema „Präskriptive Big-Data-Analyse". *WIRTSCHAFTSINFORMATIK, 56*(5), 329–331.

Kuß, A., Wildner, R., & Kreis, H. (2014). *Marktforschung: Grundlagen der Datenerhebung und Datenanalyse* (5th ed.). Springer Gabler. https://doi.org/10.1007/978-3-658-01864-1

Küsters, U. (2001). Data Mining Methoden: Einordnung und Überblick. In H. Hippner & K. Wilde (Eds.), *Business Computing. Handbuch Data Mining im Marketing: Knowledge Discovery in Marketing Databases* (1st ed., pp. 95–130). Vieweg.

Küsters, U. (2012). Evaluation, Kombination und Auswahl betriebswirtschaftlicher Prognoseverfahren. In P. Mertens & S. Rässler (Eds.), *Prognoserechnung* (pp. 423–467). Physica-Verlag HD.

Loebich, M. (2014). *Überbestände und Retourenwahnsinn machen Onlinehandel unprofitabel*. BearingPoint. Retrieved October 1, 2015, from https://www.bearingpoint.com/de-de/7-9697/ueberbestaende-und-retourenwahnsinn-machen-onlinehandel-unprofitabel/

Petersohn, H. (2009). *Data Mining: Verfahren, Prozesse, Anwendungsarchitektur*. Oldenbourg. https://doi.org/10.1524/9783486593334

Ratner, B. (2012). *Statistical and machine-learning data mining: Techniques for better predictive modeling and analysis of big data* (2nd ed.). CRC Press.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Rosentreter, J. (1977). *Prognosen in der industriellen Planung: Eine Analyse von Prognoseverfahren und Anwendungsproblemen unter besonderer Berücksichtigung von Verfahren auf der Basis der exponentiellen Glättung. Reihe Wirtschaftswissenschaften: Vol. 127*. Deutsch.

Schröder, M. (2012). Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren. In P. Mertens & S. Rässler (Eds.), *Prognoserechnung* (pp. 11–45). Physica-Verlag HD.

Toktay, L. B. (2001). *Forecasting Product Returns*. INSEAD working paper. http://www.prism.gatech.edu/~bt71/articles/forecasting.pdf

Weber, J., & Wallenburg, C. M. (2010). *Logistik- und Supply Chain Controlling* (6. vollständig überarbeitete Auflage). Schäffer-Poeschel Verlag.

Wenzel, E., & Dziemba, O. (2014). *#Wir: Wie die Digitalisierung unseren Alltag verändert* (1.th ed.). Redline-Verl.

Werner, H. (2013). *Supply Chain Management.* Springer. https://doi.org/10.1007/978-3-8349-3769-8

VI Big data analytics in returns management – Are complex techniques necessary to forecast consumer returns properly?

Authors: Björn Asdecker, David Karl

Publication Outlet: Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics

DOI: 10.4995/CARMA2018.2018.8303

published as:

Asdecker, Björn; Karl, David (2018): Big data analytics in returns management – Are complex techniques necessary to forecast consumer returns properly? In: *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics*, Valencia, Spain, pp. 39-46. DOI: 10.4995/CARMA2018.2018.8303

Abstract

The more people shop online, the more consumer returns e-tailers face. In order to plan the returns management process capacity adequately, it is necessary to forecast the expected amount of returned parcels. Big data analytics provides a vast number of methods to perform such tasks. However, it should be noted that particularly small- and medium-sized e-tailers lack the capabilities and resources to employ such complex techniques. Against this background, this paper analyzes the performance of several data analysis methods that differ in application complexitiy using real data from an apparel e-tailer. On the one hand, we find that –as expected– complex methods outperform simple ones. On the other hand, and from a practitioner's perspective probably even more interesting, we also conclude that a binary logistical regression as the simplest analyzed method may already provide satisfactory results. The findings indicate that the use of big data analytics is of great value to effectively and efficiently manage consumer returns – even if not the most sophisticated state-of-the-art method is used.

Keywords: Returns management, product returns, e-commerce, forecast models

VI.1 Introduction and motivation

In today's retailing world, more people shop online more frequently. Consequently, e-commerce revenues have been skyrocketing in the last two decades and an end to this development is not foreseeable. In 2015, for instance, US consumers spent more than $340.41 billion online; up from $4.98 billion in 1998 (United States Census Bureau, 2018). While e-tailing clearly offers numerous advantages over traditional brick-and-mortar retailing, there is one major disadvantage: supply and demand are geographically separated. Therefore, consumers are unable to see, feel, and test the products before purchase. In other words, they can not try before they buy, which almost inevitably leads to consumer returns.

From a business perspective consumer returns are a major cost driver and pose a serious threat on an e-tailer's profitability. According to Stock et al. (2006), expenditures can be as high as $30–35 per return, while return rates may well exceed 50 % for fashion items (Asdecker et al., 2017). These already staggering numbers still reflect only part of the problem. In addition to the direct costs there are indirect effects that influence customer value.

Existing research shows that the returns process is part of the post purchase-experience and herein influences customer satisfaction and retention (Petersen & Kumar, 2009). Laseter and Rabinovich (2012) argue that the improvement of the product return experience is based on the following three principles: (1) lower customer efforts to return the product, (2) offer customized solutions that fit the customers' needs, and (3) exceed customer expectations when processing returns. The third principle specifically refers to the desired outcome (e.g., compensation of made payments), ease of contact, and recovery responsiveness (Mollenkopf et al., 2007). The latter can be operationalized with the time necessary to process a return. To speed returns up, operations require the most accurate capacity planning, which, in turn, is based on forecasts. The importance of this task cannot be underestimated: The better the forecasts are, the more effective and efficient will cosumer returns be processed.

Literature provides various econometric, statistical and/or data mining methods that can be employed to predict returns. Some are higly complex while others are more straightforward and easier to apply. In a world where many decision makers strive for the one optimal alternative, complex state-of-the-art methods seem to be the best choice. However, they also demand sophisticated skills, additional capabilities and financial resources, which are confined, particularly in small- and medium-sized e-tailing companies (Coleman et al., 2016). In a challenging paper, Banks (1993, p. 360) concludes: “I would guess that intelligent use of simple tools will achieve 95 % of the knowledge that could be obtained through more sophisticated techniques, at much smaller cost. Also the simple tools can be applied more quickly to all problems, whereas the complex tools are unlikely to be ubiquitously used.” While the paper was written in a different era the general message remains. Against this background, we address the following research-leading question:

- *How are simple data analysis methods performing compared to complex, more sophisticated ones when predicting consumer returns?*

Unlike other publications that try to identify factors that influence the likelihood of consumer returns (e.g., Asdecker et al., 2017; Smriti, 2018; Toktay et al., 2004), this paper compares the performance of different forecasting methods. To the authors’ best knowledge, the publication that is the closest to the one at hand originates from Urbanke et al. (2015). They developed a decision support system that identifies transactions with a high likelihood to return before the actual sale takes place and demonstrated the approach’s applicability using a large dataset from a German fashion e-tailer. Within their study they compare seven forecasting techniques, namely the principal component analysis, linear discriminant analysis, randomized truncated singular value decomposition, feature selection based on univariate chi-squared statistic, random projection, non-negative matrix factorization, and a specific feature extraction technique that ignores nominal indicators. While they search for the technique with maximum precision, they do not compare simple with complex approaches.

The remaining article will provide an overview of different forecasting approaches, followed by a report on the performance of the previously introduced techniques. Finally, we conclude with a summary and an outlook on future research.

VI.2 Theoretical background and return forecasting techniques

The data science and statistics literature provides a variety of different methods or techniques that can be used to to predict consumer returns. Since consumers decide to either return or keep the delivered items the dependent variable is binary in nature. This article considers five approaches that are briefly described as follows.

VI.2.1 Binary logistic regression

The binary logistic regression is the simplest method to be taken into account. It is an extension of the linear regression, where the dependent variable is binary (1=return, 0=keep). The independent variables can be either continuous (interval/ratio) or categorial (ordinal/nominal) in nature. For each observation, the binary logistic determines the probability that the dependent variable takes value “1” (Hastie et al., 2009).

VI.2.2 Linear discriminant function analysis

The linear discriminant function analysis, which was first performed by Fisher (1936), shares great similarities with the logistic regression. The analysis requires at least two a priori known groups to which observations shall be assigned. The basic idea is to create a linear combination of independent variables, which best classifies the available data. Thereby, it determines a score for each observation which is then compared to a critical discriminant score in order to carry out the classification (return vs. keep).

VI.2.3 Artifical neuronal network: multiylayer perceptron

Artificial neuronal networks are based on a set of connected nodes, the so called artificial neurons, which are organized in layers. Connected artificial neurons can exchange signals with each other. The receiving artificial neuron will then process it and, in turn, signal artificial neurons connected to it. The ultimate goal is to find a function that best assigns input data to the correct output. To achieve this goal in the returns management context, this study uses multilayer perceptrons, a class of feedforward neural networks. Herein, the information flows exclusively from the input layer through hidden layers with a certain amount of units to the output layer with no feedback flow. Training is done through backpropagation, which is a supervised learning technique that compares the outputs of the network with the known actual values (Hastie et al., 2009).

VI.2.4 Decision tree learning: C5.0 algorithm

Decision trees are hierarchical structures of branches, representing conjunctions of certain characterisics, and leafs, representing class labels. The goal of this technique is to create a decision tree that best classifies the available observation. For this purpose many decision tree algorithms have been presented. This analysis refers to the C5.0 algorithm. C5.0 is the faster, more efficient successor of the widely-employed C4.5 algorithm (Pandya & Pandya, 2015).

VI.2.5 Ensemble learning technique

The ensemble learning method uses several algorithms to improve predictive performance. It determines the result of every single algorithm and interprets it as a hypothesis for the final verdict (Polikar, 2006). In this study, we used the training data to determine the three most accurate techniques from the following selection: artificial neuronal network, different decision trees (C5.0, QUEST, CART, CHAID), binary logistic regression, linear discriminant analysis, Bayesian network, and nearest neighbor. This resulted in three hypotheses that were given a vote proportional to the confidence, which equals the probability that the postulated hypothesis of a specific algorithm is accurate.

With regard to the required expertise and software, the binary logistic regression and the two-group discriminant analysis are quite straightforward and implemented in common statistical programs with simple and intuitive user interfaces. The remaining three are more complex and therefore require more data mining knowledge as well as less intuitive software (e.g., R, Python, MATLAB, or specific data mining tools such as the IBM SPSS Modeler).

VI.3 Comparison of the different forecasting techniques

The introduced prediction techniques will be tested using real data from a medium-sized German fashion e-tailer that requested confidentiality concerning its name. We will first describe the characteristics of the provided data and then determine the predictive performance of the previously introducted techniques.

VI.3.1 The dataset

The dataset contains shipment and returns information from April 2012 to April 2013. More up-to-date data would be desirable but is unrealistic because many merchants consider returns data as proprietary and are unwilling to share this kind of information. During the analyzed period, the e-tailer sent 220,474 shipments and received 131,907 returns, which equals an α-returns rate of 59.8 % (Asdecker, 2015). While the returns rate might change over time due to different customer behavior or successful avoidance strategies, it should be noted that data age is irrelevant to this type of comparative investigation. The e-tailer shared the following information:

- Package ID: unique ID of the shipment
- Price: Total value of shipped goods in Euro
- Number of articles: total number of articles in the shipment
- Delivery time: time between order placement and delivery in days
- Customer type: categorial variable that indicated the customer type (1=female, 2=male, 3=family, 4=company, 5=unknown)
- Customer age: customer age at the time of order in years
- Account age: age of the customer account at the time of order in days
- Return: binary variable to indicate whether the enclosed return label had been used (1=return, 0=no return)

VI.3.2 Predictive performance

We divided the shared dataset in two parts. The first twelve months, that is, the data from April 2012 to March 2013, were used to derive the prediction model. The last month, April 2013, is the actual test data. Within that period, the e-tailer sent 22,069 shipments and received 13,166 returns. Consequently the α-returns rate is 59.7 %. The criterion used to assess the quality of the forecast is the total absolute error (TAE), which is the summed absolute difference between the predicted and the actual value. This equals the total number of cases that were mispredicted, which means they were either classified as a return even though nothing was returned or vice versa. All five predicting methods were implemented in IBM SPSS Modeler 18, leading to the following results:

- Binary logistic regression, TAE=7,329 (33.21 %)
- Linear discriminant function analysis, TAE=7,372 (33.40 %)
- Artifical neuronal network I (MLP, one hidden layer, six units/layer), TAE=7,364 (33.37 %)
- Artifical neuronal network II (MLP, two hidden layers, six units/layer), TAE=7,146 (32.38 %)
- Decision tree learning: C5.0 algorithm, TAE=6,973 (31.60 %)
- Ensemble learning technique, TAE=6,963 (31.55 %)

Accordingly, the ensemble learning technique provides the best results and predicts 68.45 % of the analyzed cases correctly. The three techniques derived from the training data and used in the ensemble were C5.0, CHAID, and QUEST, which are all decision trees. This finding was expected because the utilization of multiple algorithms within an ensemble usually provides better predictive results than any of the algorithms separately (Polikar, 2006). More surprising is the good performance of the simpler, less sophisticated techniques. Binary logistic regression and linear discriminant function analysis correctly predict 66.79 % and 66.60 % of cases, respectively. This is equivalent to the artificial neuronal network with one hidden layer (66.63 %) and only slightly worse than the neuronal network with two hidden layers (67.62 %).

On the one hand, this generally shows that data mining can be helpful when it comes to plan the return processes and determining the necessary capacity in the returns department. On the other hand, this study also highlights that it does not always have to be the most sophisticated method to generate an acceptable consumer return forecast. In fact, a cost-benefit analysis might favor simple methods, since more sophisticated and complex methods require more resources and data mining

knowledge. This holds particulary true for the binary logistic regression, which is not only easy to conduct but also allows for a detailed analysis regarding the factors that affect consumer return behavior. Table VI.1 summarizes the SPSS report for the statistically significant binary logistic regression ($p=0.000$, Nagelkerke's $R^2=0.212$) derived from the twelve months training data.
In the model, the probability of a return increases with the total value of the shipped goods, the number of items in a shipment and the account age, whereas it decreases with the delivery time. Packages that are delivered to women have the highest return probability, followed by families, companies, men and unknown recipients. The standardized coefficients b and the Wald statistics show the factor's relative impact: the biggest effect has the price, followed by the number of items and the customer type.

Variable	b	SE	Exp(b) (Odds Ratio)	Wald
Constant*	0.655	0.006	1.926	12,514.985
Price*	1.267	0.010	3.550	15,552.654
Nr. of articles*	0.325	0.006	1.384	2,532.507
Delivery time*	-0.026	0.005	0.975	23.871
Sent to: Mr.[a]*	-0.437	0.026	0.646	272.223
Sent to: Family [a]	-0.139	0.086	0.870	2.597
Sent to: Company[a]	-0.303	0.178	0.739	2.904
Sent to: Unkown [a]*	-0.554	0.190	0.574	8.525
Customer age*	-0.065	0.005	0.937	172.983
Account age*	0.036	0.005	1.036	48.129

Table VI.1. Results of the binary logistic regression model.

VI.4 Summary and outlook

Overall, the good classification performance of the parametric methods (binary logistic regression and linear discriminant analysis) surprises. In fact, their performance is only 1.66/1.85 percentage points worse than the ensemble technique as the best nonparametric method. However, their results are easy to interpret and understand. Therefore, simple models such as the binary logistic regression might be the better choice in business practice, especially for small and medium-sized e-tailers that face

limited data mining capabilities and financial resources. This holds particularly true because they can also be used for the initiation of preventive returns management measures.
This study is based on real data provided by a German e-tailer. Nevertheless, the scope of the analyzed data was very limited. It would be desirable if future studies had access to additional information, e.g., a customer's order and return history, shopping basket composition, to substantiate the presented results. With a larger amount of data, it is very likely that the investigated complex techniques can better exploit their advantages. Moreover, this analysis focused on the prediction of return shipments which is important to plan the reverse logistics process. In further research, it may be of interest to take a closer look inside the shipments to extend the analysis to single articles/items.

References

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Banks, D. (1993). Is Industrial Statistics Out of Control? *Statistical Science, 8*(4).

Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Reis, M. S. (2016). How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International, 32*(6), 2151–2164.

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics, 7*(2), 179–188.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The Elements of Statistical Learning*. Springer New York. https://doi.org/10.1007/978-0-387-84858-7

Laseter, T. M., & Rabinovich, E. (2012). *Internet retail operations: Integrating theory and practice for managers. Supply chain integration modeling, optimization and application: Vol. 5*. CRC Press.

