ESSAYS ON CONSUMER RETURN POLICY DESIGN

A Dissertation

by

HUSEYN ABDULLA

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michael Ketzenberg
Co-Chair of Committee,	James D. Abbey
Committee Members,	Gregory R. Heim
	Christina Kan
Head of Department,	Rich Metters

May 2022

Major Subject: Business Administration

ABSTRACT

Due to the ongoing and dramatic growth in the volume of consumer returns, retailers struggle with the trade-off in returns service strategies between implementing stricter return policies to lower operational costs and environmental footprint versus providing customers with lenient return policies to positively stimulate customers' value perceptions and patronage intentions. This dissertation contributes to knowledge by providing theoretical and practical insights on managing this trade-off. In particular, the dissertation offers three essays on consumer return policy design. The first essay reviews and classifies the inter-disciplinary and multi-method research on consumer return policy design through a holistic conceptual framework and identifies relatively under-explored as well as unexplored research areas. The second essay investigates, through randomized experiments, how return policy leniency across five leniency levers available to retailers affects consumers' purchase intentions and proposes a causal mechanism to explain these effects. The results suggest that monetary and exchange levers are the most effective levers in influencing purchase intentions, whereas time, scope, and effort levers are significantly less so. Further, the findings suggest that perceived service quality and perceived transaction costs in parallel and perceived service value in series mediate the effect of return policy leniency across the levers on purchase intentions. The third essay examines how restrictive changes to long-established lenient return policies impact consumer trust in retailers and the resultant favorable behavioral intentions, and how managerial transparency moderates this impact. Results from randomized experiments suggest that restrictive changes to long-established lenient return policies generally result in decreased consumer trust in the retailer and favorable behavioral intentions. This negative effect becomes stronger in the severity of the restriction. However, managerial transparency in the form of communicating the rationale for the change can help to mitigate this negative effect.

DEDICATION

In loving memory of my brother Ali Abdulla.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisors, Dr. James D. Abbey and Dr. Michael Ketzenberg, for their excellent mentorship and guidance throughout my doctoral education. They helped me not only develop the skills set to succeed in the academic field, but also remain resilient and focused while going through challenging times. I hope that I can be a great mentor like my advisors in the future.

I would like to thank Dr. Gregory R. Heim and Dr. Christina Kan for their excellent service as a member in my doctoral committee and for their insightful and constructive comments on my dissertation. I am looking forward to collaborating with them on many more research projects in near future. I would like to also extend my gratitude to Dr. Rogelio Oliva for providing me valuable opportunities for collaboration in research and helping me to broaden my vision as an academic researcher.

I want to express my sincere gratitude to Dr. Rich Metters for his leadership as the Department Head and his efforts to make the life of doctoral students more enjoyable. Furthermore, I want to thank Ms. Donna Shumaker, Ms. Tammy Louther, Ms. Theresa Ralston, Ms. Veronica Stilley, and the rest of the department for the help I got so freely from them.

Many thanks to my Ph.D. program colleagues at the department for always being so helpful in many matters. Special thanks go to Han K. Oh, Seulchan Lee, and Mayukh Majumdar, with whom I walked the way through the Ph.D. program for five years.

Finally, I would like to express my deepest gratitude to my wife Leyla, my parents Khalida and Museyib, and my brother Vagif who have always been sources of inspiration and motivation for me. Their unconditional love and support helped me to overcome all the hurdles, and they never let me feel alone in this journey.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Michael Ketzenberg [advisor], James D. Abbey [co-advisor], Gregory R. Heim of the Department of Information and Operations Management and Professor Christina Kan of the Department of Marketing at University of Connecticut.

The first and second essays in this dissertation have been published as original research articles at the Journal of Operations Management and Production and Operations Management, respectively. The essays have benefited from feedback provided by the editors and anonymous reviewers who handled them.

Funding Sources

Graduate study was supported by fellowships from Mays Business School and the Department of Information and Operations Management at Texas A&M University. There are no outside funding contributions to acknowledge related to the research and compilation of this document.

NOMENCLATURE

RP	Return Policy
RM	Return Management
CLSC	Closed-Loop Supply Chain
CB	Consumer Behavior
PE	Planning & Execution
FV	Product Fit and Valuation Uncertainty
CH	Customer Heterogeneity
TC	Transaction Costs
TP	Trial Period Consumption
RT	Return Timing
DR	Demand and Return Sensitivity
RA	Risk Aversion
CR	Cognitive Responses
AR	Affective Responses
BI	Behavioral Intentions
BA	Behavioral Actions
G	Money Back Guarantee Adoption
M	Monetary Leniency
T	Time Leniency
F	Effort Leniency
S	Scope Leniency
X	Exchange Leniency

O	Overall Leniency
PR	Pricing
I	Inventory
SC	Supply Chain Coordination
FD	Forward Channel Design
C	Competition
PC	Product Characteristics
PA	Product Assortment
A	Acquisition
P	Processing
D	Disposition
CDF	Cumulative Distribution Function
MBG	Money Back Guarantee
CT	Cognitive Trust
PD	Product Dissonance
AA	Assortment Attractiveness
PRPF	Perceived Return Policy Fairness
PVRP	Perceived Value of Return Policy
PSQ	Perceived Service Quality
PPQ	Perceived Product Quality
PRD	Perceived Return Difficulty
ED	Emotional Dissonance
LI	Loyalty Intention
PI	Purchase Intention
WTPR	Willingness-to-Pay for Return Option

RI	Return Intention
OI	Order Intention
KI	Keep Intention
WTPP	Willigness-to-Pay for Product
OF	Order Frequency
AVO	Average Value of Order
AVPI	Average Value of Purchased Items
AVRI	Average Value of Returned Items
TVO	Total Value of Orders
SV	Sales Volume
OP	Overall Profitability
MVRP	Monetary Value of Return Policy
PTC	Perceived Transaction Costs
PSV	Perceived Service Value
TCE	Transaction Cost Economics
OM	Operations Management
EFA	Exploratory Factor Analysis
CFA	Confirmatory Factor Analysis
CI	Confidence Interval
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Consumer returns have become an intrinsic element of the U.S. retail industry. The annual value of returned products in the U.S. has grown to approximately $400 billion, representing more than 10% of total sales (National Retail Federation, 2021). Product proliferation, heterogeneity in customer expectations and valuations, and the rise of e-commerce can all be listed among the drivers for this growth (Cheng, 2015). Yet another key contributor of the dramatic increase in the volume of consumer returns is the trend of offering generous return policies. Fierce competition, decreased consumer switching costs, and a “Customer is the King” mantra lead many retailers to offer overly lenient return policies, bearing returns related operational costs that exceed $100 billion per year (Blanchard, 2007). Overly lenient return policies also engender moral hazard issues and result in the rise of fraudulent and opportunistic returns, which recently surpassed $24 billion per year (National Retail Federation, 2021). Given the rapid growth of consumer returns and their corresponding importance in the retail marketplace, academic research in the area of consumer return policy design is of significant practical relevance.

Return policy design poses an interesting operations-marketing interface problem. From the marketing perspective, more leniency can stimulate customer purchases through positive product and service quality signaling and a reduction of purchase related risks, thereby enhancing the consumer value proposition. From the operations perspective, return policies constitute an important strategic lever in the front-end of reverse supply chains that influences the volume, timing, and quality of commercial product returns (Guide & Van Wassenhove, 2009) The choice of return policies, through documented influences on consumer purchase and return behaviors, is also intertwined with and carries significant implications for other retail planning and execution activities ranging from product pricing to assortment decisions to inventory management. From this perspective, the reciprocal relationship between managerial decisions with respect to return policy design and the resultant consumer cognitive, affective, and behavioral reactions also constitutes a significant problem of behavioral operations management (Donohue et al., 2018).

This dissertation offers three essays that contribute to the academic literature on consumer return policy design. In the first essay (§2), entitled "Taking Stock of Consumer Returns: A Review and Classification of the Literature," I provide an in-depth review of the multi-disciplinary and multi-method research on consumer return policy design in order to identify generalizable insights and critical research needs. The second essay (§3) titled "How Consumers Value Retailer's Return Policy Leniency Levers: An Empirical Investigation," provides the first joint assessment of the relative effectiveness of five return policy leniency levers that are available to retail managers in terms of their influence on consumer perceptions with respect to return service and the resultant purchase intentions. Finally, in the third essay (§4) titled "Restrictive Changes to Long-Established Lenient Return Policies and Consumer Reactions to Them," I empirically examine the psychological contract violation and negative signaling effects of restrictive changes imposed on long-established lenient return policies—a recent trend in the retail industry. In particular, I study how restrictive changes to return policies and severity of the restrictions influence consumer trust perceptions toward a retailer and resultant behavioral intentions. I also examine how managerial transparency, operationalized as an explicit communication of the decision rationale, moderates the potential negative effects of such decisions.

Theoretically, I draw upon multiple theories of psychology and behavioral economics and postulate process models (i.e., mediation, moderation, and conditional processes) to explain the cognitive-affective mechanisms that transmit the causal effect of managerial decisions with respect to return policy design on consumer responses. Methodologically, I employ various experimental designs (i.e., full-factorial between-subjects and pre-test, post-test control group designs) that allow identification of causal effects, and I use diverse U.S.-based consumer samples recruited through online crowd-sourcing platforms to collect experimental data. I perform rigorous measurement scale development and construct validation procedures to define latent variables in the process models and use the state-of-the-art regression-based mediation, moderation, and conditional process analysis techniques to test the main research hypotheses.

2. TAKING STOCK OF CONSUMER RETURNS: A REVIEW AND CLASSIFICATION OF THE LITERATURE*

2.1 Introduction

Consumer returns continue to grow and recently surpassed $643 billion globally (Cheng, 2015) and $400 billion in the US alone (National Retail Federation, 2021). Academics have taken notice of the increase and significance of returns as evidenced by a recent surge in research on the topic. While research on returns extends back several decades in the economics literature on money back guarantees and product warranties (Heal, 1977), the preponderance of the recent research largely has shifted to the disciplines of operations and marketing. Because of the rapid growth, the state of the literature regarding consumer returns remains unclear. To this end, we systematically review the literature and classify the contributions according to a holistic conceptual framework—a framework that clarifies the state of the research and also identifies theoretical gaps and opportunities for future research.

Research on consumer returns examines a wide range of managerial issues. The return policy establishes restocking fees, length of time allowed to make returns, channel restrictions, and more. Further, operational planning and execution activities (e.g., inventory management, product pricing, and assortment planning) influence return policy decisions or are themselves influenced by return policies. Of course, research on consumer returns also involves the management of the returned products themselves, which includes their acquisition, processing, and disposition. Return policies also influence consumer perceptions and behaviors. The consumer is intimately involved with purchase and return processes. As such, the sociological and psychometric factors that influence and explain consumer behavior play a significant role in setting policy and developing managerial insights. Consequently, the validity and relevance of research in this area necessitates a strong empirical foundation. Yet, as will become clear in the comprehensive review, relatively

*Reprinted with permission from "Taking Stock of Consumer Returns: A Review and Classification of the Literature" by Huseyn Abdulla, Michael Ketzenberg, James D. Abbey, 2019. *Journal of Operations Management, 65(6), 560–605, Copyright (2019) by John Wiley & Sons - Books.*

little is empirically established regarding consumer behavior in relation to return policies.

The review process began with an identification of search terms (consumer returns, return policy, money-back guarantee) with a resulting database (Google Scholar, ScienceDirect, Web of Science, JSTOR) search and abstract evaluation. To be included in the review, papers had to consider the return policy decisions of a retailer or manufacturer making direct sales. Any works not meeting the return policy decision criterion, such as those solely focused on returns from retailers to manufacturers due to buyback contracts or those that only addressed process aspects of consumer returns, were excluded. At this stage, each paper meeting the inclusion criteria also required an additional search of all works cited (i.e., citation snowballing), which led to the final set of 100 works spanning multiple decades. In sum, our review's scope includes, to the best of our knowledge, all research contributions that have been published until the end of 2018, as peer reviewed journal articles, along with any working papers cited therefrom, that specifically address 1) managerial decision-making related to return policies or 2) consumer behavior in response to such decision-making. Our inclusion criteria for the working papers are that 1) they should not be active (i.e., not undergoing peer-review but publicly available on Social Science Research Network) and 2) they should be cited in published works and therefore have influenced the subsequent research in the area. In total, the papers included in our review have been published in 43 different academic journals from various disciplines, including operations management, operations research, information systems, marketing, economics, and management.

We begin in §2.2 by introducing a conceptual framework on consumer returns. In §2.3, we transform the conceptual framework into a classification framework. We then proceed to our review and classification in §2.4 and §2.5, organized separately according to the natural split that arises between analytical modeling and empirical papers. In §2.6 we conclude with a comprehensive discussion that integrates the insights from the conceptual framework, summarizes the state of the literature, characterizes the ongoing momentum of the domain, and reveals opportunities for future research.

2.2 Conceptual Framework

Investigating the anatomy of a purchase and return transaction serves as a starting point to develop a conceptual framework for classifying and reviewing the literature. Figure 2.1 provides a simplified view of the purchase and return process, from pre-purchase to post-return, and distinguishes epochs by the consumer's purchase and return decisions. Figure 2.1 is notable for providing a clear, integrated, and parallel view of the return process from both consumer and retailer perspectives, emphasizing the significant role of consumer behavior in managing consumer returns. Chircu and Mahajan (2006) conceptualize the retail transaction as a sequence of steps, including store access, search, evaluation and selection, ordering, payment, order fulfillment, and post-sales service. Each of these steps has associated costs: price-type costs (credit charges, taxes, etc.), time-type costs (waiting time, delivery time), or psychological-type costs (perceived ease of use, convenience, frustration, annoyance, anxiety, etc.).

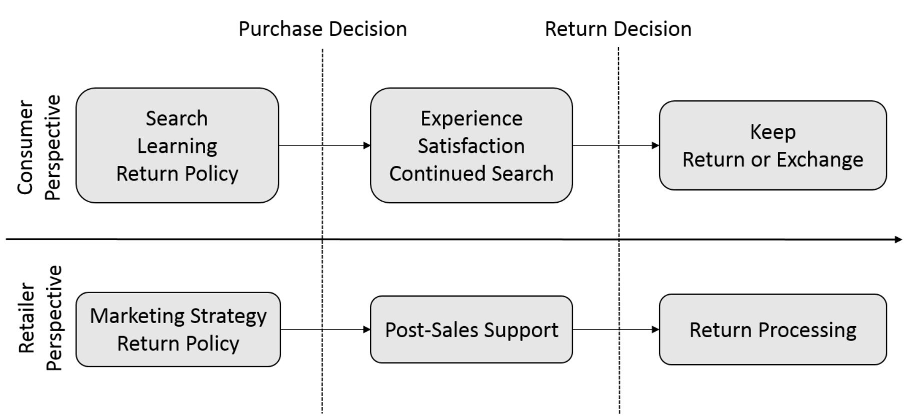

Figure 2.1: Anatomy of a Purchase and Return Transaction

Customers evaluate a return policy in terms of convenience for product evaluation, hassles for potential returns, learning through product experience during the return window, and the possibility of a need for continued search. Returns arise due to a multitude of factors that include uncertainties regarding product fit and valuation, along with product defects and opportunistic behavior, among others. Even pre-purchase, a retailer's return policy can have a demonstrable effect on consumer behavior. More than 70% of online consumers consider return policies before making a purchase decision (Su, 2009a). Managerial decision-making with respect to the return policy, along with many other operational considerations, may affect consumer behavior at each step of the process.

From the retailer's perspective there are distinct sets of decisions and activities to be managed pre-purchase, post-purchase, and post-return. Pre-purchase, the retailer must design a return policy, provide information, set pricing, and perform all the other operational planning activities necessary for retail execution. Of course, retailer performance also continues in the context of post-sales support that includes customer relationship management, warranty services, and physical management of returned products.

Collectively, from both the consumer and retailer perspectives, four broad and inter-related domains of research emerge: return policy, consumer behavior, planning and execution, and return management. These four domains are illustrated graphically in Figure 2.2 to form a conceptual framework. Each domain appears as a circle with interrelationships among domains denoted by their overlapping regions. We proceed by defining and characterizing the scope of research that pertains to each of the four domains, beginning with the focal return policy domain.

The return policy (RP) domain pertains to research that designs, describes, and prescribes return policies in the context of managerial decision-making. Return policies are generally characterized in terms of their leniency, which refers to the convenience and ease with which consumers are allowed to make returns. Some retailers have remarkably lenient return policies that allow any return for any reason, at any time, and provide a full refund of the price paid. Less lenient retailers impose restrictions, such as restocking fees, limitations on the allowable time to return, or even outright denial of returns for certain products. Research on return policy design investigates return policy leniency in various forms and analyzes the operating conditions that influence leniency choices.

The return management (RM) domain concerns the efficient and effective acquisition, processing, and disposition of returns. RM falls under the umbrella of closed-loop supply chain manage-

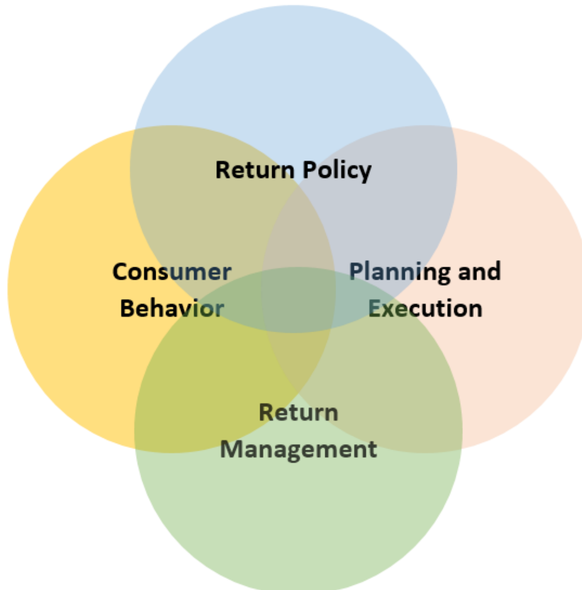

Figure 2.2: Conceptual Framework

ment (CLSC). This considerably larger body of literature is concerned with the management of product returns at their end-of-life and end-of-use, in addition to consumer returns. For reviews of the CLSC literature, we refer the interested reader to Atasu et al. (2008), Souza (2013), and Govindan et al. (2015). A retailer's return management infrastructure and practices are likely to influence return policy decisions, and vice-versa, a retailer who views its lenient return policy as an important part of the value proposition seeks ways to reduce the cost burden of returns through effective return management strategies. Thus, the inquiries in the intersection of RP and RM domains promise actionable insights for retail managers.

The consumer behavior (CB) domain explicitly considers the consumer decision-making process, with respect to both sales and returns. The primary motivation behind the study of consumer behavior in reaction to return policies is the conjecture that sales, as well as returns, can be influenced by return policy decisions. In general, research in the intersection of CB and RP domains either 1) empirically tests relationships between return policy leniency and various constructs that reflect various cognitive, affective, and behavioral responses or 2) provides return policy prescriptions by utilizing analytical models in which consumer behavior are operationalized through specific modeling constructs.

The operational planning and execution (PE) domain captures how returns influence forward logistics and supply chain management. Our review and classification highlight several distinct areas of managerial decision-making, such as pricing, inventory, channel, assortment, competition, coordination, and quality. Moreover, decisions and actions may be taken to either reduce returns or mitigate related costs. Whereas RM domain research addresses decision-making regarding management of returns in the reverse supply chain, PE domain research addresses decision-making regarding returns in the forward supply chain. In general, the research in the overlap of RP and PE domains studies interplay between return policy and other operational decisions in a broad variety of settings.

Collectively, the four domains of RP, RM, CB, and PE combine to form a conceptual framework that can meaningfully differentiate, and at times integrate, research contributions within the domain of consumer returns. From this perspective, the framework provides a foundation to classify the literature that specifically addresses consumer return policies and thereby serves as a lens to understand the state of the field.

As mentioned previously, we limit the scope of our classification solely to the focal RP domain but do include all of the intersections with the other three domains. The scope limitation is critical because incorporating literature beyond the return policy focus reveals a massive and unwieldy body of literature that would only allow a cursory examination of the literature, prohibiting a deeper understanding of the state of the research. Even with the scope restriction, this work is considerable in depth and breadth as the coming sections will demonstrate. Of course, there are many papers falling outside our scope that nonetheless make significant contributions to the study of consumer returns phenomenon. For example, there is a large body of literature on refurbishing and remanufacturing of consumer returns (Abbey & Guide Jr, 2018). This literature fits squarely in the domain of return management and closed loop supply chain management but does not involve managerial decision-making related to consumer return policies. Moreover, there are other published reviews that address this literature such as Atasu et al. (2008) and Souza (2013). In addition, there is a significant amount of published research that focuses on returns processing but does not address return policy decisions. A prime example is Griffis et al. (2012), which empirically investigates how a return experience and refund processing speed may influence post-return customer behavior.

2.3 Transforming Concept to Classification

Our objective is to review and classify the literature according to the conceptual framework. A classification will help to inform the state of the literature, clarify research contributions, and identify gaps, conflicts, and future research opportunities. This necessitates identification of various classifiers that capture key modeling choices to distinguish the contributions in the literature. In essence, the classifiers serve to transform the conceptual framework into a classification framework. A classifier plays the role of an indicator for whether a modeling construct appears within the analytical or empirical model of a given paper. Hence, the classifiers identify the modeling elements corresponding to the various domains related to consumer behavior (CB), return policy (RP), planning and execution (PE), and return management (RM). We use Su (2009a) and Bonifield et al. (2010) — two well-known and often cited contributions, one analytical, one empirical — as representative examples from the literature to facilitate the discussion that introduces and illustrates our classifiers. We introduce our classifiers in Table 2.1 for each of the four domains in our framework. Table 2.1 also presents the classification of Su (2009a) and Bonifield et al. (2010). The presence [absence] of a check mark signifies that the respective analytical and empirical models in Su (2009a) and Bonifield et al. (2010) incorporate [do not incorporate] modeling constructs corresponding to the classifier indicated by the row. We first briefly summarize Su (2009a) and Bonifield et al. (2010) and then proceed to discuss the set of classifiers for each domain, beginning with the RP domain in §3.1.

Su (2009a) examines the performance implications of return policies in a supply chain of a manufacturer and a retailer. The key contribution of this study from a return policy leniency perspective is to show that full refunds that are commonly observed in practice may be overly generous

	Classifiers	Description	Su (2009a)	Bonifield et al. (2010)
CB	Product Fit and Valuation Uncertainty (FV)	Probability that a product fits a customer's needs and ex-post realized monetary value of the purchased product to the customer.	✓	
	Customer Heterogeneity (CH)	Presence of distinct segments within a customer base in terms of product fit, valuation, hassle costs, etc.	✓	
	Transaction Costs (TC)	Customer's monetary, time, and psychological costs of returning.	✓	
	Trial Period Consumption (TP)	Using a product during a trial period and returning a fitting product after partial consumption based on expected utility of doing so.		
	Return Timing (RT)	Time between purchasing and returning.		
	Demand and Return Sensitivity (DR)	Sensitivity of aggregate demand and return volume to a retailer's decisions.		
	Risk Aversion (RA)	Customer's degree of risk aversion.		
	Cognitive Responses (CR)	Perceptive reactions resulting from a deliberate thought process (e.g., perceived value, risk, quality).		✓
	Affective Responses (AR)	Emotional reactions to situations (e.g., regret, anger, liking).		
	Behavioral Intentions (BI)	Self-reported likelihood of engaging in certain behaviors (e.g., purchase intention, return intention, stated willingness-to-pay).		✓
	Behavioral Actions (BA)	Actual outcomes of cognitive-affective responses and behavioral intentions (e.g., purchase, return, referral, manifested willingness-to-pay).		✓
RP	Money Back Guarantee Adoption (G)	Binary decision with respect to offering an MBG.	✓	
	Monetary Leniency (M)	The amount of refund or restocking fees.	✓	
	Time Leniency (T)	The length of time window allowed for returns.		
	Effort Leniency (F)	Degree of hassles imposed by a retailer.		
	Scope Leniency (S)	Degree to which a uniform return policy applies to various products/categories.		
	Exchange Leniency (X)	Whether a retailer offers cash refund or only store credit.		
	Overall Leniency (O)	An overall assessment of leniency.		✓
PE	Pricing (PR)	Retailer's product price decision.	✓	
	Inventory (I)	Retailer's inventory management decisions.	✓	
	Supply Chain Coordination (SC)	Contractual mechanisms to coordinate the supply chain.	✓	
	Forward Channel Design (FD)	Retailer's optimal sales channel decisions.		
	Competition (C)	Competition between two or multiple differentiated retailers.		
	Product Characteristics (PC)	Intrinsic product attributes such as quality, modularity, etc.		✓
	Product Assortment (PA)	Retailer's product mix decisions.		
RM	Acquisition (A)	Retailer's end-of-use product collection and reverse channel design decisions.	✓	
	Processing (P)	Retailer's returns handling, restocking, recovery, etc.	✓	
	Disposition (D)	Retailer's returns disposition options (e.g., salvaging, returning to manufacturer, selling in a secondary market).	✓	

Table 2.1: Classification Framework

and not optimal for a retailer. Instead, the author finds that the optimal refund amount should be equal to the salvage value of a returned product. In the presence of a wholesale contract between the retailer and the manufacturer, Su (2009a) shows that a partial-refund policy improves the profits of both parties relative to a full-refund policy. Next, the author studies the impact of consumer returns under different supply chain contractual mechanisms. He proves that a buy-back contract with a manufacturer may induce the retailer to adopt an overly lenient return policy. The author proposes several strategies to rectify the incentive alignment problem and achieve coordination between retailer and manufacturer.

Bonifield et al. (2010) investigate the relationship between overall leniency of a retailer's return policy and the retailer's quality through two empirical studies. In the first study, the authors content analyze the return policies of e-tailers chosen randomly from an online quality rating website, and show that for non-consumable product categories, the ratings of e-tailers are positively associated with the leniency of return policies the retailers offer. In a follow up study, the authors show that perceived control of the online shopping experience and cognitive trust for an e-tailer moderate the relationship between return policy leniency and purchase intention.

2.3.1 Return Policy

There are seven classifiers for the RP domain. The first of these classifiers, MBG adoption, identifies modeling instances that investigate a retailer's binary choice between allowing returns with a full refund or disallowing them. The next five classifiers correspond to five different dimensions or levers of return policy leniency that were originally introduced as a cohesive typology of leniency by Janakiraman et al. (2016). We also require a classifier to capture the instances in which an overall degree or perception of leniency is of concern. The seven classifiers are defined in Table 2.1. Su (2009a) investigates both MBG adoption (G) and monetary leniency (M) in the form of a refund to the customer when making a return. Bonifield et al. (2010) consider overall leniency (O) of return policies that capture leniency across multiple dimensions.

2.3.2 Consumer Behavior

We introduce distinct sets of classifiers to classify the analytical and empirical literature for ontological reasons. Specifically, the classifiers for the analytical literature are generated based on the constructs that are fundamental in modeling an individual customer's utility or market demand function — common operationalizations of consumer behavior in the analytical sense. This set of classifiers includes: 1) product fit and valuation uncertainty, 2) customer heterogeneity, 3) transaction costs, 4) trial period consumption, 5) return timing, 6) demand and returns sensitivity, and 7) risk aversion. The classifiers for the empirical literature are based on four key components of consumer behavior according to the theory of buyer behavior: 1) cognitive responses, 2) affective responses, 3) behavioral intentions, and 4) behavioral actions (Howard & Sheth, 1969; Bagozzi, 1982). We next discuss the analytical classifiers in the context of Su (2009a) and follow with a discussion of the empirical classifiers in the context of Bonifield et al. (2010).

Analytical Classifiers: Product fit uncertainty reflects the risk that a product may fail to deliver expected quality and performance or may not fit individual customer tastes and pre-purchase expectations. Valuation uncertainty relates to how much a customer values a product, typically in terms of utility, after a purchase. In the specific example of Su (2009a), customers face valuation uncertainty that is only resolved after purchase. The CH classifier serves to segment the analytical literature based on assumptions regarding the ex-ante (prior to purchase) characteristics of customers. Homogeneity refers to the case that all customers are identical in their intrinsic characteristics. Conversely, heterogeneity corresponds to cases in which customers have certain differentiating characteristics. In the case of Su (2009a), customers are assumed to be homogeneous in the base case but heterogeneous in a model extension where product valuation uncertainty is the differentiating customer characteristic. The TC classifier corresponds to the presence of time, psychological, and monetary costs incurred by a customer making a return. These costs are often referred to as hassles in the context of the return process and includes things such as filling out forms, requiring identification, and other similar costs. In the specific example of Su (2009a), the author captures return hassle with a cost parameter.

Though the remaining four classifiers are not present in Su (2009a), each represents an important aspect of research into RP. First, the TP classifier corresponds to product value consumed during the return window. This is an important element to modeling opportunism which concerns the moral hazard behavior of customers who purchase items with the full ex-ante intention of returning them. Ostensibly, the opportunistic customer is able to extract enough value from the product during the retailer's allowable return window such that the extractable value exceeds the costs associated with purchasing and returning the item. Second, the RT classifier corresponds to whether a model specifically addresses the length of time a customer will keep a product before making a return. Third, the DR classifier corresponds to whether, at an aggregate level, demand and return volumes are influenced by a return policy. This aggregate level approach contrasts with modeling customer utility at the individual level as in Su (2009a). At the aggregate level, utility is not modeled directly. Rather, demand and return functions are generally expressed in simple linear form with respect to return policy leniency. Finally, the RA classifier relates to risk attitude and captures instances in which customers or sellers are assumed to be risk averse, as opposed to risk-neutral, which is the case in Su (2009a).

Empirical Classifiers: The CR classifier addresses cognitive responses that include perceptions and reactions that are formed based on a conscious, deliberate thought process given available information. These may include perceptions related to risk, quality, value, effort, control, among others. For example, Bonifield et al. (2010) study 'perceived control of website navigation' and 'trust' to e-tailer as moderators of the relationship between return policy leniency and 'perceived e-tailer quality'. The AR classifier is predicated on emotional reactions to the objects, which may include feelings such as 'dissonance', 'liking', 'satisfaction', 'regret', and so on. Bonifield et al. (2010) do not consider any construct related to affective responses. The BI classifier relates to constructs that are based on self-reported likelihood of engaging in certain behaviors, whereas the BA classifier relates to actual, observed behavioral outcomes. In the case of Bonifield et al. (2010), the authors use customer ratings for e-tailers from an online website in the first, secondary data-based study. These customer ratings correspond to behavioral actions. In the second experimental study, the focal outcome is purchase intention, which is captured by the BI classifier.

2.3.3 Planning & Execution

The classifiers in the PE domain capture various operational planning and executional issues that are studied in the context of return policy decision-making. Our classification framework includes a set of seven classifiers. Most of these classifiers are self-explanatory. Pricing (P) concerns the determination of optimal product price. Inventory (I) addresses policies related to the management of items held in inventory and used to satisfy demand, including both the timing and quantity of replenishment, as well as storage, transportation, and handling. Supply chain coordination (SC) addresses the choice of contractual mechanisms between a retailer and a supplier or manufacturer in order to optimize supply chain performance. Forward channel design (FD) concerns the choice of one or more channels for distribution and sales, such as online and brick-and-mortar. Competition (C) addresses return policy design in the context of competing retailers or where a retailer competes with a direct-to-consumer manufacturer. Product assortment (PA) concerns the selection of products to make available for sale. Finally, product characteristics (PC) correspond to innate product characteristics such as quality, modularity, durability, etc. Overall, Su (2009a) serves as a representative modeling instance for P, I, and SC classifiers, recognizes the C and PA classifiers as future research opportunities, but does not consider the PC classifier. Bonifield et al. (2010) show that the relationship between return policy leniency and quality ratings of e-tailers are based on whether the products that the e-tailers carry are largely consumable or non-consumable. In essence, the moderating role of whether products are consumable or non-consumable corresponds to an instance of PC classifier, which is the only PE related classifier for which Bonifield et al. (2010) has an instance.

2.3.4 Return Management

The return management (RM) domain utilizes three classifiers that identify modeling instances of three aspects of return management: 1) acquisition (A), 2) processing (P), and 3) disposition (D). The A classifier concerns processes and decision-making with respect to collecting returns from customers. The P classifier concerns the processing of returns once acquired. Processing may be as simple as inspecting returned items for damage or may involve more intensive handling such as repackaging or refurbishing. Finally, the D classifier concerns the ultimate disposition of returned items, such as returning to the manufacturer, selling as open-box, reselling as new, etc. Su (2009a) addresses all three RM classifiers, whereas Bonifield et al. (2010) do not address any of the RM classifiers.

To conclude, this section introduces a set of classifiers that effectively transforms the conceptual framework into a classification framework. We use Su (2009a) and Bonifield et al. (2010) as representative examples to illustrate the operationalization of the framework. We now proceed to use the classifiers to organize our classification and review of the analytical (§4) and empirical (§5) literature. Two researchers conducted a binary coding of each paper based on whether there is an instance of a classifier in the paper. The initial inter-coder agreement rate was 87%. The cases of disagreements were then resolved by inclusion of a third coder where the final classification status is determined through majority consensus (Durach et al., 2017). The complete classification of both empirical and analytical literature, in an alphabetical order, appears in the Appendix.

2.4 Analytical Research

Analytical research on consumer return policies (RP) addresses the optimal return policy in a variety of operational settings and develops managerial insights. Such research commonly explores the operating conditions in which lenient or restrictive return policies are optimal. Leniency affects consumer behavior that influences both firm-level and typically supply chain-wide performance. The optimal decisions regarding a return policy often are made jointly with other operational decisions that may impact planning and execution as well as return management. Hence, all four domains of the framework are collectively integral to modeling research in this area. The classification and review of the analytical literature first examines the methods and approaches taken to model consumer behavior (§4.1). Next, §4.2 discusses findings regarding the RP classifiers and is then accompanied by the classification and review of contributions in the intersections of RP and PE (§4.3) as well as RP and RM domains (§4.4).

2.4.1 Consumer Behavior

This section examines the literature in the light of each CB classifier, starting with product fit and valuation uncertainty, which is then followed by a discussion of the remaining six classifiers. Though more technical in nature, the coming discussion highlights the key modeling approaches with respect to the operationalization of each classifier.

2.4.1.1 Product Fit and Valuation Uncertainty

The primary function of a return option is to reduce the risk associated with purchases. The main sources of risk are 1) product quality and performance and 2) customer tastes and pre-purchase expectations. In general, the analytical literature operationalizes this risk by two parameters: fit uncertainty and valuation uncertainty. Fit uncertainty is captured by the probability that the product does not fail, has a satisfactory level of quality, or perfectly matches the customer's taste and pre-purchase expectations. Valuation uncertainty is expressed in terms of a random variable whose value is realized post-purchase and corresponds to the monetary value (reservation price) of the purchased product to the customer. Multiple approaches are present in the literature to model fit and valuation uncertainty, largely determined by how the two parameters are related. To systematize these different approaches, we use a common set of notation motivated by the modeling approach in Yalabik et al. (2005).