Mollenkopf, D. A., Rabinovich, E., Laseter, T. M., & Boyer, K. K. (2007). Managing Internet Product Returns: A Focus on Effective Service Operations. *Decision Sciences, 38*(2), 215–250.

Pandya, R., & Pandya, J. (2015). C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning. *International Journal of Computer Applications, 117*(16), 18–21.

Petersen, J. A., & Kumar, V. (2009). Are Product Returns a Necessary Evil? Antecedents and Consequences. *Journal of Marketing, 73*(3), 35–51.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine, 6*(3), 21–45.

Smriti, K. (2018). Predicting Online Returns. In J. Wang, A. Kumar, & S. Saurav (Eds.), *Advances in Logistics, Operations, and Management Science. Supply Chain Management Strategies and Risk Assessment in Retail Environments* (pp. 181–194). IGI Global.

Stock, J. R., Speh, T., & Shear, H. (2006). Managing Product Returns for Competitive Advantage. *MIT Sloan Management Review, 48*(1), 57–62.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

United States Census Bureau. (2018). *U.S. Retail Trade Sales - Total and E-commerce (1998-2015)*. http://www2.census.gov/retail/releases/current/arts/ecommerce.xls

Urbanke, P., Kranz, J., & Kolbe, L. (2015). Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction. In *Proceedings of the 14th International Conference on Computer and Information Science,* Fort Worth, United States of America.

Part D – Contributions on Return Drivers and Return Policy

VII Examining Drivers of Consumer Returns in E-Tailing using Real Shop Data

Authors: Björn Asdecker, David Karl, Eric Sucky

Publication Outlet: Proceedings of the 50th Hawaii International Conference on System Sciences
(ranked as B in VHB-Rating 2024)

DOI: 10.24251/HICSS.2017.507

published as:

Asdecker, B.; Karl, D.; Sucky, E. (2017): Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In: *Proceedings of the 50th Hawaii International Conference on System Sciences*, Waikoloa Village, United States of America, pp. 4192–4201. DOI: 10.24251/HICSS.2017.507

Abstract

Returns management – an important component of supply chain management – is a key aspect of online retailers' business models. Despite increasing interest in this issue, few studies have published empirical results on the drivers of consumer returns in e-tailing. Because this knowledge is essential to enabling better decisions about return flows, we explored an extensive dataset from an online apparel retailer using linear and logistic regression models. This approach distinguishes our study from other empirical work, which is usually based on survey methods. Before the data analysis, previously untested hypotheses were formulated using established theories and anecdotal information.

Keywords: Big Data, Consumer Returns, E-Commerce, Retailing, Returns Management

VII.1 Introduction and motivation

The impact of digitization is particularly significant in retailing. In today's retailing world, more and more people shop online. This process can be quantified. In the US, for instance, e-commerce revenues more than quadrupled between 2004 and 2014, increasing from 72.7 to 298.6 billion dollars (United States Census Bureau, 2016). Most of this growth can be attributed to the cannibalization of other distribution channels, particularly traditional brick-and-mortar retailing. This trend will continue for two main reasons:

1. In contrast to traditional stationary merchants, online retailers (known as e-tailers) do not face shelf space limitations. As a result, online businesses can offer an almost unlimited variety of products, which is becoming more important and valuable in a world in which customer needs are increasingly diverse.
2. Historically, two of the major advantages of brick-and-mortar retailing have been immediate product availability and the opportunity to provide customers with assistance from well-trained employees. However, because logistics service providers can now offer same-day services and information technologies can provide unprecedented customer relationship management capabilities, the previous selling advantages of traditional retailers have been eroded.

E-tailing fundamentally influences retailers' business models. First, as mentioned above, distribution channels are changing. Moreover, customer relationships, core activities and cost structures are changing as a result of digitization. While e-tailing clearly offers numerous advantages, there is one major disadvantage: supply and demand are generally geographically separated. Therefore, consumers are unable to touch and try out products before purchase, which leads to higher consumer returns. Thus, returns management – as one of eight major supply chain management processes (Croxton et al., 2001) – is a key activity in e-tailing. It is becoming increasingly important as more people shop online and are more likely to return their purchases.

From a business perspective, consumer returns are a major cost driver and represent a threat to profitability. Stock et al. (2006) estimate such expenditures at $ 30-35 per return. Prior research has shown that returns rates may exceed 60 % for e-commerce fashion retailers (Foscht et al.,

2013). However, it should be noted that e-tailers use different types of returns rates. Asdecker (2015) distinguishes between the α-returns rate and the β-returns rate:
The α-returns rate is the number of returned shipments in relation to the total number of outbound shipments. In contrast, the β-returns rate is the number of returned items in relation to the number of shipped items.

These two key performance indicators may deviate substantially, as the following example illustrates. Imagine three shipments: the first one comprises three articles, of which two are returned; the second one, four articles, of which one is returned; and the third one, a single article that the customer keeps. In this example, the α-returns rate equals 2/3 = 66.7 percent, whereas the β-returns rate equals 3/8 = 37.5 percent.
From a logistical point of view, the α-returns rate is a valuable piece of information for predicting the number of parcels that must be handled. The β-returns rate is more valuable to marketing and sales because it helps decision makers to evaluate distribution success.
Asdecker (2015) presents a circular model of the sales and returns process, which reveals a disproportionate relationship between the rate of returns and the associated costs. For this reason, many e-tailers emphasize preventive returns management measures as a way of reducing the volume of products sent back. Stock and Mulki (2009) note that "[...] the best way of optimizing the product returns process is to not have returns at all [...]".
Avoiding returns before they occur requires knowledge of the drivers affecting the returns rate. These drivers may also help retailers to forecast the volume of returned products, which is essential to planning an efficient physical return process. Toktay et al. (2004) indicate that: "[...] there is little research on identifying factors that significantly influence return flow characteristics. Developing a good understanding [...] would enable better decision making for influencing return flows". Although this call for research was issued more than a decade ago, little progress has been made on this topic. We therefore address the following research-guiding question:

- *Which drivers influence the flow of returns in e-tailing and to what extent?*

The nature of online shopping generates large amounts of data, which e-tailers could profitably analyze using sophisticated big data analytics. To address the research question above, we use actual order and returns data from an online apparel retailer. This approach distinguishes our study from other empirical work, which is usually based on survey methods.
The remainder of this article is structured as follows: the next section provides an overview of the relevant literature and develops the hypotheses to be tested. Then, we introduce the chosen methodology. Next, we present the findings of the empirical models. Finally, we offer a conclusion and an outlook for future research.

VII.2 Selected literature review and development of hypotheses

Interest in the issues of consumer returns and returns management has grown over time, leading to an increase in publications on the subject (Stock & Mulki, 2009). This selective literature review focuses on conceptual/theoretical and empirical contributions, allowing us to deduce the hypotheses to be tested.
E-tailers can manage their customers' expectations by providing detailed product information, high-resolution product photography, videos, and/or balanced customer reviews. However, this addresses only part of the problem. There is also a psychological dimension to returns, which can be explained using the post-purchase dissonance (PPD) theory (Lee, 2015; Powers & Jack, 2013). Festinger (1957) describes cognitive dissonance as an uncomfortable state of mind that one experiences after choosing among a set of alternatives, each of which has both positive and negative attributes. Cognitive dissonance theory implies that individuals try to maintain a consistent set of beliefs. Any deviation from this consistency causes psychological tension, such as anxiety or uncertainty. Inconsistency may originate from the positive attributes of the rejected alternatives or the negative attributes of the chosen alternative. The theory also holds that people attempt to alleviate these negative psychological states. Lee (2015) notes that "[...] consumers increasingly use product returns to cope with PPD over other actions identified in the literature". Thus, with regard to online shopping, PPD theory suggests that consumers tend to

order more products to eliminate mental discomfort resulting from the rejection of alternatives and, hence, return more to reduce dissonant tension resulting from the chosen alternatives. We therefore assume that additionally ordered items have a positive effect on the returns rate.

- *H1.1: There is a positive relationship between the number of additional items and the α-/β-returns rate of an order.*

The argument above is particularly true for so-called "multiple-item orders". In these cases, customers purchase multiple items in different styles, sizes, or colors because they are uncertain about which article will best meet their expectations. In the end, the majority of articles are sent back. This phenomenon has been anecdotally reported by several practitioner-oriented outlets: "Paula Cuneo [...] recently ordered 10 pairs of corduroy pants in varying sizes and colors [...], only to return seven of them. Ms. Cuneo is shopping online for Christmas gifts this year, ordering coats and shoes in a range of sizes and colors. She will let her four children choose the items they want—and return the rest" (Banjo, 2013). Therefore, we specifically hypothesize as follows:

- *H1.2a: There is a positive relationship between the number of multiple-items regarding size (i.e., same style and same color) and the α-/β-returns rate of an order.*
- *H1.2b: There is a positive relationship between the number of multiple-items regarding color (i.e., same style and same size) and the α-/β-returns rate of an order.*
- *H1.2c: There is a positive relationship between the number of multiple-items regarding style (i.e., same color and same size) and the α-/β-returns rate of an order.*

Utility theory suggests that consumers decide to order online when the associated utility exceeds the utility of not ordering (McFadden, 1974). Therefore, the utility of ordering is positively influenced by the customer's expectations of the product and negatively influenced by the attached uncertainty (Minnema et al., 2016; Rust et al., 1999). To reduce the customer's perceived risk, many e-tailers offer liberal returns policies that can be considered as insurance against negative experiences (Mollenkopf et al., 2007; Padmanabhan & Png, 1995). Actively advertising these liberal policies decreases a consumer's uncertainty, thus increasing his or her overall likelihood of placing an order (Kirmani & Rao, 2000; Wood, 2001). However, because customers have only imperfect information about a product's performance prior to receiving it, these liberal policies also lead to more products being returned when customer expectations are not met. According to Zeithaml et al. (1990), a product's price can influence a customer's expectations. Teboul (1991) notes that price determines the level of quality a customer demands. In a recent interview, Mayuki Chou, founder of the Taiwanese fashion e-tailer W-style, notes that consumers have lower expectations of lower-priced items and develop a higher tolerance for disappointment (E27.co, 2015). Accordingly, Hess and Mayhew (1997) argue, "[...] that consumers will be less likely to accept a poor fit as the item becomes more expensive". They use data from an apparel marketer and find that price is significantly positively correlated with the returns rate. Therefore, we also hypothesize that higher prices lead to higher returns rates.

- *H2: There is a positive relationship between the mean item price and the α-/β-returns rate of an order.*

From an information processing perspective, a customer's return decision is a mixed one. Lynch and Srull (1982) define a mixed decision as a choice in which some information is physically present, while some is retrieved from memory. Consequently, consumers base their return decision on (1) the newly acquired information after delivery (e.g., testing product performance, other consumer opinions) and (2) the accessible and available information from the pre-order phase. Bechwati and Siegal (2005) show that the number of positive cognitive responses elicited during the pre-order phase strongly determines the ability to defend against disconfirming information that can trigger returns. In other words, e-tailers that want to reduce returns should avoid measures that lead to their customers making careless and hurried order decisions. However, many firms engage in marketing that encourages the exact opposite behavior by targeting impulse buyers: "Marketers often announce special offers for limited periods of time [...]. Our findings suggest that this approach re-

sults in high product returns because under these circumstances, customers have little opportunity to generate cognitive responses in favor of the chosen alternative [...]" (Bechwati & Siegal, 2005). Bellenger et al. (1978) show that fashion items are frequently purchased on impulse. Dawson and Kim (2009) identify coupons as an external impulse trigger on apparel websites. Therefore, redeeming coupon codes while shopping for apparel on the internet will most likely increase the returns rate.
However, it should be noted that coupons also reduce prices, which lowers expectations and generates a higher tolerance against negative experiences (Hess & Mayhew, 1997). Consequently, the negative price effect compensates for the positive impulse buying effect with an increasing relative coupon code value. We consequently hypothesize the following:

- *H3.1: The use of coupons codes has a positive effect on the α-/β-returns rate of an order.*
- *H3.2: There is a negative relationship between the relative value of a coupon and the α-/β-returns rate of an order.*

Another external impulse trigger is the offer of free gifts with purchase. Some e-tailers require customers to return any free gift received with a product, while others do not. Assuming that the gift must also be returned, the gift increases the value of the actual ordered product. If consumers may keep the gift, it takes on a psychological dimension. The idea of gifts is based on the concept of reciprocity (Schwartz, 1967). It is part of social etiquette that we return not only favors but also gifts. In some cultures, the counter-gift must even exceed the value of the initial gift (McVeigh, 2014). Therefore, accepting a gift leads to informal accountability, and the person who receives it is obliged to give back (Ferrary, 2003). Thus, keeping the gift without buying something in return leads to psychological pressure, which increases return hassles for the consumer. Davis et al. (1998) show that consumers only choose to return if the purchase price minus transaction costs (e.g., restocking fees, any type of hassle) exceeds the remaining product value (Davis, 1995). Ceteris paribus, the first option increases the remaining product value, whereas the second option increases transaction costs because it introduces more hassle. Thus, we strongly believe that free gifts reduce the returns rate.

- *H4: Adding free gifts to an order has a negative effect on the α-/β-returns rate of an order.*

Another factor that may enhance careless and hurried order decisions is the payment option. A recent Nielsen (2016) report shows that payment practices vary considerably around the world (Nielsen, 2016). In North America, consumers prefer credit cards over digital payment systems (e.g., PayPal, Alipay) and store-specific gift cards. In India, cash on delivery is the most popular option, followed by debit cards and direct debit. In contrast, the majority (58 %) of German online shoppers, representing the largest e-commerce market in central Europe, prefer to pay by invoice (Budde et al., 2013). In other words, customers receive a bill together with the ordered goods that generally needs to be settled within 14 days, which puts the merchant at risk of not receiving payment. While different regions have their preferred payment terms, we may distinguish between pre-delivery options (e.g., digital payment systems, credit cards, cash on delivery) and post-delivery alternatives (e.g., invoice). From the consumer's perspective, post-payment provides a seamless shopping experience that is both convenient and risk-free. However, several practitioners suggest that it potentially leads to a "[...] higher percentage of returned goods [...]" (About Payments, 2016) because uncertain, uninformed customers with a high likelihood to return prefer such risk-free payment options. We therefore believe that post-delivery payment options increase the returns rate.

- *H5: Post-delivery payment options such as invoicing have a positive effect on the α-/β-returns rate of an order.*

E-tailers can prevent their customers from forming excessive product expectations and stimulate cognitive processes by providing detailed product descriptions, high-resolution images, videos, and/or balanced customer reviews to help facilitate thoughtful and deliberate order decisions. While these measures may avoid some returns, others are simply inevitable (Foscht et al., 2013). Each item carries a return probability that is unknown to a retailer prior to its listing. However, as products are sold, e-tailers learn about this inherent returns rate. Thus, we conclude that the aggregated item returns rate observed in the past positively influences the likelihood to return:

- *H6: There is a positive relationship between the ordered items' aggregated historical β-returns rate and the α-/β-returns rate of an order.*

Finally, we consider general consumer behavior to be a factor. Some media outlets have reported on returnaholics or serial returners. These people have a tendency to abuse liberal policies. As a result, some merchants have changed their policies to prevent damage from this type of behavior (Tuttle, 2013). Others have used technology to track down serial returners. An Associated Press report shows that retailers such as Home Depot, Victoria's Secret, Best Buy, and Nike create customer return profiles (Kerr, 2013). If a customer's return activity is too high or suggests a questionable pattern, he or she may be banned from future returns. These approaches are built on the common-sense notion that past behavior predicts future behavior. In other words, returning becomes a habit. This phenomenon is well documented in psychology. Ouellette and Wood (1998) provide a meta-analytic synthesis and show that repeatedly performed tasks in stable contexts become habitual as conscious information processing becomes automatic. In unstable contexts, past behavior contributes to intentions, which guide future behavior. In the latter case, the effect is weaker but still observable (Ouellette & Wood, 1998). Aarts et al. (1998) examine travel mode choices and come to similar conclusions: "[...] any type of repetitive behavior requires less and less mental effort and conscious attention, and may therefore eventually become habitual. Consequently, these behaviors may no longer be guided by deliberately formed intentions, but are accompanied by a rather limited process of decision making". The conditions surrounding a return decision may be a mixture of stable and unstable contexts, as the items to be evaluated change, while the process itself remains the same. We therefore hypothesize that past return behavior positively influences the returns rate:

- *H7: There is a positive relationship between the customer's historical β-returns rate and the α-/β-returns rate of an order.*

VII.3 Methodology

The hypotheses derived above are tested using comprehensive data from a German apparel e-tailer that operates exclusively online and specifically targets women. The company requested confidentiality concerning its name. On its website, the merchant employs several stimuli that increase the consumers' likelihood of placing an order, such as coupon codes, free gifts, and invoicing as a post-delivery payment option. The company offers free returns within 14 days after delivery with no questions asked (Directive 2011/83/EU on consumer rights, 2011). The database contains sale and returns information over 18 months (from April 2014 to December 2015). During that time, the e-tailer served almost 300,000 customers who chose from approximately 3,300 articles and placed almost 650,000 orders. Net sales totaled 30.28 million Euro. The average shopping basket contained 3.15 items, of which 1.64 were returned. On average, the shipped merchandise was worth 108.95 Euro, and returns were valued at 62.49 Euro. In sum, the retailer shipped 2,049,853 items, of which 1,069,008 were returned. This corresponds to a β-returns rate of 52.15 %. Because the merchant has an order consolidation policy in place, it can be assumed that each order was delivered in one shipment. Cases in which orders must be split due to significantly different delivery times may be a problem in product categories other than apparel. In the fashion industry, goods are ordered at the beginning of a selling season and put on display until they are sold out. According to the data, 651,658 outbound shipments resulted in 412,584 returns. That is, the α-returns rate was 63.31 %. The dataset provided by the merchant contained the information documented in Table VII.1.

Variable (Mean / SD)	**Description**
OrderID	Unique identification number of an order.
OrderDate	Date of the order.
CustomerID	Unique identification number of the customer placing the order.
AddItems (2.15 / 2.53)	Total number of additionally ordered items (=Total number of ordered items–1).
OrderValue (110.43 / 92.31)	Total retail price of all the items ordered in Euro.
Coupon (.14 / .35)	Dummy variable indicating whether a coupon code has been used (=1) or not (=0).
CouponValueAbs (1.48 / 4.08)	Absolute value of the coupon code in Euro.
CouponValueRel (.02 / .05)	Relative value of the coupon in percent (=CouponValAbs/ OrderValue).
MeanPrice (39.10 / 22.61)	Mean retail price of the ordered items (=OrderValue/AddItems+1) in Euro.
FreeGift (.02 / .14)	Dummy variable indicating whether the customer received a free gift (=1) or not (0).
MultItemsSize (.15 / .36)	Number of multiple items with different size but same color and same style.
MultItemsColor (.24 / .52)	Number of multiple items with different color but same size and same style.
MultItemsStyle (.19 / .49)	Number of multiple items with different style but same color and same size.
PayByInvoice (.74 / .43)	Dummy variable indicating whether the customer paid by invoice (=1) or by another payment method (=0).
Return (.63 / .48)	Dummy variable indicating whether a return was made (=1) or not (=0).
ItemsReturned (1.6 / 2.10)	Total number of items returned.
RefundValue (62.49 / 82.29)	Total value of the returned items in Euro.
HistNrOrders (4.06 / 11.33)	Number of total order the customer has placed before.
HistBetaCust (47.73 / 17.47)	The customer's historic β-returns rate up to this order in percent.
HistBetaItemsOrd (80.71 / 17.58)	Aggregated historic β-returns rate of the items ordered in percent.

Table VII.1. Description of the dataset variables.

In general, missing values were not an issue, except for the historic β-returns rate of new customers. In these cases, we inserted the average customer returns rate observed in the first quarter of 2014 (47.52 %), which is the period before the dataset starts. This procedure was applied by the merchant when calculating the variable “HistBetaItemsOrd” just in case recently listed articles with no historical data were part of that order.

We use regression models as our data analysis method because they are easy to interpret and are therefore most suitable for this contribution. Herein, the α-/β-returns rate of an order serves as the regressand. As the α-returns rate of an entire order can take only two values (0 % if the consumer keeps the entire order or 100 % if the consumer returns one or more items), we refer to a binary logistic regression, whereas traditional multivariate linear regression is used to explain the β-returns rate. Before performing the calculations using SPSS 23, the data were visualized to check for U-shaped or curvilinear relationships. We recognized a conspicuous relation between the regressand and the variables “HistBetaCust” and “HistBetaItemsOrd” but attribute this to the regressor’s learning character; this leads to some distortion.

VII.4 Findings

We first estimate the linear regression model with the β-returns rate as the dependent variable. The independent variables include the total number of additionally ordered items (AddItems), the number of multiple-items concerning size (MultItemsSize), color (MultItemsColor) and style (MultItemsStyle), the average retail price (MeanPrice), the use of a coupon (Coupon), the relative coupon value (CouponValRel), the addition of a free gift (FreeGift), the payment method (PayByInvoice), the customer’s historic β-returns rate (HistBetaCust), and the aggregated historic β-returns rate of all the items ordered (HistBetaItemsOrd). The predictors “PayByInvoice”, “Coupon”, and “FreeGift” are modeled as dummy variables. It should be noted that the p-values should not be misinterpreted in a way that smaller values indicate more relevant results (Wasserstein & Lazar, 2016). In our model, even small effects become highly significant due to the large sample size. Instead, the reported standardized coefficients are more suitable to assess the relative importance of an effect.

In general, the model is significant (p=.000, F=14764.970, R=.447, R-squared=.200) and includes all exogenous variables mentioned above. Concerning the quality of the model, a large R-squared indicates a good fit. In this case, the employed independent variables explain 20.0 % of the response variable variance. This seems acceptable for a complex behavioral problem with many other dimensions (e.g., customer demographics, logistical performance) that are not part of this analysis (Cohen, 1992). Table VII.2 shows the results in detail.

The findings support most of the hypotheses, except for H1.1 and H1.2b. This is notable because it indicates that, in fashion retailing, ordering more does not automatically result in more items being returned. The effect is small (b(AddItems)=-.004, p=.000), but based on these findings, e-tailers trying to reduce returns may be ill-advised to limit the maximum number of items a customer can order, as stipulated by some e-commerce solutions (Shopify, 2016).

Variable	Coefficient b	Standard error	Stand. coefficient Beta	T-value	Sig.
Constant	-.432	.003		-148,986	.000
AddItems	-.004	.000	-.028	-14,782	.000
MultItemsSize	.141	.001	.127	105,240	.000
MultItemsColor	-.012	.001	-.016	-12,266	.000
MultItemsStyle	.026	.001	.032	23,867	.000
MeanPrice	.002	.000	.130	112,527	.000
Coupon	.012	.002	.011	5,388	.000
CouponValRel	-.003	.000	-.036	-18,295	.000
FreeGift	-.081	.003	-.028	-25,120	.000
PayByInvoice	.100	.001	.108	95,601	.000
HistBetaItemsOrd	.006	.000	.258	172,487	.000
HistBetaCust	.005	.000	.211	187,845	.000

Table VII.2. Results of the linear regression model.

Even more striking is the fact that, if customers purchase the same items in different colors, their returns rate decreases (b(MultItemsColor)=-.012, p=.000). That is, the majority of customers are not ordering multiple colors due to uncertainty but for the sake of owning different variants of the product. Consequently, category managers should not shy away from increasing the number of color variants out of fear of higher returns. However, if consumers order a product in multiple sizes (b(MultItems-Size)=.141, p=.000) and styles (b(MultItemsStyle)=.026, p=.000), the likelihood of returns increases. The effect of size exceeds the effect of style. This observation supports some of the actions taken by the German fashion shoe retailer Mirapodo (www.mirapodo.de). If a customer orders an item in different sizes, a note is displayed at the beginning of the checkout procedure that says: “Do you really need multiple sizes? Please note: Any return generates costs and pollutes the environment.”

Price (b(MeanPrice)=.002, p=.000) is also a significant factor. The results indicate that the likelihood of returns increases with the mean price of the ordered items. This replicates the results of Hess and Mayhew (1997), who perform their study using data from a fashion mail-order company. Our study complements their findings in the sense that the observed price effect has entered into e-commerce. Consequently, e-tailers with a high-priced product range face higher costs of product returns. This reality should be considered when making listing and pricing decisions.

As hypothesized, the results suggest that coupons cause two opposing effects. In general, coupon codes trigger impulse buyers to order, leading to higher returns. The coefficient of the dummy variable (b(Coupon)=.012, p=.000) indicates that the β-returns rate of an order is 1.2 percentage points higher than it is when a coupon code is not redeemed during checkout. However, this effect is mitigated by the costs savings the coupon provides, which reduces the price to be paid (b(CouponValRel)=-.003, p=.000). In other words, as the relative value of a coupon increases, the likelihood of returning an ordered item decreases. Within the examined data, the negative price effect compensates for the positive impulse-buying effect if the relative value of the coupon exceeds 4 percent. While previous studies have addressed only the sales-promoting effect of coupon codes (Edelman et al., 2016), our study is the first to include the impact on product returns. The findings suggest that rebate coupons are an appropriate marketing instrument when merchants are interested not only in boosting sales but also in avoiding returns. This may be particularly true at the end of selling seasons, when it becomes harder to resell potentially returned products. Other coupons with a low relative value, e.g., codes that only compensate for shipping costs, may increase the probability of orders, but the overall financial effect remains unclear. Because returns incur significant costs, e-tailers should weigh the pros and cons of such coupons thoroughly.

Another marketing measure is free gifts. We hypothesized that free gifts would increase the perceived order value or increase psychological hassle, both of which reduce the likelihood of returns. The regression coefficient of the dummy variable supports this point of view (b(FreeGift)=-.081, p=.000). The β-returns rates of orders accompanied by a gift were 8.1 percentage points lower than those without a freebie. If these results can be confirmed by other studies, e-tailers should look into providing small gifts with their deliveries, such as small product samples or small packages of candy, because they may not only help build a sustainable customer relationship but also reduce consumer returns.

Additionally, the influence of the payment option should not be underestimated. In this study, we specifically looked into the impact of invoicing as a post-delivery payment option. So far, invoicing is a German peculiarity, but emerging service providers such as Klarna are currently trying to make it available internationally. Such service providers attempt to eliminate this risk by paying the retailer immediately and taking over the debt collection for a percentage fee. Their main argument is that invoices notably decrease barriers to ordering and, thus, increase sales. The business news network CNBC calls Klarna one of the world's most disruptive companies and predicts that it will be highly successful (CNBC, 2016). However, post-delivery payment options seem to be a mixed blessing, according to the present study. Compared with the pre-delivery alternatives (b(PayByInvoice)=.100, p=.000), invoicing increases the β-returns rate by 10.0 percentage points. Therefore, before buying into the promising sales effects of invoicing, e-tailers should seriously evaluate the negative cost effects of invoicing that result from higher returns rates. Introducing invoicing should be considered only if the overall financial impact is positive. In addition, these results provide e-tailers that have introduced invoicing and that are trying to cut back on returns with a viable strategy: they might restrict the payment options available to those customers with histories of conspicuous or unethical returns behavior.

Variable	Coefficient b	Standard error	Wald	Sig.	Exp(b)
Constant	-6.017	.023	66049.932	.000	.002
AddItems	.145	.003	2092.468	.000	1.156
MultItemsSize	2.466	.023	11159.336	.000	11.776
MultItemsColor	-.242	.007	1081.295	.000	.785
MultItemsStyle	.201	.010	410.051	.000	1.223
MeanPrice	.009	.000	4209.977	.000	1.010
Coupon	.420	.021	383.201	.000	1.523
CouponValRel	-3.257	.141	534.816	.000	.039
FreeGift	-.029	.023	1.582	.208	.971
PayByInvoice	.681	.007	9461.774	.000	1.976
HistBetaItemsOrd	.049	.000	30397.016	.000	1.050
HistBetaCust	.032	.000	27698.918	.000	1.032

Table VII.3. Results of the binary logistic regression model.

Finally, we examined the influence of item characteristics and past customer behavior, both of which are described by the order and return history. Our results suggest that the historical returns rate of the ordered articles (b(HistBetaItemsOrd)=.006, p=.000) and the historical customer returns rate (b(HistBetaCust)=.005, p=.000) positively influence the returns rate. In fact, the standardized coefficients (b'(HistBetaItemsOrd)=.258, b'(HistBetaCust)=.211) show that the historical return rates have the greatest relative effect on the dependent variable. These results lead to an important managerial implication. E-tailers should collect and use returns data to learn from the past. Items with returns rates so high that they may not be profitably distributed should be discontinued. Moreover, it can be beneficial to segment customers based on their past returns behavior and to close unprofitable accounts. Amazon is known for such measures. They justify their actions with the following explanation: "[...] a careful review of this account and related ones shows you've requested refunds and replacements on a majority of your orders for a variety of reasons. In the normal course of business, we expect there may be occasional problems. However, the rate at which such problems have occurred on your account is extraordinary, and it cannot continue. Your Amazon.com account has been closed, and you will no longer be able to shop in our store" (Amazon.com, 2012).

The multivariate linear regression was complemented by a binary logistic model to explain the factors influencing the α-returns rate as the regressand. To ensure comparability, the independent variables are consistent

with those of the first model. Again, the model is significant (p=.000, Cox & Snell R-squared=.309, Nagelkerke's R-squared=.423) and includes all exogenous variables. Binary logistic regression models provide only pseudo R-squared statistics to evaluate the goodness of fit. According to Eßig and Glas (2015), acceptable values for Cox and Snell's R-squared and for Nagelkerke's R-squared are greater than .2, while good values start at .4. Thus, the model's fit is acceptable-to-good, depending on the measure applied. Table VII.3 presents the detailed results.
The binary logistic regression strongly supports the findings of the linear regression analysis, with two exceptions. First, the sign of b(AddItems) changed to positive. This indicates that, while ordering more does not automatically result in more items being returned, it obviously leads to more return shipments that must be transported and handled. Second, the influence of b(FreeGift) is no longer significant (p=.208). That is, the psychological effects of free gifts are strong enough to reduce the number of items being returned but are insufficient to prevent additional return shipments.
In conclusion, it should be noted that the more-than-acceptable goodness of fit indicators show that using data analysis techniques not only may help in identifying influencing factors but also may be used to predict future return shipments, which helps the returns department with better capacity planning. This shows that data mining and big data analytics support better decision making and create value in returns management.

VII.5 Summary and future research

Returns management – as an important component of supply chain management – is a key activity in an online retailer's business model. Prior research has shown that fashion e-tailers experience returns rates of up to 60 % (Foscht et al., 2013). While successful returns may stimulate customer satisfaction and retention, they also incur costs (Asdecker, 2015). It is therefore essential to understand the drivers influencing the flow of returns. Only with these insights can consumer returns be effectively managed.
Prior empirical research in this area is usually based on survey data. This paper contributes to the literature by testing several hypotheses using real shop data from an apparel e-tailer. This type of study is rare because many

merchants consider order and returns data as proprietary and are usually unwilling to share this information (Stock & Mulki, 2009).
The hypotheses were formulated based on several established theories – such as post-purchase dissonance theory – and anecdotal information. Methodologically, we used linear and logistic regression models to analyze the data at hand. The results supported most of the hypotheses. Table VII.4 provides a summary.

	Hypothesis	**Results**
H1.1	There is a positive relationship between the number of additional items and the α-/β-returns rate of an order.	α: Supported β: Not supported
H1.2a	There is a positive relationship between the number of multiple-items regarding size (i.e., same style and same color) and the α-/β-returns rate of an order.	α: Supported β: Supported
H1.2b	There is a positive relationship between the number of multiple-items regarding color (i.e. same style and same size) and the α-/β-returns rate of an order.	α: Not supported β: Not supported
H1.2c	There is a positive relationship between the number of multiple-items regarding style (i.e. color and same size) and the α-/β-returns rate of an order.	α: Supported β: Supported
H2	There is a positive relationship between the mean item price and the α-/β-returns rate of an order.	α: Supported β: Supported
H3.1	The use of coupons codes has a positive effect on the α-/β-returns rate of an order.	α: Supported β: Supported
H3.2	There is a negative relationship between the relative value of a coupon and the α-/β-returns rate of an order.	α: Supported β: Supported
H4	Adding free gifts to an order has a negative effect on the α-/β-returns rate of an order.	α: Not supported β: Supported
H5	Post-delivery payment options such as invoicing have a positive effect on the α-/β-returns rate of an order.	α: Supported β: Supported
H6	There is a positive relationship between the ordered items' aggregated historical β-returns rate and the α-/β-returns rate of an order.	α: Supported β: Supported
H7	There is a positive relationship between the customer's historical α-/β-returns rate and the α-/β-returns rate of an order.	α: Supported β: Supported

Table VII.4. Summary of hypotheses test results.