Let us denote the probability of product fit for a given customer with p. The existing literature refers to this parameter with a variety of names such as probability of a working product, probability of a match, likelihood of satisfaction, etc. We will use probability of product fit for consistency. Next, let V be a random variable with cumulative distribution function (CDF) $F(v)$, whose realization corresponds to the customer's post-purchase valuation (i.e., reservation price) of the product. The conditional CDF of the customer's valuation conditioned on the probability of fit is denoted with $F(v|p)$. Finally, let $\bar{v}$ be the upper limit of the possible realizations of V.

There are three general approaches to model product fit and valuation uncertainty. First, a stream of studies assumes $F(v|p)$ follows a Bernoulli distribution, where $V = \bar{v}$ with probability

p and $V = 0$ with probability $1 - p$. Hence, a customer either receives a deterministic level of valuation if the product is a match or a zero valuation otherwise. Some contributions following this approach are Davis et al. (1995), Moorthy and Srinivasan (1995), Fruchter and Gerstner (1999), Hsiao and Chen (2012), and the base model in Yalabik et al. (2005).

The second approach assumes that $F(v|p) = F(v)$, in which the valuation of product follows a general distribution that is independent of the fit probability p. The studies that follow this assumption usually exclude the probability of product fit as a parameter. Also, a special case of this approach is to assume $F(v|p) = {}^{v}/_{\bar{v}}$, meaning that valuation follows a general uniform distribution (i.e., $V \sim U[0, \bar{v}]$) or a standard uniform distribution in the case of $\bar{v} = 1$. Su (2009a), Akçay et al. (2013), Huang et al. (2014), Altug and Aydinliyim (2016), and J. Chen and Chen (2016) use a general distribution function for valuation, whereas Davis et al. (1998) make the standard uniform distribution assumption to derive key insights.

The third approach separates fit uncertainty from valuation uncertainty in a similar vein to the independence approach but retains both of the parameters for explicit use in the model. These studies assume that with probability p the product is a perfect fit for the customer and that product valuation follows a uniform distribution (i.e., $V \sim U[0, \bar{v}]$), whereas with probability $1 - p$ the product is not a fit and the customer's valuation is zero. This assumption allows for the fact that a product may perfectly match customer tastes and expectations, however, the customer may still have zero valuation for the product. For example, such a situation can occur if the product is no longer needed by the time the product is received. Using this approach, Hess et al. (1996), Chu et al. (1998), and Ülkü et al. (2013) assume that product valuation follows a general uniform distribution, whereas J. Chen and Bell (2012), McWilliams (2012), B. Chen and Chen (2017a), B. Chen and Chen (2017b), and Yang et al. (2017) assume a standard uniform distribution.

A vast majority of studies assume that the return policy does not have a direct impact on a customer's post-purchase valuation of the transaction. Yet, exceptions exist, such as Ülkü et al. (2013) and Ülkü and Gürler (2018), that assume consumer valuation is distributed uniformly over $[0, \bar{v}]$, where $\bar{v}$ is equal to the sum of intrinsic valuation and positive extrinsic valuation received from the length of return time window set by the retailer. A similar approach is taken in Xu et al. (2015), where the incremental valuation received from return policy time leniency is added to both lower and upper bounds of the uniformly distributed intrinsic valuation parameter. Overall, this approach facilitates the assumption that customers are more likely to buy a product if the return window is longer. In J. Chen and Bell (2012) and J. Chen and Chen (2016), the authors model valuation uncertainty as the product of $V \sim U[0,1]$ and a parameter $0 < \theta < 1$ that captures the degree that a customer accepts a non-returnable channel (i.e., impact of a no-returns policy on valuation).

Finally, yet another approach to modeling valuation uncertainty holds that product valuation is captured as a sum of ex-ante known and ex-post realized components. For examples, see Q. Liu and Xiao (2008), Shulman et al. (2009), Y.-J. Chen (2011), and Inderst and Tirosh (2015).

2.4.1.2 Customer Heterogeneity

In general, analytical studies characterize the customer base by assuming that it is either homogeneous or heterogeneous. The homogeneity assumption occurs in the majority of the studies and signifies that all customers are identical in their behaviors. In contrast, literature assuming heterogeneity uses a variety of approaches in defining the nature of a heterogeneous customer base.

Moorthy and Srinivasan (1995) appear to be the first to distinguish between homogeneous and heterogeneous customer bases. In the homogeneous case, customers product valuations follow the same Bernoulli distribution as described earlier. To reflect customer heterogeneity, the authors assume that the market is segmented according to a uniform distribution $\Omega \sim U[0,1]$, and $\Omega = \omega$ denotes the segment. Then, customer product valuations within each segment are given by ωv, where v is the reservation value of a working product. Fruchter and Gerstner (1999) model customer heterogeneity in terms of both product fit probability (in the base model, with a unique p_i defined for each customer) and valuation (using two customer segments with high and low product valuations).

Several recent studies use the term heterogeneous customers to highlight the assumption that customers hold uncertain valuations of the product drawn from the same probability distribution

(J. Chen & Bell, 2012; McWilliams, 2012; Ülkü et al., 2013). Yet, many other studies do not label mere valuation uncertainty based on a single distribution as heterogeneity. To illustrate, in their base models, Su (2009a) and Akçay et al. (2013) make a similar assumption as the papers above regarding customer valuation uncertainty but label such cases as homogeneous.

Overall, we find that segmenting the customer base into low and high types in terms of *ex-post* product valuation is a commonly used representation of heterogeneity in the literature (Su, 2009a; Hsiao & Chen, 2012; Akçay et al., 2013; Hsiao & Chen, 2014, 2015). We can refer to this representation as imperfect ex-post heterogeneity as customers are still homogeneous within each segment. Additionally, there are a few studies which represent ex-post customer heterogeneity in terms of a unique distribution function for product valuation of each individual customer (Anderson, Hansen, & Simester, 2009; Y.-J. Chen, 2011; Zhang, 2013). Under this assumption, we can say that customers are perfectly heterogeneous in terms of ex-post valuation. In addition to heterogeneity in ex-post valuation, several authors use low and high types to segment customers based on their return hassle costs (Fruchter & Gerstner, 1999; Hsiao & Chen, 2012).

Several studies assume that customers can also be *ex-ante* heterogeneous on various dimensions. Fruchter and Gerstner (1999) capture ex-ante heterogeneity in terms of individual probabilities of satisfaction with a purchase (i.e., product fit probability). Shulman et al. (2009) segment the customer base based on an individual taste parameter and use another parameter to reflect the likelihood that customers in a given segment know their product valuation ex-ante. Shang, Ghosh, and Galbreth (2017) assume that ex-ante, the customer base consists of ordinary and opportunistic customers. Finally, Xu et al. (2018) assume that the customer base consists of two segments. The first segment includes customers that have a higher return likelihood when they observe others who return, whereas the second segment is immune to this effect.

2.4.1.3 Transaction Costs

Table 2.2 lists the articles that explicitly include transaction costs in their models (unless otherwise stated, the papers presented in all tables follow a chronological order). The 2.2 identifies the type of transaction cost modeled and whether it is used as a parameter or decision variable. The table also reveals that considerable differences exist in the conceptual definitions of transaction costs. From a modeling standpoint, transaction costs are expressed either as a parameter or as a decision variable, with the majority of the papers falling into the former category. To the best of our knowledge, Davis et al. (1998) and Hsiao and Chen (2014) are the only two studies that model hassle cost as a decision variable, whose level is to be determined by the retailer. Several studies consider heterogeneity in hassle costs within the customer base, either based on a high- and low-type segmentation (Fruchter & Gerstner, 1999; Hsiao & Chen, 2012, 2014), or based on individual hassle cost for each customer (Anderson, Hansen, & Simester, 2009; Y.-J. Chen, 2011), whereas the majority of the studies assume homogeneous transaction costs for returning.

2.4.1.4 Trial Period Consumption

Lenient return policies allow customers to purchase and try products during a specified time window and return the product if it does not fit or meet expectations. Trial period consumption occurs when customers extract a fraction of the product value prior to returning the product, which is key to modeling opportunism. A modeling framework addressing opportunism also needs to allow for the case that the product may be returned by a customer even when the product perfectly matches the customer's expectations. The customer may act opportunistically and return a perfectly matching product after some use if returning it for a refund provides more utility than the residual consumption value of the product beyond the trial period. Overall, this is the most common modeling paradigm to deal with opportunistic return behavior in the literature (Davis et al., 1995; Hess et al., 1996; Davis et al., 1998; Chu et al., 1998; Yalabik et al., 2005; Ülkü et al., 2013). A slightly differing approach to model opportunism is to assume that not all customers consider opportunism and there is an ex-ante known segment of customers who may act opportunistically, such as in Shang, Ghosh, and Galbreth (2017). Several papers assume that such value consumption occurs only when the product is a fit (Davis et al., 1995; Hess et al., 1996; Chu et al., 1998; Ülkü et al., 2013). Although, there are a few studies that allow for some value consumption, even when there is not a fit (Yalabik et al., 2005; Hsiao & Chen, 2012).

Article	Term (Definition, if provided)	Use	Heterogeneity
Davis et al. (1995)	Transaction cost (travel cost, lost time, mental anxiety)	Parameter	
Moorthy and Srinivasan (1995)	Transaction cost	Parameter	
Chu et al. (1998)	Complaining cost	Parameter	
Davis et al. (1998)	Hassle (retailer imposed restrictions on returning customers)	Decision variable	
Fruchter and Gerstner (1999)	Hassle cost (traveling and confrontation costs)	Parameter	✓
Heiman et al. (2002)	Return cost (inconvenience costs related to returns)	Parameter	
Yalabik et al. (2005)	Customer logistics cost	Parameter	
Matthews and Persico (2007)	Customer's return cost	Parameter	
Anderson, Hansen, and Simester (2009)	Customer's return cost (both monetary and psychological)	Parameter	✓
Su (2009a)	Hassle cost (time, effort, and monetary resources to make a return trip to store)	Parameter	
Shulman et al. (2009)	Hassle cost (time spent in line for customer service, traveling to store)	Parameter	
Xiao et al. (2010)	Consumer's cost of returning	Parameter	
Y.-J. Chen (2011)	Hassle cost (transportation, time value, and in-store waiting costs)	Parameter	✓
Ofek et al. (2011)	Cost of returning a mismatched product	Parameter	
Swinney (2011)	Hassle cost (travel cost of returning to a store)	Parameter	
McWilliams (2012)	Transaction cost	Parameter	
Hsiao and Chen (2012)	Hassle cost (opportunity cost of time and effort)	Parameter	✓
Ülkü et al. (2013)	Hassle cost (psychological and physical costs related to returns)	Parameter	
Zhang (2013)	Shipping, time, and effort cost	Parameter	
Hsiao and Chen (2014)	Hassle cost (transportation costs, cost of keeping an unwanted product, time values, and in-store waiting costs)	Decision variable	✓
Alptekinoğlu and Grasas (2014)	Disutility of returning (hassle cost to make a trip to store)	Parameter	
Altug and Aydinliyim (2016)	Transaction cost of returning (transportation to store, packaging and shipping, disutility from return process)	Parameter	
B. Chen and Chen (2017a)	Hassle cost (cost of returning a product to a store)	Parameter	✓
Yang et al. (2017)	Customer's cost of returning (transportation or shipping)	Parameter	
J. Chen et al. (2018)	Customer's cost of returning	Parameter	✓
Ülkü and Gürler (2018)	Hassle cost	Parameter	

Table 2.2: Return Policy Literature Involving Transaction Costs

2.4.1.5 Return Timing

Hess and Mayhew (1997) argue that the timing of returns is an important component to modeling consumer returns. Despite this early emphasis, however, time-related aspects in consumer return modeling are largely missing, in part due to analytical complexities that a time dimension may introduce. Hess and Mayhew (1997) propose a functional relationship between the return rate and the time-to-return, based on a hazard rate model, which is a popular method in duration modeling in marketing (Helsen & Schmittlein, 1993). Following the approach of Hess and Mayhew (1997), Jalil and Shahzad (2013) model the impact of a return policy time limit on return volumes and their variability. The authors choose a Weibull form hazard function, as the Weibull allows the authors to model returns with a truncated distribution and is flexible enough in parameters to facilitate different assumptions regarding return behaviors. By assuming an exponential form hazard function (i.e., constant hazard rate), Xu et al. (2015) model the expected utility of a customer from returning and not returning a product until the return deadline. Alternatively, Heiman et al. (2002) employ a learning function to capture the likelihood that a customer discovers a product misfit by the return deadline. The authors assume that customer learn the most about the product earlier in the return window and the rate of learning decreases over time. They also assume a that the more time a customer has to return a product, the lower the customer's return transaction cost.

2.4.1.6 Demand and Return Sensitivity

Several studies in the literature model aggregate market demand and return volumes as a function of a retailer's decisions, typically in a linear form. These decisions commonly include monetary leniency and price, and depending on the context, others such as product quality and degree of product modularity. The construction of aggregate market models is achieved through a set of parameters to capture the sensitivity of demand and returns to the decisions of interest. Table 2.3 specifies the collective papers using this assumption and details whether demand or returns (or both) arrive at an aggregate level, as well as whether the model is stochastic or deterministic. Whether return sensitivity is modeled under our classification depends on the presence of a function involving one or several retailer decisions with an associated sensitivity coefficient. For example, Mukhopadhyay and Setaputra (2004) model returns as $R = \phi + \psi r$, where ϕ is a base return amount and ψ denotes customer sensitivity to monetary leniency. Several works such as G. Li et al. (2017), Heydari et al. (2017), and Batarfi et al. (2017), express returns volume only as a deterministic fraction of demand without a return-specific sensitivity parameter, and we therefore classify them as not having a return sensitivity.

Article	Demand	Return	Sensitivity Parameters	Stochastic
Mukhopadhyay and Setaputra (2004)	✓	✓	Price, refund amount	
Mukhopadhyay and Setaputra (2005)	✓	✓	Price, refund amount, product modularity	
Ketzenberg and Zuidwijk (2009)	✓	✓	Price, policy restrictiveness	
Xiao et al. (2010)	✓		Refund amount	✓
N. Liu et al. (2012)	✓	✓	Price, refund amount, product modularity	✓
Choi (2013)	✓	✓	Price, refund amount, product modularity	✓
Choi et al. (2013)	✓	✓	Price, refund amount (full or none), product modularity	✓
Y. Li et al. (2013)	✓	✓	Price, refund amount, quality	
W. Hu et al. (2014)	✓		Price	✓
Yoo et al. (2015)	✓	✓	Price, refund amount	
Altug and Aydinliyim (2016)	✓		Refund amount	✓
G. Li et al. (2017)	✓		Price, MBG adoption	
Heydari et al. (2017)	✓		Price, MBG adoption	✓
Batarfi et al. (2017)	✓		Price, refund amount, migration (i.e., cannibalism)	
Difrancesco et al. (2018)	✓		Return window	

Table 2.3: Return Policy Literature Involving Demand and Return Sensitivity

In general, demand and return sensitivity across papers use consistent assumptions regarding the relationship between demand and return volumes and retailer decisions. Specifically, most works assume that demand decreases in price and increases in adoption of MBG, return policy leniency, and product modularity, whereas returns increase in return policy leniency, and decrease in product quality. Among the works that are classified with a return sensitivity, none consider price sensitivity of returns, despite the earlier empirical evidence suggesting a positive association between the product price and return likelihood (Hess & Mayhew, 1997; Anderson, Hansen, Simester, & Wang, 2009). In addition, a relationship between demand and returns, as empirically identified in Anderson, Hansen, Simester, and Wang (2009), is also overlooked in these works.

2.4.1.7 Risk Aversion

Only three instances in the literature assume that consumers are risk averse. Welling (1989) shows that if consumers have a positive degree of risk aversion, higher-income consumers purchase higher-quality goods, pay a higher price, and receive a higher refund if the product does not work. However, the author shows that the relationship between the optimal price-refund pair and the degree of risk aversion is ambiguous. Che (1996) demonstrates that MBG adoption is not optimal when customers are risk neutral or when retail costs are low. If the degree of customer risk aversion and high retail costs favor MBG adoption, then the retailer's profit is not affected by the degree of risk aversion. As suggested in Su (2009a), Samatli-Pac et al. (2018) show that when customers are loss averse in addition to being risk averse, the optimal refund amount increases above the unit salvage value.

Assumptions regarding the risk attitude of consumers are typically not explicitly stated in the literature. The risk averse consumers assumption is rare in the existing literature, appearing in only three instances. Welling (1989) shows that if consumers differ in income and have a positive degree of risk aversion, higher-income consumers purchase higher-quality goods, pay a higher price, and receive a higher refund should the product does not work. However, the author shows that the relationship between the optimal price-refund pair and the degree of risk aversion is ambiguous. Che (1996) demonstrates that MBG adoption is not optimal when customers are risk neutral or when retail costs are low. Whereas, customers' risk aversion and high retail costs favor MBG adoption and in this case, the retailer's profit does not get affected by the degree of risk aversion. As suggested in Su (2009a), Samatli-Pac et al. (2018) show that when customers are loss averse in addition to being risk averse, the optimal refund amount increases above the unit salvage value.

2.4.2 Return Policy

This section classifies the research and discusses key findings with respect to return policy design. Overall, the literature shows that the degree of monetary leniency, typically expressed as a refund amount or a restocking fee, is the most extensively studied return policy decision. Only a few works investigate each of the remaining leniency levers, which leaves significant opportunities for future research. The literature classified in the RP domain appears in Tables A.1 and A.2 within the Appendix. The classification includes only those works that involve decisions regarding the return policy of a retailer. The decision(s) can be either endogenous, expressed as a decision variable, or exogenous, expressed as a modeling parameter. For the latter, we consider only those papers that present a sensitivity analysis with respect to the exogenously set leniency lever.

2.4.2.1 Money Back Guarantee (MBG) Adoption

In the early development of the analytical RP research, a common line of inquiry is the investigation of conditions under which it is optimal for a retailer to offer an MBG (i.e., a full-refund policy) rather than to disallow returns (i.e., a no-refund policy). For example, Davis et al. (1995) show that retailers should prefer adopting an MBG only when the customer's transaction cost of returning is low or the salvage value for returned products is high. Che (1996) finds that retailers tend to offer an MBG when customers are highly risk averse or when retail costs are high. In a more general modeling framework, Su (2009a) shows that these findings are robust with respect to assumptions regarding customer valuation uncertainty. Moorthy and Srinivasan (1995) study how a high-quality retailer can adopt an MBG to signal quality and to differentiate from a low-quality retailer. More recently, McWilliams (2012) extends upon Moorthy and Srinivasan (1995) and Davis et al. (1995) to study MBG adoption as quality signal in a duopoly. The author shows that contrary to Davis et al. (1995), a retailer can lose profit by adopting an MBG. McWilliams (2012) also shows that a low-quality retailer captures the primary benefit of quality signaling, while Moorthy and Srinivasan (1995) suggest that it is the high-quality retailer that benefits from adopting an MBG. More recently, the literature addressing the MBG decision is often coupled with other operational decisions. As such, we address such couplings within our classification at the intersection of the RP and PE domains (§4.3).

2.4.2.2 Monetary Leniency

Monetary leniency is the most commonly studied leniency lever. Modeling contexts and approaches vary significantly across studies, which makes deriving generalizable insights a challenge. In most cases, monetary leniency is defined in the context of a decision variable that represents a refund amount of the purchase price (Mukhopadhyay & Setaputra, 2004; Yalabik et al., 2005; Su, 2009a). Monetary leniency can also be expressed as a restocking fee or non-refundable charge (Hess et al., 1996; Shulman et al., 2009, 2011; Swinney, 2011). Yet another approach treats monetary leniency as a rate or percentage refund relative to the purchase price (Chu et al., 1998; Mukhopadhyay & Setaputra, 2007; Alptekinoğlu & Grasas, 2014).

A distinct feature of monetary leniency is the allowance for partial-refund policies, as opposed to the binary choice between a full refund and no refund as in the MBG classifier. Under a typical modeling framework that accounts for individual valuation uncertainty, such as in Su (2009a), Shulman et al. (2009), and Akçay et al. (2013), monetary leniency affects both purchase and keep or return decisions. In effect, a customer keeps the product if her valuation is greater than the refund amount offered for a return. This implies that during a purchase decision, the customer relies on her expected utility, which is a function of the product valuation and refund amount. Under this approach, higher monetary leniency also implies a higher return probability.

A rather consistent result reported in the literature is that full-refund policies may be overly generous and are often not optimal in a broad range of operational settings. However, the profit maximizing level of monetary leniency is open to debate and findings are context dependent. Su (2009a) finds that with homogeneous customers, the optimal refund amount is equal to the salvage value, whereas heterogeneity in valuations results in more lenient refund policies. Samatli-Pac et al. (2018) demonstrate that in the case of loss-averse customers, a retailer would offer a refund that is greater than the salvage value. Xu et al. (2018) find that if the salvage value is smaller than the lower-bound of customer's product valuation, the optimal refund amount should equal the lower-bound, not the salvage value. Akçay et al. (2013) consider a case where a retailer sells both new and open-box items that have been returned previously. In contrast to Su (2009a), they find that monetary leniency decreases if the retailer serves both high and low customer segments.

Chu et al. (1998) investigate partial-refund policies as a mechanism to discourage opportunism. The authors find that if customers can extract significant consumption value during the trial period and the salvage value to the retailer is low, then restocking fees are optimal. Shang, Ghosh, and Galbreth (2017) show that when opportunism is not present, the optimal refund amount equals the salvage value, as in Su (2009a) and Akçay et al. (2013). However, in the presence of opportunism, the optimal refund amount is a non-increasing function of both the extent of opportunism and the consumption benefit from opportunism.

Shulman et al. (2009) and Shulman et al. (2011) extend the inquiry of optimal monetary leniency in a competitive environment with horizontally differentiated products available for exchange. In Shulman et al. (2009), the authors reveal that, under different values of retailer attributes and customer behaviors, the optimal return policy of a monopolist may involve a full or partial refund. Shulman et al. (2011) find that, compared to a monopolist scenario, a duopoly with a product exchange option may increase the optimal restocking fee.

Overall, the majority of the articles suggest that a partial-refund policy is optimal for retailers, though this seemingly robust finding is not universal in the literature. For example, Altug and Aydinliyim (2016) discuss several conditions that may result in a full-refund equilibrium. Moreover, many mainstream retailers continue to offer full refund policies. If in addition to individual utility benefits, lenient return policies can substantially stimulate the aggregate demand (as in demand sensitivity functions), then full-refunds may be optimal. Another observation from practice is that many retailers such as Nordstrom and Macy's also offer free return shipping, which corresponds to a refund amount that is greater than the price paid for a product. Several studies, such as Fruchter and Gerstner (1999), Yalabik et al. (2005), and Matthews and Persico (2007), find such generous return policies to be optimal under certain conditions. Of course, several other studies address monetary leniency. Most of these works have unique modeling contexts tailored to investigate various PE domain trade-offs, and we discuss these works in §4.3 with respect to the PE classifiers.

2.4.2.3 Time Leniency

Our classification of the literature reveals only five studies (Heiman et al., 2002; Jalil & Shahzad, 2013; Ülkü et al., 2013; Xu et al., 2015; Difrancesco et al., 2018) in which the length of the return period is a return policy decision variable. Another study (Ülkü & Gürler, 2018) considers time leniency as a modeling parameter. Overall, the modeling contexts and trade-offs studied vary substantially across the works and often have little apparent cohesion. The unpublished working paper by Jalil and Shahzad (2013) investigates risk pooling benefits through the time leniency lever. This study shows that increasing time leniency results in reduced variability in the volume of product returns by changing customers' return behaviors.

Ülkü et al. (2013) model a retailer's price and return deadline decisions in the presence of consumer opportunism. The authors introduce an iterative procedure to jointly determine the optimal price and time leniency. Comparative statics reveal that higher time leniency leads to a higher probability of a legitimate return but also to a higher probability of an opportunistic return if the product price is below a threshold. Further, the authors show that the optimal return period may be shorter than the maximum product lifetime if the additional utility received by customers due to time leniency is below a threshold.

Xu et al. (2015) consider a retailer deciding on a return deadline, along with refund amount, price, and order quantity. This study shows that the length of the product life-cycle and consumer return rate play crucial roles in the optimal time leniency decision. When the return rate is low, the retailer offers an indefinite return deadline. Otherwise, the optimal length of the return window depends on the product life-cycle. In particular, for products with a short life-cycle, the retailer chooses an indefinite return window, whereas for products with a moderately long life-cycle, a fixed return deadline that is shorter than the life-cycle is optimal. For products with very long life-cycles, the optimal return time window is either indefinite or a fixed deadline.

Difrancesco et al. (2018) develop a model to optimize the return time window decision of a fashion retailer that has a closed-loop supply chain with refurbishing. The authors find that the optimal return window decreases as the average time until return decreases. Furthermore, the optimal return window is very large for lower values of the return rate, whereas the window sharply decreases in the return rate. The rate of decrease in the optimal return time window is high when products have a lower number of life-cycles.

Ülkü and Gürler (2018) reveal insights regarding the relationship between a retailer's exogenously set time leniency and the retailer's order quantity, return rate, and profit, in the presence of both legitimate and opportunistic returns. First, the authors show that the retailer's order quantity and profit increase in time leniency. Second, as the return policy becomes more lenient in time, both legitimate and opportunistic return rates decrease, along with an increase in average salvage value.

2.4.2.4 Effort, Scope, and Exchange Leniency

We combine our discussion of the remaining three forms of leniency due to the dearth of research. As for effort leniency, we are aware of only three studies in the literature: Davis et al. (1998), Hsiao and Chen (2014), and C. H. Lee and Rhee (2018). Davis et al. (1998) are the first to acknowledge that most retail stores avoid partial-refund policies and restrict returns by imposing some degree of return hassle. The authors show that under certain conditions, a retailer is more likely to offer a return policy that is lenient in effort. These conditions are 1) customers cannot receive significant short term benefit from the product, 2) cross-selling margins are high, and 3) salvage values of the products returned are high.

Hsiao and Chen (2014) compare the performance of various return policy settings in the presence of product quality risk, heterogeneous valuations, and hassle costs for returning. The authors show that a return policy with a full refund and imposed hassles versus a hassle-free return policy with a restocking fee lead to the same consumer behavioral outcomes for different customer segments. However, the optimal hassle-free policy always dominates the optimal full-refund policy. Although the quality signaling view would predict that convenience in returning would be a strong signal for the seller's product quality, this study finds that optimal effort leniency decreases as the product quality increases. Finally, under certain cases, a moderate level of leniency through both monetary and effort levers is optimal. However, when customer hassle costs vary substantially across segments or when the valuation gap between the segments is large, the optimal policy does not call for the manipulation of both levers—either a full-refund or no-hassle policy will be optimal.

C. H. Lee and Rhee (2018) consider both effort and monetary leniency as decisions of a retailer who runs a resale market in addition to a conventional sales channel. The authors show that both in the presence and absence of a resale market, the optimal return policy is characterized by a hassle-free policy with a partial refund that is equal to the salvage value.

As for scope leniency, Yang et al. (2017) is the only study of which we are aware. In this study, the authors model a retailer who decides on the optimal product assortment by choosing either or both of the brands supplied by two manufacturers as well as whether to offer an MBG for each brand. The choice of return policy for the supply chain depends only on whether or not the retailer can efficiently recover the value of returned products. An MBG is optimal for both brands in the assortment provided that they have positive net salvage value.

Finally, no literature of which we are aware addresses exchange leniency. Although exchange leniency is not considered as a decision, Shulman et al. (2009), Shulman et al. (2010), and Akçay et al. (2013) model a retailer that allows customers to either return a product for a partial refund or exchange for an alternate product, which effectively corresponds to high exchange leniency.

2.4.3 Return Policy and Planning & Execution

This section examines the research on consumer returns that resides at the intersection of the RP and PE domains of our framework.

2.4.3.1 Pricing

Other than return policy levers, pricing is the most prevalent decision in the analytical literature on RP. This section provides a brief overview of the key findings with respect to a retailer's pricing and return policy decisions. Early papers that jointly consider RP and pricing decisions focus on product quality signaling. For example, Shieh (1996) shows that price and MBGs together can completely reveal a monopolistic firm's private information on product quality at no additional cost. However, this study assumes that a retailer honoring an MBG does not incur any costs for handling returned products. Moorthy and Srinivasan (1995) demonstrate that an MBG can efficiently signal quality and improve profit only when a retailer's handling cost for processing returns is low. Otherwise, signaling through a higher price can be less costly.

The existing research demonstrates that a monopolistic retailer can command a price premium by offering an MBG and that this premium depends on product valuation uncertainty (Heiman et al., 2002; Matthews & Persico, 2007; Su, 2009a). Fruchter and Gerstner (1999) find that in a duopoly where only one retailer offers an MBG, the retailer who offers an MBG will command a higher price and earn higher profit. McWilliams (2012) compares equilibrium prices with and without an MBG in a duopoly of quality differentiated retailers. The author finds that compared to the no-MBG case, adopting an MBG results in a price increase for both high quality and low quality retailers.

The vast majority of the literature considers uniform pricing decisions for a single product. Yet, a few derive insights in a multiple product context. For example, Shulman et al. (2011) model two competing retailers, each of whom sell two horizontally differentiated products. The authors show that uncertainty in valuations of the products and differentiation between the products are key determinants of the equilibrium prices in the presence of consumer returns and exchanges. In particular, for products with little differentiation, product returns can increase prices due to limited ability to price discriminate, whereas with high differentiation, product returns reduce equilibrium prices as a result of increased restocking fees. Akçay et al. (2013) model a retailer who sells both new and open-box products that are returned and restocked. The authors show that compared to when no returns are allowed, offering an MBG without restocking returns increases new product price. Also, when the retailer resells the returned products as open-box items at a discount, the new product price increases further. However, the retailer's secondary sales opportunity lowers restocking fees.

Several works consider retailer pricing and return policy decisions for both regular and opportunistic customers. Ülkü et al. (2013) demonstrate that an increase in product price decreases the probability that a customer will keep the product, but increases the probability of an opportunistic return. Further, the authors show that when price is set lower than a threshold, increasing return time leniency also increases opportunistic returns. Shang, Ghosh, and Galbreth (2017) also consider opportunism and find that the optimal price for a product decreases with the benefit from opportunism but is non-increasing in the extent of opportunism.

2.4.3.2 Inventory

The existing literature on consumer returns has many instances that study inventory management in the presence of returns. The typical modeling framework is similar to that of a classical newsvendor problem. Extending the problem to the consumer returns context brings in additional factors, such as return policy leniency and the salvage value of returned products, which may affect the optimal ordering strategy. Under this framework, a retailer sets price and return policy leniency and determines the order quantity in the beginning of a single-period selling horizon based on a probability distribution for uncertain demand. Once sales and returns are realized, the retailer salvages leftover new and returned products. From an inventory optimization standpoint, the objective is to obtain a closed-form solution for the optimal order quantity minimizing total cost.

We classify the RP works that address inventory management and present the resulting classification in Table 2.4. For the fourteen inventory papers, Table 2.4 notes the planning horizon, describes demand characteristics, and identifies both revenue and cost streams. Because a retailer's demand model, cost, and revenue streams are the key elements to formulate and solve a newsvendor problem and modeling frameworks vary significantly in terms of these elements, we focus on these aspects in our classification. We do not include works that omit return policy decisions, such as Mostard and Teunter (2006) and Vlachos and Dekker (2003), even though these studies address inventory optimization in the context of consumer returns. Further, the classification table focuses on each paper's base model(s) and does not include modeling extensions.

Su (2009a) and Ketzenberg and Zuidwijk (2009) are seminal in considering inventory optimization in the context of RP decisions. Using a generic probability distribution to express demand uncertainty, Su (2009a) finds that, compared to a full-refund return policy, a partial-refund

Article	Event Horizon	Demand	Revenue Streams	Cost Streams
Su (2009a)	Single selling season	Random with a generic distribution	Sales of new products, salvaging unsold and returned products	Production cost, refund for returns
Ketzenberg and Zuidwijk (2009)	Single selling season with two periods (i.e., regular sales and resale)	Linear in price and return policy leniency	Sales of new products and first-period returns, salvaging first and second period returns	Procurement, holding, and recovery costs, refund for returns
Xiao et al. (2010)	Single selling season	Increasing function of refund amount with a random error	Sales of new products, buyback credit for unsold products,	Procurement and returns handling costs
J. Chen (2011)	Single selling season	Random with a base demand plus an error term	Sales of new product, salvaging unsold and returned products	Procurement, shortage, and returns handling costs
Akçay et al. (2013)	Single selling season	Random with a generic distribution	Sales of new and open-box products, salvaging unsold products	Procurement cost, refunds for returns
W. Hu et al. (2014)	Single selling season	Decreasing function of price with a random multiplier	Sales of new products, salvaging unsold and returned products	Procurement cost, refunds for returns
Huang et al. (2014)	Single selling season	Random with a generic distribution	Sales of new products, buyback credit for unsold products	Procurement cost, refunds for returns
Xu et al. (2015)	Single selling season	Random with a generic distribution	Sales of new products, salvaging unsold and returned products	Procurement cost, refunds for returns
Altug and Aydinliyim (2016)	Single selling season with two periods (i.e., regular sales and clearance)	Random with a generic distribution	Sales of new products, salvaging returned and unsold products	Procurement cost, refunds for returns
Batarfi et al. (2017)	Infinite planning horizon	Deterministic with a given rate	Sales of new and refurbished products	Procurement, ordering, storage costs and refunds for returns
Heydari et al. (2017)	Single selling season	Random with a generic distribution	Procurement cost, storage costs for returned and unsold products	Sales of new products, buyback credit for unsold and returned products, refunds for returns
Xu et al. (2018)	Single selling season	Random with a generic distribution	New product sales, salvaging returned and unsold products	Procurement cost, refunds for returns
C. H. Lee and Rhee (2018)	Single selling season with two periods (i.e., regular sales and resale)	Random with a generic distribution	Sales of new products, salvaging unsold or returned products	Procurement cost, refunds for returns, resale allowance paid to customers
Ülkü and Gürler (2018)	Single selling season	Random with a generic distribution	Sales of new products, salvaging legitimate and abusive returns, and unsold products	Procurement cost, loss of goodwill costs, refunds for returns

Table 2.4: Return Policy Literature Involving Inventory

policy induces a higher order quantity and service level, which increases a retailer's profit. As an alternative to Su (2009a), Ketzenberg and Zuidwijk (2009) employ an aggregate demand model in which the demand is sensitive to price and return policy leniency. They consider a two-period selling horizon in which a retailer has an opportunity to resell returned products before salvaging at the end of the horizon. The optimal order quantity substantially decreases in price, return policy sensitivity of demand, per unit production cost, and shows only a slight decrease in the return rate and product recovery costs.