The findings enable better decision making by identifying several factors that significantly influence the returns rate. We distinguish between shopping basket-related (H1-H1.2c), sales/marketing-related (H2-H5), product-related (H6), and customer-related factors (H7).

Our results indicate that the composition of the customer's shopping basket influences the returns rate. It should be noted that more ordered items do not necessarily lead to more returned items. Moreover, it is interesting to see that multiple-item orders regarding color have a negative effect on the returns rate. Consequently, product returns should not be an issue when deciding on additional color variations. In contrast, multiple-item orders in terms of size and style significantly increase the returns rate.

This article also investigates the effects of sales stimuli (such as coupon codes, free gifts, and payment methods) on product returns. We find that coupons increase the returns rate in general. However, this effect is reversed in conjunction with increasing relative coupon value. Moreover, free gifts may reduce the likelihood of returning an item. Nonetheless, this finding may not be applied to return shipments. It seems as if free gifts contribute to a customer's bad conscience if the entire order is supposed to be sent back. In other words, freebies increase the likelihood that a customer will keep at least one item. The data also indicate that allowing customers to pay after delivery influences the returns rate. It shows that adding invoicing to the existing slate of payment options is a double-edged sword. Decision makers have to carefully balance revenue against cost effects to ensure a positive financial outcome.

On the product and customer end, we find that it can be beneficial for e-tailers to perform big data analytics on their order and return histories because analyzing the past may help to predict future behavior. Such analyses may show that it is better to close customer accounts and discontinue products with extremely high returns rates, true to the motto that "better an end with pain, than pain without end".

The present study uses data from a single e-tailer. Although an extensive dataset with more than 650,000 orders over the course of 18 months was used, the generalizability of the presented findings is limited. Therefore, this publication should be complemented by additional future research. To gain a more accurate picture, we call for similar studies using data from e-tailers with different target groups (e.g., apparel for men or other age groups) or product groups (e.g., consumer electronics, books). There are also concerns regarding endogeneity (Guide & Ketokivi, 2015). While trying to provide strong theoretical arguments when deducing the hypotheses, it must be acknowledged that the secondary data provided did not allow for the development and testing of potential instrumental variables. We therefore call for complementary experimental research to establish causality.

Moreover, it would be interesting to compare gender-specific results. Other factors, such as age or location, could also be taken into account. An international comparison of identified factors would also be a valuable contribution to the literature. Furthermore, in the context of controversial same-day delivery and anticipatory shipping services (Spiegel et al., 2013), returns rates need to be more closely investigated in terms of delivery modes and times.

References

Aarts, H., Verplanken, B., & Knippenberg, A. (1998). Predicting Behavior From Actions in the Past: Repeated Decision Making or a Matter of Habit? *Journal of Applied Social Psychology, 28*(15), 1355–1374.

About Payments. (2016). *Germany. E-Commerce Market Profile*. Retrieved June 7, 2016, from https://www.about-payments.com/knowledge-base/market/germany

Amazon.com. (2012). *Amazon.com General Help Forum - My Account Has Been Closed*. Retrieved June 10, 2016, from http://www.amazon.com/gp/help/customer/forums/ref=help_forum_tft_tp?ie=UTF8&cdForum=Fx2NFGOONPZEXIP&cdThread=Tx1IS0VNGMKXPXQ

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Banjo, S. (2013). Rampant Returns Plague E-Retailers. *Wall Street Journal - Eastern Edition, 22*(148), B1–B2.

Bechwati, N. N., & Siegal, W. S. (2005). The Impact of the Prechoice Process on Product Returns. *Journal of Marketing Research, 42*(3), 358–367.

Bellenger, D. N., Robertson, D. H., & Hirschman, E. C. (1978). Impulse buying varies by product. *Journal of Advertising Research, 18*(6), 15–18.

Budde, L., Hampe, K., Puppe, M., & Shahd, M. (2013). *Trends im E-Commerce*. Retrieved June 1, 2016, from https://www.bitkom.org/Publikationen/2013/Studien/Trends-im-ECommerce/BITKOM-ECommerce-Studienbericht.pdf

CNBC. (2016). *Klarna: No sale left behind*. Retrieved May 24, 2016, from https://www.cnbc.com/2016/06/07/klarna-2016-disruptor-50.html

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Croxton, K. L., García-Dastugue, S. J., Lambert, D. M., & Rogers, D. S. (2001). The Supply Chain Management Processes. *The International Journal of Logistics Management, 12*(2), 13–36.

Davis, S. (1995). Money back guarantees in retailing: matching products to consumer tastes. *Journal of Retailing, 71*(1), 7–22.

Davis, S., Hagerty, M., & Gerstner, E. (1998). Return policies and the optimal level of "hassle", *50*(5), 445–460.

Dawson, S., & Kim, M. (2009). External and internal trigger cues of impulse buying online. *Direct Marketing: An International Journal, 3*(1), 20–34.

E27.co. (2015). *5 factors behind high e-commerce return rate*. Retrieved June 5, 2023, from https://e27.co/5-factors-behind-high-e-commerce-return-rate-20150721/

Edelman, B., Jaffe, S., & Kominers, S. D. (2016). To Groupon or not to Groupon: The profitability of deep discounts. *Marketing Letters, 27*(1), 39–53.

Eßig, M., & Glas, A. H. (2015). Considering Small and Medium-Sized Suppliers in Public Procurement—The Case of the German Defence Sector. In *Logistics Management: Proceedings of the 9th Conference "Logistikmanagement",* Braunschweig, Germany.

Directive 2011/83/EU on consumer rights (2011). https://eur-lex.europa.eu/eli/dir/2011/83/2022-05-28

Ferrary, M. (2003). The Gift Exchange in the Social Networks of Silicon Valley. *California Management Review*, *45*(4), 120–138.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press. https://doi.org/10.1515/9781503620766

Foscht, T., Ernstreiter, K., Maloles, C., Sinha, I., & Swoboda, B. (2013). Retaining or returning? *International Journal of Retail & Distribution Management, 41*(2), 113–134.

Guide, V. D. R., & Ketokivi, M. (2015). Notes from the Editors: Redefining some methodological criteria for the journal⋆. *Journal of Operations Management, 37*(1), v–viii.

Hess, J. D., & Mayhew, G. E. (1997). Modeling merchandise returns in direct marketing. *Journal of Direct Marketing, 11*(2), 20–35.

Kerr, J. C. (2013, August 12). Serial returners, beware: Retailers are tracking you. *TODAY*. https://www.today.com/money/serial-returners-beware-retailers-are-tracking-you-6C10900265

Kirmani, A., & Rao, A. R. (2000). No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality. *Journal of Marketing, 64*(2), 66–79.

Lee, D. H. (2015). An Alternative Explanation of Consumer Product Returns from the Postpurchase Dissonance and Ecological Marketing Perspectives. *Psychology & Marketing, 32*(1), 49–64.

Lynch, J. J. G., & Srull, T. K. (1982). Memory and Attentional Factors in Consumer Choice: Concepts and Research Methods. *Journal of Consumer Research, 9*(1), 18–37.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics, 3*(4), 303–328.

McVeigh, B. J. (2014). *Interpreting Japan: Approaches and applications for the classroom.* Routledge. https://doi.org/10.4324/9781315850559

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To Keep or Not to Keep: Effects of Online Customer Reviews on Product Returns. *Journal of Retailing, 92*(3), 253–267.

Mollenkopf, D. A., Rabinovich, E., Laseter, T. M., & Boyer, K. K. (2007). Managing Internet Product Returns: A Focus on Effective Service Operations. *Decision Sciences, 38*(2), 215–250.

Nielsen. (2016). *Global Connected Commerce*. Retrieved May 25, 2016, from https://www.nielsen.com/de/content/dam/corporate/us/reports-downloads/2016-reports/connected-commerce-report-jan-2016.pdf

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin, 124*(1), 54–74.

Padmanabhan, V., & Png, I. (1995). Returns policies: Make money by making good. *Sloan Management Review, 37*(1), 65–72.

Powers, T. L., & Jack, E. P. (2013). The Influence of Cognitive Dissonance on Retail Product Returns. *Psychology & Marketing, 30*(8), 724–735.

Rust, R. T., Inman, J., Jia, J., & Zahorik, A. (1999). What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions. *Marketing Science, 18*(1), 77–92.

Schwartz, B. (1967). The social psychology of the gift. *AJS; American Journal of Sociology, 73*(1), 1–11.

Shopify. (2016). *Limiting Purchase Quantities*. Retrieved June 13, 2016, from https://help.shopify.com/manual/products/inventory/limiting-purchase-quantities

Spiegel, J. R., McKenna, M. T., Lakshman, G. S., & Nordstrom, P. G. (2013). Method and System for Anticipatory Shipping(US 8,615,473).

Stock, J. R., & Mulki, J. P. (2009). Product Returns Processing: An Examination of Practices of Manufacturers, Wholesalers/Distributors, and Retailers. *Journal of Business Logistics*, *30*(1), 33–62.

Stock, J. R., Speh, T., & Shear, H. (2006). Managing Product Returns for Competitive Advantage. *MIT Sloan Management Review*, *48*(1), 57–62.

Teboul, J. (1991). *Quality management*. Prentice-Hall.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

Tuttle, N. (2013). *Shoppers Face End of Generous Retailer Return Policies*. Retrieved June 11, 2016, from https://business.time.com/2013/09/19/shoppers-face-end-of-generous-retailer-return-policies/

United States Census Bureau. (2016). *Estimated Annual U.S. Retail Trade Sales - Total and E-commerce: 1998-2015*. Retrieved May 29, 2016, from https://www2.census.gov/retail/releases/current/arts/ecommerce.xls

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p -Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129–133.

Wood, S. L. (2001). Remote Purchase Environments: The Influence of Return Policy Leniency on Two-Stage Decision Processes. *Journal of Marketing Research*, *38*(2), 157–169.

Zeithaml, V., Berry, L. L., & Parasuraman, A. (1990). *Delivering quality service: Balancing Customer Perceptions and Expectations*. The Free Press.

VIII The Impact of Displaying Quantity Scarcity and Relative Discounts on Sales and Consumer Returns in Flash Sale E-Commerce

Authors: David Karl, Björn Asdecker, Christine Feddersen-Arden

Publication Outlet: Proceedings of the 55th Hawaii International Conference on System Sciences
(ranked as B in VHB-Rating 2024)

DOI: 10.24251/HICSS.2022.556

published as:

Karl, D.; Asdecker, B.; Feddersen-Arden, C. (2022): The Impact of Displaying Quantity Scarcity and Relative Discounts on Sales and Consumer Returns in Flash Sale E-Commerce. In: *Proceedings of the 55th Hawaii International Conference on System Sciences*, Maui, United States of America, pp. 4569–4578. DOI: 10.24251/HICSS.2022.556

Abstract

Flash sale e-commerce is a very competitive business with low margins due to the high discounts granted. Against this background, merchants pursue the goal of generating as many orders as possible. To achieve this, techniques that stimulate impulse buying behavior are often used. This paper examines two specific instruments that have the potential to contribute to impulse buying, namely, the on-site display of (1) scarcity notifications and (2) the relative discount provided. We use real-world data from a flash sale e-tailer to analyze the impact on customer sales and returns. In that regard, this study is the first to focus specifically on fashion, which is the product category most affected by returns. Furthermore, to synthesize both perspectives, a quantitative model is presented to serve as the basis for a decision support system that enables managers to better deal with the underlying trade-off.

Keywords: Electronic Marketing, discounts, e-commerce, flash sales, returns, scarcity

VIII.1 Introduction

Imagine the following self-talk of someone who spotted a great offer in a flash sale community:
"*I really like that brand... Let's have a quick look at what they are offering... Wow, 60 % off, that's quite a deal! Hmm, but only two left. I have to act fast.*"
A week later, after the parcel carrier delivered the order:
"*Well... that's not really what I expected. I think maybe I rushed it. But, no, I don't like it... It definitely needs to go back. Great that they offer free returns.*"
When reading these lines, many passionate online shoppers might experience a déjà-vu. This example illustrates the weal and woe of every e-tailer. First, customers have to be persuaded to order. For that purpose, e-tailers use on-site marketing tools to stimulate the customer's impulse buying behavior and thus increase order likelihood. These include, among others, displaying price discounts and disclosing the limited remaining stock. However, such measures may also affect the probability that consumers develop buyers' remorse. Since customers cannot physically assess the quality and fit of products online, returns are an inevitable part of the business model (De et al., 2013; Hong & Pavlou, 2014). Merchants of fashion products with return rates of 50 % and more are particularly affected (Asdecker et al., 2017). Although relative return rates declined slightly during the COVID-19 pandemic, the absolute number of consumer returns continued to rise due to the massive growth in e-commerce (Karl & Asdecker, 2021).
With such high return figures, ordinary e-tailers already face problems of operating profitably. For flash sale e-tailers, however, the challenge is even more pressing. Specialized flash sale e-tailers such as HauteLook, Gilt, BestSecret, VIP.com – some of them among the leading fashion e-commerce companies worldwide (Statista, 2021) – build a closed shopping community for which they organize private sales with highly discounted products of popular brands for a limited period. The products are on offer until sold out or the campaign period ends.
Flash sale shopping club platforms are a popular distribution channel among well-known brands since the closed community prevents the cannibalization of their primary markets while having the opportunity to market the excess stock at the end of the selling season. In general, flash sales are a commission business. This means that the flash sale e-tailer

receives a percentage of the sales price for organizing the sale and processing the orders and returns. Any unsold items are returned to the brand manufacturers. The margins of the business model are significantly lower than those in regular retail due to the high discounts. For this reason, it is essential for flash sale e-tailers to generate as many orders as possible. At the same time, however, the return rates must not be too high. Against this background, this paper addresses the following research question:

- *To what extent do the display of quantity scarcity messages and price discounts impact sales and consumer returns in a flash sale e-commerce environment, and how does it influence profitability?*

To answer this research question, we use real-world data from a leading specialized flash sale e-tailer, as Pavlou et al. (2007) recommended. This study contributes to the digital commerce and electronic marketing literature by examining the impact of on-site communication efforts on both dimensions of the underlying problem: (1) sales and (2) returns. Furthermore, we present a quantitative model that integrates both perspectives and may serve as the basis for a decision support system that puts decision-makers in the position to better deal with the underlying trade-off.
The remainder of this paper is organized as follows: The following section presents backgrounds, the relevant literature, and the development of our hypotheses. The methodology and the data used are presented in Section VIII.3. Section VIII.4 contains the analysis of the data and a discussion of the results, while Section VIII.5 illustrates a generalized approach for assessing the profitability of interventions. Finally, we conclude with an outlook and future research opportunities in Section VIII.6.

VIII.2 Background, literature review, and derivation of hypotheses

VIII.2.1 Background

As described in the introduction, triggering impulse purchases is one of the most effective instruments to increase orders. Impulse buying "[...] describes any purchase which a shopper makes but has not planned in advance" (Stern, 1962). Often, visuals at the point-of-sale (or on the respective website) trigger impulse purchases. Major influences for impulse buying are, for example, low prices, marginal need for an item, or prominent store display (Stern, 1962). Thus, impulse buying can be stimulated by instruments such as communication or price stimuli (Iyer et al., 2020). Pricing-oriented instruments have long been used in marketing to attract customers and increase sales, such as time-limited discounts or VIP customer treatment (Li et al., 2019).

Another measure to attract e-commerce customers could be the limited availability of products. Scarcity messages increase perceived arousal and thus positively impact impulse purchases (Wu et al., 2021). This effect can be attributed to limited time or limited quantity (Aggarwal et al., 2011; Cialdini, 2014). Flash sale e-tailers use both types. Time scarcity refers to an offer becoming unavailable after a certain period. Limited quantity scarcity refers to an offer only being available as long as supply lasts. As soon as the retailer displays the remaining stock on their webshop, the latter kind of scarcity creates uncertainty and the urge to order quickly because the behavior of unknown others determines one's own outcome. Such "[...] use of user-interface design elements to guide people's behavior in digital choice environments [...]" (Weinmann et al., 2016) is referred to as digital nudging. User-interface elements support and influence decision-making without imposing significant economic incentives or restricting the freedom of choice (Meske & Potthoff, 2017). The recent discussion about digital nudging includes increasing personal or social welfare (Abdulla et al., 2019; Gray et al., 2018) to better distinguish these interventions from marketing techniques such as dark patterns, undermining the consumers' best interest (Gray et al., 2018). Against this backdrop, information on e-commerce websites such as messages about product scarcity or the display of relative price discounts acts as a digital nudge, at least in a broader sense.

In general, but even more so in the flash sale context, purchases in online retailing entail uncertainties, which can be categorized into product quality/fit and valuation uncertainty (Abdulla et al., 2019). Impulse buying with little time to inform oneself about the product decreases the probability of a good fit or developing realistic expectations about the product (Hong & Pavlou, 2014). In contrast, low prices can reduce valuation risk by lowering the expectations to be met. To minimize the perceived uncertainties, create trust, and encourage consumers to order, companies are granting liberal return policies, leading to more returns than in traditional brick-and-mortar retailing (e.g., Xia & Zhang, 2010). Since the costs caused by returns do not rise linearly but disproportionately with the return rate, returns management is considered a critical success factor for overall profitability (Asdecker, 2015). In flash sale e-commerce, returns are even more crucial because of the below-average margins.

VIII.2.2 Relevant literature

One strain of literature this work relates to is the **design and information value of retailers' shopping websites** (e.g., the use of product-related content such as zoom features (De et al., 2013) or customer product reviews (Sahoo et al., 2018)). Scarcity or discount messages act as instruments to catch customers' attention and influence their behavior. This has been studied by Luo et al. (2019) for cart targeting. They found that scarcity and price incentives can promote sales if specific requirements are met. According to their field experiment, price incentives were relatively ineffective before an item was shortlisted to a user's cart. In contrast, costless scarcity messages significantly boosted purchases and were superior to price incentives.

The second stream of literature specifically investigates **scarcity and its effect on sales**. Following signaling theory, scarcity messages signal product quality for consumers with limited information (Stock & Balachander, 2005). Although the low stock boosts sales and consumer competition, sold-out products are a reason for consumer dissatisfaction (Peinkofer et al., 2016). Cui et al. (2019) identify flash sale customers from Amazon Lightning Deals and show that low stock situations increase their order likelihood. Calvo et al. (2020) replicate this finding for another flash-sale retailer, where the disclosure of scarcity increases hourly sales by 13.6 %. In contrast to most literature, Park et al. (2020) observe decreased sales of durable goods when partial stock information ("less than five items left") is displayed.

Concerning the **display of price discounts to promote sales**, Lehtimäki et al. (2019) and Choi and Coulter (2012) discuss different variants of displaying price discounts (e.g., absolute and relative discounts under different price levels). While some papers examine how the price level influences return behavior, we do not know of any previous work investigating the effects of visual cues regarding the granted discount on consumer returns.

Concerning the reverse part of the e-commerce supply chain, few scholars have investigated the **effects of scarcity on returns**. Rao et al. (2014) analyze online channel transaction data from a personal accessory retailer to show that quantity scarcity perceptions increase the likelihood of a return. Ishfaq et al. (2016) model and simulate the probability of returns under scarcity and price leadership conditions with data from an e-tailer. They conclude that products with low stock are more likely to be returned, but low stock interacts with the effects of price leadership. In their analysis, price leadership significantly reduces returns. Contrasting empirical findings from most other works, actual sale prices do not affect returns. Apart from that, Cook and Yurchisin (2017) survey young fashion retailer customers. According to the results of a structural equation model, perceived scarcity and low prices increase impulse buying, which relates to negative postpurchase emotions and thus an increased probability of returning. Finally, Calvo et al. (2020) examine the effects of disclosing low stock on sales, returns, and resulting net sales with the data of a flash sales retailer that offers a wide product assortment. Disclosing product availability increases all of the variables mentioned, suggesting that additional sales effects generally outweigh losses from returned goods, while their base return rate equals just 5.4 %.

This points to our contribution: By combining the highlighting of price discounts and their effects on returns, we widen the agenda of Li et al. (2019). They aimed to study the impact of pricing-related web technologies on product returns. We also respond to the call for research by Toktay et al. (2004) and Rogers et al. (2002) to identify factors that influence returns. We complement by investigating the effect of discount disclosure not only on sales but also on consumer returns. This study also aims to resolve the reported inconsistencies regarding the impact of scarcity messages on sales (e.g., Calvo et al., 2020; Park et al., 2020). Furthermore, to check the generalizability of these results, unlike Calvo et al. (2020), who investigate data from an e-tailer with low return rates, we look into the most affected product category: apparel. In addition, we present a quantitative model to substantiate managerial decisions regarding the business case of such measures.

VIII.2.3 Hypothesis development

As consumers do not know a product's initial stock, information about low remaining stock triggers the perception of high demand. This leads to two distinct reactions. First, consumers obtain the impression of high product quality (e.g., due to a low level of information) (Stock & Balachander, 2005). Moreover, consumers conclude that an offer's price is highly attractive. This herding effect, when consumers draw inferences from the behavior of others and try to imitate their behavior regardless of their own information (Banerjee, 1992), was empirically documented by Cui et al. (2019) for retail customers. Second, perceived scarcity urges consumers to decide whether to order in a limited amount of time. This can be attributed to the scarcity effect, which describes the occurrence of buying frenzy by incompletely informed customers in anticipation of out-of-stock situations: Any customer who waits too long will find no more units available (DeGraba, 1995). Lower product availability is then accompanied by the perception of a higher product value making the product more attractive (Cialdini, 2014).
Scarcity also increases the customer's attention to scarce products (Zhu & Ratner, 2015) compared to highly available products, which applies especially in flash sales with mostly restricted stock, leading to more impulse purchases due to higher arousal. Furthermore, fashion is a product category strongly associated with impulse purchases (Bellenger et al., 1978). Therefore, we hypothesize:

- *H1. Displaying product scarcity increases sales in a flash sale e-commerce environment.*

It is well accepted that price savings create an incentive for purchasing (Alba et al., 1999; Devaraj et al., 2002; Kotler, 2008). Price and discount communication frame price evaluations as well as the subsequent purchase decision (Tversky & Kahneman, 1981). Customers implicitly estimate the difference between price quotes in relative terms (approximately 50 percent off) (Monroe, 1973). That is why a price reduction from $50 to

$30 is perceived as more beneficial than a price reduction from $100 to $80, although the absolute discount is identical (Choi & Coulter, 2012). When the value of goods is low, indicating a relative discount promotes the perception of the offer as favorable (Lehtimäki et al., 2019). This applies to the present case, as clothing is generally in a low- to medium-price segment. Besides, the display of a relative discount in addition to recommended retail prices and reduced prices lowers the cognitive effort required to estimate and classify the discount following mental accounting theory (Choi & Coulter, 2012). Therefore, impulsive buying is promoted since the information evaluation phase before the purchase decision is shortened. Consequently, we hypothesize:

- *H2. Displaying relative price discounts increases sales in a flash sale e-commerce environment.*

Herding and scarcity effects drive impulse buying behavior. Impulse purchases are made unintended, unreflective, and immediate, which means that not all consequences of buying are taken into account (Rook, 1987). Liberal return policies support impulse ordering, as they detach the final buying decision from the ordering decision. At the moment of ordering, product valuation is increased by an emotional surplus (Rao et al., 2014). After receiving the product, consumers assess the utilitarian and hedonic performance of the product, which might not meet expectations.

Furthermore, costs can outweigh the benefits after an impulse purchase episode, leading to negative emotions such as regret and dissatisfaction, although the product has no objectively identifiable shortages (Bayley & Nancarrow, 1998). A way out of these negative emotions is returning the product (D'innocenzio). This aligns with Gardner and Rook (Gardner & Rook, 1988), who describe impulse buyers as less happy with their purchases. Furthermore, impulsive purchases are usually less aligned with real tastes and needs (Hong & Pavlou, 2014) due to the ordering pressure. Therefore, we conclude in agreement with (Calvo et al., 2020; Chen et al., 2020; Cook & Yurchisin, 2017; Ishfaq et al., 2016; Rao et al., 2014) that impulse buying behavior triggered by scarcity messages could increase the probability of returning a product:

- *H3. Displaying product scarcity increases returns in a flash sale e-commerce environment.*

Highlighting relative price discounts leads to different, opposite effects. On the one hand, price discounts reduce customer expectations, as a lower quality will be considered acceptable (Zeithaml et al., 1990). On the other hand, high discounts promote impulse purchases with a higher uncertainty regarding quality and fit. Additionally, as justified previously, impulse purchases are more likely to cause negative emotions or regret. We believe that the effect of impulse buying predominates, and therefore, we conclude:

- *H4. Displaying relative price discounts increases returns in a flash sale e-commerce environment.*

VIII.3 Methodology

VIII.3.1 Data source and structure

For our study, we collaborated with a large flash sale e-tailer located in the European Union. The e-tailer provided comprehensive campaign-level real-world data on time-limited discounted sale campaigns. This approach surpasses the most frequent approach in returns management, consumer self-reports: Many empirical problems such as social desirability or recall bias do not apply, as actual purchases are observed. Sample selection bias cannot be completely ruled out, but the collaborating retailer assured us that no particular data set was picked, so we assume this selection is random.

The retailers' assortment mainly focuses on fashion products such as apparel, shoes, and accessories. Sale campaigns are only accessible by a closed community. This means no external customers can be attracted, e.g., via price search engines. Sales are brand-specific and are scheduled in advance. The retailer offers free returns within 14 days with no questions asked.

The dataset used for this study (see Table VIII.1) consists of sales level data of 20 flash sale campaigns in the product categories apparel/shoes/accessories during the year 2019. To reduce noise, we filtered the data to apparel products only, which was the overwhelming majority of products sold. The campaigns were exclusively directed to the German market, the largest national e-commerce market in the European Union.

If a product was sold out during a campaign, reordering and restocking were impossible for the retailer. The 20 campaigns contain 5,826 unique products sold at least once with an initial stock of at least five items. Variants (i.e., different sizes or colors) are each counted as independent products. The average number of items sold for a product equals 4.76, which results in 27,725 items sold with a total order value of 490,992 €. In sum, 6,005 items were later returned (21.66 %).

Variable (Mean / SD)	**Description**
RetPrice (69.08 / 58.77)	Recommended retail price (RRP)
DiscPrice (27.50 / 22.85)	Discounted product price (DP)
AbsDisc (39.11 / 36.27)	Difference between RRP and DP
RelDisc (58.78 % / 7.53 %)	Discount in percent
RelDiscDisp (92.76 % / 25.92 %)	Dummy variable for products with the relative discount displayed (above 50 %)
RelDisc*RelDiscDisp (55.63 % / 16.58 %)	Interaction: value of the displayed discount
ItemsSold (4.76 / 6.29)	Number of sold items per product
ItemsRet (1.03 / 1.60)	Number of returned items per product
ReturnRate (21.66 % / 26.71 %)	Return rate (weighted by ItemsSold)
ProdRev (84.28 / 114.23)	Revenue of a product (before returns)
InitialStock (45.41 / 51.90)	Initial stock level of the product
StockDisp (.18 / .38)	Dummy variable for products with at least 1 item sold disclosing low stock
StockDispProp (.13/.31)	Proportion of sold items per product disclosing low stock
Female (.60/.49) Male (.31/.46)	Dummy variables for female- and male-targeted products (reference: children and others)
CloseFit (.32/.47)	Dummy variable: close-fitted product

Table VIII.1. Dataset description.

VIII.3.2 Operationalization

Methodologically, this paper is based on parametric statistical methods (t-test and multivariate regression). The dependent outcome variables for sales are the number of ordered items (Model 1a) or the revenue of a product before returns (i.e., number of ordered items * discounted price, Model 1b). For returns, we use the relative return rate of a product (Model

2), which is calculated by the number of returned items in proportion to the number of sold items.

Our study aimed to examine two independent factors in detail: displaying information about low stock and the relative price discount. Both informational interventions are designed to be subtle enough to act as digital nudges according to the nudging categories identified by Meske and Potthoff (Meske & Potthoff, 2017). Scarcity messages ("Only x items left") inform the customer about the products' availability without clearly urging customers to order quickly, as there was no message such as "Order now before it gets unavailable". These messages were displayed from the point when the available stock dropped to 5 items. The message contained the actual number of items left in stock, which means that the retailer did not fake low stock. This variable was operationalized by a dummy, indicating whether at least one product was sold when the stock was displayed (i.e., remaining stock < 5). For closer examination, we modeled the proportion of items sold with the stock message displayed to the customer. If a product with an initial stock of 10 items and 7 items sold had 2 items sold with a low stock message displayed, *StockDispProp* is calculated as 2/7 = 0.29, whereas a product with an initial stock of 5 and 4 items sold leads to *StockDispProp*=1.

Displaying the relative discount frames the customer's decision to price aspects and acts as a simplification for customers since they do not need to determine the relative discount themselves. It served as additional promotion as it was only displayed when above 50 %. The recommended retail price and discounted price were always visible to the customer. We operationalized the conditional display as an interaction term of the relative discount and the dummy variable whether the information was shown or not. This term is either 0 or ranges between 0.5 and 1.

First, we controlled for the product price, as the empirical evidence that a higher price induces more returns is consistent (Asdecker et al., 2017; De et al., 2013; Hess & Mayhew, 1997; Sahoo et al., 2018). Second, we controlled for the effects of the product target group (female/male/children and others). Women are supposed to be the customer group with the highest return rate (Asdecker et al., 2017), so we expect women's products to be returned more often as well. Next, we controlled for close-fitted apparel such as lingerie, pajamas, or leggings, as fitting issues are the most important reason for returning an apparel purchase (Hong & Pavlou, 2010). There is still a higher degree of fit uncertainty in close-fitting products because the customer cannot try those on before delivery. Finally, we controlled for a product's initial stock and the absolute discount, which customers can easily estimate since both recommended retail and discounted prices are displayed.

VIII.4 Data analysis results

VIII.4.1 Preliminary analysis

To test our hypotheses, we first used t-tests to compare products for which at least one item was sold with a low stock message and those for which the message was never displayed (Table VIII.2).

	No low stock display	**Low stock display**	**T (df)**	**p**
ItemsSold	4.68	5.11	2.14	.03
	(6.39)	(5.77)	(1649)	
ProdRev	69.93	150.08	15.69	<.01
	(96.18)	(158.77)	(1214)	
ReturnRate	19.28 %	31.66 %	27.76	<.01
	(25.24)	(30.15)	(7212)	
	No discount display	**Discount display**	**T (df)**	**p**
ItemsSold	5.09	4.73	-1.11	.27
	(5.95)	(6.31)	(5824)	
ProdRev	74.02	85.08	2.64	<.01
	(79.50)	(116.47)	(572.71)	
ReturnRate	16.50 %	22.09 %	11.95	<.01
	(20.23)	(27.14)	(2836)	

Table VIII.2. Two-sample t-test analysis.

For the relative return rate, we weighted the data by the number of items sold. Regarding H1, we observe a significant increase in sales in terms of items sold and product revenue. This suggests a substantial influence on the product price and leads to biased results, so in the next step, the price

needs to be controlled for. Regarding H3, the return rate is significantly larger for products bought with a low stock message displayed.
Comparing the groups that show the relative discount or not, an increase in sales is inconclusive (H2). Contrary effects of the number of sold items and product revenue emphasize the need for an in-depth analysis. However, the return rate is significantly higher for products with a displayed discount (H4).

VIII.4.2 Regression analysis of sales

First, we estimated two linear regression models (1a and 1b) with the items sold and product revenue as dependent variables. The independent variables included two nudges (proportion of sold items when low stock is displayed and relative discount display) and all previously introduced controls (Table VIII.3). In general, both models are significant. Concerning the quality of the model, the employed independent variables explained 7.5 % and 25 % of the response variable variance.

	Model 1a (ItemsSold)			**Model 1b** (ProdRev)			**Model 2** (Return-Rate, weighted by items sold)		
Variable	**B(SE)**	**Beta**	**p**	**B(SE)**	**Beta**	**p**	**B(SE)**	**Beta**	**p**
StockDispProp	.75 (.28)	.04	.01	13.22 (4.52)	.04	<.01	3.79 (.62)	.04	<.01
RelDisc* RelDiscDisp	.02 (.01)	.05	<.01	-.13 (.09)	-.02	.15	.02 (.01)	.01	.08
Controls									
DiscPrice	-.07 (.01)	-.25	<.01	1.62 (.12)	.32	<.01	.48 (.02)	.38	<.01
AbsDisc	.02 (.00)	.11	<.01	.76 (.07)	.24	<.01	.01 (.01)	.01	.67
InitialStock	.02 (.00)	.18	<.01	.27 (.03)	.12	<.01	-.01 (.00)	-.03	<.01
Female	-.65 (.32)	-.05	.04	-13.20 (5.21)	-.06	.01	7.25 (.58)	.13	<.01
Male	-.26 (.34)	-.02	.45	-18.71 (5.56)	-.08	<.01	-1.72 (.62)	-.03	.01
CloseFit	-.03 (.19)	.00	.86	13.10 (3.07)	.05	<.01	-.30 (.32)	.01	.34
(Constant)	4.28 (.36)		<.01	13.09 (5.83)		.03	6.94 (.63)		<.01
F (df, p value)	59.90 (8, <.01)			242.19 (8, <.01)			865.05 (8, <.01)		
R^2	.075			.250			.200		

Table VIII.3. Results of the linear regression models.