The other studies in this area draw upon and extend the works just discussed. These studies reveal that the relationship between optimal order quantity and return policy leniency may not be always negative as suggested by Su (2009a). For example, restocking returns and selling them as open-box, which is not considered by Su (2009a), Akçay et al. (2013) show that the optimal order quantity may decrease as well as increase in monetary leniency. C. H. Lee and Rhee (2018) consider several disposition options for customers: complete consumption, immediate return, and resale in an external or retailer-run secondary market after some use. The authors show that the optimal order quantity increases when a resale market is available. Ülkü and Gürler (2018) consider time leniency and opportunism and show that the optimal order quantity increases in the former and decreases in the latter. Using a time contingent return rate model, Xu et al. (2015) conclude that the optimal order quantity increases in time leniency only when the corresponding return rate is low, whereas the order quantity does not increase in case of a high return rate.

2.4.3.3 Supply Chain Coordination

Supply chain coordination research related to consumer returns provides novel insights to the well established supply chain contracting and coordination literature (Cachon, 2003). In particular, the research shows that traditional contracts may fail to achieve supply chain coordination when consumer returns are allowed by a retailer. Typically, these works involve a supply chain of a manufacturer and a retailer in which the manufacturer sets the terms of a contract. The retailer then makes decisions on price, return policy, and order quantity. All of the works consider an endogenous return policy decision, except Huang et al. (2014) that assumes an exogenous refund amount expressed as a parameter that is then examined through sensitivity analysis.

Su (2009a) is the first to recognize that traditional buyback contracts with wholesale pricing may not be able to coordinate a two-echelon supply chain with consumer returns. Moreover, buyback contracts may result in inefficiently generous return policies. Two different contracts are proposed as a remedy to the coordination problem and their effectiveness are analytically demonstrated. Samatli-Pac et al. (2018) show that the contracts proposed in Su (2009a) can still coordinate the supply chain when customers are loss-averse. Xiao et al. (2010) employ a demand model in which demand is sensitive to the refund amount to study a similar coordination problem and characterize the terms of a buyback contract that can coordinate the supply chain. The authors find that in the coordinated setting with a given buyback price, increasing the refund amount first increases both players' expected profits and order quantities, but subsequently decreases both. Xu et al. (2015) show that the differentiated buyback contract proposed in Su (2009a), which sets different buyback prices on unsold and returned units, may also fail to coordinate a supply chain if the salvage value of returned products depends on the time-to-return. To achieve coordination, the authors introduce a novel differentiated buyback contract that is contingent on the return deadline set by the retailer. Heydari et al. (2017) find that wholesale pricing and buyback contracts cannot simultaneously achieve Pareto improving coordination for both parties in a two-echelon supply chain. The authors show that a differentiated buyback contract can help achieve this goal.

Yoo et al. (2015) compare optimal prices and refunds under three contracts that may coordinate a two-echelon supply chain where the supplier has a bargaining power over the retailer. The authors find that selling prices are highest in the case of a wholesale pricing contract, followed by a buyback contract and then a quantity discount contract. Whereas, monetary leniency is greatest when a quantity discount contract is adopted, followed by a wholesale pricing contract and then a buyback contract. Overall, the results of this study suggest that a quantity discount contract that coordinates the supply chain is the most favorable from the perspective of the supplier, as this contract leads to the lowest prices and highest return policy leniency that also stimulates sales.

2.4.3.4 *Forward Channel Design*

In general, the analytical RP studies that address forward channel design investigate the conditions under which various channel structures are optimal. A typical decision model considers the choice of offering a single or dual channel (e.g., adding an online sales channel to an existing brick-and-mortar channel). In addition, channel specific return policy leniency, as well as other operational decisions such as pricing and order quantity, are considered in these studies. No studies were found within the scope of review that considers an omni-channel environment with features such as “buy online, pick up in store” or “buy online, return to store”. For each contribution, Table 2.5 reports study context, decision variables, and key findings from six studies under this classifier.

2.4.3.5 *Competition*

Research addressing competition typically compares a monopoly to a duopoly. Generally, a retailer in these studies sells a single, unique product and competes with another retailer through price and return policy. Return policy decisions are usually modeled in terms of MBG adoption or the monetary leniency lever. An important assumption in modeling a competitive environment is related to the differentiation of the competing parties. Such differentiation includes product quality, market presence (i.e., established or entrant), cost advantage in returns handling or salvaging, etc. Both sequential (i.e., Stackelberg) and simultaneous (i.e., Nash) games are considered in these studies. Table 2.6 lists the relevant studies, including information regarding nature of competition, number of players, differentiating factors, and major findings.

2.4.3.6 *Product Characteristics*

Two common product characteristics studied in the analytical literature are product quality and product modularity. Retailers use return policies both as a mechanism to decrease purchase risk associated with products and as a means to signal product quality. As such, product quality choices, made by either a retailer or a manufacturer, have an impact on return policy decisions. The early literature on RP considers product quality as a source of information asymmetry between a retailer and customers and studies the quality signaling function of MBGs (Heal, 1977; Welling, 1989;

Article	Context	Decisions	Key Results
Ofek et al. (2011)	Two retailers, with the options to operate either brick-and-mortar only or through dual channels.	Price and store assistance level. Restocking fee in an extension.	There is an asymmetric equilibrium where only one firm operates an online channel yet makes lower profits than brick-and-mortar only rival. Also, there is an equilibrium where both rivals operate an online arm even if the total profits decrease. Retailers may charge restocking fees up to a maximum allowable amount.
Batarfi et al. (2017)	Single brick-and-mortar channel versus a dual-channel strategy in a closed-loop supply chain of a manufacturer, a remanufacturer, and a retailer.	Price and order quantity. Return policy is exogeneous.	When the sensitivity of demand to return policy leniency increases, demand, prices, and total profits in the supply chain increase in both single- and dual-channel systems.
B. Chen and Chen (2017b)	A retailer that considers adding an online channel to an existing brick-and-mortar channel.	Whether and when a dual-channel system is optimal. Whether to adopt an MBG or personalized pricing strategy (PPS) for each channel.	The retailer offers an MBG for a channel if the net salvage value of a returned product is positive. The retailer's optimal return policy decision depends only on the efficiency in salvaging the returned products. In addition, PPS makes the online channel and MBGs more attractive.
G. Li et al. (2017)	Dual-channel supply chain of a manufacturer and a retailer. Manufacturer can sell directly or through retailer.	Whether or not to offer full refund in each channel.	The manufacturer opts for a full-refund policy in both channels when the return rate is low. Otherwise, the manufacturer chooses a no-refund policy for both channels. For the retailer, the optimal strategy is to have a full-refund (no-refund) policy in the indirect channel and no-refund (full-refund) policy in the direct channel when the return rate is low (high).
W. Li et al. (2018)	A dual-channel supply chain consisting of a retailer and a manufacturer. The manufacturer sells a high-quality product indirectly through a retailer and considers opening a second, direct channel to sell relatively lower quality items.	Whether to adopt an MBG or personalized pricing strategy (PPS) for each channel.	Regardless of the pricing strategy, the retailer should offer an MBG if the net salvage value of a returned product is positive. The manufacturer benefits from having a direct channel with uniform pricing, and similar to the retailer, should offer an MBG when the net salvage value is positive.
Letizia et al. (2018)	A single manufacturer.	Whether to operate through online (i.e., direct), brick-and-mortar (i.e., through retail stores), or dual-channels. Return policy is exogeneous.	A dual-channel structure may be the optimal choice that maximizes the manufacturer's profits, due to the flexibility that the dual-channel structure provides for the manufacturer to reduce returns without charging a high restocking fee as in the case of online only channel.

Table 2.5: Return Policy Literature Involving Forward Channel Design

Article	Game Type	Players and Differentiating Factors	Key Findings
Fruchter and Gerstner (1999)	Nash	Two sellers, one offering MBG, the other without an MBG.	The seller with MBG charges a higher price and extracts all the expected surplus of its customers. Both sellers earn positive profits, but the seller with MBG earns a higher profit.
Shulman et al. (2011)	Nash	Two firms, each selling two products with maximal differentiation.	Equilibrium restocking fees in a competitive environment can be higher than those charged by a monopolist. The equilibrium restocking fees increase in the extent of perceived differentiation between the products.
Ofek et al. (2011)	Nash	Two retailers, with the options to operate either brick-and-mortar only or through dual channels, choosing price and store assistance levels. Restocking fee decision is considered in an extension.	There is an asymmetric equilibrium where only one firm operates an online channel yet makes lower profits than brick-and-mortar only rival. Also, there is an equilibrium where both rivals operate an online arm even if the total profits decrease. Retailers may charge restocking fees up to a maximum allowable amount.
McWilliams (2012)	Nash	Two retailers, differentiated in quality (high vs low) and marginal costs per sale for a product.	Duopoly profits increase with MBGs. Low-quality retailer increases profits while high-quality retailer decreases profits, compared to the case where neither retailer offers an MBG.
J. Chen and Grewal (2013)	Nash and Stackelberg	Two retailers, one established and offering an MBG, the other a new entrant and considering whether to offer an MBG.	When the customer acceptance of a no-refund product is low (high), the entrant retailer should offer a full-refund policy (no-refund) and set a higher (lower) price.
Inderst and Tirosh (2015)	Nash	Two firms, differentiated in quality (high vs low).	In equilibrium, the low-quality firm offers a lower refund, even though salvage value for the products are the same and product quality is observable to customers.
J. Chen and Chen (2016)	Nash	Multiple retailers offering full-refund return policies but differing in other leniency levers.	As a retailer's return rate increases, it becomes less competitive as a result of increased prices and decreased market share. In equilibrium, the retailers with higher leniencies serve the customers with higher valuations, resulting in a market segmentation.
Altug and Aydinliyim (2016)	Nash	Multiple retailers with different clearence period stock availability.	If salvage value for returned products is below certain threshold, a retailer facing competition should not allow returns.
Yang et al. (2017)	Stackelberg	Two manufacturers, supplying a common retailer and offering two distinct brands with different production costs and fit probabilities.	If the retailer's net salvage value for a returned product supplied by either manufacturer is positive, the retailer accepts the wholesale prices set by the manufacturers and offer an MBG for both brands. This equilibrium improves profits for all parties.
B. Chen and Chen (2017a)	Nash	Two retailers with differing customer satisfaction rates.	Competition leads both retailers to employ an MBG and PPS. MBG adoption softens price competition and result in Pareto profit improvement, but PPS leads to intensified price competition.
J. Chen et al. (2018)	Nash and Stackelberg	Two retailers, differing in service quality (high vs low) and unit costs for procurement and returns processing.	If the sum of a retailer's and customer's logistics costs is lower than the salvage value, both low and high quality retailers offer an MBG. In this case, the cost-efficient retailer will be the price leader and the competitor will be the follower, and both retailers will be better off.

Table 2.6: Return Policy Literature Involving Competition

Moorthy & Srinivasan, 1995; Shieh, 1996). However, this line of research assumes product quality is exogenously given and does not consider joint optimization of product quality and return policy decisions. Several studies, however, do investigate the relationship between quality decisions and monetary leniency.

Decisions regarding the degree of product modularity (i.e., customization) are commonly studied with clear trade-offs. Sellers try to attract more customers by offering customization and lenient return policies. However, returned items are typically of low value to the seller due to their customization. Several studies investigate the trade-off between modularity and return policy leniency together with pricing decisions. Table 2.7 lists the studies that have instances for either of two product characteristics, identifies both the decisions that they address, and key findings. By key findings, we refer to the analytical insights regarding various trade-offs between the return policy leniency decision and other decisions, such as product quality, modularity, and price.

Article	Decisions	Key Findings
Welling (1989)	Price, refund amount, product quality	Both price and refund are increasing in quality. The higher the quality, the smaller the difference between price and refund amount.
Mukhopadhyay and Setaputra (2005)	Refund amount, degree of product modularity	When customer demand is more sensitive to i) the return policy leniency, ii) the degree of product modularity, the seller offers a greater refund and higher degree of product modularity.
Mukhopadhyay and Setaputra (2007)	Product price, refund rate (as a percentage of price), product quality	Earlier in the product life-cycle, quality has a steep increase and higher refund rate is provided. As price decreases over the life-cycle, the return policy becomes less lenient.
N. Liu et al. (2012)	Product price, refund amount, degree of product modularity	For a risk averse retailer, when baseline demand increases, optimal product price, refund rate, and degree of product modularity decreases.
Hsiao and Chen (2012)	Product price, refund amount, degree of product risk (as an extension)	When the product quality is low, the optimal refund can exceed the product price. When the quality is relatively high, further improvement in product quality may not induce a higher product price and refund amount.
Y. Li et al. (2013)	Product price, refund amount, product and service quality	Pricing, refund amound, and quality decisions are mutually complementary. If the seller is a price-taker, a higher refund amount should be complemented by a higher quality.
Choi (2013)	Product price, return service charge, degree of product modularity	Optimal return service charge increases if degree of product modularity decreases but is independent of the sensitivity of demand to the modularity.
Choi et al. (2013)	Product price, return service charge, degree of product modularity	When a seller offers returns, product price and modularity decisions are linearly related to each other, and price is a decreasing function of return service charge.
Yoo (2014)	Product price, refund amount, product quality	Increasing monetary leniency cannot be optimal without enhancing product quality regardless of risk attitude. However, quality enhancement without setting a lenient return policy can be optimal.
Hsiao and Chen (2015)	Product price, refund amount, product quality	Returns are accepted by the retailer only when the product reliability is moderate and the valuation of the low segment is moderate. Without consumer returns, the optimal product quality always increases in the probability of a fit. With consumer returns, this relation can be reversed.

Table 2.7: Return Policy Literature Involving Product Characteristics

2.4.3.7 Product Assortment

Research addressing product assortment considers how the assortment decision is affected by the return policy decision and vice-versa. We are, however, aware of only four such contributions (Shulman et al., 2009, 2011; Alptekinoğlu & Grasas, 2014; Yang et al., 2017). Note that the key findings of the most recent contribution are already discussed in the context of scope leniency in §4.2.4.

Shulman et al. (2009) are the first to address the multi-product (2) case and investigates the implications with respect to monetary leniency and product pricing decisions. Customers arbitrarily pick between the two products only when the expected utility of making a purchase is positive and then make one of the three decisions: 1) return the product and leave the market, 2) exchange the purchased product with the other product, or 3) keep the purchased product. The authors show that monetary leniency in the two product case may be different from a single product case. Shulman et al. (2011) extend Shulman et al. (2009) to the case of a duopoly and show that competition leads to less generous return policies if customers perceive strong differences between the products.

Alptekinoğlu and Grasas (2014) model a retailer who makes the assortment decision by choosing between popular, highly attractive products and eccentric, less attractive products under an exogenously given return policy. With a lenient return policy, the authors find that a retailer should decrease product variety and make the assortment mix based on popular products that have a lower probability of being returned. However, if the retailer has a restrictive return policy, the optimal assortment may consist of both popular and eccentric products.

2.4.4 Return Policy and Return Management

In general, the RM domain falls under a broader body of CLSC literature. RM research primarily investigates various trade-offs in reverse channel operations, including acquisition, processing, and disposition of consumer returns. Studies also examine decisions with respect to reverse channel design in order to execute these operations, such as the allocation of returns acquisition and salvaging responsibility amongst the parties of a supply chain. Although the amount of analytical research in this domain is significant, only a few of the published studies address return policy decisions.

At a minimum, the studies at the intersection of the RP and RM domains take into account salvaging, which concerns the disposition of returns. Taking salvaging into account typically does not go beyond incorporating a single modeling parameter for the salvage value of returned products, whereas a detailed inquiry of options for and execution of salvaging is largely absent. Numerous studies show that a product's salvage value is a critical determinant of a retailer's optimal return policy choices. In particular, these studies find that retailers with more efficient salvaging options can offer more lenient policies (Davis et al., 1995; Su, 2009a; Altug & Aydinliyim, 2016).

Shulman et al. (2010) focus on reverse channel design for salvaging. In particular, they investigate the optimal return policy in the context of a supply chain in which the manufacturer chooses the salvaging party: either the manufacturer itself or the retailer. The authors show that if the retailer can extract higher value from salvaging compared to the manufacturer, retailer can charge higher restocking fees than in the case in which the manufacturer is the salvaging party. Overall, this study reveals that a higher salvage value generally leads to a more lenient return policy, as also demonstrated in the earlier literature. This effect may be reversed by the manufacturer's incentive to coordinate the reverse supply chain via wholesale and buyback pricing.

There are several other return disposition issues considered in the existing literature. Two consider reselling returns. Ketzenberg and Zuidwijk (2009) investigate the case where returns are recovered at a cost and resold as new products. Alternatively, Akçay et al. (2013) consider reselling returned products as open-box (without recovery). Hsiao and Chen (2015) study the case where the retailer either salvages the returns or sends them back to the manufacturer. Non-uniform salvage values are also studied. For example, Xu et al. (2015) consider salvage value that is contingent on time-to-return and is the only study in the RP literature that considers time-decay of products. Shang, Ghosh, and Galbreth (2017) and Ülkü and Gürler (2018) assume that salvage value depends on whether or not returns arise due to opportunism.

In general, returns processing in analytical models is limited to a cost parameter that captures a retailer's per unit returns processing cost. Moorthy and Srinivasan (1995) find that absent a positive salvage value, the return processing cost is a key determinant of the MBG adoption decision in order to efficiently signal product quality. In particular, when a retailer bears high processing costs, they should avoid offering an MBG and rely on pricing to signal quality. Swinney (2011) also considers a return processing cost and shows that the optimal restocking fee should be at least as high as this cost. Su (2009a) considers both a positive salvage value and a return processing cost and demonstrates that the optimal refund amount is equal to the salvage value minus the processing cost.

On the returns acquisition side, Su (2009a) shows how an arrangement where consumer returns are collected directly by a manufacturer instead of a retailer can be used to achieve supply chain coordination in a supply chain with consumer returns. Samatli-Pac et al. (2018) show that the findings in Su (2009a) are robust with respect to customer loss-aversion. J. Chen and Bell (2012) study a retailer's decision to offer a returnable, non-returnable, and dual-channel for the reverse flow of products, characterized by an MBG adoption decision. The key finding is that, for certain ranges of customer sensitivity to the prohibition to make a return, the retailer can earn a profit with a dual-channel structure that is at least equal to the profits from a single, returnable or non-returnable channel. The authors show that under a dual-channel structure, the retailer can segment the market by setting different prices for each channel.

2.5 Empirical Research

The classification and review of the empirical literature reveals that more than half of the empirical works have been published in the last six years, which highlights the nascent and evolving state of the literature. §2.5.1 outlines the theoretical and methodological foundations of empirical research, which provides the basis for the discussion of key findings from the literature in §2.5.2 and §2.5.3. §2.5.2 focuses on the literature that resides at the intersection of RP and CB domains, in which the vast majority of the literature is concentrated. §5.3 is dedicated to the research in the intersection of RP with PE and RM domains.

2.5.1 Theoretical and Methodological Foundations

This section begins with brief descriptions of the theories used in the literature to develop hypotheses. The section then discusses data collection and analysis methods to test such hypotheses.

2.5.1.1 Theoretical Foundations

Empirical research in the RP domain draws upon several well-established theories from economics and psychology. Table 2.8 provides an easy reference for the categorization of studies based on the theoretical lenses used. To maintain an impartial approach, the categorization relies only upon the *explicitly* stated theories used in each paper's hypothesis development.

Theory	Based on	Associated RP Literature
Signaling	Spence (1973)	Wood (2001); d'Astous and Guèvremont (2008); Bonifield et al. (2010); Pei et al. (2014); Rao et al. (2017); Zhang et al. (2017); Oghazi et al. (2018)
Prospect (including endowment effect)	Kahneman and Tversky (1979)	Wood (2001); Wang (2009); Heiman et al. (2015)
Distributive (Equity) and Procedural Justice	Adams and Freedman (1976); Lind and Tyler (1988)	Suwelack et al. (2011); Bower and Maxham III (2012); Pei et al. (2014)
Attribution	Kelley (1967)	Bower and Maxham III (2012)
Construal Level	Liberman and Trope (1998)	Janakiraman and Ordóñez (2012)
Norm	Kahneman and Miller (1986)	Kim and Wansink (2012)
Cognitive Dissonance	Festinger (1957)	Powers and Jack (2013)

Table 2.8: Theoretical Basis for Empirical Research on Return Policy

Signaling theory appears to be the most commonly employed theory. As initially proposed by Spence (1973), signaling theory addresses the reduction of information asymmetry between two parties involved in a transaction via costly signals. The theory provides a lens to study whether return policies can reduce information asymmetry between the retailer and the customer with regard to the seller quality and the product quality. Return policy leniency is a costly signal as such leniency generates significant operational costs and losses for retailers and thus serves as a valid signal in signaling theory.

Prospect theory is a behavioral economics theory that helps to explain how people choose between alternatives bearing risk when outcome probabilities are known (Kahneman & Tversky, 1979). The fundamental tenet of the theory is that people make heuristic decisions based on the potential value of losses and gains rather than on the final outcome. With its emphasis on loss aversion, this theory provides an important theoretical lens to study cognitive and behavioral responses of customers toward return policies that typically lead to monetary losses due to low monetary and/or exchange leniency (Heiman et al., 2015). Another interesting phenomenon studied under prospect theory is the endowment effect in which an individual values something already owned more than something that is not yet owned. In the context of time leniency, the endowment effect suggests that providing a longer return time window should increase a customer's valuation of a product and result in a lower return likelihood (Wood, 2001).

Two justice theories from social and organizational psychology provide a foundation to study retailer decisions relative to a customer's fairness perceptions. Distributive justice theory, also known as equity theory, posits that people value fair treatment which motivates them to maintain relationships with organizations (Adams & Freedman, 1976). Procedural justice theory suggests that individuals are motivated not only by fairness in the distribution of outcomes, but also the processes and rules that lead to particular outcomes (Lind & Tyler, 1988). In essence, return policies are procedural statements that determine potential outcomes in case of a return event. Therefore, distributive and procedural justice theories provide a basis to study how customers perceive return policy fairness.

Attribution theory provides a framework to understand how people explain the causes of behavior and events. Although several theories of attribution exist, Kelley's Covariation Model is the most popular, which posits that people judge events caused by a particular action based on whether the action should be attributed to some characteristic of the actor or the environment (Kelley, 1967). For example, a consumer's perceptions regarding who should be held responsible for a need to return a product can moderate his or her perceptions of a retailer's return policy. Alternatively, attribution theory may also be helpful to explain customer reactions to restrictive return policy changes.

Construal level theory, developed by Liberman and Trope (1998), suggests that the temporal distance to an event alters the mental representation of events and can affect individual decisions. In particular, the theory posits that as the temporal distance gets shorter, an individual's thinking about objects and events shifts from being abstract to concrete. For example, using different scenarios of consumer choice Liberman and Trope (1998) find that under a shortened time window, individuals rate the feasibility attributes of a service (e.g., delivery method, prices, etc.) as highly important factors, reflecting a concrete thinking. However, with a longer time window, individuals tend to focus more on desirability attributes that are more abstract (e.g., the design and color fit, user experience, etc.). In the study of return policies, a return deadline that is specified by a retailer's time leniency decision, represents a temporal distance. Under high time leniency, a customer may focus more on product fit and experiential benefits, which may result in a different post-purchase outcome compared to a low leniency scenario, in which the customer would focus more on aspects such as product price and return effort. Therefore, the theory provides a basis to study how return rates are affected by the interaction of time and effort leniency, as well potential moderators in this relationship, such as product characteristics and price.

Norm theory in psychology posits that after experiencing an event, individuals develop unrealized alternative versions of the event (i.e., what the event would, could, or should have been) as a direct response (Kahneman & Miller, 1986). Kahneman and Tversky (1982) initially termed the unrealized versions of an event as counterfactuals and the mental activity of recruiting counterfactuals as counterfactual thinking. Epstude and Roese (2008) develop the functional theory of counterfactual thinking to study how the counterfactual thinking process leads to behavioral actions, based on norm theory. These theories provide a basis to develop hypotheses regarding how lenient return policies would influence responses of customers in the case of an imperfect match of a purchase that is made under lenient versus restrictive return policies.

Cognitive dissonance theory posits that inconsistencies with the personal beliefs or expectations result in cognitive dissonance. Individuals are motivated to reduce the dissonance and actively avoid situations that increase dissonance (Festinger, 1957). Cognitive dissonance theory suggests that lenient return policies can reduce the intensity of post-purchase dissonance since the purchase decisions under lenient return policies would be reversible with lower transaction costs. The study of return policies through the cognitive dissonance theoretical lens is relevant to understanding pre- and post-purchase, as well as post-return behaviors of customers under high valuation uncertainty.

As seen in Table 2.8, less than half of the studies ground hypotheses on formal theories. Further, the existing empirical literature largely draws upon theories from behavioral economics and psychology. Therefore, significant opportunities remain for expanding the theory base, such as incorporating organizational and management theories to study the retail perspective of return policy decision-making.

2.5.1.2 Methodological Foundations

The empirical research on RP incorporates a variety of methods for data collection and analysis. Table 2.9 presents a methodological categorization. Works that involve simple descriptive statistics (Hawes & Lumpkin, 1986), meta-analysis (Janakiraman et al., 2016), and literature driven theory development (Heiman et al., 2001), or works that do not explicitly specify the method used (Petersen & Kumar, 2010) appear in the "Other" column.

With respect to data collection methods, lab experiments are most frequently used. The following observations arise from the review of experimental literature on RP: 1) all leniency levers are studied, 2) time leniency is the most commonly manipulated lever with eight studies, 3) none of the studies include manipulations of all available levers, and 4) only one study (Janakiraman & Ordóñez, 2012) explicitly studies the interaction (i.e., moderation) effect between two leniency levers (i.e., effort and time). Further, we note that most of the experimental studies follow a between subject design with random assignment, with a few exceptions that use within subject (i.e., repeated measures) designs, such as Wood (2001) and a field experiment in Bower and Maxham III (2012).

Survey-based research is the next most commonly used technique and is largely explanatory. In most studies, the unit of analysis is at the retailer or customer level, with the latter being dominant.

Data Collection / Data Analysis	Analysis of Variance	Regression	Structural Equation Modeling	Other
Survey & Interview		Davis et al. (1998); Autry (2005); Heiman et al. (2015)	Mollenkopf et al. (2007); Hsieh (2013); Pei et al. (2014); Oghazi et al. (2018)	Hawes and Lumpkin (1986)
Lab Experiment	Van den Poel and Leunis (1999); Wood (2001); d'Astous and Guèvremont (2008); Wang (2009); Bonifield et al. (2010); Suwelack et al. (2011); Bahn and Boyd (2014); Seo et al. (2016)	Janakiraman and Ordóñez (2012)	Zhang et al. (2017); Jeng (2017)	
Field Experiment		Bower and Maxham III (2012); Lantz and Hjort (2013); Rao et al. (2017)		Petersen and Kumar (2010)
Secondary Data		Mixon (1999); Heim and Field (2007); Posselt et al. (2008); Anderson, Hansen, and Simester (2009); Bonifield et al. (2010); Zhou and Hinz (2016); Hjort and Lantz (2016); Shang, Pekgün, et al. (2017)		
Literature Content				Heiman et al. (2001); Janakiraman et al. (2016)

Table 2.9: Data Collection and Analysis Methods in the Empirical Literature

In fact, only (Davis et al., 1998) investigates the retailer perspective with a sample of 143 retailers. The use of field experiments in RP literature is rare. Among the few contributions, Petersen and Kumar (2010) report on a field experiment that studies how several retail performance measures changed over time once an apparel and footwear retailer loosened its return policy. Lantz and Hjort (2013) conduct an online field experiment through an e-retailer's website to study customers' behavioral responses to free delivery and return shipping. As a final example, Rao et al. (2017) conduct real eBay auctions to investigate the impact of return time leniency on willingness-to-pay. As an alternative to the primary data methods outlined above, secondary data continues to gain popularity in the operations management field and research in the RP domain is no exception. The RP literature involves secondary data from a variety of sources, such as retailer transaction records (Hjort & Lantz, 2016; Zhou & Hinz, 2016; Shang, Pekgün, et al., 2017), rating websites (Heim & Field, 2007; Bonifield et al., 2010), and advertisements (Mixon, 1999).

Regarding data analysis methods, our review reveals that regression and analysis of variance techniques are most commonly used. In general, secondary data studies employ regression methods, such as ordinary least squares regression (Hjort & Lantz, 2016), logistic regression (Shang, Pekgün, et al., 2017), and multilevel modeling (Rao et al., 2017). Structural equation modeling SEM methods are commonly used in survey and interview research (Mollenkopf et al., 2007; Powers & Jack, 2013), but also are used to analyze experimental data (Zhang et al., 2017; Jeng, 2017). Analysis of variance techniques (e.g., ANOVA, MANOVA) are the most commonly used data analysis methods in experimental research on RP. We conjecture that the study of more complex mediation and moderation processes related to the impact of return policy leniency on various consumer behavior outcomes may increase applications of regression based techniques to analyze experimental data. Further, we can expect that repeated measures and longitudinal experimental studies are likely to increase applications of advanced SEM techniques.

2.5.2 Return Policy and Consumer Behavior

Several empirical studies in the RP literature address how return policy leniency influences constructs related to the four empirically defined classifiers for the CB domain, that include cognitive responses (CR), affective responses (AR), behavioral intentions (BI), and behavioral actions (BA). Table 2.10 categorizes the literature based on RP and CB classifiers. Predominantly, CR, BI and BA constructs are the focus of the literature, whereas a few works consider AR constructs. In particular, our review shows that perceived product quality is the most commonly studied cognitive response, which is expected given the large number of studies that draw upon signaling theory. Liking and regret are two affective responses that are studied in two papers each. Purchase intention is the most frequently studied construct of behavioral intention. Whereas, common behavioral actions studied are average value of orders and average value of returns.

Tables 2.11 and 2.12 provide an easy reference that summarizes the key findings in the literature, in terms of relationships between different leniency levers and the constructs that operationalize the CB classifiers. The tables note the theoretical constructs that relate to each of the four empirical CB classifiers and form the basis of explicitly stated hypotheses. The last column includes the statistically significant findings related to these hypotheses. In addition, we include those statistically insignificant findings that either 1) fail to confirm earlier findings, 2) provide somewhat contradictory evidence to earlier findings, or 3) suggest that a given leniency lever may be less effective than another lever in influencing a construct. Note that two works are excluded

RP Classifiers / CB Classifiers	Cognitive Responses	Affective Responses	Behavioral Intentions	Behavioral Actions
MBG adoption	Van den Poel and Leunis (1999)		Hawes and Lumpkin (1986)	Anderson, Hansen, and Simester (2009); Shang, Pekgün, et al. (2017)
Monetary leniency	Wood (2001); Bower and Maxham III (2012); Pei et al. (2014); Zhang et al. (2017)	Posselt et al. (2008); Bower and Maxham III (2012)	Wood (2001); Wang (2009); Pei et al. (2014); Janakiraman et al. (2016); Gelbrich et al. (2017); Zhang et al. (2017)	Wood (2001); Heim and Field (2007); Bower and Maxham III (2012); Lantz and Hjort (2013); Hjort and Lantz (2016)
Time leniency	d'Astous and Guèvremont (2008); Wang (2009); Suwelack et al. (2011); Janakiraman and Ordóñez (2012); Bahn and Boyd (2014); Zhang et al. (2017)	Posselt et al. (2008); Suwelack et al. (2011)	Wang (2009); Suwelack et al. (2011); Janakiraman and Ordóñez (2012); Heiman et al. (2015); Janakiraman et al. (2016); Rao et al. (2017); Zhang et al. (2017)	Heim and Field (2007)
Scope leniency	Wood (2001); Kim and Wansink (2012)		Wood (2001)	Wood (2001)
Effort leniency	Mollenkopf et al. (2007); Suwelack et al. (2011); Janakiraman et al. (2016)	Mollenkopf et al. (2007); Suwelack et al. (2011)	Mollenkopf et al. (2007); Suwelack et al. (2011); Janakiraman and Ordóñez (2012); Janakiraman et al. (2016)	Heim and Field (2007)
Exchange leniency	d'Astous and Guèvremont (2008)		Heiman et al. (2015); Janakiraman et al. (2016)	Heim and Field (2007)
Overall leniency	Bonifield et al. (2010); Hsieh (2013); Powers and Jack (2013); Jeng (2017); Oghazi et al. (2018)	Powers and Jack (2013)	Bonifield et al. (2010); Hsieh (2013); Powers and Jack (2013); Jeng (2017); Oghazi et al. (2018)	Petersen and Kumar (2010); Zhou and Hinz (2016)

Table 2.10: Classification of the Empirical Research in the Intersection of RP and CB Domains

from the tables. Heiman et al. (2001) is a conceptual piece that does not test any hypotheses that link return policy leniency to consumer behavioral constructs. Seo et al. (2016) also does not consider return policy leniency as a focal variable in their hypotheses but do involve MBG adoption as a moderator in the link between type of buying motivation (hedonic vs. utilitarian), purchase planning behavior (planned vs. unplanned), and purchase intention.

As an illustrative example for understanding Tables 2.11 and 2.12, consider Zhang et al. (2017) from Table 2.12. This study investigates how monetary and time leniency influence an important BI construct, purchase intention, through a mediation path of two CR constructs—perceived return difficulty and perceived service quality. From the last column, we find that higher leniency in either monetary or time leniency is associated with an increase in purchase intention. This happens through a serial mediation, in which leniency decreases perceived return difficulty and lower perceived return difficulty results in a higher perceived service quality that subsequently results in a higher purchase intention ($\text{M, T} \xrightarrow{-} \text{PRD} \xrightarrow{-} \text{PSQ} \xrightarrow{+} \text{PI}$). Further, this study finds that although monetary leniency can effectively signal perceived product quality ($\text{M} \xrightarrow{+} \text{PPQ}$), time leniency fails to do so (T $\xrightarrow{\text{n.s.}}$ PPQ). We now proceed to discuss key insights from the literature in the crossover of RP and CB domains, organized in terms of the RP classifiers, beginning with the MBG adoption classifier.

Five studies investigate the relationship between the MBG adoption and various constructs classified under CR. These studies suggest that MBG is an effective mechanism to reduce perceived purchase risk (Hawes & Lumpkin, 1986) and increases self-reported purchase likelihood (Van den Poel & Leunis, 1999). Regarding the monetary value of an MBG, Anderson, Hansen, and Simester (2009) find considerable variation in the value of returns across customers and categories and show that average purchase rates increase with the value of a return option. Shang, Pekgün, et al. (2017) show that offering an MBG, compared to a no-refund policy, increases the value of a return policy. However, the results suggest that the incremental monetary value of an MBG beyond the product price may not be as high as sellers may assume or as shown by Anderson, Hansen, and Simester (2009).

The impact of monetary leniency on consumer behavior is tested in ten studies, using both primary and secondary data methods. For the CR classifier, several studies lend support to signaling theory by finding empirical evidence that monetary leniency signals product and service quality resulting in higher perceived value (Wood, 2001; Wang, 2009; Zhang et al., 2017). Moreover, research shows that monetary leniency positively influences perceived cost fairness (Bower & Maxham III, 2012) and policy fairness (Pei et al., 2014). For the AR classifier, research suggests that monetary leniency can stimulate liking toward a retailer's return policy (Posselt et al., 2008) and decrease post-return regret (Bower & Maxham III, 2012). In general, there is strong empirical evidence that higher monetary leniency results in favorable customer behavioral intentions and actions from a retailer's perspective (Lantz & Hjort, 2013; Pei et al., 2014; Zhang et al., 2017; Gelbrich et al., 2017).

Ten studies explore the impact of time leniency with respect to various CB classifiers. With respect to the CR classifier, a consistent finding is that higher time leniency is associated with a higher willingness-to-pay for product and return option, however, the relationship shows a diminishing

Article	Leniency Levers	Cognitive Responses	Affective Responses	Behavioral Intentions	Behavioral Actions	Key Findings
Hawes and Lumpkin (1986)	G			Perceived purchase risk (PPR)		$G \xrightarrow{+}$ PPR
Van den Poel and Leunis (1999)	G	Purchase likelihood (PL)				$G \xrightarrow{-}$ PL
Wood (2001)	M, S	Perceived product quality (PPQ)		Continued search intention (CSI)	Return deliberation time (RDT), Purchase delibaration time (PDT)	M, S $\xrightarrow{-}$ PDT M, S $\xrightarrow{+}$ PPQ M, S $\xrightarrow{\text{n.s.}}$ RDT, M $\xrightarrow{-}$ CSI
Heim and Field (2007)	M, T, F, X				Customer ratings for ease of returns and refunds (CRR)	M, F $\xrightarrow{+}$ CRR T, X $\xrightarrow{\text{n.s.}}$ CRR
Mollenkopf et al. (2007)	F	Perceived value of returns offering (PVRO)	Return satisfaction (RS)	Loyalty intentions (LI)		F $\xrightarrow{-}$ PVRO $\xrightarrow{+}$ LI F $\xrightarrow{-}$ RS $\xrightarrow{+}$ LI
d'Astous and Guèvremont (2008)	T, X	Perceived product quality (PPQ), Perceived retailer image (PRI), Need for information (NI)				X, T $\xrightarrow{+}$ PRI X, T $\xrightarrow{\text{n.s.}}$ PPQ, NI
Posselt et al. (2008)	M, T		Liking (L)			M, T $\xrightarrow{+}$ L
Anderson, Hansen, and Simester (2009)	G				Monetary value of return option (MVRO), Average category purchase rate (ACPR)	G $\xrightarrow{+}$ MVRO G $\xrightarrow{+}$ ACPR (when MVRO is high)
Wang (2009)	M, T	Endowment effect (EE), Product valuation (PV)		Willingness-to-pay for product (WTPP), Return intention (RI)		M, T $\xrightarrow{+}$ PV T $\xrightarrow{+}$ EE, WTPP
Bonifield et al. (2010)	O	Perceived website control (PC)		Purchase intention (PI)		O $\xrightarrow{+}$ PI (when PC is high) O $\xrightarrow{\text{n.s.}}$ PI (when PC is low)
Petersen and Kumar (2010)	O				Average purchase value (APR), Average product return value (APRV), Average profit per customer (APC), Average number of referrals (ANR)	O $\xrightarrow{+}$ APR, APRV, APC, ANR
Suwelack et al. (2011)	T, F	MBG credibility (MBGC), Performance risk (PR), Financial risk (FR)	Anticipated regret (AR), Liking (L)	Willingness-to-pay for product (WTPP), Purchase intention (PI)		T $\xrightarrow{+,+,\text{n.s.}}$ MBGC, WTPP, PI F $\xrightarrow{+}$ MBGC, WTPP, PI MBGC $\xrightarrow{-,+}$ PR, L $\xrightarrow{-,+}$ WTPP MBGC $\xrightarrow{-}$ PR, FR, AR $\xrightarrow{-}$ PI MBGC $\xrightarrow{+}$ L $\xrightarrow{+}$ PI
Bower and Maxham III (2012)	M	Perceived Cost Fairness (PCF), Self-Attribution (SA), Retailer attribution (RA)	Post-return regret (PR)		Post-return spending (PS)	M $\xrightarrow{+}$ PCF (SA and RA positively moderate) M $\xrightarrow{+}$ PCF $\xrightarrow{-}$ PR $\xrightarrow{-}$ PS
Janakiraman and Ordóñez (2012)	T, F	Perceived cognitive effort (PCE), Perceived physical effort (PPE)		Return propensity (RP)		T $\xrightarrow{-}$ RP (when F is low) T $\xrightarrow{\text{n.s.}}$ RP (when F is high) T $\xrightarrow{+}$ PCE, T $\xrightarrow{\text{n.s.}}$ PPE, F $\xrightarrow{-}$ PPE
Kim and Wansink (2012)	S	Upward counter-factuals (UC), Downward counter-factuals (DC), Product valuation (PV)				S $\xrightarrow{+}$ UC, S $\xrightarrow{+}$ DC S $\xrightarrow{+}$ PV (under pre-purchase recommendation) S $\xrightarrow{-}$ PV (under no recommendation)

RP Classifiers: G: MBG Adoption, M: Monetary, T: Time, F: Effort, X: Exchange, S: Scope, O: Overall leniency

Table 2.11: Empirical Literature on Return Policy and Consumer Behavior - Part A

Article	Leniency Levers	Cognitive Responses	Affective Responses	Behavioral Intentions	Behavioral Actions	Key Findings
Hsieh (2013)	O	Perceived retailer opportunism (PRO), Cognitive trust (CT)		Loyalty intention (LI)		O $\xrightarrow{-}$ PRO $\xrightarrow{-}$ CT $\xrightarrow{+}$ LI
Lantz and Hjort (2013)	M				Order frequency (OF), Average value of orders (AVO), Average value of purchased items (AVPI), Average value of returned items (AVRI), Return probability (RP)	M $\xrightarrow{+}$ OF, RP M $\xrightarrow{-}$ AVO, AVPI M $\xrightarrow{\text{n.s.}}$ AVRI
Powers and Jack (2013)	O	Product dissonance (PD)	Emotional dissonance (ED)	Return propensity (RP)		O $\xrightarrow{-}$ ED $\xrightarrow{+}$ RP O $\xrightarrow{-}$ PD $\xrightarrow{+}$ RP
Bahn and Boyd (2014)	T	Assortment attractiveness (AA)				T $\xrightarrow{+}$ AA (under low variety assortment) T $\xrightarrow{-}$ AA (under high-variety assortment)
Pei et al. (2014)	M	Perceived return policy fairness (PRPF)		Purchase intention (PI)		M $\xrightarrow{+}$ PI, M $\xrightarrow{+}$ PRPF
Heiman et al. (2015)	T, X			Willingness-to-pay for return option (WTPR)		T $\xrightarrow{+}$ WTPR, X $\xrightarrow{+}$ WTPR
Hjort and Lantz (2016)	M				Average value of orders (AVO), Total value of orders (TVO)	M $\xrightarrow{-}$ AVO, TVO
Janakiraman et al. (2016)*	M, T, F, S, X			Purchase intention (PI), Return intention (RI)		M, F $\xrightarrow{+}$ PI, M, F $\xrightarrow{\text{n.s.}}$ RI T, X $\xrightarrow{\text{n.s.}}$ PI, T, X $\xrightarrow{-}$ RI S $\xrightarrow{+,\text{n.s.}}$ RI, PI
Zhou and Hinz (2016)	O				Sales volume (SV), Number of returns (NR), Overall profitability (OP)	O $\xrightarrow{+}$ SV, NR, OP (moderated by reputation)
Gelbrich et al. (2017)	M			Order intention (OI), Keep intention (KI)		M $\xrightarrow{+}$ OI M $\xrightarrow{+}$ KI (under low shopping frequency) M $\xrightarrow{-}$ KI (under high shopping frequency)
Jeng (2017)	O	Perceived value of return policy (PVRP)		Purchase intention (PI)		O $\xrightarrow{+}$ PVRP $\xrightarrow{+}$ PI (stronger for lesser known retailers)
Shang, Pekgün, et al. (2017)	G				Monetary value of return policy (MVRP)	G $\xrightarrow{+}$ MVRP
Rao et al. (2017)	T			Willingness-to-pay for product (WTPP)		T $\xrightarrow{+}$ WTPP
Zhang et al. (2017)	M, T	Perceived service quality (PSQ), Perceived product quality (PPQ), Perceived return difficulty (PRD)		Purchase intention (PI)		M, T $\xrightarrow{-}$ PRD $\xrightarrow{-}$ PSQ $\xrightarrow{+}$ PI M $\xrightarrow{+}$ PPQ, T $\xrightarrow{\text{n.s.}}$ PPQ
Oghazi et al. (2018)	O	Cognitive trust (CT)		Purchase intention (PI)		O $\xrightarrow{+}$ CT $\xrightarrow{+}$ PI

RP Classifiers: G: MBG Adoption, M: Monetary, T: Time, F: Effort, X: Exchange, S: Scope, O: Overall leniency

* Janakiraman et al. (2016) is a meta-analytic review paper.

Table 2.12: Empirical Literature on Return Policy and Consumer Behavior - Part B

return to scale (Wang, 2009; Suwelack et al., 2011; Heiman et al., 2015; Rao et al., 2017). Research on signaling quality suggests that time leniency can signal retailer service quality but may not signal product quality (d'Astous & Guèvremont, 2008; Zhang et al., 2017). In terms of the AR classifier, research shows that time leniency can stimulate liking towards a retailer and reduce anticipated regret (Posselt et al., 2008; Suwelack et al., 2011). Regarding the BI classifier, Zhang et al. (2017) suggest that time leniency can increase purchase intention, whereas Suwelack et al. (2011) cannot find a statistically significant relationship between time leniency and purchase intention. To our knowledge, no research examines in isolation how time leniency influences behavioral actions.

Four studies address effort leniency. In terms of the CR classifier, effort leniency positively influences customer ratings for ease of returns service (Heim & Field, 2007), perceived value of returns service in general (Mollenkopf et al., 2007), and MBG credibility (Suwelack et al., 2011), but reduces performance and financial risks perceptions (Suwelack et al., 2011). Also, the studies suggest that effort leniency can also influence certain AR constructs. In particular, effort leniency can increase return satisfaction and liking but can reduce anticipated regret (Mollenkopf et al., 2007; Suwelack et al., 2011). One study investigates the interaction between effort and time leniency and finds that higher time leniency can reduce return propensity when effort leniency is also high, but may not significantly impact return propensity when effort leniency is low (Janakiraman & Ordóñez, 2012).

The relationship between scope leniency and various consumer behaviors is not studied extensively. In fact, only Wood (2001) and Kim and Wansink (2012) examine scope leniency in different ways. Wood (2001) operationalizes scope leniency in terms of whether discounted products are allowed to return. She shows that higher scope leniency increases perceived product quality and decreases purchase deliberation time. Kim and Wansink (2012) operationalize scope leniency in terms of whether products are allowed to be returned for any reason or only defective products can be returned. They show that higher scope leniency leads to a higher product valuation when a retailer offers a pre-purchase recommendation but results in a lower product valuation when the retailer does not offer a recommendation.

Research that investigates exchange leniency is also limited with only three representative studies. Heim and Field (2007) show that exchange leniency is not a significant predictor of a customer's rating for a retailer's returns service convenience. d'Astous and Guèvremont (2008) conclude that exchange leniency may be effective in increasing perceived retailer image but may not signal product quality. Heiman et al. (2015) find evidence that customers value policies that offer cash refund (i.e., high exchange leniency) as opposed to store credit only (i.e., low exchange leniency).

Several studies investigate the effect of return policy leniency on consumer behavior by operationalizing leniency as an indistinguishable combination of multiple levers or as an overall leniency measure. Experimental and survey based studies show that overall leniency of a return policy can influence various constructs associated with CR and BI classifiers. In particular, these studies demonstrate that lenient return policies are associated with greater trust to a retailer (Hsieh, 2013; Oghazi et al., 2018), purchase and loyalty intention (Bonifield et al., 2010; Hsieh, 2013; Jeng, 2017; Oghazi et al., 2018), and that return policy leniency can reduce dissonance and return intention (Powers & Jack, 2013). Two studies look at the relationship between overall leniency and various behavioral actions that relate to retailer performance. Zhou and Hinz (2016) find that higher leniency is associated with an increase in both sales and returns. Overall profit seems to improve, but this benefit is likely moderated by a retailer's reputation. In a field experiment, Petersen and Kumar (2010) find that increasing return policy leniency leads to an increase in average purchase and return value, average profit, and average number of referrals, all measured on a yearly per customer.

To conclude, research at the intersection of RP and CB domains provides ample empirical evidence that return policy leniency significantly impacts various cognitive and affective responses along with behavioral intentions that are antecedents of actual behavioral actions. Collectively, these findings suggest that retailers should view return policies as a key part of their overall service offering that may impact the bottom-line. Finally, except for the meta-analytic review of 21 studies by Janakiraman et al. (2016), we are unaware of any study that investigates the influence of different leniency levers on consumer perceptions and/or behaviors within a unified empirical framework.

2.5.3 Return Policy, Planning & Execution, and Return Management

The empirical literature at the crossover of RP domain with the PE and the RM domains is rather scarce despite a few attempts in the earlier development of the literature that links return policy leniency to PE and RM concepts. Table 2.13 reports the few relevant studies, their classification in terms of RP, PE, and RM classifiers, and key findings.

At the RP and PE intersection, few contributions address PR, PA, PC, and C classifiers of the PE domain and MBG adoption, M, T, and O classifiers of the RP domain. As can be seen from the table, there is virtually no research that considers empirical relationships between return policy design and I, FD, and SC issues. Research in the intersection of RP and PM domains is even scarcer. To this end, we could find only three works that address the D and P classifiers of the RM domain and the G and O classifiers of the RP domain.

2.6 Conclusion

This review and classification of the analytical and empirical literature on consumer return policies demonstrates that, although the count of research contributions continues to grow rapidly, the literature is still nascent and evolving, particularly within the empirical stream. Though this manuscript provides a detailed state of the literature and numerous insights of both academic and practical relevance, our conceptual framework and classification may be even more notable for signaling what remains to be addressed in future research. In particular, this section will demonstrate and stress the importance of continued triangulation in terms of theoretical, analytical, and empirical evidence as well as continued observations from practice. To this end, this section highlights essential open questions and proposes a number of possible future research directions. To streamline the discussion and provide momentum going forward, the remainder of this section examines essential topics that capture the state of the literature. §6.1 discusses the need for a broader

Article	RP							PE							RM			Key Findings
	G	M	T	F	S	X	O	PR	I	SC	FD	C	PC	PA	A	P	D	
Davis et al. (1998)						✓		✓						✓			✓	There is a positive relationship between assortment variety and overall return policy leniency. There is no statistically significant relationship between retailer's salvage advantage and return policy leniency.
Mixon (1999)	✓							✓										Sellers are more likely to offer an MBG for experience goods, and the likelihood of MBG is even higher for high-price experience goods.
Van den Poel and Leunis (1999)	✓							✓										Offering both a price discount and MBG has a greater impact in terms of purchase risk relief than offering either of these individually. The authors also find that on average, MBG adoption is a more effective risk relief mechanism compared to price reduction, in terms of a positive impact on purchase likelihood.
Autry (2005)						✓											✓	There is a positive relationship between overall return policy leniency and return handling capabilities.
Posselt et al. (2008)		✓	✓					✓										A higher degree of leniency in both monetary (in the form of free return shipping) and time levers is significantly associated with higher prices.
Anderson, Hansen, and Simester (2009)	✓							✓					✓					The value of a return option is contingent on the product category and vary significantly across different categories.
Bonifield et al. (2010)						✓						✓						Overall return policy leniency can signal e-tailer quality, however, this effect is contingent on product characteristics. In particular, signaling effect is significant only for sellers of non-consumable products.
Bahn and Boyd (2014)			✓										✓					Facing a high-variety assortment, time leniency is negatively associated with assortment attractiveness, whereas facing a low variety assortment, leniency is positively associated with assortment attractiveness.
Shang, Pekgün, et al. (2017)	✓											✓					✓	Salvage value (using store status and feedback volume as proxies) is a major driver behind MBG adoption decisions of eBay sellers. Competitive intensity is not a major predictor of MBG adoption decisions of eBay sellers.

Table 2.13: Empirical Research in RP–PE and RP–RM Intersection

conceptualization of customer heterogeneity and related implications for both research and practice. From there, we delve into theory building and the drivers of return policy design in §6.2, both in terms of consumer behavior and in terms of managerial decision-making. §6.3 then addresses examples of how to expand insights in the domain of research on return policy design, particularly through the opportunity to extend beyond monetary leniency. Next, in §6.4, we address specific research opportunities that lie at the intersection of return policies with the other domains in our framework and finish in §6.5 by addressing needs with respect to integrating the analytical and empirical literature.

2.6.1 Customer Heterogeneity and Implications for Research and Practice

The literature convincingly demonstrates that customers are heterogeneous in their return behaviors and some current and forthcoming contributions have gone so far as to classify different customer segments based on their return propensities (Foscht et al., 2013; Ketzenberg et al., 2018). While the literature maintains a distinct focus on the differences among customers that make returns, an interesting empirical finding is the presence of a large proportion of non-returners (Petersen & Kumar, 2010; Ketzenberg et al., 2018). If a customer never returns products, the literature posits that such a customer has a very high return cost, a strong preference for a retailer's products (Anderson, Hansen, Simester, & Wang, 2009), or perhaps other psychological factors are at play (Dahl, 2016). In fact, the studies by Foscht et al. (2013) and Ketzenberg et al. (2018), suggest that retailers could employ differentiated return policies for different customers, penalizing heavy returners and offering incentives to infrequent returners (Gelbrich et al., 2017). In practice, we already see that some retailers are moving in this direction, as evidenced by Jet.com that offers an upfront discount to customers that opt out of free returns (Jet.com, 2019) and others such as Amazon who have taken action to ban customers for excessive returns (Picchi, 2018).

Excessive returners include those customers who abuse return policies in the form of opportunism or even return fraud. The difference between fraud and opportunism is not just one of legality. At the heart of opportunism is the consumer's full intention to return a product prior to purchase and thereby obtain value at the expense of a retailer. Unlike an opportunistic return, a fraudulent return occurs when a person engages in criminal activity such as returning stolen goods—a product returned without an associated purchase transaction. Most of the academic interest has focused on opportunism. We are aware of only one active working paper (Akturk et al., 2019) that addresses both fraud and opportunism. According to a recent survey by the National Retail Federation, opportunism, often referred to as friendly fraud, reached $5.2 billion in 2017 (National Retail Federation, 2017). In that same report, actual return fraud constitutes an even greater amount at $17.6 billion (National Retail Federation, 2017).

As stated by Petersen and Kumar (2009) and Hjort et al. (2013), if different customers return different amounts, retailers need to understand the underlying relationship among customer purchase behavior, customer characteristics, and return behavior. In effect, various customer segments require differentiated managerial consideration and treatment. Fortunately, with today's technologies and loyalty cards, return policies effectively can be customized for specific customer segments and even individual customers. Best Buy currently differentiates customer classes by time leniency, offering customers that purchase higher volumes a longer period of time to make a return. In addition, several mainstream retailers contract with third parties, such as The Retail Equation, to track individual consumer return behavior in order to prevent return abuse (Safdar, 2018). 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. Customers may perceive such actions by retailers as promoting fairness (i.e., only penalizing abusive actors) if such countermeasures work reliably and accurately. Conversely, customers may perceive such data-driven countermeasures or customized policies as unfair, punitive, and overly invasive with regard to privacy.

2.6.2 Drivers and Theory of Return Policy Design

Returns often stem from risks associated with uncertainties regarding product fit and valuation, particularly for remote purchases. Of course, a primary function of return policies is to reduce such uncertainty and perceived risk. Yet, little research examines how customers manage these risks. Indeed, the simplest means to avoid risk is to make no purchase at all. That said, if a purchase is necessary, consumers may intensify pre-purchase search and learning as well as continued search and product evaluation post-purchase. Customers may also buy a basket of items, evaluate the items within the return time window, and return the items that are not satisfactory. How consumers assess return policies in relation to such risks largely remains an open empirical issue. Further, another key driver of return policy design is simply to provide a return option to enable service recovery (Mollenkopf et al., 2007). Service recovery provides the retailer with a second opportunity to satisfy customer needs and a second opportunity to generate profit (Etzel & Silverman, 1981; McCollough & Bharadwaj, 1992). Our review reveals that there is little research on service recovery options or how such service recovery relates to the return policy decision. Though not within the scope of our review, the paper by Ertekin (2018) suggests potential service recovery opportunities.

Though a fair amount of empirical work investigates consumer behavior, there is little that addresses the retailer side of the equation. Specifically, we are unaware of any research that focuses on *how* return policy decisions are made by managers in practice. In other words, what internal and external factors play a key role in return policy decision-making, either in designing new return policies or changing existing policies? Of course, the analytical research speaks to these factors, such as salvage value and product price, that influence optimal return policy design. Moreover, a number of studies links these factors to return policy leniency using secondary data, such as Davis et al. (1998) and Shang, Pekgün, et al. (2017). Yet, an empirical grounding in practice is largely absent. Further, we know little about how important the return policy decision is with respect to firm performance. How do managers align return policy decisions with their firm's overall competitive strategy? In what ways should operational and tactical decisions regarding returns and returns processes be made so that they are in alignment with the return policy and firm strategy? These represent significant opportunities for continued research.

Our classification also emphasizes the need for theory-driven empirical research on return policy design. Less than half of the already small subset of published empirical research explicitly draws upon, extends, or references established theory. Moreover, we observe that the theory base of RP research largely consists of theories from psychology and behavioral economics, which is primarily due to the dominance of studies focusing on consumer behavioral aspects of return policies. Future research, particularly in the direction of managerial decision-making, can contribute to this growth by transferring management and organization theories, as well as transaction cost economics theories, into the domain. We conjecture that theories of contingency, organizational information processing, organizational learning, and psychological contracts can be particularly useful in this regard. By drawing upon such theories researchers can build novel conceptual and theoretical frameworks to identify key internal factors and operating conditions that underlie return policy decisions. In turn, these frameworks can build a foundation for understanding how these factors can be linked to managerial and organizational cognition mechanisms and information processing capabilities.

2.6.3 Return Policy Design and Extending Beyond Monetary Leniency

Analytical modeling research on return policies has a significant body of work regarding money-back guarantees and monetary leniency. Though seemingly studied at length, the review and classification reveals that significant opportunities still remain for research on monetary leniency. As an example, more research is needed to understand under what conditions a full versus partial refund is optimal. For instance, Altug and Aydinliyim (2016) attempt to demonstrate optimality of a full-refund equilibrium by revising several key modeling assumptions as discussed in §4.2.2. Yet, many studies conclude that partial-refund policies are optimal, as full refunds are overly generous and can subject a retailer to costly opportunistic behavior. Even without consideration of opportunism, the findings from several studies (e.g., Su (2009a)) indicate partial refunds are optimal. Yet, in practice, most mainstream retailers offer full refunds for at least some extended period of time after purchase.

Continued research could incorporate empirically observed consumer behavioral reactions to return policy decisions related to full versus partial refunds when implemented in practice. One such reaction relates to unfairness perceptions and resulting outcomes due to restrictive policy changes. The importance of fairness aspects in the study of microeconomic exchanges has been long pointed out, perhaps most famously by Kahneman et al. (1986). Empirical research, through distributive and procedural justice theoretical lenses, shows that customers perceive partial refund policies very negatively and such policies may decrease customer's positive behavioral intentions and outcomes (Bower & Maxham III, 2012; Pei et al., 2014). Further, there is significant anecdotal evidence regarding negative attitude changes toward retailers who charge restocking fees or impose restrictions to long-time lenient policies (Narvar, 2017; UPS, 2017; Price, 2018). Hence, one potential reason why major retailers maintain full-refund policies might be the potentially detrimental effect from adopting a partial refund policy in lieu of a long-established full-refund policy. In fact, many retailers seem to be aware that their overly lenient return policies hurt their bottom-line but still consider such policies an integral part of their value proposition (Dennis, 2018). In effect, retailers may choose a sub-optimal policy due to competitive or strategic considerations, though continued research needs to establish the underlying reasons.

An interesting observation from Tables A.1 and A.2 (see Appendix) is that out of the 67 analytical studies included in our review, only 14 involve time, effort, scope, and exchange leniency combined. Researchers often use single-period, single product, stylized game-theoretical models that focus on optimal return policy decisions where the interplay of several operating parameters drive main results. Of course, such contributions are essential to building a foundation to the field and developing rigorous, key managerial insights. At the same time, these assumptions often are not conducive for conducting research on some of the non-monetary return policy levers, most notably scope and exchange leniency. Consider that most modeling literature usually focuses on the single product case. By definition, scope leniency addresses the issue of return policy complexity that arises from product assortments in which characteristics across products may necessitate different return policy considerations. Relatedly, there is an apparent need for research regarding drivers of different return policies for different product categories. Regardless of the rationale, many retailers have complex return policies. For instance, the return policies for a group of 12 prominent retailers that includes Amazon, Kohl's, Best Buy, Walmart, Target, etc., span some 67 pages, totaling over 30,000 words (ConsumerWorld.org, 2017). Hence, return policy scope represents a significant, but largely unexplored, domain in the extant research.

The time leniency lever could also be examined in a multiple product category setting in which the objective is to develop an optimization framework to determine an optimal return policy portfolio by assigning discretized return time windows to various product categories. Shang et al. (2018) demonstrate category-level heterogeneity in return timing and discuss potential benefits of a category-based return policy portfolio with varying time leniency. Thus, significant opportunities remain for research on non-monetary policy levers. Research that addresses more than a single return policy lever is rare, providing opportunities to explore potential interactions amongst the different leniency levers as well as establishing their relative importance in setting policy.

More generally, research is needed on the analysis of product assortments and return policy design given the many varied consumer behavioral reactions. For example, consumers typically buy baskets of goods and may have different return propensities for different items in a basket. How does the return policy affect basket size and product selection? For instance, retailers often offer free shipping for a minimum purchase size. Consumers may fill the cart with items in order to meet the threshold and then later return those "filler" items. Cachon et al. (2018) is an active working paper on this topic that highlights the need to also address return policy design issues. As another example, consider that customers often engage in sampling behavior in which they will buy multiple products with the express intent of trying them all and returning all but the one or two that fit best—a practice encouraged and exemplified by Amazon's Prime Wardrobe service. Though these types of behaviors are common and prevalent, the current modeling paradigms are unable to capture these effects when examining a single product.

Overall, a fundamental premise underlying virtually all models addressing return policy decisions is that returns are costly (e.g., requiring that the salvage value of a return is less than the procurement cost). Of course, returns can indeed be quite costly as evidenced by a plethora of industry sources (Enright, 2013; National Retail Federation, 2017; Boyajian, 2018; Dennis, 2018). Nevertheless, growing evidence shows that returning customers are more profitable than non-returning customers (Hjort et al., 2013; Ketzenberg et al., 2020). In fact, customer profitability can actually increase as returns propensity increases, at least up to a threshold (Petersen & Kumar, 2010; Ketzenberg et al., 2020). Apparently, customers who value return policy leniency will purchase more and exhibit greater loyalty than their non-returning counterparts. Thus, there is an opportunity for analytical research focused at the customer level that can provide important insights not available at the transaction level. As noted by Mollenkopf et al. (2011), Petersen and Kumar (2010), and Ketzenberg et al. (2020), managing costs as well as customer relationships are both essential to effectively managing returns. Extending the modeling paradigm to include multiple periods or otherwise allowing multiple purchases and multiple returns is a key means to enable analysis at the customer level and for managing customer relationships over a lifetime of transactions. Moreover, such consumers appear to continually repeat such return transactions—behavior that has inspired the label *serial returner* (Samorani et al., 2016). Yet, the single period models commonly used in the analytical modeling literature are not conducive to studying such characteristics. As such, the modeling of serial return behavior represents a motivating example to extend the modeling paradigm to include multi-period models. Similarly, time leniency, time-to-return, as well as the decay in value of product returns over time, also serve as additional catalysts for the study of multi-period models on consumer returns. Such expanded analyses may provide new insights into the optimality of partial or full refunds.

2.6.4 Research at the Intersection of Domains

There is a paucity of research on return policy design that intersects the RM and PE domains. Consider for example, research at the intersection of the RM domain. The empirical results in the literature indicate that post-return handling and management is an important element to return policy design (Petersen & Kumar, 2010; Bower & Maxham III, 2012). In the modeling literature, however, return management is typically abstracted to simply assuming a fixed salvage value for returns. Shulman et al. (2010), Akçay et al. (2013), and Hsiao and Chen (2015) are notable exceptions that delve into decision-making aspects of return processing or disposition.

The cost and timeliness of acquiring and processing returns should have implications with respect to return policy design. For example, research has shown that the value of product returns can decay over time. Hence, one would expect time leniency to be affected by the rate of the decay, as modeled in Xu et al. (2015), but there may also be implications with respect to the other leniency levers. Hence, there remains an opportunity to empirically investigate the relative importance and influence that these levers have on consumer behavior, particularly with regard to return timing. These latter observations also indicate additional research opportunities arising at the intersection of the CB and RP domains.

Research on disposition options for returns in the context of return policy decision-making is limited. Many returns are never opened, remain effectively as good as new, and can simply be put back on the shelf. Yet, other returns can only be liquidated. In fact, a variety of disposition options exist, including: resell as new, warranty replacement, resell in secondary markets, resell as open-box item, salvage, dispose, and liquidate to third parties. Indeed, many retailers employ a variety of methods to dispose of their returns. Joint examination of a firm's disposition capabilities and return policy decisions, perhaps also incorporating planning and execution insights, offers yet another research opportunity (Autry, 2005). For example, when returns can be resold as new or open-box, the required quantity of new products is reduced (Ketzenberg & Zuidwijk, 2009; Akçay et al., 2013). When returns can only be salvaged, the required quantity of new products increases (Su, 2009a). This simple comparison illustrates different outcomes for planning and execution (inventory optimization) with respect to differing disposition options and highlights the need for research into how return policy interacts with these operational elements.

As with return management, there remains a need for research at the intersection of the RP and PE domains. Indeed, there is a growing body of literature that considers numerous PE related issues in the context of product returns, but most of this literature does not address decision-making regarding the return policy itself. Consider for example that there are a number of operational tools, relevant to planning and execution activities, whose effectiveness in reducing purchase risks (i.e., increasing sales) and mitigating returns are empirically documented. For example, Gallino and Moreno (2018) demonstrate in a randomized field experiment that virtual fitting-room technologies can increase sales and decrease likelihood of returns, mostly due to decreasing the number of multiple ordering customers. De et al. (2013) investigate the effect of different product-oriented web technology features (i.e., zooming, alternative product photos, color swatch) on returns and conclude that not all types of product information provision techniques are likely to reduce returns. In practice, the implementation of such technologies requires significant investments from retailers. Hence, future research could assess the conditions under which firms can generate value from such technology investments and how such investments enable changes in return policies.

2.6.5 Integrating Empirical and Analytical Research

A final point of our study relates to opportunities for methodological triangulation and integrating analytical and empirical research (Singhal & Singhal, 2012). Integration is not merely about citing earlier analytical (empirical) findings in a new empirical (analytical) study. Integration also involves 1) conducting empirical-analytical hybrid studies, such as Davis et al. (1998), 2) empirical testing of insights from analytical literature, and 3) incorporating new findings from empirical literature in developing new analytical models. We elaborate on points 2) and 3) below.

Our classification is noteworthy by identifying several new and significant findings from the empirical literature on consumer behavior that have not yet been integrated into analytical studies. For instance, though a large proportion of customers apparently do not make a return (Petersen & Kumar, 2009; Ketzenberg et al., 2018), many customers value the return policy itself and are therefore generally willing to pay more for return policy leniency (Wang, 2009; Heiman et al., 2015; Rao et al., 2017). Further, an item will tend to exhibit a lower return rate when offered at a lower price than at a higher price (i.e., a within-item price effect) (Shang et al., 2018). Similarly, items in an assortment offered at higher prices will tend to also exhibit higher return rates than other items in the same assortment offered at lower prices (i.e., a between item price effect). We also note that the probability of a return decreases in the time since purchase (i.e., the endowment effect). Finally, as mentioned previously, customers who return more are *generally* more profitable than customers who return less. Addressing these empirical findings in the context of analytical models holds much promise for enhancing our knowledge concerning optimal return policy decision-making.

Just as empirical research is capable of informing analytical modeling research, insights generated from the modeling literature can also benefit from research triangulation from the empirical literature. Consider the relationship between competition and return policy leniency. Shulman et al. (2011), Ofek et al. (2011), and Altug and Aydinliyim (2016) suggest that competition may result in lower policy leniency, whereas B. Chen and Chen (2017a) and J. Chen et al. (2018) suggest otherwise. Alternatively, Shang, Pekgün, et al. (2017), using eBay data, do not find a relationship between competitive intensity and MBG adoption. Hence, empirical investigation through behavioral experiments or other methods may help to clarify how retailers integrate competitive issues into their return policy decisions and help to identify the significant moderating factors.

As this concluding section makes clear, though the extensive classification and review revealed a significant body of research, numerous triangulation opportunities across the analytical and empirical domains remain. The above represents a mere sampling of the myriad of future directions available to researchers in the return policy domain. Nonetheless, we hope that this discussion will be thought provoking and provide a first step toward understanding and seizing upon these future research opportunities.

3. HOW CONSUMERS VALUE RETAILER'S RETURN POLICY LENIENCY LEVERS: AN EMPIRICAL INVESTIGATION*

3.1 Introduction

Consumer returns are endemic to U.S retail practice. The annual value of returned products in the U.S. has grown to over $400 billion, representing 10.6% of total sales (National Retail Federation, 2021). Product proliferation, heterogeneity in customer expectations and valuations, and the rise of e-commerce can all be listed among the drivers for this growth (Cheng, 2015). Another key contributor is the proliferation of generous return policies (Shang, Pekgün, et al., 2017). Fierce competition, decreased consumer switching costs, and a "Customer is the King" mantra lead many retailers to offer overly lenient return policies. Yet, these retailers bear returns related operational costs that exceed $100 billion per year (Blanchard, 2007). Generous return policies also engender moral hazard issues and result in the rise of fraudulent and opportunistic returns, which recently surpassed $25 billion per year (National Retail Federation, 2021). More dramatically, excessive returns driven by such lenient policies pose significant environmental problems. For instance, products returned to U.S. retailers in a year generate 5 billion pounds of landfill which is equivalent to trash produced by 5 million Americans (Constable, 2017). Handling processes of these returns annually generate 15 million metric tons of carbon dioxide (Optoro, 2018).

Lenient return policies have a long history in the U.S., starting with J.R.Watkins' "Satisfaction guaranteed, or your money back!" policy in 1868. Today, consumer return policies come in a myriad of forms. Some retailers offer a no-questions-asked, full-refund return policy, whereas others try to disincentivize returns by charging restocking fees, setting strict return deadlines, and imposing hassles. Recently, the tide seems to be turning against leniency. Many large retailers, such as Best Buy, Macy's, and Bed Bath and Beyond, have tightened their return policies over the last couple of years to reign in the cost of returns, though the specifics of how they choose to do so has

*Reprinted with permission from "How Consumers Value Retailer's Return Policy Leniency Levers: An Empirical Investigation" by Huseyn Abdulla, James D. Abbey, Michael Ketzenberg, 2021. *Production and Operations Management, n.a.(n.a.), 1–33, Copyright (2021) by John Wiley & Sons - Books.*

significantly varied (ConsumerWorld.org, 2018). Some retailers imposed restocking fees on certain product categories, whereas others shortened the allowable time window for making returns. The variety of changes observed may be in part due to the lack of comprehensive research into the impact that these changes have on consumer responses—a primary purpose of this research.