The findings support H1: *StockDispProp* significantly increases sales, as indicated by Models 1a and 1b. The promotional effect is moderate in comparison to the control variables (Beta = .04). If a product shows the remaining stock for the whole campaign, it increases the number of items sold by 0.75, and the product revenue increases by 13.22€ compared to a product that does not show the scarcity message. Against this, H2 needs to be rejected. Although Model 1a shows a minimal and significant positive effect of the displayed relative price discount, this variable is insignificant in Model 1b and changes direction. Managerially, it can be seen that the display of discounts has a much smaller influence on the order probability than the amount of the discount.

Regarding the control variables, products for women or men are sold less frequently than those for children or an undefined target group. Nevertheless, the differences between the target groups are low and not consistent in Models 1a and 1b. Close-fitted products result in higher revenue, but this variable is not significant for the number of items ordered. The most considerable relative influence is the price and the amount of the discount. The initial stock level correlates with sales because low stock levels limit the maximum possible sales of a product.

VIII.4.3 Regression analysis of returns

For testing H3 and H4, we estimated a linear model with a product's relative return rate as the dependent variable (Model 2 in Table VIII.3). In general, the model is significant and explains 20 % of the return rate's variance. Table VIII.3 presents the results in detail.

Since *StockDispProp* is significant with a positive algebraic sign, H3 is supported: Products with a scarcity message displayed throughout the campaign increase the return rate by 3.79 percentage points. This is in line with H1, which suggests an increase in impulse purchases, leading to a higher level of order regret or more unfulfilled expectations. For this reason, the positive sales effect of this nudge must be carefully weighed against the negative impact due to the associated higher returns. In addition, customers may order a product of different sizes to ensure that one item fits well, especially for highly attractive products with low remaining stock (Asdecker et al., 2017), resulting in an increased likelihood of returns.

The absolute discount, which was not directly displayed to the customer, and the dummy for close-fitted fashion is not significant at the alpha=.05 level. Additionally, the display of a relative discount above 50 % is only significant at the alpha=.1 level. When the relative discount is displayed, the return rate is significantly increased by .02 percentage points, which is a very limited effect. This is not surprising, as no increase in impulse buying behavior can be demonstrated by the discount display (see H2). Another explanation is that flash sale customers are used to high discounts, and this nudge does not change the valuation uncertainty. Moreover, it is worth noting that this type of nudge can prevent customer dissatisfaction, as Peinkofer et al. found that heavily discounted products are more likely to be accepted as sold out (Peinkofer et al., 2015).

The most influential control variable is discounted price (B=.48; Beta=.38). This indicates that the likelihood of returns increases with the price level of an ordered item, replicating the results from empirical returns management literature (e.g., De et al., 2013; Hess & Mayhew, 1997). Based on our data, we quantify that a 10€ higher price increases the relative article-based return rate by approximately 1 percentage point. Regarding the other control variables, products for women have a 7.25 percentage point higher return rate than products for children or others. In comparison, the return rate of products for men is 1.72 percentage points lower. With the slightly lower number of women's products ordered, discounts for men's products can eventually be calculated more tightly from a managerial point. These products are ordered more often and returned less frequently and thus entail less unwanted expenses. In contrast to the models for sales, the absolute amount of the discount is not a significant factor and does not exert a relevant influence on the return rate despite its significance.

VIII.5 Synthesizing the sales and returns perspective: A quantitative model

The preceding data analysis demonstrated that advertising artifacts such as the display of low available stock affect both sales and returns, that is, the marketing and logistics/operations perspective. However, do such measures pay off, or would e-tailers be better off without them? Mollenkopf et al. (2011) emphasized that it is crucial for the success of an e-tailer not to consider only one view in isolation but to integrate both.

To this end, a comparative quantitative model is introduced that allows an objective assessment of whether such measures are beneficial. The analysis is based on the realized contribution margins and draws on the general conditions and relationships described in Asdecker (2015). Table VIII.4 provides an overview of the parameters used.

The model compares two scenarios. In scenario 1, the status quo, the company tries to promote impulse purchases with the measures examined in this paper and accepts the known return rate. In scenario 2, it is assumed that the company forgoes these measures, which reduces the number of returns and orders.

Symbol	Description
TCMSQ	Total contribution margin with electronic marketing measures (e.g., display of low stock) → status quo
TCMNO	Total contribution margin without electronic marketing measures
ORD	Number of ordered items in the planning period in the status quo
ΔORD	Percentage change in ordered items (%) after omission of the marketing measures
RR	Average return rate (%) of an ordered item in the status quo
ΔRR	Percentage change in the average return rate (%) after omission of the marketing measures
P	Average selling price of the ordered items
C	Average wholesale price of the ordered items
RC	Average cost of a returned item
DC	Average distribution costs of the ordered items (e.g., shipping, picking, packing)

Table VIII.4. Model Parameters.

The omission of the advertising measures is beneficial if $TCM_{SQ} < TCM_{NO}$. The total contribution margin in the status quo (TCM_{SQ}) consists of the number of ordered items multiplied by the per item contribution margin, which in turn depends on the share of retained items (1-RR), the retail margin (P-C), the return rate (RR), the return costs (RC), and the distribution costs (DC):

$$TCM_{SQ} = ORD \cdot \left((1 - RR) \cdot (P - C) - RR \cdot RC - DC \right)$$

If the electronic marketing measures are omitted (TCM_{NO}), the percentage changes regarding the number of ordered items (ΔORD) and the return rate (ΔRR) need to be taken into account:

$$TCM_{NO} = ORD \cdot (1 + \Delta ORD) \cdot \left(\left(1 - \left(RR \cdot (1 + \Delta RR)\right) \cdot (P - C)\right) - RR \cdot (1 + \Delta RR) \cdot RC - DC \right)$$

By comparing the two scenarios and solving the inequation $TCM_{SQ} < TCM_{NO}$ for ΔORD, we can now determine the maximum acceptable reduction in demand if omitting the advertising measures reduces the return rate by ΔRR:

$$\Delta ORD > -\frac{RR \cdot \Delta RR \cdot (P - C + RC)}{(P - C) \cdot (1 - RR \cdot (1 + \Delta RR)) - DC - RC \cdot RR \cdot (1 + \Delta RR)}$$

For a specific product on sale, the average return rate is 30.77 %. According to the regression model, not displaying information about a low stock can reduce the return rate by 1.89 percentage points, corresponding to a relative decrease of ΔRR=-6.16 %. Furthermore, the flash sale e-tailer that provided the dataset indicated that the following parameter values appear realistic for their business model: P=27.50 €; C=13.75 € (equivalent to a 50 % commission); RC=5 €; DC=3 €.
For these values, it follows that ΔORD>-6.64 % (i.e., the omission of the displayed information is only justified if the number of orders decreases by a maximum of 6.64 %). The regression for the specific numerical example predicts a decrease in orders by 8.9 %, which indicates that the nudge should remain active. However, if return costs were close to 7.50 €

(see Figure VIII.1 for sensitivity with varying return costs), the decision would have to be reversed.

Figure VIII.1. Sensitivity analysis.

It shows that a decision for or against a measure always has to be made on a case-by-case basis, that it depends on the respective conditions, and that it should not be driven by a functional perspective (sales vs. returns). The model may serve as a basis for a decision support model that objectifies the underlying issue. It mainly fits flash sale e-tailers since no additional costs due to remarketing occur in their business model. Unsold stock is returned to the brand manufacturers. However, simple adjustments can be made to account for these costs as well.

VIII.6 Conclusion

Motivated by the hypothetical but typical situation of a flash sale customer, we addressed whether displaying quantity scarcity and price discounts affect sales and returns and how this impacts profitability. This paper delivers three findings:

- *Displaying low stock messages in a flash sale e-commerce environment promotes sales and increases the rate of returns.*

- *Highlighting the relative price discount neither boosts sales nor does it imply relevant higher returns.*
- *The presented quantitative model can evaluate this trade-off.*

Thus, what is the theoretical contribution of this paper? According to Whetten (1989), contributions arise from four categories: (1) factors to explain a phenomenon (what?), (2) the relationship between those factors (how?), (3) logical justifications for altered views (why?), and (4) conditions that limit generalizability.

We contribute primarily to the factors for sales and returns and their relationship by confirming the positive effect of low stock messages (i.e., quantity scarcity) on sales (H1 accepted) in flash sale e-commerce (Calvo et al., 2020; Cui et al., 2019). Unlike previous contributions, this is the first study to focus on fashion flash sales and the peculiarities of this specific product category. The same contribution type (what and how?) applies to the increase of the relative return rate (H3 accepted).

The results strengthen the recently published study by Calvo et al. (2020) that investigates the effect at a flash sale e-tailer with a broad product assortment from toys to home appliances. They also assess the overall impact on profitability by referring to net sales as a proxy. Net sales are defined as the number of sold items minus returns: "[B]ecause the firm's returns are [...] slightly above 5 % [...], and the cost of a return compares with the margin of a sale [...], the average treatment effect is very large (+12.5 % in net sales) [...]" (Calvo et al., 2020). They furthermore conclude that "[t]he positive effects on [...] profitability amplify over wide assortments [...]." However, as we show, it is necessary to integrate the financial perspective, namely, contribution margins and costs, to draw conclusions about profitability. To our understanding, net sales are a measure with limited suitability for assessing overall profitability. Instead of relying on net sales, we contribute a generalizable quantitative model assessing actual profitability based on cost and revenue parameters to better substantiate the decision-making. In our numerical example, the sales loss forecast by the regression outweighs the savings due to fewer returns. Nevertheless, these results may vary depending on the prerequisites of the dealer, the sales campaign, and the product. By this generalization, we contribute to the "why". Although this research specifically relates to flash sales, the results may also apply to a broader context, including shopping situations that entail a purchase decision under pressure (e.g., Amazon Prime Day or regular "deals of the day").

Furthermore, to the best of our knowledge, this study is the first to investigate the impact of relative discount disclosure on both sales and returns. The results show that the effect of this nudge (H2 and H4 rejected) remains negligible compared to the display of the stock level. This adds to the findings of Luo et al. (2019), who investigated this effect only concerning sales.

Managerially, this research highlights that decision-makers must keep in mind both sides of the medal when designing nudges for e-commerce websites, that is, sales and returns. At first glance, integrating such nudges appears attractive as they are quick and easy to implement at virtually no cost. However, the devil is in the details. High return rates and costs can destroy the desired profitability of an instrument (Hewitt, 2008). The introduced quantitative model can objectify such decisions. The conducted sensitivity analysis shows that it is always case-sensitive, which calls for machine-learning approaches that identify products for which the display or omission of such nudges is beneficial.

Our study is not free of limitations. It is based on data from a single flash-sale e-tailer with data from their German customer community, limiting generalizability. Furthermore, we cannot determine whether the display of the scarcity message exclusively causes the observed effects or if unobserved factors exaggerate these effects. Furthermore, the additional variance explained by the display of low stock or relative discounts remains modest. This hints at a repeated analysis with more detailed data, for instance, by revealing the timing of sales and integrating customer- and product-based attributes for the sales and returns model. The analysis also calls for a replication using data from different countries since German customers are characterized by unique high-returning behavior (Nicola, 2019). The observed return rates (< 25 %) are lower than those found in the literature for "classical" fashion e-commerce (up to 60 %; Foscht et al., 2013). Nevertheless, in combination with the synthesized quantitative model for assessing profitability, we are confident that our findings are helpful for practitioners from e-tailers and other scholars researching this topic.

References

Abdulla, H., Ketzenberg, M. E., & Abbey, J. D. (2019). Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management, 65*(6), 560–605.

Aggarwal, P., Jun, S. Y., & Huh, J. H. (2011). Scarcity Messages: A Consumer Competition Perspective. *Journal of Advertising, 40*(3), 19–30.

Alba, J. W., Mela, C. F., Shimp, T. A., & Urbany, J. E. (1999). The Effect of Discount Frequency and Depth on Consumer Price Judgments. *Journal of Consumer Research, 26*(2), 99–114.

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Asdecker, B., Karl, D., & Sucky, E. (2017). Examining Drivers of Consumer Returns in E-Tailing with Real Shop Data. In *Proceedings of the 50th Hawaii International Conference on System Sciences,* Waikoloa Village, United States of America.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics, 107*(3), 797–817.

Bayley, G., & Nancarrow, C. (1998). Impulse purchasing: a qualitative exploration of the phenomenon. *Qualitative Market Research: An International Journal, 1*(2), 99–114.

Bellenger, D. N., Robertson, D. H., & Hirschman, E. C. (1978). Impulse buying varies by product. *Journal of Advertising Research, 18*(6), 15–18.

Calvo, E., Cui, R., & Wagner, L. (2020). Disclosing Product Availability in Online Retail. *Manufacturing & Service Operations Management*. Advance online publication.

Chen, W.-K., Chen, C.-W., & Lin, Y.-C. (2020). Understanding the influence of impulse buying toward consumers' post-purchase dissonance and return intention: An empirical investigation of apparel websites. *Journal of Ambient Intelligence and Humanized Computing,* 1–14.

Choi, P., & Coulter, K. S. (2012). It's Not All Relative: The Effects of Mental and Physical Positioning of Comparative Prices on Absolute versus Relative Discount Assessment. *Journal of Retailing, 88*(4), 512–527.

Cialdini, R. B. (2014). *Influence: Science and practice* (5. ed., Pearson new international ed.). Pearson.

Cook, S. C., & Yurchisin, J. (2017). Fast fashion environments: consumer's heaven or retailer's nightmare? *International Journal of Retail & Distribution Management, 45*(2), 143–157.

Cui, R., Zhang, D. J., & Bassamboo, A. (2019). Learning from Inventory Availability Information: Evidence from Field Experiments on Amazon. *Management Science, 65*(3), 1216–1235.

De, P., Hu, Y., & Rahman, M. S. (2013). Product-Oriented Web Technologies and Product Returns: An Exploratory Study. *Information Systems Research, 24*(4), 998–1010.

DeGraba, P. (1995). Buying Frenzies and Seller-Induced Excess Demand. *RAND Journal of Economics, 26*(2), 331–342.

Devaraj, S., Fan, M., & Kohli, R. (2002). Antecedents of B2C Channel Satisfaction and Preference: Validating e-Commerce Metrics. *Information Systems Research, 13*(3), 316–333.

D'innocenzio, A. *Holiday sales up, but returns may set a record.* Retrieved June 3, 2021, from https://www.nbcnews.com/id/wbna45673527

Foscht, T., Ernstreiter, K., Maloles, C., Sinha, I., & Swoboda, B. (2013). Retaining or returning? *International Journal of Retail & Distribution Management, 41*(2), 113–134.

Gardner, M. P., & Rook, D. W. (1988). Effects of Impulse Purchases on Consumers' Affective States. *Advances in Consumer Research, 15*(1), 127–130.

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The Dark (Patterns) Side of UX Design. In *Proceeding of the CHI Conference on Human Factors in Computing Systems.*

Hess, J. D., & Mayhew, G. E. (1997). Modeling merchandise returns in direct marketing. *Journal of Direct Marketing, 11*(2), 20–35.

Hewitt, M. (2008). Retailers: time to shout about your returns policy. *Logistics & Transport Focus, 10*(8), 58–61.

Hong, Y., & Pavlou, P. A. (2010). Fit Does Matter! An Empirical Study on Product Fit Uncertainty in Online Marketplaces. In *Proceedings of the 31st International Conference on Information Systems,* St. Louis, United States of America.

Hong, Y., & Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research, 25*(2), 328–344.

Ishfaq, R., Raja, U., & Rao, S. (2016). Seller-induced scarcity and price-leadership. *The International Journal of Logistics Management, 27*(2), 552–569.

Iyer, G. R., Blut, M., Xiao, S. H., & Grewal, D. (2020). Impulse buying: a meta-analytic review. *Journal of the Academy of Marketing Science, 48*(3), 384–404.

Karl, D., & Asdecker, B. (2021). How Does the Covid-19 Pandemic Affect Consumer Returns: An Exploratory Study. In *Proceedings of the 50th European Marketing Academy Conference,* Madrid, Spain.

Kotler, P. (2008). *Principles of marketing* (5. European ed.). Financial Times Prentice Hall.

Lehtimäki, A.-V., Monroe, K. B., & Somervuori, O. (2019). The influence of regular price level (low, medium, or high) and framing of discount (monetary or percentage) on perceived attractiveness of discount amount. *Journal of Revenue and Pricing Management, 18*(1), 76–85.

Li, M., Li, H., Fang, Y., Wang, Y., & Ming, Q. (2019). Pricing-oriented Web Technologies and Product Returns in the E-marketplace: The Moderating Role of Seller Reputation. In *Proceedings of the 40th International Conference in Information Systems,* Munich, Germany.

Luo, X., Lu, X., & Li, J. (2019). When and How to Leverage E-commerce Cart Targeting: The Relative and Moderated Effects of Scarcity and Price Incentives with a Two-Stage Field Experiment and Causal Forest Optimization. *Information Systems Research, 30*(4), 1203–1227.

Meske, C., & Potthoff, T. (2017). The DINU-Model – A Process Model for the Design of Nudges. In *Proceedings of the 25th European Conference on Information Systems,* Guimarães, Portugal.

Mollenkopf, D. A., Frankel, R., & Russo, I. (2011). Creating value through returns management: Exploring the marketing-operations interface. *Journal of Operations Management, 29*(5), 391–403.

Monroe, K. B. (1973). Buyers' Subjective Perceptions of Price. *Journal of Marketing Research, 10*(1), 70–80.

Nicola, S. (2019). *Germany's Returning Customers*. Retrieved November 15, 2020, from https://www.bloomberg.com/news/articles/2019-08-08/germany-s-returning-customers

Park, S., Rabinovich, E., Tang, C. S., & Yin, R. (2020). The impact of disclosing inventory-scarcity messages on sales in online retailing. *Journal of Operations Management, 66*(5), 534–552.

Pavlou, P. A., Liang, H., & Xue, Y. (2007). Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. *MIS Quarterly, 31*(1), 105–136.

Peinkofer, S. T., Esper, T. L., & Howlett, E. (2016). Hurry! Sale Ends Soon: The Impact of Limited Inventory Availability Disclosure on Consumer Responses to Online Stockouts. *Journal of Business Logistics, 37*(3), 231–246.

Peinkofer, S. T., Esper, T. L., Smith, R. J., & Williams, B. D. (2015). Assessing the Impact of Price Promotions on Consumer Response to Online Stockouts. *Journal of Business Logistics, 36*(3), 260–272.

Rao, S., Rabinovich, E., & Raju, D. (2014). The role of physical distribution services as determinants of product returns in Internet retailing. *Journal of Operations Management, 32*(6), 295–312.

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The Returns Management Process. *The International Journal of Logistics Management, 13*(2), 1–18.

Rook, D. W. (1987). The Buying Impulse. *Journal of Consumer Research, 14*(2), 189–199.

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The Impact of Online Product Reviews on Product Returns. *Information Systems Research, 29*(3), 723–738.

Statista. (2021). *Top fashion e-commerce companies by market cap 2020 | Statista*. Retrieved June 15, 2021, from https://www.statista.com/statistics/245648/leading-fashion-e-commerce-companies-by-market-cap/

Stern, H. (1962). The Significance of Impulse Buying Today. *Journal of Marketing, 26*(2), 59–62.

Stock, A., & Balachander, S. (2005). The Making of a "Hot Product": A Signaling Explanation of Marketers' Scarcity Strategy. *Management Science, 51*(8), 1181–1192.

Toktay, L. B., van der Laan, E. A., & de Brito, M. P. (2004). Managing Product Returns: The Role of Forecasting. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. van Wassenhove (Eds.), *Reverse Logistics* (pp. 45–64). Springer.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Weinmann, M., Schneider, C., & vom Brocke, J. (2016). Digital Nudging. *Business & Information Systems Engineering, 58*(6), 433–436.

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution? *Academy of Management Review, 14*(4), 490–495.

Wu, Y., Xin, L., Li, D., Yu, J., & Guo, J. (2021). How does scarcity promotion lead to impulse purchase in the online market? A field experiment. *Information & Management, 58*(1), 103283.

Xia, Y., & Zhang, G. P. (2010). The Impact of the Online Channel on Retailers' Performances: An Empirical Evaluation. *Decision Sciences, 41*(3), 517–546.

Zeithaml, V., Berry, L. L., & Parasuraman, A. (1990). *Delivering quality service: Balancing Customer Perceptions and Expectations*. The Free Press.

Zhu, M., & Ratner, R. K. (2015). Scarcity Polarizes Preferences: The Impact on Choice among Multiple Items in a Product Class. *Journal of Marketing Research, 52*(1), 13–26.

IX Return Policy Leniency Impacting Customers' Purchase Intention – A Viable Strategy for E-Tailers?

Authors: David Karl, Kilian Vornberger, Björn Asdecker

Publication Outlet: Proceedings of the European Marketing Academy

published as:

Karl, D.; Vornberger, K.; Asdecker, B. (2022): Return Policy Leniency Impacting Customers' Purchase Intention – A Viable Strategy for E-Tailers? In: *Proceedings of the 13th European Marketing Academy Regional Conference*, Kaunas, Lithuania.
https://proceedings.emac-online.org/pdfs/R2022-111790.pdf

Abstract

Return policy can reduce e-commerce consumer returns by subjecting high-returning customers to a stricter return policy. Besides return behavior, purchase intention is affected. In an online survey of 197 participants, return policy leniency strongly influences purchase intention. Other variables, such as perceived trust, show a weaker impact on purchase intention than return policy directly. Managerially, this paper improves companies' understanding of how different return policies affect customer behavior. Academically, the research on return policy and purchase intention is complemented by examining three different return policy manifestations under control of trust, fairness, opportunism, and return difficulty.

Keywords: Consumer returns, return policy leniency, purchase intention

IX.1 Introduction

The importance of consumer returns is increasing due to the steady growth of the online B2C market. The measures associated with the COVID-19 pandemic further accelerated the expansion of e-commerce to new firms, business areas, and customers (OECD, 2020). Growing e-commerce also increases consumer returns (Xia & Zhang, 2010). Consumer returns cause high costs (Asdecker, 2015) and impact emissions (Khusainova, 2019).

Nevertheless, to increase customer satisfaction, retailers usually offer return policies (RP) characterized as lenient. Leniency means how conveniently a customer can return an item (Abdulla et al., 2019). We use two of the dimensions developed by Janakiraman et al. (2016): (1) Time leniency (how long can items be returned), and (2) monetary leniency (fees for shipping or returning), and add one more dimension, payment leniency.

RP influences customer behavior. Past research suggests that a generous return policy leads to more returns and orders du to more impulsive purchases (Lantz & Hjort, 2013). About 86 % of customers report that return policies influence their purchase decisions (Olick, 2019). Accordingly, return policy can reduce returns but equally increase sales (Bahn & Boyd, 2014). Abbey et al. (2018) show that a small proportion of customers are responsible for a large share of the total return volume. They advocate categorizing customers according to their return behavior and tailoring return policies accordingly. Nevertheless, how do individualized return policies influence purchase behavior? We aim to improve this understanding, taking into consideration confounding variables, with this research question:

- *How does return policy leniency influence an e-commerce customer's purchase intention?*

IX.2 Literature Background

For B2C returns management in general, we refer to a review by Abdulla et al. (2019), who pointed out return policy as an essential research subject. The following studies have investigated return policy in the context of purchase intention (PI). According to Bonifield et al. (2010), customers exposed to more lenient return policies rate the retailer's quality higher and show increased PI. Hsieh (2013) find that lenient return policies and information credibility negatively impact perceived opportunism and positively impact trust, while perceived opportunism negatively affects trust, which influences stickiness intention positively. Pei et al. (2014) state that return policy positively influences PI and perceived fairness, moderated by a higher reputation or lower competition among e-tailers. Perceived fairness positively affects perceived trust, which in turn has a positive effect on PI. According to Zhang et al. (2017), consumers perceive a return under a lenient policy as easier than under a more strict return policy, and thus, perceived return difficulty and perceived service quality positively influence PI. Oghazi et al. (2018) show that perceived return policy leniency positively influences PI with trust as a mediator, while a direct influence of leniency on PI cannot be confirmed. Wang et al. (2020) conclude that leniency positively affects perceived fairness, perceived return service quality, and repurchase intention; perceived fairness and perceived service quality also positively impact consumers' repurchase intention.

The selected papers cover most of the return policy dimensions identified by Janakiraman et al. (2016) and suggest a direct influence of return policy on PI while uncovering other indirect relationships. However, leniency dimensions are primarily examined in separate studies or as different variables. We consider the return policy integrally and integrate the payment dimension, which has not been mentioned in this research strand so far. E.g., paying by credit card elicits less pain than paying in cash because payment is decoupled from the timing of consumption (Prelec & Loewenstein, 1998). Garnefeld et al. (2017) show that payment after receiving the goods increases returns compared to payment before delivery. Based on these considerations, we investigate the RP's influence on PI, perceived trust, fairness, return difficulty, and their interrelationships.

IX.3 Hypotheses

In e-commerce, information asymmetries exist because physical distance causes uncertainties, and customers cannot evaluate items before purchase. Retailers can reduce asymmetries through a signal (Spence, 2002). According to Signaling Theory (Spence, 1973), signaling describes a signal sent by an agent observable by a principal to reduce pre-contractual information asymmetries (Kirmani & Rao, 2000). Since return policy leniency acts as an information mechanism in the relationship between online retailers and customers (Wang et al., 2020), this could reduce information asymmetries: For example, lenient return policies signal customers being able to act flexibly because they can avoid costs of a wrong purchase decision (Wood, 2001). Thus, leniency could positively influence PI. Pei et al. (2014), and Wang et al. (2020) support this assumption. Therefore, we hypothesize:

- *H1: Customers' purchase intention is positively associated with return policy leniency.*

According to Equity Theory, perceived fairness results from the ratio between profit and investment in an exchange (Adams, 1965). Concerning Procedural Justice Theory as part of Equity Theory, people are interested in fair distribution and fair processes (Lind & Tyler, 1988). People prefer their own advantage or positive inequality (Bower & Maxham, 2012). We assume that customers value fair treatment and prefer a customer-friendly return policy. Customers feeling mistreated are less likely to shop at a retailer in the future and vice versa (Bower & Maxham, 2012). Pei et al. (2014) and Wang et al. (2020) indicate that return policy leniency positively influences perceived fairness, promoting PI. Accordingly, we hypothesize:

- *H2a: Customers' perceived fairness is positively associated with return policy leniency.*
- *H2b: Customers' purchase intention is positively associated with perceived fairness.*

Trust is crucial to reduce uncertainties in e-commerce (Hsieh, 2013). Trust is the willingness of a party to expose itself to the actions of a second party, anticipating that the second party will fulfill the expectations of the first party without control (Mayer et al., 1995). According to Agency Theory, in a relationship between two or more economic entities in which a principal instructs an agent to perform a service, information asymmetries exist between buyers and sellers (Jensen & Meckling, 1976). By reducing incomplete information through a deliberate signal, higher trustworthiness could be achieved (Spence, 2002). Return policy leniency could represent this kind of signal. Oghazi et al. (2018) and Hsieh (2013) show a relationship between perceived trust and return policy. Therefore, we hypothesize:

- *H3a: Customers' perceived trust is positively associated with return policy leniency.*

A lack of trust can harm attitudes toward e-commerce (McKnight et al., 2002). Conversely, Kim and Peterson (2017) show that trust promotes PI. Pei et al. (2014) and Oghazi et al. (2018) confirm this relationship for return policies. Accordingly, it seems essential to foster trust for increasing future purchases. We hypothesize:

- *H3b: Customers' purchase intention is positively associated with perceived trust.*

Adherence to fairness positively impacts trust (Bies & Tripp, 1995; Pei et al., 2014). Accordingly, a signal of fairness can reduce information asymmetries, leading to increased trust (Waterman & Meier, 1998). We hypothesize in the context of trust:

- *H3c: Customers' perceived trust is positively associated with perceived fairness.*

In internet-based exchange relationships, online retailers may behave opportunistically (Liang et al., 2005). Opportunistic behavior describes the lack of honesty as well as pronounced self-interest in transactions (Williamson, 1975). In a buyer-seller relationship, the seller puts his own goals above the buyer's benefit (Hsieh, 2013). Information asymmetries between buyer and seller facilitate opportunistic behavior (Mishra et al., 1998; Waterman & Meier, 1998). Hsieh (2013) shows that a lenient return policy contributes to mitigating perceived opportunism. Accordingly, this study conjectures that return policy leniency can counter perceived opportunism:

- *H3d: Customers' perceived opportunism is negatively associated with return policy leniency.*

Li et al. (2006) argue that credible behavior is perceived as reliable, but unmet expectations damage trust. Opportunistic behavior can be understood as an unmet expectation. Moreover, the retailer is assumed to behave opportunistically (Eisenhardt, 1989; Mishra et al., 1998). The signal sent to the customer to reduce information asymmetries may also be harmful (Connelly et al., 2011). Thus, this study assumes that opportunistic behavior harms trust. Li et al. (2006) and Hsieh (2013) describe a negative relationship between opportunism and trust. Consequently, we hypothesize:

- *H3e: Customers' perceived trust is negatively associated with perceived opportunism.*

Perceived return difficulty is the customer's perceived inconvenience in returning an item to receive a refund (Zhang et al., 2017). Both return depth and return time impact the perceived return difficulty. For example, if customers perceive a potential return as difficult, they perceive an increased risk of unpredictable costs. Since customers tend to avoid wrong decisions preventively (Mitchell, 1999), we hypothesize:

- *H4a: Customers' perceived return difficulty is negatively associated with return policy leniency.*
- *H4b: Customers' purchase intention is negatively associated with perceived return difficulty.*

IX.4 Methodology

IX.4.1 Survey Description

For data collection, this research used an online survey conducted in February 2021. The questionnaire consists of three parts. Before the actual questionnaire, a virtual cover letter informs the participants about the survey's background and assures them anonymity. Next, we queried essential characteristics of the respondents. In the central part, each participant goes through two scenarios. The participants are asked to imagine purchasing an item from a fictitious online fashion retailer and to answer several items on PI, perceived fairness, perceived trust, perceived return difficulty, and perceived opportunism. Scenario 1 is balanced characterized by neither particularly strict nor lenient return policy elements. Scenario 2 involves one of three randomly assigned manipulations, i.e., either a strict, balanced, or lenient scenario (Garnefeld et al., 2017; Lantz & Hjort, 2013; Raghubir & Srivastava, 2008; Wood, 2001) (Table IX.1).

Scenario	Strict	Balanced	Lenient
Shipping costs	Yes	Yes	No
Return costs	Yes	No	No
Payment period	Immediate	14 days	30 days
Return period	14 days	30 days	100 days

Table IX.1. Randomly assigned return policies.

1,214 participants started the survey, of which 302 subjects completed the questionnaire (24.9%). After removing 105 samples due to missing return experience or inconsistent responses, the final sample consists of 197 participants, almost all from Germany. The average age is 29.4; 58.9% had at least a college degree, 71.6% were female. The average completion time was 6.5 minutes. Regarding the gender imbalance in our sample, no significant differences for the mean and variance of the PI were observed.

IX.4.2 Manipulation Check

A one-factor ANOVA checks the manipulation by the scenarios. In addition, post hoc tests provide information about which groups differ from each other, using the mean values of PI. The Levene test indicates that equality of variance between the groups can be assumed ($p>.05$). Significant differences in the mean values exist between all groups ($F=193.345$; $p<.001$). The Bonferroni posthoc test and the Scheffé procedure confirm the manipulation functionality. Thus, the subjects show a significantly different PI depending on the scenario.

IX.4.3 Operationalization of Constructs

We tested the hypothesized relationships using structural equation modeling (SEM) to integrate multiple exogenous and endogenous latent and manifest variables (Ullman & Bentler, 2013). The focal constructs of our study, namely purchase intention (PI), perceived fairness (FA), perceived opportunism (OPP), perceived trust (TR), and return difficulty (DI), were operationalized with multi-item scales. We adopted them from existing studies showing statistical validity and reliability of these constructs (Table IX.2). All items were measured on a 5-point Likert type scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree").

Construct	Source	Items used	Reliability	AVE	Squared Multiple Correlation
PI	Kukar-Kinney et al. (2007); Wang et al. (2020)	6	.98	.87	.945
FA	Pei et al. (2014); Kukar-Kinney et al. (2007)	4	.93	.77	.622
OPP	Hsieh (2013)	3	.88	.71	.319
TR	Hsieh (2013)	4	.95	.83	.674
DI	Jarvenpaa et al. (2000)	3	.96	.90	.542

Table IX.2. Measurement scales and summary statistics.