In general, return policy design poses an interesting operations–marketing interface problem in today's retail environment. From the marketing perspective, more leniency can stimulate customer purchases through positive product and service quality signaling and a reduction of purchase related risks and thereby enhance the consumer value proposition. From the operations perspective, return policies constitute an important strategic lever in the front-end of closed-loop supply chains that influences the volume, timing, and quality of product returns (Guide & Van Wassenhove, 2009). Further, the choice of return policies, through documented influences on consumer purchase and return behaviors, is intertwined with and carries significant implications for other retail planning and execution activities. Such activities range from product pricing to assortment decisions to inventory management (Abdulla et al., 2019). Thus, retailers face a challenge to find a solution that balances the fundamental operational cost versus consumer value proposition trade-off. In this paper, we contend that effective management of this trade-off requires an understanding of the process through which consumers perceive, value, and react to different return policies.

From the customer's standpoint, a return policy is an important element of the returns service offered by a retailer and plays a role in the valuation of the returns service (Davis et al., 1998). Anecdotal evidence from popular press often suggests that "the best" return policies share some common characteristics that signal leniency (Kirkham, 2015; Mash, 2017). These characteristics include a full refund, a long return window, low return effort, among others. Janakiraman et al. (2016) develop a typology of return policies that consists of five leniency dimensions: monetary, time, effort, scope, and exchange. We note that all five dimensions of return policy leniency can be strategically chosen by retailers to influence consumer perceptions and behaviors. Following Abdulla et al. (2019), we refer to these dimensions as return policy leniency levers. Table 3.1 provides definitions for each of the five leniency levers.

Leniency Lever	Definition and Explanation
Monetary	Amount of refund as percentage of the price paid. Return policies offering a full refund are more lenient. Less lenient policies impose "restocking fees" or a "non-refundable shipping and handling fee."
Time	Duration allowed for a return. Return policies that allow a longer return period are considered as more lenient.
Effort	Consumer effort required to execute return. Effort leniency varies based on the hassles imposed by a retailer. Hassles take various forms, such as filling out forms, providing receipts, requiring original packaging/tags, ID verification, etc.
Scope	The extent of return-eligible products (i.e., the degree of exclusivity). For example, less scope lenient policies may completely exclude sales and clearance items from being returned.
Exchange	Whether the retailer offers cash refund or only allows store credit/exchange. A more lenient policy allows a cash refund as an option.

Note: Adapted from Janakiraman et al. (2016) and Abdulla et al. (2019).

Table 3.1: Conceptual Definitions of Return Policy Leniency Levers

Consumer return policies are studied extensively in the analytical operations management (OM) literature as a strategic instrument that can be used to stimulate sales (Ketzenberg & Zuidwijk, 2009), control opportunistic returns (Shang, Ghosh, & Galbreth, 2017), signal quality (Moorthy & Srinivasan, 1995), compete (Shulman et al., 2011), and coordinate a supply chain (Su, 2009a). However, these papers predominantly focus on the monetary leniency lever with only rare examples studying the other levers (for a list, see Abdulla et al., 2019). Empirically, several studies examine the antecedents (Shang, Pekgün, et al., 2017) and performance impacts (Ertekin & Agrawal, 2021) of return policy leniency. Though valuable contributions, these studies leverage unique contexts involving a single retailer or eBay auctions and use quasi-experimental or non-experimental approaches to estimate the impact of return policy leniency based on observed transactional data. As such, these works do not propose or empirically test causal consumer cognitive mechanisms that potentially drive the observed behaviors.

Empirical research in the Operations–Marketing and Operations–Information Systems interfaces focuses on the effect of return policy leniency on various cognitive, affective, and behavioral reactions, such as purchase risk, willingness-to-pay, product and service quality perceptions, dissonance and regret, among others (e.g., Mollenkopf et al., 2007; Suwelack et al., 2011). The vast majority of the studies operationalize and manipulate return policy leniency through only one or two levers or use an overall leniency measure. In general, the collective evidence from the existing literature signals that different leniency levers may have varying effects on these perceptual and behavioral constructs. Nevertheless, the existing studies vary significantly in terms of their empirical models and study designs. As such, making comparisons on the relative effects of these levers is not feasible. In fact, there are inconclusive and even conflicting findings regarding the effect of a particular leniency lever when results across studies are cross-examined.

A comparative analysis of the relative effectiveness of the leniency levers in terms of influencing consumer cognitive and behavioral responses has not been conducted to date. A limited number of works that discuss consumer valuation of a returns service have acknowledged this gap and made a call for future research in this direction (Mollenkopf et al., 2007; Griffis et al., 2012; Jeng, 2017). Janakiraman et al. (2016) conduct a meta-analytical review of 21 studies to derive insights regarding the effectiveness of each of the five leniency levers. Though a valuable contribution, the considerable variation in the theoretical frameworks, research designs, and empirical contexts of the previous studies, as well as the exclusion of several more recent studies from the analysis, limits generalizability and applicability of the results (Abdulla et al., 2019). Therefore, there is a need for a systematic examination of all return policy leniency levers in the context of a unified empirical framework, in order to reconcile the earlier findings and obtain more reliable and actionable insights for retailers. Our research addresses this need. We theoretically and empirically explore how consumers perceive and value return policy leniency levers available to retailers and how this valuation process influences a consumer's intention to purchase from a retailer. In particular, we investigate how different levers influence consumer purchase intentions through a cognitive process that involves perceived service quality, perceived transaction costs, and perceived value of a returns service.

We make several contributions to the literature on consumer return policies with managerial implications. Using general merchandise stores (i.e., department stores, big-box stores, variety stores) as the empirical context, first, we explain how return policy leniency across different levers indirectly influences a consumer's purchase intention through a cognitive process model predicated on a mental accounting framework that combines transaction cost economics and signaling theoretical lenses. Second, we demonstrate that the policy aspects of a service (i.e., attributes pertaining to terms and conditions), rather than solely the process aspects (i.e., attributes pertaining to service encounter and experience), can generate service quality and transaction cost perceptions. Further, we find that these perceptions can trigger service value assessment and influence subsequent behavioral intentions and outcomes. In particular, we show that return policy leniency positively affects the valuation of a returns service and that the perceived value of the returns service significantly impacts a consumer's purchase intention. This finding supports the view that a return policy, which is fundamental to a retailer's returns service, is an important element of the overall service bundle and may influence patronage decisions of consumers.

The cognitive process model also provides a foundation for limited theorizing, through deductive reasoning, on the relative influence of the levers. In particular, we hypothesize and our empirical findings support that consumer purchase intentions are more sensitive to levers that pose direct financial risks to both retailers and consumers (monetary and exchange) than those that do not (effort, time, scope). Essentially, we find that monetary leniency, followed by exchange leniency, are the two most influential levers. Time, effort, and scope leniency show significantly lower impact on the perceived service value and purchase intention of a consumer. These findings imply that retailers who consider their return policies to be overly lenient and unsustainable from an operational perspective should consider restricting their return policies not through monetary and exchange leniency levers. Rather, the retailers should prioritize the remaining three levers in order to alleviate the cost burden of returns, since these levers minimally impact perceived service value and purchase intention. Overall, our paper constitutes an important inquiry into sustainable consumer return policies and can motivate future research to understand how retailers can design return policies that still provide value to customers (people), do not hurt the financial bottom-line of retailers (profit), and reduce the volume of returns, more than half of which end up in landfills (planet) (Constable, 2017; Howland, 2017).

The remainder of this paper is organized as follows. In §3.2, we set forth the conceptual background, theory, and hypotheses. §3.3 introduces the empirical methodology to test the hypotheses and presents the data analysis and results. Finally, §3.4 is dedicated to theoretical and managerial insights, along with future research opportunities and some limitations of our work.

3.2 Theory Development and Hypotheses

We start with a brief conceptual background for the key constructs and provide several propositions based on theories from behavioral economics, cognitive psychology, and marketing as well as empirical findings from the existing return policy literature. We then combine these propositions to theorize on a cognitive process model that explains how return policy leniency influences a consumer's purchase intention through a parallel-serial mediation mechanism which involves perceived service quality (PSQ), perceived transaction costs (PTC), and perceived service value (PSV). We propose a set of hypotheses for testing this parallel-serial mediation mechanism. Then, through deductive reasoning, we hypothesize that leniency levers with direct financial risks to both retailers and consumers (i.e., monetary and exchange) have a stronger effect on purchase intentions than the other three levers. The cognitive process model is introduced in Figure 3.1.

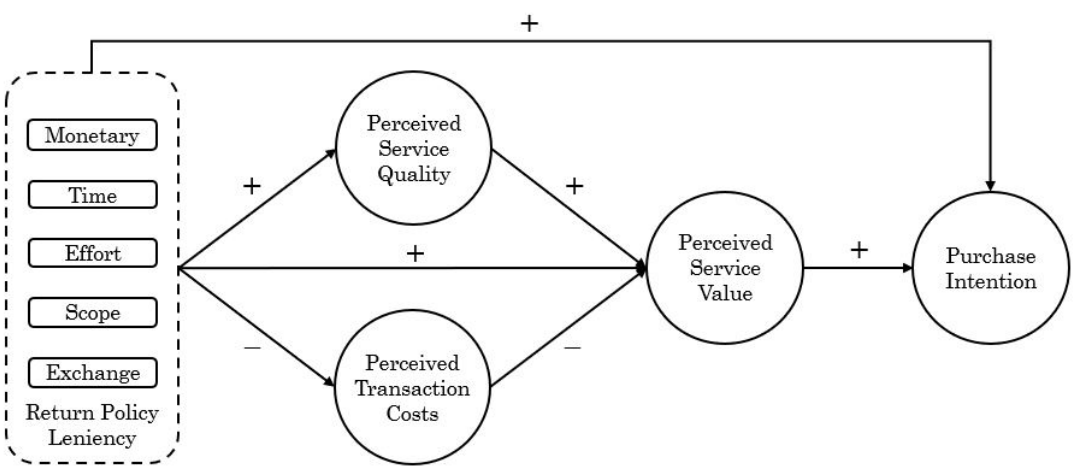

Figure 3.1: Cognitive Process Model and Empirical Directionalities

Service quality is the degree of discrepancy between the customer expectations for a service and the actual perceptions of performance (Parasuraman et al., 1985). Expanding on this definition, service quality is the overall evaluation of service performance, compared against the customer's general expectations of what a service should offer (Parasuraman et al., 1988). In general, researchers agree that expectations have a strong effect on the perceptions of service quality (e.g., Boulding et al., 1993; Cronin & Taylor, 1994). Customers may have different sources of information that lead to expectations about a potential service encounter with a particular provider. Among the listed sources, exposure to similar services of competitors, word-of-mouth, and company controlled communications are most relevant to the returns service context (Parasuraman et al., 1991).

We draw upon the signaling theory to establish the relationship between return policy leniency and perceived service quality (Spence, 1973). Signaling theory posits that sellers can use costly mechanisms to reduce information asymmetry between buyers and sellers, signaling positive characteristics such as quality. An efficacious signal is one that is costly to the sender and observable to the receiver (Connelly et al., 2011). Return policy leniency embodies both characteristics. Lenient return policies are costly to retailers. For example, higher time leniency allows consumers to return products after a potentially long trial period, resulting in revenue loss, decreased salvage value, and increased return handling costs. Offering full refunds (high monetary leniency) or making returns process hassle-free (high effort leniency) may stimulate convenience and opportunistic returns, resulting in revenue loss and increased operational costs. Further, offering cash refunds as opposed to store credits or exchange-only policies (high exchange leniency) leads to significant opportunity costs, as retailer becomes unable to secure repeat transactions and exchanges. More generally, lenient return policies make retailers vulnerable to opportunistic and even fraudulent return behaviors, which further increases the costs of returns (National Retail Federation, 2021). With respect to observability, return policies are observable to consumers as retailers are required by law to clearly communicate their return policies in stores or on websites.

To reinforce the signaling theoretical perspective, we employ the will-should expectations framework to explain the ex-ante impact of return policy leniency on perceived service quality (Boulding et al., 1993). The framework proposes two expectations standards—will and should expectations—and look at their influences on perceived service quality in a behavioral process model. Will expectations are those expectations that pertain to what will happen during a service contact with a provider. Will expectations can arise due to a company's service process descriptions, terms, and conditions. Prior experience with the same service provider can also generate such expectations. Should expectations are formed on the basis of a customer's perceptions regarding what is feasible and reasonable to receive in a service encounter. Should expectations can also arise due to experiencing or being told of competitors' services. Through field studies, Boulding et al. (1993) find strong evidence that both classes of expectations influence a customer's perceptions of service quality. Specifically, the study reveals that will expectations have a positive impact, whereas should expectations have a negative impact on perceived service quality.

The will-should expectations framework of service quality has implications for the returns service context. A clearly stated return policy sets customer will expectations, as the policy specifies what customers will go through if they have to use a retailer's returns service. Meanwhile, prior return experience or familiarity with the return policies of other retailers engender should expectations. When a customer faces the return policy of the retailer under consideration, we expect the customer's will and should expectations to activate. As a result, different levels of perceived service quality emerge, depending on the retailer's return policy leniency across different levers. Therefore, combining the signaling theory and will-should expectations framework leads us to the following proposition:

Proposition 1 (a–e)**.** *Higher a) monetary, b) time, c) effort, d) scope, e) exchange leniency leads to higher perceived quality of a returns service.*

The concept of transaction costs has its roots in Transaction Cost Economics (TCE) theory, which posits that buyers and sellers experience costs embedded within different aspects of transactions (Coase, 1937; Williamson, 1989). The theory suggests that the price of goods and services is not the sole criterion that influence a buyer's decision. The buyer tends to have an overall assessment of non-price transaction costs before deciding whether or not to engage in a particular exchange. Originally employed to examine inter-firm exchanges, the TCE perspective has been applied by researchers in a multitude of exchange contexts between economic agents, including between consumers and retail firms (e.g., Grønhaug & Gilly, 1991; Griffis et al., 2012). Asset-specificity, uncertainty, and transaction frequency are important characteristics of a transaction that generate transaction costs (Williamson, 1989).

In the retail exchange and consumer returns context, transaction costs encompass monetary-type, time-type, and psychological-type costs (Chircu & Mahajan, 2006). Monetary-type costs may include restocking fees, non-refundable forward or return shipping fees, transportation costs, return packaging costs, etc. Time-type costs capture time spent on initiating a return order, commuting to store, waiting for refund processing, searching for a new product to make an exchange, etc. Psychological-type costs are hard to quantify yet are believed to have a strong impact on customer's perceived sacrifice while engaging in an exchange (Woodall, 2003). Costs of this type may pertain to mental and physical effort, stress, inconvenience, frustration, and annoyance experienced in the post-purchase period.

In line with the TCE perspective, consumers incur transaction-specific costs (i.e., asset specificity) described above when searching for, deciding on, and purchasing a product from a retailer and when returning the purchase to the retailer. Significant uncertainty exists in such transactions due to uncertainties related to product fit, quality, and personal valuation (Abdulla et al., 2019). Return policies can be viewed as contractual mechanisms that govern the allocation of transaction-specific costs generated due to these uncertainties between retailers and consumers. Customers who have prior experiences of purchasing and returning to a particular retailer (i.e., transaction frequency) are likely to have an overall assessment of transaction costs that they incur. We posit that regardless of transactional history, retailers, by offering a lenient return policy, can reduce ex-ante perceived transaction costs of consumers which can subsequently increase their purchase intentions.

For instance, low monetary leniency (e.g., a 15% restocking fee) may increase the perceived risk of losing a certain amount of money in case of a product mismatch. Low time leniency implies opportunity costs to consumers who could otherwise procrastinate returning and may generate time pressure in assessing the fit and value of the product. Low exchange leniency (i.e., a store credit or exchange only policy) may increase psychological costs associated with the possibility of getting locked-in with the retailer by imposing a requirement to choose one of the products offered by the retailer. Low effort leniency may trigger consumers to mentally simulate the return process and perceive high time-, monetary-, and psychological-type transaction costs. Low scope leniency, such as being disallowed to return discounted products, may generate psychological costs such as anticipated regret. Thus, we have the following proposition:

Proposition 2 (a–e)**.** *Higher a) monetary, b) time, c) effort, d) scope, e) exchange leniency leads to lower perceived transaction costs.*

Mental accounting theory explains consumers' patronage and purchase decisions under risk and uncertainty (Thaler, 1985). The theory posits that consumers evaluate a transaction with a party in two stages: 1) evaluating the potential transaction (judgment process) and 2) approving or disapproving the transaction (decision process). To evaluate the transaction, consumers weight the perceived utility against perceived disutility, which encompasses all types of transaction costs discussed above. Based on this mental accounting process, consumers realize a perceived net utility (i.e., value). The perceived value of the transaction then leads to behavioral intentions and outcomes of consumers.

Perceived value is defined as a consumer's overall assessment of the utility of a product or service based on perceptions of what is received and what is given (Zeithaml, 1988). The construct of perceived value has been studied extensively in both product (e.g., Simpson et al., 2019) and service (e.g., Buell & Norton, 2011) contexts in the OM domain. Many researchers report a strong impact of perceived service value on consumer behavioral outcomes such as loyalty, (re)purchase intentions, and positive word-of-mouth (Cronin et al., 2000; Kuo et al., 2009).

There is a significant body of service literature that provides evidence of a positive relationship between perceived service quality and perceived service value. For example, Bolton and Drew

(1991) develop a conceptual model for assessing service performance, quality, and value. The authors apply the model to residential telephone services and find a significant, positive association between service quality and service value. Gooding (1995) finds that perceived quality is a key antecedent to perceived service value, where the latter largely determines the choice of a healthcare service provider. Andreassen and Lindestad (1998) report the same effect in the context of package tour services.

Research also documents the negative effect of perceived transaction costs on perceived service value. In fact, following the standard definition of service value as the difference between what is received and what is given, perceived transaction costs stand for the sacrifice involved in a service exchange. Most of the works on this theme use the constructs of perceived sacrifice, perceived risk, or perceived cost to represent perceived transaction costs in our model. For example, Spreng et al. (1993) report that a consumer's anticipation of future sacrifice, including purchase, psychological, and time costs, has significant effect on the ex-ante perceived service value. Gooding (1995) operationalizes perceived sacrifice through distance, transportation time to a hospital and out-of-pocket costs, providing empirical evidence that perceived sacrifice (i.e., transaction costs) decreases perceived service value. By defining perceived transaction costs as a sum of price, time, and effort, Brady and Robertson (1999) show a negative association between perceived transaction costs and perceived service value. Finally, the marketing literature provides significant evidence from broad service contexts that perceived service quality and perceived transaction costs are key antecedents of perceived value (Dodds et al., 1991; Teas & Agarwal, 2000; Cronin et al., 2000). Combining the mental accounting perspective and perceived value framework, we have the following proposition:

Proposition 3. *Perceived service quality is positively, and perceived transaction costs are negatively, associated with the perceived value of a returns service.*

A return service is a post-sales service included in a retailer's overall service bundle and consists of policy and process aspects (Mollenkopf et al., 2007). Policy aspects include the terms and conditions of the returns service offering and are communicated through formalized return policy statements. Therefore, a return policy is a fundamental element of the returns service design and is subject to consumer evaluation (Davis et al., 1998). In addition, a return policy is a means of marketing communication that informs consumers what to expect in case they need to return a purchase. As discussed earlier, we can characterize any return policy by the degree of leniency offered across five levers. We contend that all five leniency levers are important design levers for a retailer's return service and can be used to influence the perceived value of the returns service. These arguments together with a logical combination of Propositions 1–3 lead us to the following proposition:

Proposition 4 (a–e)**.** *Higher a) monetary, b) time, c) effort, d) scope, e) exchange leniency leads to higher perceived value of a returns service.*

Perceived value is found to be the most important predictor of purchase intentions in a variety of service settings (Parasuraman, 1997; Parasuraman & Grewal, 2000; Baker et al., 2002). With respect to our context, research on return policies lends support that the overall leniency of a return policy is positively associated with purchase intention (Bonifield et al., 2010; Oghazi et al., 2018). Jeng (2017) finds that the perceived value of a return policy mediates the relationship between overall return policy leniency and purchase intention. Therefore, we state the final proposition:

Proposition 5 (a–e)**.** *Higher a) monetary, b) time, c) effort, d) scope, e) exchange leniency leads to a higher intention to purchase from a retailer, by increasing the perceived value of the returns service that the retailer offers.*

Based on the theoretical grounding above, we now present the formal hypotheses regarding the proposed cognitive process for a consumer's returns service valuation and the resultant purchase intention by anchoring on the degree of leniency across five levers. In particular, we posit that given two return policies, *ceteris paribus*, the policy with a higher leniency in one of the levers results in higher perceived service quality and lower perceived transaction costs, which then leads to a higher perceived service value, and ultimately results in a higher purchase intention. More formally, we hypothesize:

Hypotheses (1-5). *Perceived service quality and perceived transaction costs in parallel and perceived service value in series mediate the relationship between 1) monetary, 2) time, 3) effort, 4) scope, 5) exchange leniency and purchase intention, such that higher leniency leads to a higher purchase intention through increased perceived service quality, decreased perceived transaction costs, and increased perceived service value.*

We hypothesize on the positive effects of leniency across all five return policy levers on purchase intentions through the cognitive process model described above. However, the overarching theoretical model also leads us to expect significant heterogeneity among the levers in terms of their relative effectiveness. As signals of service quality, leniency levers would have different signal strengths or salience (Connelly et al., 2011). In particular, unlike high time, effort, and scope leniency, high monetary and exchange leniency involves direct financial risks and constitute *cost-risking* signals for the retailers (Kirmani & Rao, 2000). The implied costs to low-quality retailers, therefore, are higher when monetary and exchange leniency is high relative to high leniency across the other three levers. Consequently, high monetary and exchange leniency would become a stronger, more salient quality signal and increase perceived service quality of the retailer. Further, consumers would perceive greater transaction costs when monetary and exchange leniency is low due to direct financial losses, in addition to associated psychological transaction costs (Chircu & Mahajan, 2006). Due to their potentially stronger influences on perceived service quality and transaction costs, we expect a stronger effect on perceived service value and subsequently on purchase intention from monetary and exchange leniency levers relative to time, effort, and scope levers. Hence, we hypothesize:

Monetary and exchange leniency levers have stronger effects on a consumer's purchase intention than time, effort, and scope levers.

3.3 Empirical Method

To test the research hypotheses, we use primary data collected from six experimental studies involving U.S.-based consumers. The first five of these studies are pre-studies that are conducted to 1) develop robust measurement scales, 2) empirically validate the constructs in our model, 3) pre-test the experimental vignettes and manipulations, and 4) inform various design tradeoffs and choices for the main study (Eckerd et al., 2021). In sections §3.3.1 and §3.3.2 we provide background regarding the first two purposes. Detailed discussions that are related to the purposes 3 and 4 are provided in Appendix B.

All pre-studies and the main study are designed using Qualtrics and are conducted via Amazon Mechanical Turk (MTurk) online crowdsourcing platform. MTurk enables data collection from a diverse and representative consumer population, which is important in terms of the external validity of the empirical findings (Berinsky et al., 2012; Paolacci & Chandler, 2014; Goodman & Paolacci, 2017). The platform also allows participants to complete studies in their natural living or working environments—which mitigates concerns regarding observer effects—and provides a lab-in-the-field setting. Each participant received a compensation of $1.00 for completing a study, which took, on average, 7 minutes. On an hourly basis, the payment is significantly higher than average incentives paid to MTurk workers ($1.66, Paolacci & Chandler, 2014) as well as above the federal minimum wage ($7.25, U.S. Department of Labor, 2019).

All experimental studies use vignette-based methods. Experimental vignette methods, particularly when combined with diverse samples such us ours, are viewed as an effective, balancing solution to the methodological dilemma in choosing between conventional lab experiments with questionable external, but high internal validity and non-experimental or quasi-experimental methods that provide greater external validity, but engender many threats to internal validity (Aguinis & Bradley, 2014). Experimental vignettes are particularly relevant when the goal is to study cognitive-affective perceptions (Eckerd et al., 2021). Service contexts provide an ideal setting for vignette-based experiments, because participants tend to be familiar with the contexts and can easily engage with the described situations or, more applicable to our context, innately comprehend the information provided in the vignettes (Rungtusanatham et al., 2011; Eckerd et al., 2021).

3.3.1 Measurement Scale Development and Construct Validation

We go through a rigorous scale development and construct validation process through a multi-sample, experimental approach with a replication logic (Pagell, 2020) for two key reasons. First, there are not well-established measurement scales readily available in the extant return policy literature to measure the constructs of interest in our empirical model. Second, our research objective is to provide a comparative assessment of different leniency levers and robust measurement scales are crucial to conduct such an analysis with reliable results. To this end, the five pre-studies involve joint experimental manipulations of different leniency levers, enabling a cross-validation and stress testing of measurement scales.

In developing the measurement scales and validating the constructs, we follow the methodological practices recommended by O'Leary-Kelly and Vokurka (1998) and MacKenzie et al. (2011). We start by generating measurement items to capture the domain of the constructs of interest. To do so, we analyze the existing return policy literature and theoretical research in marketing to generate an initial set of measurement items. These measurement items are either adapted from the scales in the existing literature to the context of our research or developed, as needed, based on theoretical and conceptual foundations of the constructs. All measurement items use a 7-point Likert scale. We empirically validate our constructs by assessing four key components: unidimensionality, reliability, convergent validity, and discriminant validity. The full set of the measurement items and details of our construct validation process are provided in Appendix B.

3.3.2 Main Study

Following the pre-studies that establish empirically validated constructs and measurement scales, this section presents the main study in which we formally test our hypotheses. The next section discusses the details of the experimental procedure and sample characteristics, while §3.3.2.2 presents the data analysis and results.

3.3.2.1 Experimental Procedure and Sample

The main study has a completely randomized, full-factorial between-subject design (i.e., 5 levers manipulated at low vs high leniency levels constituting 32 cells). In designing the experimental manipulations of each leniency lever, careful consideration was given to achieve a balance among the following: 1) the operationalization of low and high levels of leniency should be realistic and actionable from a managerial perspective (Bachrach & Bendoly, 2011), 2) the low and high levels of leniency need to constitute significant contrasts to be salient (Rungtusanatham et al., 2011), and 3) the nature of manipulations are aligned with the existing literature (Abdulla et al., 2019). To this end, the pre-study phase provided significant conceptual knowledge and preliminary empirical insights. In addition, we made a number of observations from practice to inform the design of the main study. We provide details for the design process in Appendix B.

A total of 840 participants (45.1% female) were randomly assigned to one of the 32 treatment conditions. Participants were asked to view a vignette designed as a website of retailer "ABC," which was presented as one of the large retailers in the U.S. who sell products in multiple categories through online and offline channels. Our choice of a multiple-category retailer based in the U.S. as the vignette context for this study is predicated on the following. First, the choice allows estimating the average treatment effects of different return policy leniency levers in a broad context (i.e., without priming participants on a narrow set of product types or price levels). Indeed, understanding the effect of different leniency levers averaged across all potential product categories and/or price levels would be of practical significance to the large retailers (i.e., big-box stores, department stores) who may have negative connotations about offering complex return policies with many category-based exclusions. Second, we surveyed the top 20 U.S-based retailers by sales revenue. We found that 16 of them were general merchandise stores (i.e., department stores, big-box stores) and e-tailers that carried multiple-category assortments sold through online and offline channels (the remainder was supermarkets specialized in grocery). Considering that the sales volume of these retailers constitute a significant percentage of all U.S. retail sales (i.e., $1.4 of $5.5 trillion as of 2019), at an average return rate of 10%, these retailers alone would account for approximately $140B out of $369B of annual returns in the U.S. (National Retail Federation, 2019). Therefore, among a myriad of other options, choosing a multiple-category retailer selling through both online and offline channels for the vignette served the purpose for broader managerial relevance.

The vignette is designed to include common features in a typical retailer website (e.g., search box, product categories, store finder, member login, etc.) in addition to the return policy statement. This is to increase the realism and ecological validity of the experimental environment (Aguinis & Bradley, 2014). To check the perceived realism of the experimental vignettes by participants, we asked two realism check questions. In particular, on a 7-point Likert scale with 1 (Strongly Disagree) to 7 (Strongly Agree), the participants indicated a score with mean = 5.110 (s.d. = 1.350) to the statement “The ABC website presented in the study carried most of the elements that I find in other retailers’ websites.” Second, the participants indicated a score with mean = 5.255 (s.d = 1.100) to the statement “ABC return policy was realistic in its wording and format considering other return policies I have seen.” The vignettes for the highest and the lowest leniency (across all levers) treatment conditions are provided in Appendix B.

Demographic characteristics of the sample were comparable with those from pre-studies (see Appendix B). At the end of the study, participants were required to answer two attention check questions, asking the restocking fee amount and return time window indicated in the return policy (Abbey & Meloy, 2017). These attention check questions qualify as factual manipulation checks (Kane & Barabas, 2019), meaning 1) they have objective, correct answers and 2) they are directly related to the experimental manipulations used. Factual manipulation checks are considered more effective compared to other common types of manipulation checks, such as instructional or subjective manipulation checks (Oppenheimer et al., 2009). Of the 840 participants who completed the study, 650 participants (77.4%) answered the two attention checks (factual manipulation checks) correctly and whose locations were verified to be in the U.S. are included in the final analysis reported in the next section. In the final sample, we did not observe significant imbalances in terms of the number of observations per treatment condition, which ranged between 19–22.

3.3.2.2 Analysis and Results

To test the hypotheses, we perform a regression-based mediation analysis with the non-parametric bootstrapping approach using the PROCESS macro for SPSS (Hayes, 2018). This is a modern method to reliably estimate mediation effects with adequate statistical power (Rungtusanatham et al., 2014; Hayes, 2018).

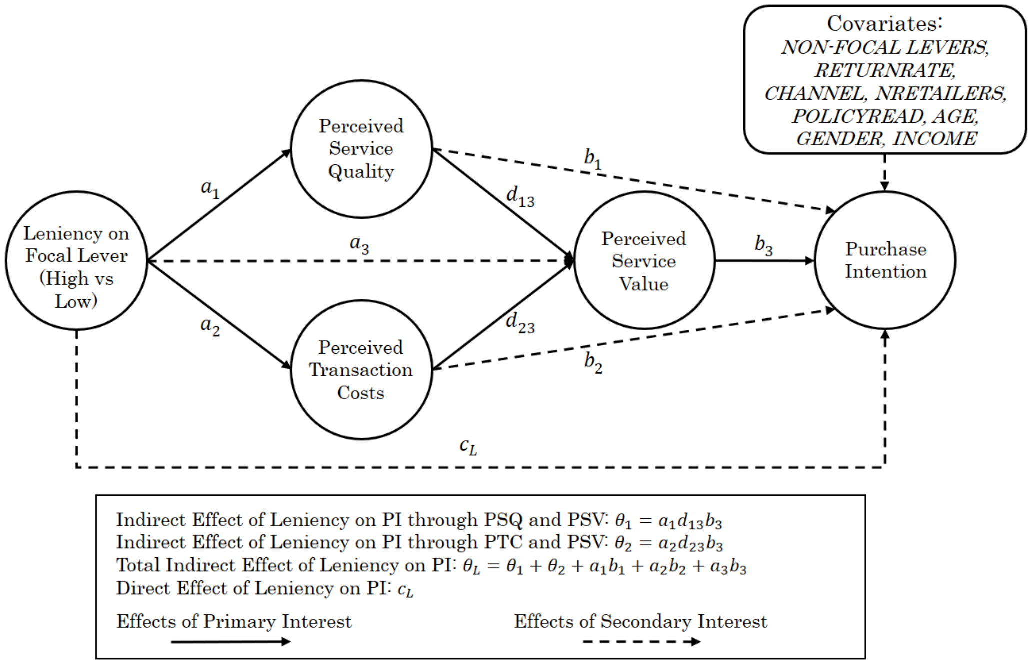

Figure 3.2: Statistical Diagram and Notations for Mediation Analysis

We use 10,000 bootstrap re-samples for the analysis. In all regression equations estimating the path coefficients, we also include several covariates to further improve the statistical power and precision of the estimates (i.e., to reduce the standard errors). These covariates and definitions are reported in Table 3.2. Latent variable scores are calculated as averages of the respective measurement items. Again, the measurement scales and constructs are successfully validated via an EFA, using the same two factor extraction-rotation methods as in the pre-studies, and via CFA (see Appendix B for details). Descriptive statistics for the latent variables are provided in Table 3.3 and the inter-variable correlations are reported in Appendix B. Figure 3.2 shows the statistical diagram for the mediation analysis and notations to facilitate the discussion.

Table 3.4 reports the results of the analyses and includes the estimated individual path coefficients, direct effects, total indirect effects, and all path-specific indirect effects with 95% percentile bootstrap confidence intervals (henceforth CI for brevity). A significant mediation effect exists when the CI for the estimate of an indirect effect does not contain zero. The analyses provide sup-

Covariate	Type	Definition
$RETURNRATE$	Continuous	Return to purchase ratio on an annual basis, based on the self-reported average number of non-grocery purchases made in a month and the number of returns made in the last 12 months.
$POLICYREAD$	Ordinal (7-point Likert)	To what degree a consumer reviews return policies before purchasing from a retailer.
$NRETAILERS$	Count	Number of distinct retailers that a consumer shops during a typical year.
$CHANNEL$	Categorical	Consumer's channel preference (brick-and-mortar vs. online vs. no strict preference (both)).
AGE	Ordinal	Participant's age.
$GENDER$	Categorical	Participant's gender.
$INCOME$	Ordinal	Participant's household income level.

Table 3.2: Covariates Included in the Mediation Analysis

Variable	PSQ	PTC	PSV	PI
Mean [s.d.]	5.031 [1.437]	4.300 [1.750]	4.113 [1.804]	4.823 [1.570]

Table 3.3: Descriptive Statistics for Focal Variables

port for Hypotheses 1–5. In particular, we find statistically significant between-subject mediation effect for all five levers. Further, we find that PSQ, PTC, and PSV *fully* mediate the relationship between return policy leniency across a given lever and PI, evidenced by the fact that the direct effects of leniency on PI are not statistically significant when PSQ, PTC, and PSV are accounted for. By estimating the contrast between the two indirect effects (through PSQ vs PTC), we find that in transmitting the effect of monetary leniency onto PI, PTC is a significantly stronger mediator ($\theta_1 - \theta_2 = -0.263, \text{CI} = [-0.441, -0.100]$). Meanwhile, for the remaining leniency levers, PTC and PSQ demonstrate statistically equivalent strengths in transmitting the effect of leniency onto PI. Among the covariates included in the mediation analysis, the demographic variables (i.e., age, gender, income) did not explain significant variability in the purchase intentions. Individuals with higher self-reported return rates and who had higher tendency to review return policies before purchasing expressed lower purchase intentions, on average, controlling for the treatment effects of return policy leniency levers.

We next compare the total indirect (mediation) effects of the five levers on PI and find support for Hypothesis 6. In particular, we find that monetary leniency—operationalized through practically common cases of a full-refund vs. 15% restocking fee—proves to be the most effective lever in influencing purchase intentions through the mediators. The second most effective lever is the exchange lever, which is manipulated through offering a cash refund vs. store credit only. The effort lever comes third, with a slightly higher total indirect effect compared to the time lever. Scope leniency, operationalized through whether sales items are allowed to be returned or not, demonstrates the smallest effect. Figure 3.3 provides a visual representation of the total indirect effects and CIs for each lever. In order to test whether the differences are statistically significant, we estimate the differences between total indirect effects of adjacent levers in the effect size rank ordering (i.e., adjacent contrast effects), using non-parametric bootstrapping. The point estimates for these differences and the 95% bootstrap CIs are provided in Table 3.5. Overall, while we find that there are statistically significant differences among the mediation effects observed, the practical significance of the differences among effort, time, and scope leniency levers in particular is not very high.

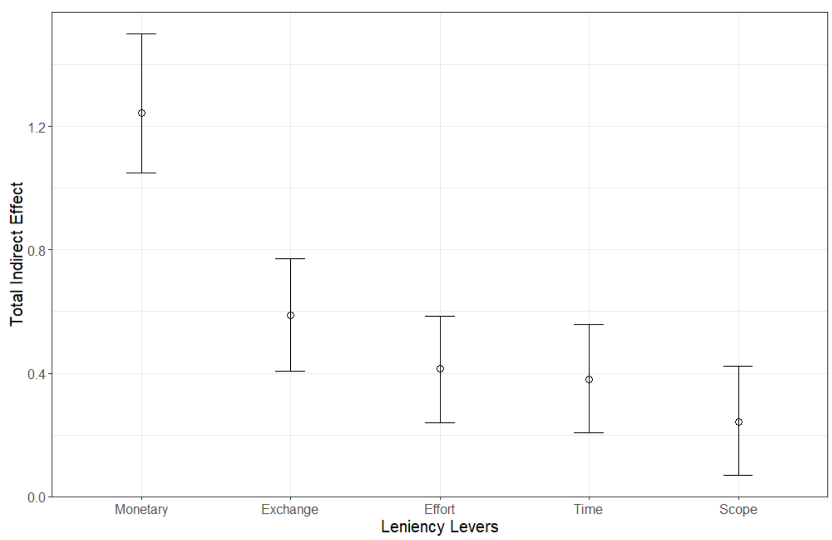

Figure 3.3: Total Indirect Effects of the Leniency Levers

Finally, we also analyze the interactions between different leniency levers in their impact on

Focal Lever	Direct Effect [CI]	Total Indirect Effect [CI]	Mediation Paths/Sub-Paths with Effect Sizes	Indirect Effects	95% CIs
Monetary	0.084 $[-0.075, 0.242]$	1.244 $[1.048, 1.449]$	M $\xrightarrow{1.014}$ PSQ $\xrightarrow{0.491}$ PSV $\xrightarrow{0.605}$ PI	0.301	$[0.217, 0.396]$
			M $\xrightarrow{-1.791}$ PTC $\xrightarrow{-0.521}$ PSV $\xrightarrow{0.605}$ PI	0.564	$[0.426, 0.721]$
			M $\xrightarrow{0.276}$ PSV $\xrightarrow{0.605}$ PI	0.167	$[0.051, 0.289]$
			M $\xrightarrow{1.014}$ PSQ $\xrightarrow{0.213}$ PI	0.216	[0.113, 0.338]
			M $\xrightarrow{-1.791}$ PTC $\xrightarrow{0.003}$ PI	-0.005	$[-0.118, 0.114]$
Time	-0.089 $[-0.225, 0.047]$	0.381 $[0.206, 0.558]$	T $\xrightarrow{0.444}$ PSQ $\xrightarrow{0.491}$ PSV $\xrightarrow{0.605}$ PI	0.132	$[0.068, 0.202]$
			T $\xrightarrow{-0.600}$ PTC $\xrightarrow{-0.521}$ PSV $\xrightarrow{0.605}$ PI	0.189	$[0.116, 0.272]$
			T $\xrightarrow{-0.055}$ PSV $\xrightarrow{0.605}$ PI	-0.033	$[-0.126, 0.055]$
			T $\xrightarrow{0.444}$ PSQ $\xrightarrow{0.213}$ PI	0.095	$[0.041, 0.161]$
			T $\xrightarrow{-0.600}$ PTC $\xrightarrow{0.003}$ PI	0.002	$[-0.041, 0.039]$
Effort	-0.252 $[-0.387, 0.116]$	0.414 $[0.240, 0.585]$	F $\xrightarrow{0.512}$ PSQ $\xrightarrow{0.491}$ PSV $\xrightarrow{0.605}$ PI	0.152	$[0.087, 0.223]$
			F $\xrightarrow{-0.683}$ PTC $\xrightarrow{-0.521}$ PSV $\xrightarrow{0.605}$ PI	0.215	$[0.135, 0.304]$
			F $\xrightarrow{-0.100}$ PSV $\xrightarrow{0.605}$ PI	-0.060	$[-0.155, 0.030]$
			F $\xrightarrow{0.512}$ PSQ $\xrightarrow{0.213}$ PI	0.109	$[0.050, 0.181]$
			F $\xrightarrow{-0.683}$ PTC $\xrightarrow{0.003}$ PI	-0.002	$[-0.046, 0.042]$
Scope	0.021 $[-0.117, 0.160]$	0.243 $[0.071, 0.424]$	S $\xrightarrow{0.361}$ PSQ $\xrightarrow{0.491}$ PSV $\xrightarrow{0.605}$ PI	0.107	[0.044, 0.174]
			S $\xrightarrow{-0.373}$ PTC $\xrightarrow{-0.521}$ PSV $\xrightarrow{0.605}$ PI	0.118	[0.047, 0.195]
			S $\xrightarrow{-0.095}$ PSV $\xrightarrow{0.605}$ PI	-0.058	$[-0.149, 0.034]$
			S $\xrightarrow{0.361}$ PSQ $\xrightarrow{0.213}$ PI	0.077	[0.026, 0.143]
			S $\xrightarrow{-0.373}$ PTC $\xrightarrow{0.003}$ PI	-0.001	$[-0.027, 0.024]$
Exchange	-0.024 $[-0.162, 0.115]$	0.588 $[0.407, 0.772]$	X $\xrightarrow{0.441}$ PSQ $\xrightarrow{0.491}$ PSV $\xrightarrow{0.605}$ PI	0.131	$[0.067, 0.203]$
			X $\xrightarrow{-0.620}$ PTC $\xrightarrow{-0.521}$ PSV $\xrightarrow{0.605}$ PI	0.195	$[0.119, 0.282]$
			X $\xrightarrow{0.280}$ PSV $\xrightarrow{0.605}$ PI	0.169	$[0.075, 0.267]$
			X $\xrightarrow{0.441}$ PSQ $\xrightarrow{0.213}$ PI	0.094	[0.041, 0.162]
			X $\xrightarrow{-0.620}$ PTC $\xrightarrow{0.003}$ PI	-0.002	$[-0.041, 0.040]$

M: Monetary, T: Time, F: Effort, S: Scope, X: Exchange leniency.
Covariates: Non-focal Levers, $RETURNRATE$, $POLICYREAD$, $NRETAILERS$, $CHANNEL$, AGE, $GENDER$, $INCOME$.

Table 3.4: Mediation Results

$\Delta\theta_L$	Point Estimate	95% CI
$\theta_L^M - \theta_L^X$	0.660	[0.658, 0.663]
$\theta_L^X - \theta_L^F$	0.175	[0.172, 0.177]
$\theta_L^F - \theta_L^T$	0.033	[0.030, 0.035]
$\theta_L^T - \theta_L^S$	0.135	[0.133, 0.137]

Table 3.5: Statistical Tests for the Significance of the Differences between Total Indirect Effects

PI through the mediation mechanism that we study. To do so, we estimate an ANCOVA model that includes all leniency levers as fixed factors, PI as the outcome variable, and the remaining variables included in the mediation analysis above as covariates. We find no statistically significant (at $p < 0.05$ level) two-way or higher-order interaction effects among the five leniency levers.

3.3.3 Robustness and Generalizability of Findings

Results of the main study suggest statistically significant and heterogeneous causal effects of the five leniency levers on purchase intentions through their influences on perceived service quality, perceived transaction costs, and perceived service value. One may wonder about the role of contextual dependencies (i.e., moderators) in the relative effects of the five levers in order to gauge the generalizability of the findings. Relevant contextual dependencies would be the assortment focus of the retailer (e.g., general merchandise versus specialized stores), sales channel (i.e., online versus offline), product category (e.g., electronics versus apparel), and price (i.e., high versus low) of products being considered for purchase. We know from the meta-analysis conducted by Janakiraman et al. (2016) that the assortment focus of a retailer and the sales channel do not moderate the effect of return policy leniency on purchase intentions. Further, our empirical results from pre-study 1, where we manipulated the product category as electronics versus apparel (two product categories with the highest return rates and volumes), suggest that product category may not be a significant moderator. However, this is by no means conclusive and warrants further examination.

We experimentally examined the moderating role of price in determining the effect of monetary and exchange leniency (two levers with a significant financial risk) on consumers' purchase intentions. The experiment had a six-cell (partial factorial) design, price-level (low = \$50, high = \$500), monetary leniency (low = 15% restocking fee, high = no restocking fee), and exchange leniency (low = no cash refund/store credit only, high = both cash refund and store credit as options) as the manipulated factors. We excluded low monetary and exchange leniency conditions from this study because they (i.e., charging a restocking fee on a store credit) are unrealistic from a practical standpoint. We chose desk chairs for the purchase scenario, since they have considerably high price variability in the market and significant product fit/quality uncertainty. We used experimental vignettes that are consistent with our earlier studies. As a result, we did not find a significant moderating effect of the price level of the product being considered for purchase, for either leniency levers that we tested (see Appendix B for detailed results). As such, we cannot reject the null hypothesis that price too, is not a significant moderating factor.

The pre-study phase provides supporting empirical evidence that consumers do have a *general* perception of what leniency levers matter most and perceive a consistent hierarchy of importance. Specifically, we asked participants to indicate on a Likert scale of (1) Most Important to (7) Least Important, how important they find leniency across different return policy levers in choosing which retailers to shop (i.e., no restocking fees (monetary), a long return time window (time), easy and hassle-free return process (effort), not only store-credit/exchange, but also cash refund option (exchange), sale and clearance items allowed to be returned as regularly-priced items (scope)). The statements associated with the five leniency levers were provided in a *randomized order* to each participant. Here, we found consistent evidence across all five pre-studies (which manipulated different subset of these five levers) that on average, consumers have an order of importance of leniency across different levers (see Appendix B). In particular, all five pre-studies revealed that on average, monetary leniency was by far the most important, followed by exchange leniency, followed by effort, which was followed by scope and time leniency. Multiple t-tests showed that monetary, exchange, and effort leniency were considered to be significantly more important relative to time and scope leniency. The difference in the importance of time versus scope leniency was not significant. Moreover, monetary leniency was considered as significantly more important than exchange leniency, which in turn was considered as more important than effort leniency. This empirical evidence suggests that the relative effects that we identify in the main study are fairly generalizable to different retail contexts, contextual dependencies, and to the presentation ordering of the return policy levers.

3.4 Conclusion

In light of the empirical findings from the main study, this section discusses managerial and theoretical implications of our research and highlights a number of limitations and future research opportunities.

Overall, our findings support the contention that the value of a retailer's returns service is an important determinant of a consumer's purchase intention. We also demonstrate that each of the five levers that constitute return policy leniency is a significant antecedent of the perceived value of a returns service, through each lever's ability to ex-ante signal perceived service quality and perceived transaction costs. However, our investigation further reveals heterogeneous effects of different leniency levers. This implies that retailers need to be careful in choosing the right leniency levers to tighten the return policies in order to reduce the operational cost burden, and our findings provide new insights regarding this decision as we discuss below.

We find that monetary leniency is the most effective lever in influencing the cognitive perceptions that ultimately impact purchase intentions. The positive effect of monetary leniency on purchase intentions has been documented in the prior literature. Extending this, we show that the effect of monetary leniency dominates the effects of the other four levers. Indeed, monetary leniency has the greatest salience among all levers in terms of financial risk. Consequently, anticipated regret can also be expected to be the greatest when leniency along this particular lever is low, increasing perceived transaction costs (Inman & Zeelenberg, 2002). From the retailer's perspective, offering a full refund as a part of the returns service offering would send strong signals regarding the retailer's understanding toward consumer needs and willingness to absorb the consumer's risk regarding product fit and valuation uncertainty (Abdulla et al., 2019) and thereby significantly stimulate returns service quality. The financial costs that a retailer bears from offering a full refund for all returns is also more salient to the consumer compared to cost of offering leniency across other levers. This may also explain the dominance of the monetary leniency lever over the other levers.

A direct managerial implication of this finding is that retailers that impose or consider imposing restocking fees should be aware of strong negative perceptions and subsequent consumer behaviors (e.g., decreased purchase intention, negative word-of-mouth, switching). For example, many retail stores, such as Best Buy, Macy's, and Sears, currently charge restocking fees on returns made in select product categories (ConsumerWorld.org, 2018). Our results suggest that in these product categories, such retailers may be losing customers to their competitors who offer a full refund. Anecdotal evidence supports the view that customers typically avoid purchasing from retailers that charge restocking fees, finding such fees unfair.

A more interesting finding is the second strongest impact of exchange leniency, previously unexplored in a causal framework, on purchase intentions through its impact on perceived service quality, perceived transaction costs, and perceived service value. Research on mental accounting has established that restricted-use funds, such as store credits, are evaluated and spent differently than equivalent cash, even in the context of the same retailer (Reinholtz et al., 2015). Our research suggests that offering a return policy that allows customers only to make an exchange or receive store credit in case of a return, instead of getting a cash refund, can be detrimental to the perceived value of a returns service and reduce purchase intentions ex-ante. The strong impact of the exchange lever would be due to financial risks associated with customer lock-in with the retailer (E. J. Johnson et al., 2003; Zauberman, 2003). In fact, this explanation is supported by the anecdotal accounts from several participants at the end of the main study.

The overall conclusion is that retailers should avoid restricting their return policies through monetary and exchange levers in order to reduce the cost burden of returns and make the returns service more sustainable from an operational standpoint. Instead, retailers should consider opportunities across the remaining three levers. For example, our results imply that a long return time window may not have a strong, positive effect on a consumer's ex-ante returns service value perception and purchase intention. In fact, almost 60% of the participants in the aggregated sample of the pre-studies indicated in a survey question that they would typically need less than a week to make a keep/return decision, with only 10% of the participants indicating that they would need more than two weeks. Therefore, providing excessively long return time windows, such as several months, may not provide any significant advantage to retailers. Rather, driven by customer inertia and procrastination, a time-to-return distribution with a long tail may cause unsustainable losses in recoverable value (Ferguson et al., 2006; Su, 2009b). This is due to the fact that the longer the time it takes for a product to be returned after purchase, the larger the deterioration in salvage value of the product (Blackburn et al., 2004). Moreover, products that are returned late tend to have fewer disposition options that can generate value to retailers (Shang et al., 2019). Thus, by reducing the time window, retailers can better reduce the operational cost of returns. From this perspective, it is noteworthy that many of the recently restricted return policies involve tightening return time windows, such as the policies of Macy's, L.L.Bean, and Bed Bath & Beyond (ConsumerWorld.org, 2018).

Decreasing effort leniency by imposing additional hassles, such as tag and original packaging requirements or asking customers to fill a return authorization form, may also help prevent some of the returns ex-post, with a relatively smaller negative impact on purchase intentions. As a case in point, Nordstorm, well-known for its generous return policy, has recently imposed tag requirements for special-occasion dresses and designer items. The smaller effect of effort leniency on ex-ante purchase intentions relative to monetary and exchange leniency is likely to be due to expected transaction costs of non-monetary nature (i.e., psychological, time, physical) that loom less than financial costs implied by monetary and exchange levers.

Further, we show that consumers may not dramatically decrease their value perceptions and purchase intentions if a retailer disallows discounted (i.e., sales or clearance) products being returned relative to when it applies a standard return policy for discounted products. The practically small effect of the particular operationalization of scope leniency in our context mirrors the dual entitlement principle (Kahneman et al., 1986). The dual entitlement principle posits that most consumers believe that they are entitled to a reasonable price and that firms are entitled to a reasonable profit. Knowing that offering products at discounted prices implies giving up on the usual sales revenue the retailer would be entitled to, consumer may feel that the retailer can also fairly disallow returns on discounted products, in order to attain a reasonable profit margin. Thus, decreasing scope leniency, particularly by disallowing returns for discounted products, provides another viable opportunity for retailers to decrease the burden of returns while keeping negative reactions at a minimum. Such a strategy, particularly when combined with effective pricing tactics (i.e., seasonal discounts, individualized pricing) can prove a better alternative to imposing restocking fees and other non-refundable charges with similar or a greater positive financial impact. Many apparel retailers, such as Gap, Tommy Hilfiger, and Michael Kors, do not allow returns or exchanges for final sale items. Dillard's does not allow returns for clearance sales and also the products sold under stacked discounts (e.g., items marked down 20% plus an additional 30% discount for a limited time). Overall, our research provides actionable guidelines to retail managers regarding how to make return policies more sustainable from an operational perspective, without significantly deteriorating consumer value perceptions and patronage intentions.

Our research has multiple theoretical contributions to research on consumer return policy design and, more generally, on service design. First, while the previous literature on consumer return policies has predominantly focused on the quality signaling aspect of return policy leniency, our research postulates and demonstrates that return policy leniency can also influence transaction cost perceptions. In fact, we demonstrate that the strength of influence on purchase intentions through transaction cost perceptions is statistically no different than through service quality perceptions for all leniency levers, except the monetary lever. Thus, accounting for both service quality and transaction cost perceptions provides a more complete cognitive process model predicated on a perceived service value framework. In turn, this provides greater explanatory power to understand the relationship between return policy leniency and consumer perceptions and purchase intentions. A broader theoretical implication is the importance of recognizing that service design decisions of firms may not only influence consumer behavioral outcomes through the service quality mechanism, but also through the transaction cost mechanism. As a result, different service policy levers may influence both service quality and transaction costs perceptions, resulting in heterogeneous effects on behavioral outcomes.

Second, we test value perceptions regarding a service and the resultant purchase intentions solely based on the service policy itself without any exposure to the process aspects of the service. An important theoretical implication is that certain attributes of a service—as indicated in service terms and conditions—can ex-ante stimulate quality, transaction costs, and value perceptions that influence a consumer's patronage decisions. This highlights the importance of empirical research with respect to policy aspects of service design, in addition to the process aspects that are more commonly studied in the OM domain. Continued research can extend the boundaries of our cognitive model to other service contexts commonly studied in the OM literature, such as credit card, insurance, travel, and ticketing services.

Third, our cognitive process model is predicated on multiple theoretical perspectives and provides a generalizable framework to examine, in a comparative sense, the influence of different service design levers on the consumer valuation process and subsequent behavioral outcomes. Our model, combined with the study design, is a particularly good fit to study service contexts where levers, attributes, or strategies of interest involve trade-offs from the firm's perspective (e.g., retailers typically do not simultaneously manipulate multiple levers while changing a return policy but rather try to choose which lever to optimize) and not necessarily from the consumer's perspective (e.g., participants did not have to make a trade-off between leniency across different levers and could value leniency across all levers). To illustrate a future application of our approach in a different service operations context, researchers can study how different omni-channel capabilities available to today's retailers would compare in their effects on service quality, transaction cost perceptions, and resultant behavioral intentions and outcomes.

Abdulla et al. (2019) call for continuing analytical research that studies non-monetary leniency levers, decisions across multiple levers, as well as design of return policies that involve the interaction amongst levers (e.g., time-based restocking fees). Our empirical findings have implications for continuing and growing analytical OM research involving consumer return policies. First, as we find no significant interaction effects across multiple pairs of leniency levers, the utility gain due to return policy leniency can be reasonably modeled through an additive, rather than a multiplicative, functional form. Second, it is important to recognize the differences across levers in terms of their effect on purchase intentions while modeling aggregate demand or individual consumer purchase decisions. For instance, the marginal impact of monetary and exchange leniency on aggregate market demand should be modeled to be greater than the marginal impact of scope and effort leniency.

Similarly, in terms of individual consumer utility, the ex-ante utility gain from purchasing due to monetary or exchange leniency of return policies needs to be modeled as greater, by a factor, than that due to time and effort leniency.

There are a number of limitations to our work that we believe can motivate future research. First, we examine cognitive perceptions and purchase intentions but not manifest behavioral outcomes such as actual purchases. This allowed us to study consumer attitudes toward a retailer based on the return policy leniency operationalized through five levers available to the retailer. The theory of planned behavior and empirical evidence from marketing literature suggest that purchase intentions can reasonably predict actual purchase decisions (Ajzen, 1985; Chandon et al., 2005). Still, testing how return policy leniency across the five levers influences different behavioral outcomes when real monetary stakes are involved (i.e., purchase and return transactions) can be an interesting future research. This would strengthen the findings and implications by establishing predictive validity.

As a matter of scope, we propose a high-level, parsimonious theory and document consistent empirical evidence for heterogeneous average treatment effects of leniency across different levers on purchase intentions. We do so by focusing on a broadly applicable context motivated from practice. Though we empirically examine product category in a pre-study and price level in a post-study and find no statistically significant effect of these two contextual factors on the average treatment effects, our inquiries along these lines are not exhaustive. Future research can examine more systematically and exhaustively the role of different practically relevant contingencies on the overarching treatment effects of each leniency lever on purchase intentions.

Another line of fruitful research would be to investigate consumer perceptions of the complexity of return policies (e.g., the number of category-specific policies in the overall return policy offering of a retailer) could also generate interesting insights. Here, researchers can note that imposing category-based exclusions to a standard return policy would be another form of low scope leniency. Given that many retailers have tremendous amount of transactional data on consumers' purchase and return behavior, an examination of the feasibility of and consumer reactions to personalized return policies (in a similar vein to personalized pricing) would also be a fruitful direction for future research.

In this paper, we provide insights for retail practitioners on how to make return policies more sustainable from an operational perspective, while minimizing damage to the value proposition offered to customers. Future research is needed to understand if there might be negative implications of restricting a long-established lenient return policy due to negative signaling (Connelly et al., 2011). Another interesting future research avenue would be to investigate how retail managers evaluate potential costs and benefits of leniency across each levers, to understand whether the managerial perceptions are aligned with those of consumers.

4. RESTRICTIVE CHANGES TO LONG-ESTABLISHED LENIENT RETURN POLICIES AND CONSUMER REACTIONS TO THEM

4.1 Introduction

The increased cost burden of managing consumer returns in the retail industry pushes retailers to rethink the generous return policies that have been pervasively offered in the U.S. over the last few decades (Abdulla et al., 2019, 2021). Consequently, many retailers have restricted their long-established return policies through the use of different levers (ConsumerWorld.org, 2018). Some retailers shortened return time windows that have long been open-ended. Others decided to charge restocking fees for returns in various product categories or imposed hassles in the forms of item tag, customer ID, or original receipt requirements. Table 4.1 presents notable cases of restrictive return policy changes from the U.S. retail industry.

Retailer	Restrictive Policy Change
REI	Restricted the open-ended return window to 365 days in 2013.
Macy's	Restricted the open-ended return window to 365 days in 2016, to 180 days in 2017, to 90 days in 2019.
L.L. Bean	Restricted the open-ended return window to 365 days in 2018, also imposed a receipt requirement.
Kohl's	Restricted the open-ended return window to 180 days in 2019.
Bed, Bath, and Beyond	Restricted the open-ended return window to 180 days in 2019, and to 90 days in 2021.
Athleta	Restricted the open-ended return window to 60 days in 2020.

Table 4.1: Examples for Restrictive Return Policy Changes to Long-Established Lenient Return Policies

Restrictive changes to long-established lenient return policies can result in consumer backlash and negative consumer sentiment. For example, L.L.Bean faced four lawsuits claiming a one-sided consumer contract breach, creating negative publicity (Gintzler, 2018). In general, restrictive return policy changes by popular retailers often spark heated discussions on consumer forums and social media: a search on Reddit for "return policy change" lists more than 200 discussion threads. Based on these discussions, consumers clearly hold divergent opinions regarding restrictive return policy changes by retailers aimed at reducing the operational cost of handling returns, loss in salvage values, and cases of opportunism. Some consumers argue that abusive consumers are to blame and that retailers have no choice but to restrict their long-established lenient return policies. Other consumers contend that firms unjustly penalize all customers for the negative actions of a few, and that the offering of a generous return policy is an important part of the value proposition offered by retailers. Hence, without such a lenient policy, they would go elsewhere. Yet some other consumers express that any restrictive changes to long-established lenient return policies are a leading indicator of a retailer's downfall in the marketplace. Given the divergence of consumer reactions, it is important for retailer managers to understand 1) whether making restrictive changes to long-established lenient return policies to deal with the increased cost of consumer returns can be detrimental to their businesses and 2) how to mitigate the potential negative business effects of such changes. Our research aims to provide insights regarding these issues.

Prior academic research shows that return policy leniency significantly impacts consumer intentions to purchase from a retailer (Abdulla et al., 2019). Abdulla et al. (2021) find that return policy leniency across five levers available to retailers can significantly impact consumer purchase intentions by increasing perceived quality of the returns service, reducing perceived transaction costs, and increasing perceived value of the returns service. One key managerial implication of this research is that retailers should focus on leniency levers that have relatively small impact on consumer purchase intentions when considering return policy restrictions, such as return time window, in order to balance between the cost burden of returns and consumer value proposition. Meanwhile, Ertekin and Agrawal (2021) report that shortening the return time window from 100 days to 60 days resulted in a decrease in the annual sales of a jewelry retailer.

Our research contributes to this line of research by theoretically and empirically examining how consumers tend to react if a retailer with a long-established lenient return policy (e.g., with an open-ended return window) changes its return policy with varying levels of restriction severity.

Theoretically, we argue that when a long-established lenient return policy is restricted by a retailer, consumers will perceive a breach of psychological contract with the retailer, hurting their trust in the retailer and motivating them to decrease their patronage or retaliate through negative word-of-mouth. The greater the restriction severity, the greater should be the damage to consumer trust and favorable behavioral intentions. Empirically, we demonstrate that restrictive changes made by retail managers to long-established lenient return policies tend to result in a psychological contract violation, resulting in loss of trust in the retailer and negatively impact favorable behavioral intentions of the customers (Zeithaml et al., 1996). We also find support for a positive association between the severity of the restriction and the magnitude of the negative reactions.

With respect to potential strategies to mitigate the negative effect of a restrictive policy change, this research focuses on managerial transparency, in the form of communicating the rationale for the change to the consumers. We are motivated by our observations of different approaches taken by the retailers with respect to communicating return policy changes. For instance, L.L. Bean announced its decision to restrict its long-established lenient return policy alongside with a managerial rationale in a letter by the CEO to customers. Meanwhile, retailers such as Macy's and Bed, Bath and Beyond remained silently about the changes and simply updated the return policies listed on their websites. In general, retailers face salient trade-offs when deciding whether to announce their restrictive return policy changes through official communication channels and whether to provide a rationale for the policy change. On one hand, an official announcement of a restrictive return policy change along with a rationale for the change would make the decision more concrete in the consumer or financial market, potentially amplifying the negative signal with respect to retailer's performance. On the other hand, not announcing the change may lower the retailer's trustworthiness in the minds of consumers who may find out about the change through alternate channels such as online consumer forums, social media, word-of-mouth, or even at the retailer when subsequently returning a product. To provide guidance for handling this managerial trade-off, we examine whether communicating the policy decision, along with a rationale, can mitigate the negative effect generated by restrictive return policy changes or perhaps exacerbate it.

Overall, our findings suggest that providing a rationale for restrictive return policy changes directly to the customers can mitigate the negative effects of such changes.

Finally, in order to extend the scope of our findings to the broader context of service policies and explore the generalizability of our theory, we examine restrictive changes to long-established, complementary credit card benefits, motivated by the recent trend in the credit card industry that parallels with restrictive return policy changes. Again, our findings support the argument that restrictive changes to the long-established complementary benefits offered by a service provider can reduce trust toward the service provider and favorable behavioral intentions. However, we also find that communicating the managerial rationale in this context, at least using a rationale that is based on several real-life cases, will not mitigate, but rather exacerbate the negative effect. This finding implies that providing a rationale for restrictive service policy changes may not be a readily-available tactic for managers to attenuate the negative reactions and that the nature of the restrictive change as well as the perceived adequacy of the provided management rationale makes a difference.

The rest of the paper is organized as follows. In the next section, we provide a theoretical background and put forth the research hypotheses. In §4.3, we present the experimental studies that test these hypotheses. We conclude with §4.4, where we discuss the implications of our research, limitations, and future opportunities.

4.2 Theoretical Background and Hypotheses

Psychological contracts refer to a set of expectations, beliefs, obligations, and entitlements as perceived by the sides of a relationship. Psychological contracts tend to be broader in nature than legal contracts. Psychological contracts can include perceptual, unwritten, and implicit terms that cannot be explicitly incorporated into legal contracts. Originally developed to study employee-employer relationships within organizations (Rousseau, 1995), the scope of psychological contract theory has been extended to study buyer-seller (Pavlou & Gefen, 2005; N. Malhotra et al., 2017) and buyer-supplier relationships (Hill et al., 2009; Eckerd et al., 2013). Psychological contracts are particularly suitable to characterize buyer-seller relationships in retail. Consumers tend to be unaware of all the explicit rules written in legal contracts and they hold an implicit understanding of the retailer's service obligations that are primarily shaped by perceived norms, references to other sellers, and prior experiences (Pavlou & Gefen, 2005).

A psychological contract violation occurs when people think they are not getting what they should expect from a relationship and feel betrayed (Robinson & Rousseau, 1994; Morrison & Robinson, 1997). Psychological contract violations reduce the trust of an exchange partner (Robinson, 1996). In fact, the psychological contract theory argues that sustaining the initial trust established between the exchange parties relies heavily on avoiding violations of psychological contracts (Niehoff & Paul, 2001). Psychological contract violations also lead to negative attitudinal and behavioral consequences. For example, Pavlou and Gefen (2005) demonstrate that psychological contract violations in buyer-seller transactions in online marketplaces, such as product misrepresentation and delivery delay, lead to both decreased transaction intentions and decreased actual transactions.

Many consumers consider lenient return policies as an important part of retailers' value propositions (Davis et al., 1998; Abdulla et al., 2019). A lenient return policy is a visible quality and trustworthiness signal for a retailer, showing that the retailer stands behind its products and is willing to compensate for a consumer's post-purchase remorse (Abdulla et al., 2021). Lenient return policies became the norm in the U.S. retail industry over the last few decades. Consequently, consumers in the U.S. grew to perceive lenient return policies more as an entitlement than as a granted privilege. Retailers generally indicate in their terms of service that they can change the return policies at any time; that is, a potential policy change is a part of the *legal* contract between retailers and consumers. However, we contend that a return policy would also be considered as a *psychological* contract due to the consumer's sense of entitlement and a basis for trust in the retailer. Pavlou and Gefen (2005) refer to "offering a return or a refund policy and then failing to acknowledge product guarantees" as a psychological contract violation. We contend that a restrictive change to a long-established lenient return policy may also be perceived as a psychological contract violation. As a result, consumers would decrease trust in the retailer and subsequently decrease favorable behavioral intentions toward the retailer. Thus, we hypothesize:

Hypothesis 1. *A restrictive change to a long-established lenient return policy decreases consumer trust in a retailer, leading to a decrease in favorable behavioral intentions.*

It is not only the occurrence of a psychological contract violation, but also the severity of the violation, that determines the nature and magnitude of the resultant negative reactions (Eckerd et al., 2013; Mir et al., 2017). In our research context, by the severity of violation, we mean the extent of the potential damage to the customer, including financial, mental, and emotional damages. The degree of restriction that a retailer imposes on its long-established lenient return policy reflects the severity of the psychological contract violation. Consequently, more restrictive changes to the retailer return policies can result in a stronger reduction in consumer trust in the retailer and favorable behavioral intentions. Hence, we hypothesize:

Hypothesis 2. *Increased severity of the restrictive change to a long-established return policy results in a greater reduction in favorable behavioral intentions, through decreased consumer trust.*

Organizational and social psychology literature lends ample empirical support that providing justifications, explanations, and motive mitigate reactions to actions and decisions. For example, Lind et al. (1980) report that people show less dissatisfaction toward an unfavorable outcome when they perceive the undertaken procedure to distribute this outcome as fair. Bies and Shapiro (1987) find that providing a justification positively influences the judgments of fairness and endorsement of a decision maker's negative actions. For example, M. C. Campbell (1999) demonstrates that buyers hold sellers responsible for price increases unless there is evidence showing otherwise. However, a firm's positive motive may lessen the perceived unfairness of the price increases even when the increases are attributable to the internal cost increase of the firm. Research also suggests an action is perceived more negatively when the inferred motive of the actor is negative. A negative inferred motive leads to a greater attribution of responsibility to the actor and the behavior is perceived as more aggressive or unfair (Betancourt & Blair, 1992; Weiner, 1995). Overall, the attribution-based research is conclusive in that inferences of negative motives due to missing justifications for the actions taken lead to greater causal attributions and subsequent negative responses. Recent literature on operational transparency also suggests that providing “behind-the-scenes” information to customers regarding the firm’s operations can generate favorable outcomes, as long as such transparency does not reveal an ineffective process, poor outcomes even from the best effort, violations of implicit social norms, or simply things that customers do not find appealing (Buell et al., 2017).

Articulating the rationale for a restrictive change to a long-established lenient return policy—managerial transparency—may signal to customers that the managerial decision is due to mitigating factors beyond the direct control of the retailer (e.g., return abuse, opportunism, or fraud) and that the restrictive policy change is “a last resort.” When coupled with such communication, a restrictive change decision may not engender a sense of betrayal by the retailer or perceptions of the retailer’s incompetence, relative to the response to a silent action that completely leaves the speculative inferences regarding the motives of the change to consumers. As a result, communicating the rationale for the restrictive change should moderate any negative impact of the change on the consumer’s trust in the retailer and his or her favorable behavioral intentions toward the retailer. Hence, we hypothesize:

Hypothesis 3. *Managerial transparency—providing the rationale for a restrictive change to a long-established lenient return policy—moderates the negative indirect effect of the change on favorable behavioral intentions through consumer trust.*

People tend to feel angrier, more upset, and more punitive when a perceived wrongdoing results in severe rather than mild negative consequences (see Giordano, 1983; Miller & Vidmar, 1981, for reviews). Accounts that explain the rationale for the wrongdoing of a party would be less effective in reducing strongly negative reactions, which can hinder cognitive reflection on the provided account. For instance, T. E. Johnson and Rule (1986) find that when individuals were highly upset, rather than mildly upset, about an experimental partner’s provocative behavior, information indicating the partner’s aggression was caused by extenuating circumstances was less effective in reducing anger and retaliation. As another example, explanations for employee layoffs are less likely to mitigate negative reactions when they are provided to the victims (i.e., those fired) rather than to the survivors of a layoff by a firm (Konovsky & Folger, 1991).