IX.4.4 Reliability and Validity Check

To check the unidimensionality of the item structure, we conducted an exploratory factor analysis for each construct (principal axis analysis and Promax). As a measure of sample adequacy, the Kaiser-Meyer-Olkin criteria of each construct all show values >.6 (Kaiser & Rice, 1974). Bartlett's test can be rejected for all constructs ($p < .001$), indicating data fit for analysis (Dziuban & Shirkey, 1974). Two items showed a communality <.5 and were not further considered. The results of the individual explorative factor analyses confirm the one-dimensionality of the constructs. Cronbach's alpha indicates high reliability on the construct level (Table IX.2).
We conducted a confirmatory factor analysis for parameter estimation to ensure reliability and validity based on the second-generation quality criteria. Since no construct correlation is >.9, no parameter is excluded. Indicator reliability for all items is >.4, so we assume acceptable reliability (Bagozzi & Baumgartner, 1994). Reliability at the construct level is determined by factor reliability. Factor reliability exceeds .6 for all constructs, confirming construct reliability. Since all constructs have an AVE>.5, we assume convergence validity (Fornell & Larcker, 1981). We assume construct validity for the reflective measurement models, as the requirements for discriminant validity are met according to the Fornell/Larcker criterion.

IX.5 Results and Discussion

The SEM was estimated by the maximum-likelihood method (Table IX.3, Figure IX.1). The indices of the measurement model show an acceptable fit. All coefficients except for two are significant. 95% of the PI variance is explained by the model (Table IX.2) The standardized coefficients of DI to PI and TR to PI are significantly <.2, while all other standardized coefficients exceed this threshold for meaningfulness (Chin, 1998).

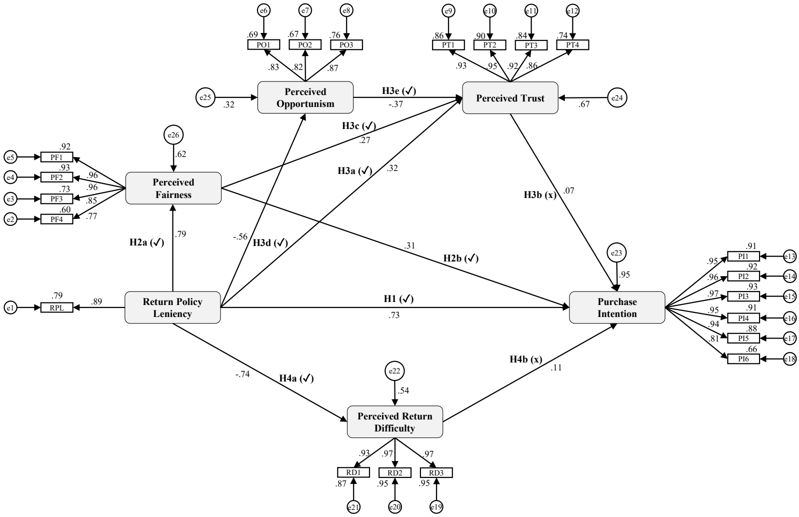

Figure IX.1. Research Model with Factor Loadings.

The largest significant positive coefficient in the model indicates that a more lenient return policy increases PI (1.323; p<.001). The data accordingly support H1. We suggest return policy leniency to signal quality and thus to reduce purchase decision conflict (Wood, 2001), which confirms the results of Pei et al. (2014) and Wang et al. (2020). The results also show that the more lenient the return policy, the fairer the customer feels treated (1.287; p<.001). In addition, higher perceived fairness positively affects PI (.341; p<.001). Thus, H2a and H2b are supported and confirm the findings of Pei et al. (2014) and Wang et al. (2020). H3a is confirmed by the data (.377; p=.003), supporting the research findings of Hsieh (2013) and Oghazi et al. (2018). Thus, lenient return policies appear to build trust. However, H3b and the results of Pei et al. (2014) and Oghazi et al. (2018) that trust positively affects PI cannot be confirmed due to a slightly positive but insignificant effect (.111; p>.05). H3c is supported (.194; p=.003) in agreement with the results of Pei et al. (2014): Customers seem to repay fair treatment with trust in the online retailer. Furthermore, the data confirm (-.643; p<.001) that a more lenient return policy makes the customer perceive less opportunism from the online retailer. Moreover, we found that perceived opportunism significantly reduces perceived trust (-.376; p<.001). Consequently, H3d and H3e are supported, consistent with Li et al. (2006) and Hsieh (2013). Perceived return difficulty decreases significantly as the return policy becomes more lenient (-1.179; p<.001). Thus, H4a is supported, confirming the results of Zhang et al. (2017). We cannot confirm the postulated negative effect of the perceived return difficulty on PI. Contrary to the conjecture, the coefficient is positive but significant (.125; p=.021). The data do not support H4b, which contradicts the research of Zhang et al. (2017).

Hypothesis	**Path**			**Coefficient**	**SE**	**CR**	**Sign.**	**Conclusion**
H1	PI	←	RP	1.323	.152	8.682	<.001	Support
H2a	FA	←	RP	1.287	.115	11.175	<.001	Support
H2b	PI	←	FA	.341	.066	5.132	<.001	Support
H3a	TR	←	RP	.377	.126	2.985	.003	Support
H3b	PI	←	TR	.111	.081	1.375	.169	Reject
H3c	TR	←	FA	.194	.065	2.967	.003	Support
H3d	OPP	←	RP	-.643	.088	-7.34	<.001	Support
H3e	TR	←	OPP	-.376	.07	-5.406	<.001	Support
H4a	DI	←	RP	-1.179	.102	-11.603	<.001	Support
H4b	PI	←	DI	.125	.054	2.301	.021	Reject
Fit indices: χ^2=368,447, df=180, χ^2/df=2.047, GFI=.836, CFI=.965, RMSEA=.074								

Table IX.3. Path coefficients and results of hypothesis tests.

IX.6 Conclusion, Contribution, and Future Research

In summary, return policy leniency strongly influences PI and, at the same time, affects other variables, which influence PI partly and with smaller effect sizes. Return policy thus represents an instrument for influencing customer behavior not only regarding return behavior but rather pre-purchase. A lenient return policy can increase trust and the fairness perceived by the customer. In turn, it reduces the perceived opportunism and the perceived difficulty of a consumer return and can thus contribute to higher customer satisfaction. On the downside, as suggested by Abbey et al. (2018), individual and strict return policies can discourage unwanted customers already from purchasing.

This paper extends previous research on consumer return policy leniency by a more holistic approach integrating time, costs, and payment modalities, rather than focussing on individual parts of return policies. Moreover, this study formulates three different return policies and thus breaks the previous dichotomous view. Using SEM, we incorporate several influencing variables, which have already partially been investigated in this context.

From a managerial point of view, this study supports e-tailers in understanding the interdependencies between return policy and PI as well as other factors important to this relationship. For reducing consumer returns, individual return policies cannot be implemented without taking PI and other variables into consideration. Following the approach of individually adjusting the return policy of customers with excessive returns (Abbey et al., 2018), retailers must balance these trade-offs to determine the suitable level of leniency and the critical thresholds. This study reveals that a stricter return policy can significantly reduce future purchases, allowing to manipulate the structure of the customer base in a smoother way than closing down customer accounts (Safdar & Stevens, 2018). Vice versa, individually adjusted, more lenient conditions might increase future revenues of low-returning customers.

Nevertheless, the results hint at some future research required. A longitudinal study could validate the results in a non-pandemic context. In addition, our sample is restricted to the European market. Furthermore, we examine only two of the five return policy dimensions identified by Janakiraman et al. (2016). Overall, an integrated analysis of return policy effects on actual purchases and returns would supplement the findings of our study.

References

Abbey, J. D., Ketzenberg, M. E., & Metters, R. (2018). A More Profitable Approach to Product Returns. *MIT Sloan Management Review, 60*(1), 71–74.

Abdulla, H., Abbey, J. D., & Ketzenberg, M. E. (2019). How Consumers Value Retailer's Return Policy Leniency Levers: An Empirical Investigation. *SSRN Electronic Journal*. Advance online publication.

Adams, J. S. (1965). Inequity In Social Exchange. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2, pp. 267–299). Academic Press.

Asdecker, B. (2015). Returning mail-order goods: Analyzing the relationship between the rate of returns and the associated costs. *Logistics Research, 8*(1), 1–12.

Bagozzi, R. P., & Baumgartner, H. (1994). The Evaluation of Structural Equation Models and Hypothesis Testing. In R. P. Bagozzi (Ed.), *Principles of Marketing Research* (pp. 386–422). Blackwell.

Bahn, K. D., & Boyd, E. (2014). Information and its impact on consumers ′ reactions to restrictive return policies. *Journal of Retailing and Consumer Services, 21*(4), 415–423.

Bies, R. J., & Tripp, T. M. (1995). The Use and Abuse of Power: Justice as Social Control. In R. Cropanzano & M. K. Kacmar (Eds.), *Organizational Politics, Justice, and Support: Managing Social Climate at Work* (pp. 131–145). Greenwood Publishing Group.

Bonifield, C., Cole, C., & Schultz, R. L. (2010). Product returns on the Internet: A case of mixed signals? *Journal of Business Research, 63*(9), 1058–1065.

Bower, A. B., & Maxham, J. G. (2012). Return Shipping Policies of Online Retailers: Normative Assumptions and the Long-Term Consequences of Fee and Free Returns. *Journal of Marketing, 76*(5), 110–124.

Chin, W. W. (1998). Commentary: Issues and Opinion on Structural Equation Modeling. *MIS Quarterly, 22*(1), 7–16.

Connelly, B., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling Theory: A Review and Assessment. *Journal of Management, 37*(1), 39–67.

Dziuban, C. D., & Shirkey, E. C. (1974). When Is a Correlation Matrix Appropriate for Factor Analysis? Some Decision Rules. *Psychological Bulletin, 81*(6), 358–361.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *The Academy of Management Review, 14*(1), 57–74.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39–50.

Garnefeld, I., Feider, L., & Boehm, E. (2017). Managing the Necessary Evil: Can Payment Methods Reduce Product Returns. In R. Chandy, J. Inman, & C. Moorman (Eds.), *Proceedings of the 28th Winter AMA Conference* (Vol. 28, F-30-F-31).

Hsieh, P.-L. (2013). Perceived opportunism (PO) in e-return service encounters. *Managing Service Quality: An International Journal, 23*(2), 96–110.

Janakiraman, N., Syrdal, H. A., & Freling, R. (2016). The Effect of Return Policy Leniency on Consumer Purchase and Return Decisions: A Meta-analytic Review. *Journal of Retailing, 92*(2), 226–235.

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an Internet store. *Information Technology and Management, 1*(1/2), 45–71.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics, 3*(4), 305–360.

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement, 34*(1), 111–117.

Khusainova, G. (2019). *There Is No Such Thing As A Free Return*. Retrieved July 25, 2021, from https://www.forbes.com/sites/gulnazkhusainova/2019/03/28/there-is-no-such-thing-as-a-free-return/

Kim, Y., & Peterson, R. A. (2017). A Meta-analysis of Online Trust Relationships in E-commerce. *Journal of Interactive Marketing, 38*(1), 44–54.

Kirmani, A., & Rao, A. R. (2000). No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality. *Journal of Marketing, 64*(2), 66–79.

Kukar-Kinney, M., Xia, L., & Monroe, K. B. (2007). Consumers' perceptions of the fairness of price-matching refund policies. *Journal of Retailing, 83*(3), 325–337.

Lantz, B., & Hjort, K. (2013). Real e-customer behavioural responses to free delivery and free returns. *Electronic Commerce Research, 13*(2), 183–198.

Li, D., Browne, G. J., & Wetherbe, J. C. (2006). Why Do Internet Users Stick with a Specific Web Site? A Relationship Perspective. *International Journal of Electronic Commerce, 10*(4), 105–141.

Liang, H., Laosethakul, K., Lloyd, S. J., & Xue, Y. (2005). Information Systems and Health Care-I: Trust, Uncertainty, and Online Prescription Filling. *Communications of the Association for Information Systems, 15*(1), 41–60.

Lind, E. A., & Tyler, T. R. (1988). *The Social Psychology of Procedural Justice*. Springer Science & Business Media.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review, 20*(3), 709–734.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research, 13*(3), 334–359.

Mishra, D. P., Heide, J. B., & Cort, S. G. (1998). Information Asymmetry and Levels of Agency Relationships. *Journal of Marketing Research, 35*(3), 277–295.

Mitchell, V.-W. (1999). Consumer perceived risk: conceptualisations and models. *European Journal of Marketing, 33*(1/2), 163–195.

OECD. (2020). *E-Commerce in the time of COVID-19*. Retrieved March 25, 2021, from http://www.oecd.org/coronavirus/policy-responses/e-commerce-in-the-time-of-covid-19-3a2b78e8/

Oghazi, P., Karlsson, S., Hellstrom, D., & Hjort, K. (2018). Online purchase return policy leniency and purchase decision: Mediating role of consumer trust. *Journal of Retailing and Consumer Services, 41*(1), 190–200.

Olick, D. (2019). *The surge in online-shopping returns has boosted the warehouse sector*. Retrieved June 15, 2021, from https://www.cnbc.com/2019/01/04/surge-in-e-commerce-gift-returns-has-boosted-the-warehouse-sector.html

Pei, Z., Paswan, A., & Yan, R. (2014). E-tailer's return policy, consumer's perception of return policy fairness and purchase intention. *Journal of Retailing and Consumer Services, 21*(3), 249–257.

Prelec, D., & Loewenstein, G. (1998). The Red and the Black: Mental Accounting of Savings and Debt. *Marketing Science, 17*(1), 4–28.

Raghubir, P., & Srivastava, J. (2008). Monopoly Money: The Effect of Payment Coupling and Form on Spending Behavior. *Journal of Experimental Psychology: Applied, 14*(3), 213–225.

Safdar, K., & Stevens, L. (2018). *Banned From Amazon: The Shoppers Who Make Too Many Returns*. Retrieved August 20, 2021, from https://www.wsj.com/articles/banned-from-amazon-the-shoppers-who-make-too-many-returns-1526981401

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(3), 355–374.

Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *The American Economic Review, 92*(3), 434–459.

Ullman, J. B., & Bentler, P. M. (2013). Structural Equation Modeling. In I. B. Weiner (Ed.), *Handbook of Psychology* (pp. 661–690). Wiley.

Wang, Y., Anderson, J., Joo, S.-J., & Huscroft, J. R. (2020). The leniency of return policy and consumers' repurchase intention in online retailing. *Industrial Management & Data Systems, 120*(1), 21–39.

Waterman, R. W., & Meier, K. J. (1998). Principal-Agent Models: An Expansion? *Journal of Public Administration Research and Theory, 8*(2), 173–202.

Williamson, O. E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications: A Study in the Economics of Internal Organization*.

Wood, S. L. (2001). Remote Purchase Environments: The Influence of Return Policy Leniency on Two-Stage Decision Processes. *Journal of Marketing Research, 38*(2), 157–169.

Xia, Y., & Zhang, G. P. (2010). The Impact of the Online Channel on Retailers' Performances: An Empirical Evaluation. *Decision Sciences, 41*(3), 517–546.

Zhang, J., Li, H., Yan, R., & Johnston, C. (2017). Examining the Signaling Effect of E-tailers' Return Policies. *Journal of Computer Information Systems, 57*(3), 191–200.

UNIVERSITY OF BAMBERG
CHAIR OF OPERATIONS MANAGEMENT & LOGISTICS

Online shopping has experienced substantial growth, with the COVID-19 pandemic further accelerating this trend. Since e-commerce customers cannot physically assess products online, returned items are part of the e-commerce business model. Consumer returns are associated with various operational challenges, are costly for retailers and impact the environment due to the additional transport emissions and resource waste they produce. Thus, retailers must actively manage returns in advance to secure profitability and limit their CO2 footprint. Various data sources (e.g., transaction data) enable retailers to generate helpful information to support these tasks.

Based on these considerations, this dissertation aims to generate data-driven insights into the field of consumer returns. First, an in-depth understanding of the phenomenon is gained through the documentation and exploration of the status quo of consumer returns in Germany. Second, data-driven analytic and predictive approaches are reviewed, applied, and evaluated. Third, influencing factors for consumers' purchase and return behavior are subsequently examined.

With a total of nine papers published in national and international journals or conference proceedings, this dissertation addresses these issues from a scientific perspective, without losing sight of practical applicability, thus contributing to a better understanding of data analysis in the context of consumer returns management.