However, prior research also provides evidence for the counter argument. For example, Brockner et al. (1990) find that explanations for layoff decisions mitigated layoff survivors' tendency to report negative work-related attitudes and behaviors when the survivors felt more, rather than less, anxious about the layoffs. Similarly, Shapiro and Buttner (1988) suggests that manager explanations for rejecting job candidates mitigated feelings of injustice when the rejection was more, rather than less, upsetting to job candidates. One can argue that a minimal level of negative reaction due to a psychological contract violation is necessary for people to care enough about or consider critically the provided rationale (Shapiro, 1991). In our context, a low-severity restriction imposed upon a long-established lenient return policy may not be detrimental enough for consumers to reflect on the provided rationale in forming a reaction to the decision. The arguments provided above thus lead us to the following competing hypotheses:

Hypothesis 4 (A)**.** *The moderating effect of managerial transparency is greater when the severity of the restrictive change to a long-established lenient return policy is lower.*

Hypothesis 4 (B)**.** *The moderating effect of managerial transparency is greater when the severity of the restrictive change to a long-established lenient return policy is higher.*

4.3 Empirical Studies

This section presents three experimental studies that we conducted to test our hypotheses. The first of these studies was a pre-study through which we developed and validated measurement scales and constructs. The pre-study also served as a pilot test for the experimental design, vignettes, manipulations, and potential treatment effect sizes, all of which informed the design and procedures of the subsequent studies. Then, Study 1 examined how restrictive changes to a long-established lenient return policy might impact consumer trust and favorable behavioral intentions (Hypothesis 1) and how this impact varies depending on the severity of the restrictive change (Hypothesis 2). This study also examined the effectiveness of managerial transparency, in the form of explaining the rationale for the restrictive change, in mitigating the associated negative impact (Hypothesis 3) and the relative effectiveness of managerial transparency based on the severity of the restrictive change (Hypotheses 4A and 4B). Finally, Study 2 examined our main hypotheses in a distinct yet related managerial context—discontinuation of long-established complementary credit card benefits—to explore generalizability of our theory to the broader context of service policies.

All experiments used vignette methodology and are designed and implemented using Qualtrics. The experiments involved samples from the Prolific Academic online social science research platform. Using this online platform enables data collection from a diverse and representative consumer population, which is important in terms of the external validity of the empirical findings (Berinsky et al., 2012; Paolacci & Chandler, 2014; Goodman & Paolacci, 2017). The Prolific Academic platform allows participation in natural living or working environments—which mitigates concerns regarding observer effects. This environment provides a lab-in-the-field setting. Each participant received a compensation of $1.50 for completing a study, which took, on average, 9 minutes. We ensured data quality from the online platform through a number of technical guards and validation procedures. In particular, we a) limited the location of participants to the U.S. using the Prolific Academic pre-screening variables (and cross-verified this with geolocation metadata collected via the Qualtrics interface), b) qualified only workers with a previous task approval rate greater than 95% to participate, c) restricted each participant to complete only one of the studies, d) asked short text entry and attention check questions, and e) required participants to manually enter a randomly-generated code at the end of each study in a designated area to receive the payment.

4.3.1 Pre-study

In the pre-study, we examined reactions to a retailer's policy change treatment. Thus, measuring pretreatment, baseline levels of trust in and favorable behavioral intentions toward the focal retailer (i.e., pretest) was important both from the perspective of realism and due to the greater statistical power to estimate the average treatment effects. Further, in order to estimate an unbiased average treatment effect of a restrictive change to the return policy, we needed a control group that would not experience the change. As such, we decided that a pretest-posttest control group design is the most appropriate experimental design for our research objective. This experimental design, when rigorously executed, is considered a powerful design in estimating the causal effect of a treatment by accounting for the pretest baseline values of the outcome variables. Despite its numerous advantages, one potential drawback of this design is that participants in the experiment may be subject to pretest sensitization (D. T. Campbell & Stanley, 1963). That is, collecting measurements during the pretest phase of the same outcome variables as in posttest may impact subsequent reactions of study participants to the treatment.

In order to ensure that the pretest-posttest control group design would not suffer from pretest sensitization effects in our research context, we conducted the pre-study using a Solomon four-group design (Solomon, 1949). In this design, participants are randomly divided into four groups and each group experiences a different measurement protocol: the first group gets the pretest, the treatment, and the posttest; the second group receives only the treatment and posttest; the third group gets the pretest, no treatment, and a posttest; and the fourth group gets only a posttest. Although the Solomon four-group design is a "data-hungry" design, with unique challenges to analyze the data in order to examine causal effects of the treatments, it is an effective design to judge whether a more efficient pretest-posttest control group design is warranted to test the causal effect of a treatment in a particular context (Spector, 1981). In particular, by examining the interaction effect between the pretest administration and treatment in a 2×2 ANOVA, one can rule out the potential sensitization effect if there is no significant interaction effect. After ruling out the sensitization effect, one can effectively proceed with a pretest-posttest control group design to examine the causal effect of a treatment (D. T. Campbell & Stanley, 1963).

One hundred sixty four participants who entered the experiment first read background information about the (fictional) focal retailer ACME, a large, U.S. based multi-channel general merchandise store. We chose multi-channel general merchandise stores as the focal context based on 1) the motivating retailer cases described in Table 4.1, and 2) an observation that these retailers constitute more than half of the retail sales and deal with approximately half of the consumer returns in the U.S. Participants were told that ACME has a large customer base and has long been known for its service quality, as evidenced by very high customer satisfaction rates indicated by *Consumer Reports* surveys. Then, participants were told to assume that they are one of the long-time customers of ACME, had purchased a variety of products over the past few years, and have consistently experienced great product and service quality. Participants were also told that a unique and frequently praised aspect of the retailer's service offering is its lenient return policy and that ACME's lenient return policy has been unchanged for the last 20 years. Overall, the role of the background information was to induce in *all* participants a high level of trust in ACME, through positive signals, and a high baseline attitude toward the retailer and to familiarize them with the long-established lenient return policy.

Following the presentation of this background information, participants were presented the snapshot of a web page that contained the long-established lenient return policy of ACME. The return policy included an open-ended return time window, in line with the retailers shown in Table 4.1, a full refund or exchange, and free return shipping for online returns (Figure 4.1). Then, the first and third experimental groups received a randomized battery of pretest measurement items, based on which baseline consumer trust in the retailer and favorable behavioral intentions were measured. Consumer trust in retailer was captured using an average of five reflective indicators adapted from Bhattacherjee (2002), and favorable behavioral intentions were captured using an average of three reflective indicators adapted from Zeithaml et al. (1996) (see Table 4.2). The second and fourth experimental groups did not receive the pretest.

To ensure that what the control groups experienced during the experiment matched that of the treatment groups except for experiencing the return policy change treatment, we used the following transition prompt that all groups read: "Imagine, a close friend of yours comes and tells you that ACME has changed its 20 years old return policy. To verify his claim, you visit ACME's website and view the return policy." Then, the participants in the control groups saw the return policy, as in Figure 4.1, while the participants in the treatment groups saw the return policy with a restrictive change. To operationalize the restrictive change, we utilized the length of the return time window,

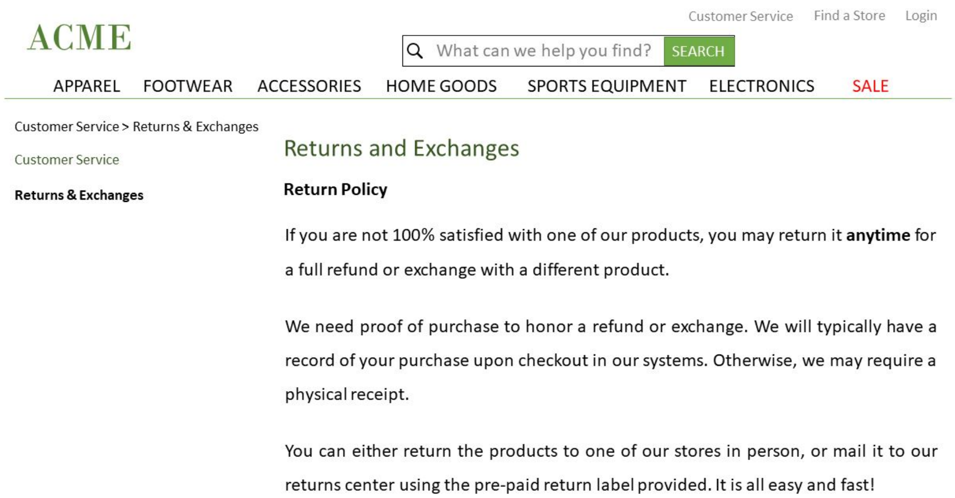

Figure 4.1: ACME's Long-Established Lenient Return Policy

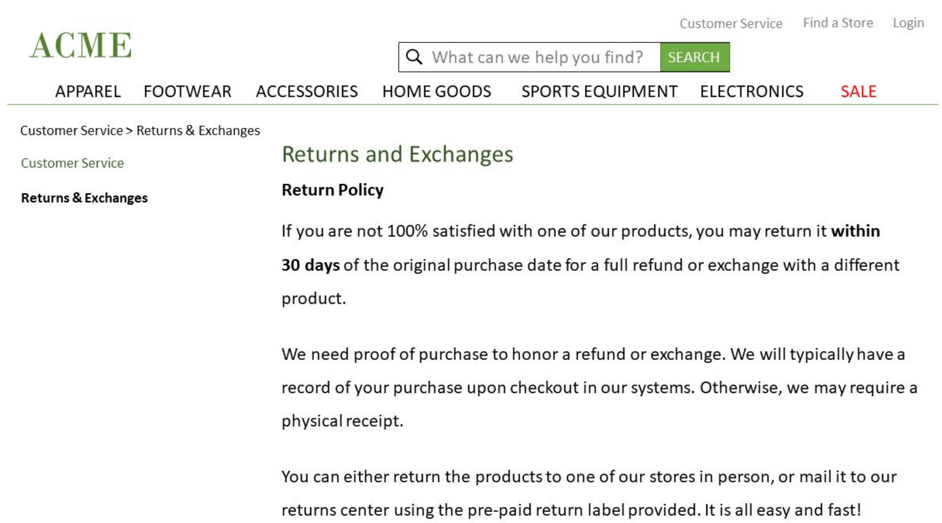

Figure 4.2: ACME's Return Policy After the Restrictive Change

Latent Variable	Reflective Measurement Item
Consumer Trust in Retailer ($\alpha = 0.84$)	1. ACME seems to have high integrity in doing retail business 2. ACME would be strongly committed to what its customers value. 3. ACME would protect its customers' best interests while doing business. 4. ACME is a caring retailer. 5. ACME has customers in mind while making business decisions.
Favorable Behavioral Intentions ($\alpha = 0.86$)	1. I would purchase from ACME. 2. I would recommend ACME to others. 3. I would use the services of ACME in the future.

Notes: All measurement items were in a 7-point Likert scale from 1–Strongly Disagree to 7–Strongly Agree

Table 4.2: Measurement Scales for the Pilot Study and Study 1

since 1) most of the recent motivating examples from retailer practice used this dimension and 2) the manipulation did not result in a substantial change in the complexity of the return policy statement. In particular, the treatment groups viewed the return policy as in Figure 4.2, where the return time window was restricted to be 30 days, which is a standard in the U.S. retail sector (Narvar, 2021).

Participants then completed the posttest survey using the same measurement items as in the pretest. At the end of the pre-study, two attention checks in the form of factual manipulation checks (Kane & Barabas, 2019) asked participants 1) whether they had observed a substantial change in the return policy of ACME, and 2) how long was the return time window when they viewed the return policy for the second time. Of 164 participants, 127 participants answered both attention check questions correctly and were included in the final empirical analysis.

A 2 $\times$ 2 ANOVA suggested no significant interaction effect between the pretest administration and treatment variables on either consumer trust ($F(1, 123) = 0.25, p = 0.622$) or favorable behavioral intentions ($F(1, 123) = 0.01, p = 0.971$). We thus concluded that there was not a significant pretest sensitization effect and continued to use a pretest-posttest control group design for the subsequent studies. Pre-study also provided initial evidence for the causal effect of a restrictive change to a long-established lenient return policy on consumer behavioral intentions. In particular, by analyzing the groups who had the pretest in an ANCOVA with behavioral intentions as the dependent variable, treatment as the factor variable, and pretest consumer trust and behavioral intentions as covariates, we found that the behavioral intentions of a consumer was negatively impacted by the restrictive change. Furthermore, using a regression-based mediation analysis with the non-parametric bootstrapping (10,000 resamples) technique (Hayes, 2018), we found that consumer trust mediates the negative impact of the restrictive change on behavioral intentions ($\theta_{X \to M \to Y} = -0.726, \text{CI} = [-1.185, -0.267]$). This provided initial evidence that supports Hypothesis 1.

4.3.2 Study 1

Using the validated experimental vignettes and measurement scales, Study 1 tested Hypotheses 1–4.

4.3.2.1 Method

Six hundred forty participants recruited through the Prolific Academic platform completed Study 1 for a compensation of $1.50 per participant. Study 1 used a partial factorial pretest-posttest control group design, with seven groups, and participants were randomly assigned into the groups. The experimental protocol was similar to the pre-study. All participants completed the pretest as well as the posttest. In line with the majority of the motivating retailer cases in Table 1 and the pre-study, the baseline return policy was stated as in Figure 4.1, with an open-ended return time window.

After completing the pretest phase and following the transition prompt, participants were randomly assigned to either the control group, that did not experience any change to the lenient return policy, or to one of the six treatment groups based on three levels for the restriction severity factor (Low/Medium/High) and two levels for the managerial transparency factor (Provided/Not Provided). Note that the resulting experimental design is a partial factorial design in the sense that there is not a managerial communication condition for the control group due to irrelevance. To operationalize the low-severity restrictive change, we manipulated the return window from open-ended to 365 days, similar to the motivating cases, such as from REI, Macy's, and L.L. Bean. We

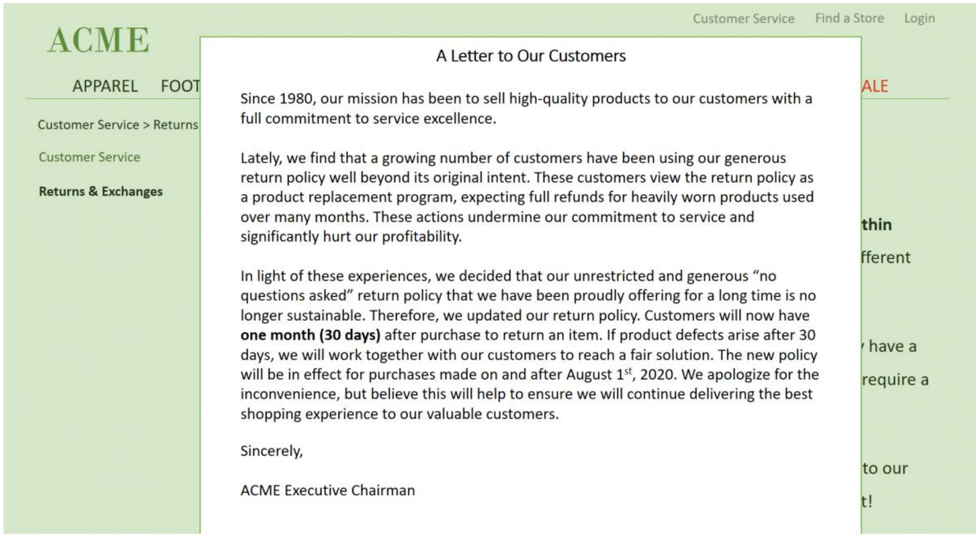

Figure 4.3: Managerial Communication to Explain the Rationale for the Restrictive Policy Change (High-Severity Condition)

implemented the medium-severity restrictive change as from an open-ended to a 180-day return window, as in the cases of Kohl's and Bed, Bath, and Beyond. Finally, the high-severity restrictive change was operationalized as a change from an open-ended to a 30-day return window, which is the most common return time window among 200 retailers in the U.S., as surveyed by Narvar (2021).

Participants who were assigned to a treatment group pertaining to managerial transparency saw, following the updated return policy statement, a managerial communication that highlighted the rationale for the restrictive return policy change. For example, participants in the high-severity condition saw the snapshot in Figure 4.3. The text of the communication was based on the letter that L.L. Bean addressed to its customers after its policy changed, as well as on popular press articles that discussed return policy changes of major retailers.

All participants completed the pretest and posttest measurements that were consistent with the pre-study. Participants also answered several post-experiment survey questions that were intended to generate covariates for the analysis, to improve the precision of the effect estimates by reducing the residual variance. These covariates and summary statistics are described in Appendix C.

4.3.2.2 Analysis and Results

Of the 640 participants, 538 participants who passed the two attention checks (factual manipulation checks), as in the pre-study, were included in the final empirical analysis. Descriptive statistics for the key variables are provided in Table 4.3.

Variable	Pretest Mean (S.D.)	Posttest Mean (S.D.)
Consumer Trust in Retailer	5.939 (0.843)	5.652 (1.051)
Favorable Behavioral Intentions	6.125 (0.743)	5.885 (0.915)

Table 4.3: Descriptive Statistics for Key Variables

A post-treatment manipulation check question asked participants' opinions, on a 7-point Likert scale from 1–Strongly Disagree to 7–Strongly Agree, regarding the statement "ACME has an extremely lenient return policy." An ANOVA confirmed that the restrictive change treatment of various severity levels was successful ($F(3, 534) = 44.524$, $p = 0.000$). A Hawthorne check asked participants' opinions, on a 7-point Likert scale from 1–Strongly Disagree to 7–Strongly Agree, regarding the statement "The wording of the return policy of ACME is very complex." An ANOVA suggested that the restrictive change treatment did not differentially impact the perceived complexity of the return policy text—a construct that was meant to not get influenced by the treatments ($F(3, 534) = 1.022$, $p = 0.382$). To see this, note that return policies in the control group and treatment groups differed only in terms of a word that stated return window but otherwise had the same content. A realism check that assessed on a 7-point Likert scale from 1–Strongly Disagree to 7–Strongly Agree "ACME return policy was realistic in its wording and format." ensured that participants found the return policy presented to them as realistic ($M = 5.77$, $SD = 1.07$).

We also assessed two manipulation checks for the managerial transparency treatment. These checks asked participants to indicate their opinions, on a 7-item Likert scale from 1–Strongly Disagree to 7–Strongly Agree, regarding the statements "I believe ACME had been losing significant revenue because of its long-lasting lenient return policy" and "I think ACME had been incurring significant operational costs because of its long-lasting lenient return policy." A MANOVA suggested that participants who were provided the rationale for the restrictive change agreed with the two statements significantly more than participants who were not provided the communication (Wilk's $\lambda = 0.975, F(2, 469) = 5.93, p = 0.003$). This finding suggested that the manipulation of managerial transparency was indeed successful. We also confirmed with another MANOVA that the pretest consumer trust and favorable behavioral intentions were not significantly different across the control and treatment groups (Wilk's $\lambda = 0.990, F(6, 1066) = 0.86, p = 0.523$).

To test the hypotheses, we used a regression-based mediation, moderation, and conditional process analysis approach (Hayes, 2018). For statistical inference, we used the non-parametric bootstrapping technique with 10,000 resamples. We utilized the pretest variables as covariates in the analysis, resulting in what Valente and MacKinnon (2017) refer to as the ANCOVA model for mediation analysis. Valente and MacKinnon (2017) compare different approaches for estimating mediation effects in a pretest-posttest control group design and conclude that the ANCOVA model generally outperforms other alternatives (residual change scores and difference score models) in terms of statistical power and Type I error rates.

In addition to the pretest variables of consumer trust in the retailer and favorable behavioral intentions, we added 1) general tendency to trust others (average of four 7-item Likert-type items, see Appendix C), 2) perceived return frequency, 3) perceived importance of lenient return policies in choosing a retailer, 4) percceived extent of return policy abuse in the U.S., as covariates. The covariates 2–4 were measured using 7-item Likert scale from 1–Strongly Disagree to 7–Strongly Agree to the statements "I end up returning products quite often," "Lenient return policies are important for me while choosing a retailer to purchase," and "In the U.S., lenient return policies are abused a lot," respectively. We chose these covariates to be included in the analysis based on their significant correlations with the mediating and outcome variables. We follow the notational conventions by Hayes (2018) in reporting the results below. We report 95% bias-corrected bootstrap confidence intervals for statistical inference; the 95% percentile-based bootstrap confidence intervals were very close.

First, we found that relative to the control group, in the absence of managerial transparency, treatment groups who experienced a medium-severity or high-severity restrictive change to the return policy significantly lowered their favorable behavioral intentions, as a result of decreased trust in the retailer. The conditional indirect (i.e., mediation) effects for the medium- and high-severity groups were $\theta_{D_2 \rightarrow M \rightarrow Y}|(W = 0) = -0.153$, CI $= [-0.256, -0.065]$ and $\theta_{D_3 \rightarrow M \rightarrow Y}|(W = 0) = -0.530$, CI $= [-0.746, -0.356]$, respectively. Meanwhile, the low-severity group did not show statistically significant conditional indirect effect relative to the control group ($\theta_{D_1 \rightarrow M \rightarrow Y}|(W = 0) = -0.067$, CI $= [-0.167, 0.017]$. Further, medium- versus low-severity ($\theta_{(D_2 \text{ vs } D_1) \rightarrow M \rightarrow Y}|(W = 0) = -0.086$, CI $= [-0.188 - 0.003]$) as well as the high- versus medium-severity ($\theta_{(D_3 \text{ vs } D_2) \rightarrow M \rightarrow Y}|(W = 0) = -0.377$, CI $= [-0.570, -0.205]$) contrasts of conditional indirect effects were both statistically significant. Thus, we found partial support for Hypothesis 1 and full support for Hypothesis 2.

Second, we found that when a restrictive change is accompanied with managerial transparency, the impact of the medium- and high-severity restrictive changes on favorable behavioral intentions, mediated by consumer trust, becomes less negative ($\theta_{D_2 \rightarrow M \rightarrow Y}|(W = 1) = -0.098$, CI $=$ $[-0.198, -0.009]$, $\theta_{D_3 \rightarrow M \rightarrow Y}|(W = 1) = -0.122$, CI $= [-0.256, -0.023]$). As expected, managerial transparency did not significantly moderate the (null) effect low-severity restrictive change on consumer trust and resultant favorable behavioral intentions. Thus, Hypothesis 3 was supported for medium- and high-severity restrictive changes. As for the competing Hypotheses 4A and 4B, our analysis reveals empirical support for the latter. In particular, the moderating effect of managerial transparency was significantly stronger under high-severity restriction than the effect under medium-severity restriction ($[\theta_{D_3 \rightarrow M \rightarrow Y}|(W = 1) - \theta_{D_3 \rightarrow M \rightarrow Y}|(W = 0)] - [\theta_{D_2 \rightarrow M \rightarrow Y}|(W = 1) - \theta_{D_2 \rightarrow M \rightarrow Y}|(W = 0)] = 0.408$, CI $= [0.225, 0.620]$).

4.3.2.3 Discussion

Our experimental investigation reveals that long-established lenient return policies of retailers are likely to act as psychological contracts between customers and retailers. As such, a restrictive change to long-established lenient return policies engenders a psychological contract violation and significantly hurts consumer trust in retailers, leading to lower favorable behavioral intentions. At the same time, our results suggest that the extent of negative consumer reactions depends significantly on the restriction severity. That is, medium- and high-severity restrictions are more likely to generate significant negative reactions relative to a low-severity restriction, such as in the case of L.L. Bean.

Further, we find that managerial transparency is an effective strategy to mitigate the negative consumer effect of such changes. Communicating the rationale for such decisions would be particularly effective with a high-severity restriction. For instance, our results indicate that when coupled with managerial transparency, the negative reaction to the high-severity restrictive change (from an open-ended to a 30-day return window) becomes statistically equivalent to the negative reaction to a medium-severity change (from an open-ended to a 180-day return window).

4.3.3 Study 2

To further explore the contextual generalizability of our theory, Study 2 investigated whether discontinuation of long-established complementary benefits offered by a service provider can trigger analogous reactions by consumers. As the context for Study 2, we focused on a discontinuation of complementary credit card benefits. Recently, several of the major credit card issuers, such as Discover and Chase announced that they discontinued several long-established complementary benefits such as price protection, extended product warranty, and purchase protection. As a rationale, the issuers cited prolonged low-usage and hassles to both card users and issuers (Karp, 2017). This recent trend in the credit card industry motivated us to extend the scope of our theory and examine whether the long-established complementary benefits are perceived by customers as a part of the psychological contract between them and the service providers. Note that the credit card benefits, such as price and purchase protection, are similar to consumer return policies in being ex-post service benefits.

4.3.3.1 Method

One hundred sixty five participants recruited via the Prolific Academic platform completed the study. The study had a pretest-posttest control group design with one control and two treatment groups. Participants were randomly assigned into the groups upon entering the study. Similar to the previous studies, participants were presented background information about the focal credit card issuer, Bank Name. Participants were told that Bank Name is one of the largest commercial banks and financial services providers in the U.S., is particularly well-known for its credit cards, has a large customer base for its credit cards, and has long been known for its excellent customer service, as evidenced by very high customer satisfaction rates indicated in surveys by *Consumer Reports*. Participants were then told to imagine that they hold the Cash Back credit card by Bank Name and had used this card for more than 10 years. Participants also read that the Bank Name Cash Back credit card has been extremely popular among the consumers for its unique complementary benefits, on top of its key benefits and that the card has offered these benefits ever since the card's introduction, 15 years ago.

After reading the background information, all participants viewed the web page that listed the long-established key and complementary benefits of the Cash Back credit card (Figure 4.4). Then, all participants completed the pretest measurements. Measurement items were adapted from the earlier studies to the specific aspects of the current context (see Appendix C). Next, all participants read the following prompt: "Imagine a close friend of yours, whom you referred to sign up for the Cash Back credit card, tells you that Bank Name must have discontinued some of its benefits for the Cash Back credit card, as he cannot find out about them on the web page. To verify this claim, you visit the Bank Name website and open the benefits page for the Cash Back credit card." Then, participants in the control group saw, again, the vignette in Figure 4.4. Participants in the first treatment group saw the vignette in Figure 4.5, while participants in the second treatment group saw, in addition to the vignette in Figure 4.5, the following managerial communication that is based on the cases of Discover and Chase in announcing their discontinuation of their respective card benefits: "We regularly evaluate our cardmember benefits to ensure that we are meeting or exceeding our cardmembers' current needs and expectations. We would like to kindly inform you that, due to prolonged low usage, effective February 28, 2021, we will discontinue Extended Product Warranty, Return Guarantee, Purchase Protection, Auto Rental Insurance and Flight Accident Insurance benefits for all Cash Back credit cards. We will continue to offer free benefits that card members use and value the most." Following the treatment administration, participants completed the posttest survey.

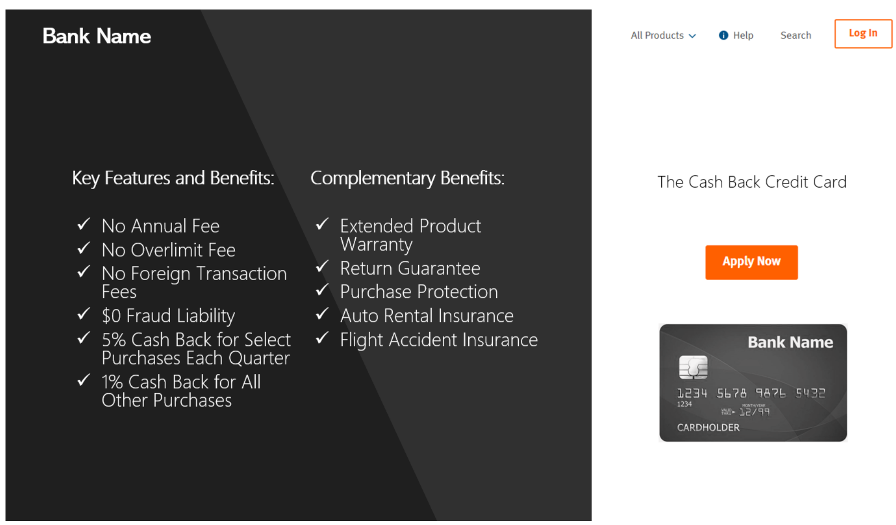

Figure 4.4: Vignette: Bank Name's Long-Established Benefits for the Cash Back Credit Card

4.3.3.2 *Analysis and Results*

One hundred fifty one participants who answered the attention check questions correctly were included in the final empirical analysis. Descriptive statistics for the key variables are provided in Table 4.4. A regression-based mediation analysis using the same non-parametric bootstrapping technique as in Study 1 with 10,000 resamples was performed. As a result, we found statistically significant and negative relative indirect effects of the restrictive change to long-established credit card benefits without ($\theta_{D_1 \to M \to Y} = -0.529, \text{CI} = [-0.890, -0.147]$) or with ($\theta_{D_2 \to M \to Y} = -0.988, \text{CI} = [-1.406, -0.559]$) managerial communication explaining the rationale. Note that, interestingly, in this credit card context, not only did the managerial communi-

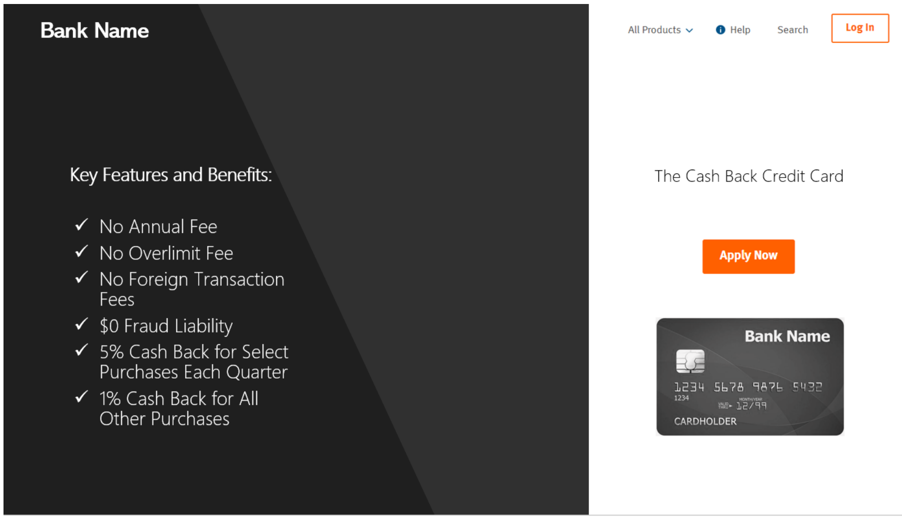

Figure 4.5: Vignette: Bank Name's Cash Back Credit Card After Discontinuation of Complementary Benefits

Variable	Pretest Mean (S.D.)	Posttest Mean (S.D.)
Consumer Trust in Provider	5.350 (1.047)	4.388 (1.377)
Favorable Behavioral Intentions	5.633 (0.920)	4.580 (1.427)

Table 4.4: Descriptive Statistics for Key Variables

cation fail to mitigate the negative effect of the restrictive change to the credit card benefits, but also exacerbated the effect. The difference between the relative indirect effects when managerial communication was provided versus not was statistically significant ($\theta_{D_2 \to M \to Y} - \theta_{D_1 \to M \to Y} = -0.459, \text{CI} = [-0.928, -0.013]$).

4.3.3.3 Discussion

Study 2 suggests that the occurrence of psychological contract violation as a result of restrictive changes to long-established service policies is generalizable to different service contexts and benefits. However, Study 2 also reveals that the effectiveness of managerial transparency in alleviating the negative impact of restrictive service policy changes may be dependent on the specific service context, nature of the change, and communicated rationale. While in the case of restricting the return time window, the communicated rationale moderated the negative impact of the decision on consumer trust and favorable behavioral intentions, in the case of discontinuation of complementary credit card benefits, the communicated rationale aggravated the negative impact. Why these opposite effects?

We know from organizational justice literature that explanation adequacy—the extent to which provided explanations are clear, reasonable, and detailed—plays a significant role in determining the effectiveness of the provided explanation in mitigating negative reactions to decisions taken by management (Bies et al., 1988; Folger & Cropanzano, 1998). Both failure to provide an explanation and providing an inadequate one can lead to adverse consequences. In fact, research suggests that the perceived adequacy of an explanation is more important than its nature (i.e., excuse or justification) when it comes to its effectiveness to mitigate negative consequences of decisions (Shaw et al., 2003). Our conjecture is that the explanation provided in the return policy change context was perceived as adequate, whereas the explanation provided in the credit card benefits change context was not.

4.4 Conclusion

Motivated by the trade-offs that U.S. retailers face in maintaining versus restricting long-established lenient return policies and increasingly common policy changes that involved restrictions to long-established lenient return policies through shortening the return time window, we investigated how such changes affect consumer trust and favorable behavioral intentions toward retailers. By employing the psychological contract and organizational trust theories as lenses, we hypothesized that restrictive changes to long-established lenient return policies can significantly hurt consumer trust and subsequently lower favorable behavioral intentions. We also hypothesized that the magnitude of the negative impact of such restrictive changes may differ based on the severity of the restriction. Testing different severity levels for the restrictive change in a randomized experiment, we found evidence that supports these hypotheses.

Noting the two distinct approaches that retailers take in terms of announcing restrictions to their long-established lenient return policy and the trade-offs involved in these approaches, we also examined whether providing the rationale for the restrictive change decision can moderate the resultant negative impact. We found that this type of managerial transparency can effectively lower the negative impact on consumer trust and favorable behavioral intentions, as long as the restrictive change has enough severity to generate significant negative reactions.

We next transferred our theory to a distinct yet related context and examined the impact of discontinuing long-established complementary credit card benefits on consumer trust and favorable behavioral intentions, as well as the moderating role of communicating the decision rationale. Here, we again found a significant negative impact of the restrictive change on consumer trust and favorable behavioral intentions. However, contrary to the case of return policy changes, however, we found in this context that managerial transparency did not moderate, but rather exacerbated the negative impact.

Our research has several managerial implications. First, retailers have to recognize that customers feel entitled to long-established lenient return policies and perceive the lenient return policies as psychological contracts. Restrictive changes to such policies will most likely hurt customers' trust and intentions to shop from them, remain loyal, and share positive word-of-mouth. The sense of entitlement to lenient return policies does not necessarily depend on the the extent to which consumers benefit from what lenient return policies offer. Abdulla et al. (2021) report from surveys that 60% of consumers think they need less than two weeks to make a return decision. However, we find that restricting an open-ended return window to even 180 days results in significantly negative consumer reactions. Thus, long-established lenient return policies are an important quality signal for retailers who offer them, and a distortion in this signal through restrictive changes can generate negative consumer reactions.

Thus, we underscore that restrictive changes to lenient return policies should only be made in light of a careful cost-benefit assessment that favors such changes. For example, restricting a return time window from open-ended to 180 days, as we found, generates significant negative consumer reactions. Would this change significantly reduce the cost burden of returns for a retailer? Shang et al. (2019) report mean return times ranging from 11 to 28 days after the purchase for a large electronics retailer that had an open-ended return window. Ketzenberg et al. (2020) report 23 days for a large department store, which is also well-known for its open-ended return window. Given the compelling empirical evidence that the vast majority of returns happen within a limited time-frame much shorter than 180 days, the net operational benefit of this decision, considering the negative impact of it on consumer trust and favorable intentions, may not be material. In fact, there is evidence that the net bottom-line impact of introducing restrictive return policy changes can be negative. For example, Ertekin and Agrawal (2021) found that a multi-channel jewelry retailer's profit significantly decreased after curbing it's 100-day return window to 60 days. Our recommendation to retailers would be to not "test the waters" when it comes to restricting long-established return policies, which seems to be the strategy followed by Macy's. A more effective strategy would be 1) to set a restriction severity whose operational cost benefits would well offset the ramifications of the restriction and 2) to set the targeted return policy in a single restrictive change.

Second, our findings imply that communicating the rationale for the restrictive return policy changes would help to mitigate the negative reactions. It is interesting that the majority of the retailers shown in Table 4.1 did not adopt this strategy. In fact, these retailers not only avoided providing a rationale for the changes they made, but also did not announce the change through their official communication channels. Furthermore, no mentioning of the return policy change appears on the retailers' annual reports for the year of the change. This evidence suggests that these retailers try to avoid drawing attention to the return policy changes and sending a negative signal to their general customer base and investors (Connelly et al., 2011). Nevertheless, popular press and consumer forums include a plethora of reports and discussions of these changes, which suggests a significant degree of press and consumer awareness. Moreover, recent research on operational transparency suggests that in certain contexts, providing transparency about relatively inferior performance can be perceived more favorably by consumers compared to no transparency at all (Buell et al., 2019). Hence, announcing and rationalizing a restrictive change through official communication channels, particularly by citing mitigating and extraneous factors that led to the change, would be a more effective tactic.

Our follow-up investigation with the rescinded complementary credit card benefits suggests that the managerial implications and recommendations documented above would be relevant to other service settings as well, with an important caveat: communicating the decision rationale can backfire and exacerbate the problem if the provided rationale is not perceived to be adequate. To improve the perceived adequacy of the communicated rationale, service providers should elaborate on and provide some data for the detrimental impacts of the status-quo policies, as well as and articulate why the status-quo is no longer sustainable for the firm. The exacerbating effect of the rationale provided in the credit card benefits context, which cited prolonged low usage of the benefits, also supports the argument that it is not the *actual use* of such benefits per se that consumers value, similar to the case of lenient return policies.

Our research provides several opportunities for future research. One can examine the role of contextual factors, such as the retailer's assortment focus and tenure on the effects that we identified in this paper. Examining return policy changes from a different baseline leniency level, with different durations for how long the baseline leniency level has been established, or examining restrictive changes across different leniency levers, such as imposing a 15% restocking fee, are also fruitful avenues. Further, a case-based research approach can be employed to examine the example retailers' actual decision-making process behind restrictive return policy changes and identify factors that may have contributed to these changes. Future research can also examine the effectiveness of different types of explanations for restrictive return policy changes in mitigating the negative consequences. One can also compare the impact of one-and-done and step-by-step restrictive policy changes with the same end level of leniency, to test the principle of "combine the pain, segment the pleasure" (Chase & Dasu, 2001).

A reasonable alternative to an across-the-board restriction of long-established lenient return policies would be to invest in big-data analytics powered information systems that can identify potentially abusive consumers on a case-by-case basis and ban these consumers from future returns. Safdar (2018) reports that retailers such as J.C. Penney, CVS, and Amazon already use such systems. Ketzenberg et al. (2020) have developed predictive models to identify abusive consumers while also considering overall profitability of consumers. Future research can compare the relative efficacy of an across-the-board policy restriction versus a more individualized approach using transactional data analytics.

5. SUMMARY AND CONCLUSION

Overly lenient return policies that have been pervasively offered by retailers in the U.S. for decades are one of the key drivers of the dramatic growth in the annual volume of consumer returns. Retail managers face an important trade-off between maintaining a lenient return policy to stimulate purchases versus restricting the return policy to reduce the cost burden of managing returns. In this dissertation, I offered three essays on consumer return policy design with theoretical and practical insights regarding this trade-off.

The first essay reviewed, classified, and synthesized the extant literature on consumer return policy design through the lens of a conceptual framework for consumer returns and a classification framework that is predicated on the key themes, modeling elements, and constructs used in the literature. The review and classification revealed a number of under-explored and unexplored areas, which motivated the research questions of the subsequent two essays.

The second essay focused on a comprehensive theoretical and empirical examination of the impact of leniency across five distinct return policy leniency levers on consumer intentions to purchase from a retailer. Here, I proposed an overarching cognitive process model to explain the effect of return policy leniency on purchase intentions and empirically tested the model through a series of randomized online experiments. The results provided support for the cognitive process model and documented the role of perceived return service quality, perceived transaction costs, and perceived return service value in mediating the effect of return policy leniency on consumer purchase intentions. Further, findings suggested that monetary, followed by exchange leniency, are two most effective levers in influencing consumers' purchase intentions. Meanwhile, the effort, time, and scope leniency levers showed statistically significant, but smaller impacts relative to the monetary and exchange leniency levers. The key practical implication of these findings is that retailers should focus on effort, time, and scope levers that are less impactful on consumers purchase intentions while considering a restrictive change to their return policies.

The third essay examined how consumers react to restrictive changes to long-established lenient return policies and how retailers can mitigate the potential negative consumer reactions to such changes. Through the lens of psychological contract theory, I hypothesized that consumers would react negatively to restrictive changes to long-established lenient return policies due to perceiving such changes as violations to the psychological contract with the retailer. Through a series of randomized online experiments, I found evidence that a restrictive change to a long-established lenient return policy through the time leniency lever results causes decreased favorable behavioral intentions, by reducing consumer trust in retailer. Further, consumers react more negatively to restrictions with greater severity. As a tactic to potentially mitigate this negative effect, I examined managerial transparency in the form of providing rationale for the restrictive change decision. I motivated this investigation by observing different approaches taken by retailers in the U.S. with respect to announcing restrictive changes to return policies. I found that providing managerial transparency results in an attenuated negative reaction to restrictive changes. This suggested that managerial transparency is a potentially effective tactic to reduce consumer backlash and negative sentiment that may arise due to restrictive changes to long-established lenient return policies. Finally, I found evidence for negative consumer reactions to restrictive changes to long-established complementary credit card benefits, transferring my theory to another service context and extending its contextual boundaries.

Overall, my dissertation research highlights the importance and relevance of examining decision-making with respect to return policies as well as service policies in general, in addition to examining service processes—a more common line of inquiry in the area of operations management. Return policies stand out as a particularly important service policy context, as highlighted throughout this dissertation, due to being intertwined with other key operational decisions in a retail environment and having a significant impact on consumer cognitive, affective, and conative responses. I hope that the research presented in this dissertation will be impactful in stimulating future empirical and analytical research on consumer return policies and other types of service policies.

REFERENCES

Abbey, J. D., & Guide Jr, V. D. R. (2018). A typology of remanufacturing in closed-loop supply chains. *International Journal of Production Research*, *56*(1-2), 374–384.

Abbey, J. D., & Meloy, M. G. (2017). Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management*, *53–56 (November)*, 63–70.

Abdulla, H., Abbey, J. D., & Ketzenberg, M. (2021). How consumers value retailer's return policy leniency levers: An empirical investigation. *Production and Operations Management*.

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

Adams, J. S., & Freedman, S. (1976). Equity theory revisited: Comments and annotated bibliography. *Advances in Experimental Social Psychology*, *9*, 43–90.

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351–371.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control* (pp. 11–39). Springer, Berlin.

Akçay, Y., Boyacı, T., & Zhang, D. (2013). Selling with money-back guarantees: The impact on prices, quantities, and retail profitability. *Production and Operations Management*, *22*(4), 777–791.

Akturk, M. S., Ketzenberg, M., & Yildiz, B. (2019). Managing consumer return abuse and an assessment of technology-enabled countermeasures. *Clemson University Working Paper*.

Alptekinoğlu, A., & Grasas, A. (2014). When to carry eccentric products? optimal retail assortment under consumer returns. *Production and Operations Management*, *23*(5), 877–892.

Altug, M. S., & Aydinliyim, T. (2016). Counteracting strategic purchase deferrals: The impact of online retailers' return policy decisions. *Manufacturing & Service Operations Management*, *18*(3), 376–392.

Anderson, E. T., Hansen, K., & Simester, D. (2009). The option value of returns: Theory and empirical evidence. *Marketing Science*, *28*(3), 405–423.

Anderson, E. T., Hansen, K., Simester, D., & Wang, L. K. (2009). How price affects returns: The perceived value and incremental customer effects. *Kellogg School of Management Working Paper*.

Andreassen, T., & Lindestad, B. (1998). Customer loyalty and complex services: The impact of corporate image on quality, customer satisfaction and loyalty for customers with varying degrees of service expertise. *International Journal of Service Industry Management*, *9*(1), 7–23.

Atasu, A., Guide Jr, V. D. R., & Van Wassenhove, L. N. (2008). Product reuse economics in closed-loop supply chain research. *Production and Operations Management*, *17*(5), 483–496.

Autry, C. W. (2005). Formalization of reverse logistics programs: a strategy for managing liberalized returns. *Industrial Marketing Management*, *34*(7), 749–757.

Bachrach, D. G., & Bendoly, E. (2011). Rigor in behavioral experiments: A basic primer for supply chain management researchers. *Journal of Supply Chain Management*, *47*(3), 5–8.

Bagozzi, R. P. (1982). A field investigation of causal relations among cognitions, affect, intentions, and behavior. *Journal of Marketing Research*, *19*(4), 562–584.

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.

Baker, J., Parasuraman, A., Grewal, D., & Voss, G. B. (2002). The influence of multiple store environment cues on perceived merchandise value and patronage intentions. *Journal of Marketing*, *66*(2), 120–141.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Batarfi, R., Jaber, M. Y., & Aljazzar, S. M. (2017). A profit maximization for a reverse logistics dual-channel supply chain with a return policy. *Computers & Industrial Engineering*, *106*, 58–82.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's mechanical turk. *Political Analysis*, *20*(3), 351–368.

Betancourt, H., & Blair, I. (1992). A cognition (attribution)-emotion model of violence in conflict situations. *Personality and Social Psychology Bulletin*, *18*(3), 343–350.

Bhattacherjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of Management Information Systems*, *19*(1), 211–241.

Bies, R. J., & Shapiro, D. L. (1987). Interactional fairness judgments: The influence of causal accounts. *Social Justice Research*, *1*(2), 199–218.

Bies, R. J., Shapiro, D. L., & Cummings, L. L. (1988). Causal accounts and managing organizational conflict: Is it enough to say it's not my fault? *Communication Research*, *15*(4), 381–399.

Blackburn, J. D., Guide Jr, V. D. R., Souza, G. C., & Van Wassenhove, L. N. (2004). Reverse supply chains for commercial returns. *California Management Review*, *46*(2), 6–22.

Blanchard, D. (2007). Supply chains also work in reverse. *IndustryWeek*.

Bolton, R. N., & Drew, J. H. (1991). A multistage model of customers' assessments of service quality and value. *Journal of Consumer Research*, *17*(4), 375–384.

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.

Boulding, W., Kalra, A., Staelin, R., & Zeithaml, V. A. (1993). A dynamic process model of sevice quality: From expectations to behavioral intentions. *Journal of Marketing Research*, *30*(1), 7–27.

Bower, A. B., & Maxham III, 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.

Boyajian, L. (2018). The real cost of returns for retailers. Retrieved from `https://retailnext.net/en/blog/the-real-cost-of-returns-for-retailers/` (Last Accessed: August 23, 2020)

Brady, M. K., & Robertson, C. J. (1999). An exploratory study of service value in the USA and Ecuador. *International Journal of Service Industry Management*, *10*(5), 469–486.

Brockner, J., DeWitt, R. L., Grover, S., & Reed, T. (1990). When it is especially important to explain why: Factors affecting the relationship between managers' explanations of a layoff and survivors' reactions to the layoff. *Journal of Experimental Social Psychology*, *26*(5), 389–407.

Buell, R. W., Kim, T., & Tsay, C.-J. (2017). Creating reciprocal value through operational transparency. *Management Science*, *63*(6), 1673–1695.

Buell, R. W., Mariadassou, S., & Zheng, Y. (2019). Relative performance transparency: Effects on sustainable choices. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*(19-079).

Buell, R. W., & Norton, M. I. (2011). The labor illusion: How operational transparency increases perceived value. *Management Science*, *57*(9), 1564–1579.

Cachon, G. P. (2003). Supply chain coordination with contracts. *Handbooks in Operations Research and Management Science*, *11*, 227–339.

Cachon, G. P., Gallino, S., & Xu, J. (2018). Free shipping is not free: A data-driven model to design free-shipping threshold policies. *SSRN*. Retrieved from `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3250971`

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin Company, Boston.

Campbell, M. C. (1999). Pricing strategy & practice "why did you do that?" the important role of inferred motive in perceptions of price fairness. *Journal of Product & Brand Management*, *8*(2).

Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2005). Do intentions really predict behavior?

self-generated validity effects in survey research. *Journal of Marketing*, *69*(2), 1–14.

Chase, R. B., & Dasu, S. (2001). Want to perfect your company's service? use behavioral science. *Harvard business review*, *79*(6), 78–84.

Che, Y.-K. (1996). Customer return policies for experience goods. *The Journal of Industrial Economics*, *44*(1), 17–24.

Chen, B., & Chen, J. (2017a). Compete in price or service? a study of personalized pricing and money back guarantees. *Journal of Retailing*, *93*(2), 154–171.

Chen, B., & Chen, J. (2017b). When to introduce an online channel, and offer money back guarantees and personalized pricing? *European Journal of Operational Research*, *257*(2), 614–624.

Chen, J. (2011). The impact of sharing customer returns information in a supply chain with and without a buyback policy. *European Journal of Operational Research*, *213*(3), 478–488.

Chen, J., & Bell, P. C. (2012). Implementing market segmentation using full-refund and no-refund customer returns policies in a dual-channel supply chain structure. *International Journal of Production Economics*, *136*(1), 56–66.

Chen, J., & Chen, B. (2016). Competing with customer returns policies. *International Journal of Production Research*, *54*(7), 2093–2107.

Chen, J., Chen, B., & Li, W. (2018). Who should be pricing leader in the presence of customer returns? *European Journal of Operational Research*, *265*(2), 735–747.

Chen, J., & Grewal, R. (2013). Competing in a supply chain via full-refund and no-refund customer returns policies. *International Journal of Production Economics*, *146*(1), 246–258.

Chen, Y.-J. (2011). Optimal selling scheme for heterogeneous consumers with uncertain valuations. *Mathematics of Operations Research*, *36*(4), 695–720.

Cheng, A. (2015). Consumers return $642.6 billion in goods each year. *MarketWatch*.

Chircu, A. M., & Mahajan, V. (2006). Managing electronic commerce retail transaction costs for customer value. *Decision Support Systems*, *42*(2), 898–914.

Choi, T.-M. (2013). Optimal return service charging policy for a fashion mass customization program. *Service Science*, *5*(1), 56–68.

Choi, T.-M., Liu, N., Ren, S., & Hui, C.-L. (2013). No refund or full refund: when should a fashion brand offer full refund consumer return service for mass customization products? *Mathematical Problems in Engineering*, *2013*.

Chu, W., Gerstner, E., & Hess, J. D. (1998). Managing dissatisfaction: How to decrease customer opportunism by partial refunds. *Journal of Service Research*, *1*(2), 140–155.

Coase, R. H. (1937). The nature of the firm. *Economica*, *4*(16), 386–405.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67.

Constable, H. (2017). Your brand new returns end up in landfill. *BBC*. Retrieved from `https://www.bbcearth.com/blog/?article=your-brand-new-returns-end-up-in-landfill` (Last Accessed: February 5, 2020)

ConsumerWorld.org. (2017). Some stores tighten the rules. Retrieved from `http://www.consumerworld.org/pages/returns.htm`

ConsumerWorld.org. (2018). Some stores tighten the rules. Retrieved from `http://www.consumerworld.org/pages/returns.htm` (Last Accessed: September 20, 2020)

Cronin, J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, *76*(2), 193–218.

Cronin, J. J., & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: Reconciling performance-based and perceptions-minus-expectations measurement of service quality. *Journal of Marketing*, *58*(1), 125–131.

Dabholkar, P. A., Thorpe, D. I., & Rentz, J. O. (1996). A measure of service quality for retail stores: Scale development and validation. *Journal of the Academy of Marketing Science*, *24*(1), 3.

Dahl, M. (2016). The weird psychology of returning stuff to the store. Retrieved from https://www.huffingtonpost.com/science-of-us/the-weird-psychology-of-r_b_9105886.html (Last Accessed: August 20, 2018)

d'Astous, A., & Guèvremont, A. (2008). Effects of retailer post-purchase guarantee policies on consumer perceptions with the moderating influence of financial risk and product complexity. *Journal of Retailing and Consumer Services*, *15*(4), 306–314.

Davis, S., Gerstner, E., & Hagerty, M. (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". *Journal of Economics and Business*, *50*(5), 445–460.

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.

Dennis, S. (2018). The Ticking Time Bomb Of E-commerce Returns. *Forbes*. Retrieved from https://www.forbes.com/sites/stevendennis/2018/02/14/the-ticking-time-bomb-of-e-commerce-returns/#3e8a6da84c7f (Last Accessed: December 15, 2018)

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

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 307–319.

Donohue, K., Katok, E., & Leider, S. (2018). *The handbook of behavioral operations*. John Wiley & Sons, Hoboken, NJ.

Durach, C. F., Kembro, J., & Wieland, A. (2017). A new paradigm for systematic literature reviews in supply chain management. *Journal of Supply Chain Management*, *53*(4), 67–85.

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, *67*(2), 261–275.

Eckerd, S., Hill, J., Boyer, K. K., Donohue, K., & Ward, P. T. (2013). The relative impact of attribute, severity, and timing of psychological contract breach on behavioral and attitudinal outcomes. *Journal of Operations Management*, *31*(7-8), 567–578.

Enright, T. (2013). *Get a grip on the hidden cost of retail returns by creating visibility and ownership* (Tech. Rep.). Gartner Group. Retrieved from `https://www.gartner.com/doc/2998121/grip-hidden-cost-retail-returns` (Last Accessed: August 20, 2018)

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*(2), 168–192.

Ertekin, N. (2018). Immediate and long-term benefits of in-store return experience. *Production and Operations Management*, *27*(1), 121–142.

Ertekin, N., & Agrawal, A. (2021). How does a return period policy change affect multichannel retailer profitability? *Manufacturing & Service Operations Management*, *23*(1), 210–229.

Etzel, M. J., & Silverman, B. I. (1981). A managerial perspective on directions for retail customer dissatisfaction research. *Journal of Retailing*, *57*(3), 124–136.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272.

Ferguson, M., Guide Jr, V. D. R., & Souza, G. C. (2006). Supply chain coordination for false failure returns. *Manufacturing & Service Operations Management*, *8*(4), 376–393.

Festinger, L. (1957). A theory of cognitive dissonance evanston. *IL: Row, Peterson*, *1*.

Folger, R. G., & Cropanzano, R. (1998). *Organizational justice and human resource management* (Vol. 7). Sage, Thousand Oaks CA.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Foscht, T., Ernstreiter, K., Maloles III, C., Sinha, I., & Swoboda, B. (2013). Retaining or returning? some insights for a better understanding of return behaviour. *International Journal of Retail & Distribution Management*, *41*(2), 113–134.

Fruchter, G. E., & Gerstner, E. (1999). Selling with "satisfaction guaranteed". *Journal of Service Research*, *1*(4), 313–323.

Gallino, S., & Moreno, A. (2018). The value of fit information in online retail: Evidence from a randomized field experiment. *Manufacturing & Service Operations Management*, *20*(4), 767–787.

Gelbrich, K., Gäthke, J., & Hübner, A. (2017). Rewarding customers who keep a product: How reinforcement affects customers' product return decision in online retailing. *Psychology & Marketing*, *34*(9), 853–867.

Gintzler, A. (2018). L.L.Bean sued (again) over return-policy change. *Outside*. Retrieved from `https://www.outsideonline.com/outdoor-gear/gear-news/ll-bean-sued-again-over-return-policy-change/`

Giordano, P. C. (1983). Sanctioning the high-status deviant: An attributional analysis. *Social Psychology Quarterly*, 329–342.

Gooding, S. K. S. (1995). Quality, sacrifice, and value in hospital choice. *Marketing Health Services*, *15*(4), 24–31.

Goodman, J. K., & Paolacci, G. (2017). Crowdsourcing consumer research. *Journal of Consumer Research*, *44*(1), 196–210.

Govindan, K., Soleimani, H., & Kannan, D. (2015). Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future. *European Journal of Operational Research*, *240*(3), 603–626.

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.

Grønhaug, K., & Gilly, M. C. (1991). A transaction cost approach to consumer dissatisfaction and complaint actions. *Journal of Economic Psychology*, *12*(1), 165–183.

Guide, V. D. R., & Van Wassenhove, L. N. (2009). OR FORUM—the evolution of closed-loop supply chain research. *Operations Research*, *57*(1), 10–18.

Hawes, J. M., & Lumpkin, J. R. (1986). Perceived risk and the selection of a retail patronage mode. *Journal of the Academy of Marketing Science*, *14*(4), 37–42.

Hayes, A. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press, New York.

Heal, G. (1977). Guarantees and risk-sharing. *The Review of Economic Studies*, 549–560.

Heim, G. R., & Field, J. M. (2007). Process drivers of e-service quality: Analysis of data from an online rating site. *Journal of Operations Management*, *25*(5), 962–984.

Heiman, A., Just, D. R., McWilliams, B. P., & Zilberman, D. (2015). A prospect theory approach to assessing changes in parameters of insurance contracts with an application to money-back guarantees. *Journal of Behavioral and Experimental Economics*, *54*, 105–117.

Heiman, A., McWilliams, B., Zhao, J., & Zilberman, D. (2002). Valuation and management of money-back guarantee options. *Journal of Retailing*, *78*(3), 193–205.

Heiman, A., McWilliams, B., & Zilberman, D. (2001). Demonstrations and money-back guarantees: Market mechanisms to reduce uncertainty. *Journal of Business Research*, *54*(1), 71–84.

Helsen, K., & Schmittlein, D. C. (1993). Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models. *Marketing Science*, *12*(4), 395–414.

Hess, J. D., Chu, W., & Gerstner, E. (1996). Controlling product returns in direct marketing. *Marketing Letters*, *7*(4), 307–317.

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

Heydari, J., Choi, T.-M., & Radkhah, S. (2017). Pareto improving supply chain coordination under a money-back guarantee service program. *Service Science*, *9*(2), 91–105.

Hill, J. A., Eckerd, S., Wilson, D., & Greer, B. (2009). The effect of unethical behavior on trust in a buyer–supplier relationship: The mediating role of psychological contract violation. *Journal of Operations Management*, *27*(4), 281–293.

Hjort, K., & Lantz, B. (2016). The impact of returns policies on profitability: A fashion e-commerce case. *Journal of Business Research*, *69*(11), 4980–4985.

Hjort, K., Lantz, B., Ericsson, D., & Gattorna, J. (2013). Customer segmentation based on buying and returning behaviour. *International Journal of Physical Distribution & Logistics Management*, *43*(10), 852–865.

Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. John Wiley & Sons, New York.

Howland, D. (2017). Study: 5B pounds of retail returns end up in landfills. *Retail Dive*. Retrieved from `https://www.retaildive.com/news/study-5b-pounds-of-retail-returns-end-up-in-landfills/513769/` (Last Accessed: February 5, 2020)

Hsiao, L., & Chen, Y.-J. (2012). Returns policy and quality risk in e-business. *Production and Operations Management*, *21*(3), 489–503.

Hsiao, L., & Chen, Y.-J. (2014). Return policy: Hassle-free or your money-back guarantee? *Naval Research Logistics (NRL)*, *61*(5), 403–417.

Hsiao, L., & Chen, Y.-J. (2015). Retailer's rationale to refuse consumer returns in supply chains. *Naval Research Logistics (NRL)*, *62*(8), 686–701.

Hsieh, P.-L. (2013). Perceived opportunism (po) in e-return service encounters. *Managing Service Quality: An International Journal*, *23*(2), 96–110.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Hu, W., Li, Y., & Govindan, K. (2014). The impact of consumer returns policies on consignment contracts with inventory control. *European Journal of Operational Research*, *233*(2), 398–407.

Huang, X., Gu, J.-W., Ching, W.-K., & Siu, T.-K. (2014). Impact of secondary market on consumer return policies and supply chain coordination. *Omega*, *45*, 57–70.

Inderst, R., & Tirosh, G. (2015). Refunds and returns in a vertically differentiated industry.

International Journal of Industrial Organization, *38*, 44–51.

Inman, J. J., & Zeelenberg, M. (2002). Regret in repeat purchase versus switching decisions: The attenuating role of decision justifiability. *Journal of Consumer Research*, *29*(1), 116–128.

Jalil, M. N., & Shahzad, S. (2013). Risk pooling in commercial returns by managing product return period. *SSRN*. Retrieved from `https://ssrn.com/abstract=2345622`

Janakiraman, N., & Ordóñez, L. (2012). Effect of effort and deadlines on consumer product returns. *Journal of Consumer Psychology*, *22*(2), 260–271.

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.

Jeng, S.-P. (2017). Increasing customer purchase intention through product return policies: the pivotal impacts of retailer brand familiarity and product categories. *Journal of Retailing and Consumer Services*, *39*(November), 182–189.

Jet.com. (2019). Return Policy. Retrieved from `https://jet.com/help-center/returns` (Last Accessed: December 15, 2018)

Johnson, E. J., Bellman, S., & Lohse, G. L. (2003). Cognitive lock-in and the power law of practice. *Journal of Marketing*, *67*(2), 62–75.

Johnson, T. E., & Rule, B. G. (1986). Mitigating circumstance information, censure, and aggression. *Journal of Personality and Social Psychology*, *50*(3), 537.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business*, *59*(4), 285–300.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, *246*(1), 160–173.

Kane, J. V., & Barabas, J. (2019). No harm in checking: Using factual manipulation checks to assess attentiveness in experiments. *American Journal of Political Science*, *63*(1), 234–249.

Karp, G. (2017). Discover will cut some card benefits in 2018. *Nerd Wallet*. Retrieved from `https://www.nerdwallet.com/article/credit-cards/discover-to-cut-some-card-benefits-2018` (Last Accessed: February 4, 2022)

Kelley, H. H. (1967). Attribution theory in social psychology. In *Nebraska symposium on motivation.*

Ketzenberg, M. E., Abbey, J., & Heim, G. (2018). Assessing return behavior through data analytics. *Texas A & M University Working Paper*.

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.

Ketzenberg, M. E., & Zuidwijk, R. A. (2009). Optimal pricing, ordering, and return policies for consumer goods. *Production and Operations Management*, *18*(3), 344–360.

Kim, J., & Wansink, B. (2012). How retailers' recommendation and return policies alter product evaluations. *Journal of Retailing*, *88*(4), 528–541.

Kirkham, E. (2015). 10 stores with the best return policies – and 5 with the worst. *Money*. Retrieved from `https://money.com/return-policies-best-worst/` (Last Accessed: December 25, 2019)

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.

Konovsky, M. A., & Folger, R. (1991). The effects of procedures, social accounts, and benefits level on victims' layoff reactions. *Journal of Applied Social Psychology*, *21*(8), 630–650.

Kuo, Y.-F., Wu, C.-M., & Deng, W.-J. (2009). The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services. *Computers in Human Behavior*, *25*(4), 887–896.

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

Lee, C. H., & Rhee, B.-D. (2018). Retailer-run resale market and optimal returns and resale policy. *European Journal of Operational Research*, *268*(2), 504–514.

Lee, Y. S., Seo, Y. W., & Siemsen, E. (2018). Running behavioral operations experiments using Amazon's Mechanical Turk. *Production and Operations Management*, *27*(5), 973–989.

Letizia, P., Pourakbar, M., & Harrison, T. (2018). The impact of consumer returns on the multi-channel sales strategies of manufacturers. *Production and Operations Management*, *27*(2), 323–349.

Li, G., Li, L., Sethi, S. P., & Guan, X. (2017). Return strategy and pricing in a dual-channel supply chain. *International Journal of Production Economics*.

Li, W., Chen, J., Liang, G., & Chen, B. (2018). Money-back guarantee and personalized pricing in a stackelberg manufacturer's dual-channel supply chain. *International Journal of Production Economics*, *197*, 84–98.

Li, Y., Xu, L., & Li, D. (2013). Examining relationships between the return policy, product quality, and pricing strategy in online direct selling. *International Journal of Production Economics*, *144*(2), 451–460.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*(1), 5.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. Springer Science & Business Media.

Lind, E. A., Walker, L., Kurtz, S., Musante, L., & Thibaut, J. W. (1980). Procedure and outcome effects on reactions to adjudicated resolution of conflicts of interest. *Journal of Personality and Social Psychology*, *39*(4), 643.

Liu, J., Mantin, B., & Wang, H. (2014). Supply chain coordination with customer returns and refund-dependent demand. *International Journal of Production Economics*, *148*, 81–89.

Liu, N., Choi, T.-M., Yuen, C.-W. M., & Ng, F. (2012). Optimal pricing, modularity, and return policy under mass customization. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *42*(3), 604–614.

Liu, Q., & Xiao, W. (2008). Selling to heterogeneous customers with uncertain valuations under returns policies. *SSRN*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1286226

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in mis and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293–334.

Malhotra, M. K., Singhal, C., Shang, G., & Ployhart, R. E. (2014). A critical evaluation of alternative methods and paradigms for conducting mediation analysis in operations management research. *Journal of Operations Management*, *32*(4), 127–137.

Malhotra, N., Sahadev, S., & Purani, K. (2017). Psychological contract violation and customer intention to reuse online retailers: Exploring mediating and moderating mechanisms. *Journal of Business Research*, *75*(June), 17–28.

Mann, D. P., & Wissink, J. P. (1988). Money-back contracts with double moral hazard. *The RAND Journal of Economics*, *19*(2), 285–292.

Mash, J. K. (2017). These 10 stores have the best return policies. *NBC Today*. Retrieved from https://www.today.com/style/stores-best-return-policies-l-l-bean-zappos-athleta-nordstrom-t106819 (Last Accessed: December 27, 2018)

Matthews, S., & Persico, N. (2007). Information acquisition and refunds for returns. *PIER Working Paper No. 07-021*.

McCollough, M. A., & Bharadwaj, S. G. (1992). The recovery paradox: an examination of consumer satisfaction in relation to disconfirmation, service quality, and attribution based theories. In C. T. Allen (Ed.), *Marketing theory and applications* (Vol. 119). Chicago: American Marketing Association.

McWilliams, B. (2012). Money-back guarantees: Helping the low-quality retailer. *Management*

Science, *58*(8), 1521–1524.

Miller, D. T., & Vidmar, N. (1981). The social psychology of punishment reactions. In *The justice motive in social behavior* (pp. 145–172). Springer, Boston.

Mir, S., Aloysius, J. A., & Eckerd, S. (2017). Understanding supplier switching behavior: The role of psychological contracts in a competitive setting. *Journal of Supply Chain Management*, *53*(3), 3–18.

Mixon, F. G. (1999). Customer return policies for experience goods: the impact of product price and consumer search costs on seller-provided informational cues. *Applied Economics*, *31*(3), 331–336.

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.

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.

Moorthy, S., & Srinivasan, K. (1995). Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science*, *14*(4), 442–466.

Morrison, E. W., & Robinson, S. L. (1997). When employees feel betrayed: A model of how psychological contract violation develops. *Academy of Management Review*, *22*(1), 226–256.

Mostard, J., & Teunter, R. (2006). The newsboy problem with resalable returns: A single period model and case study. *European Journal of Operational Research*, *169*(1), 81–96.

Mukhopadhyay, S. K., & Setaputra, R. (2004). Reverse logistics in e-business: optimal price and return policy. *International Journal of Physical Distribution & Logistics Management*, *34*(1), 70–89.

Mukhopadhyay, S. K., & Setaputra, R. (2005). Optimal return policy and modular design for build-to-order products. *Journal of Operations Management*, *23*(5), 496–506.

Mukhopadhyay, S. K., & Setaputra, R. (2007). A dynamic model for optimal design quality and return policies. *European Journal of Operational Research*, *180*(3), 1144–1154.

Narvar. (2017). Narvar Consumer Report. Retrieved from https://see.narvar.com/rs/249-TEC-877/images/Narvar_Consumer_Survey_Returns_June2017.pdf (Last Accessed: December 15, 2018)

Narvar. (2021). Returns Benchmarks Report. Retrieved from https://corp.narvar.com/returns-benchmarks-report-2021 (Last Accessed: September 21, 2021)

National Retail Federation. (2017). 2017 consumer returns in the retail industry. Retrieved from https://appriss.com/retail/wp-content/uploads/sites/4/2017/12/2017_Consumer-Returns-in-the-Retail-Industry-Report.pdf (Last Accessed: December 20, 2018)

National Retail Federation. (2018). 2017 consumer returns in the retail industry. Retrieved from https://appriss.com/retail/wp-content/uploads/sites/4/2018/12/AR3018_2018-Customer-Returns-in-the-Retail-Industry_Digital.pdf (Last Accessed: December 28, 2018)

National Retail Federation. (2019). Top 100 retailers. Retrieved from https://nrf.com/resources/top-retailers/top-100-retailers/top-100-retailers-2019

National Retail Federation. (2021). $428 billion in merchandise returned in 2020. Retrieved from https://nrf.com/media-center/press-releases/428-billion-merchandise-returned-2020 (Last Accessed: March 2, 2021)

Niehoff, B. P., & Paul, R. J. (2001). The just workplace: Developing and maintaining effective psychological contracts. *Review of Business*, *22*(1/2).

Nunnally, J. C., & Bernstein, I. (1994). *Psychometric Theory (McGraw-Hill Series in Psychology)* (Vol. 3). McGraw-Hill, New York.

Ofek, E., Katona, Z., & Sarvary, M. (2011). “bricks and clicks”: The impact of product returns on the strategies of multichannel retailers. *Marketing Science*, *30*(1), 42–60.

Oghazi, P., Karlsson, S., Hellström, D., & Hjort, K. (2018). Online purchase return policy leniency and purchase decision: Mediating role of consumer trust. *Journal of Retailing and Consumer Services*, *41*(March), 190–200.

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management*, *16*(4), 387–405.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872.

Optoro. (2018). 2018 optoro impact report. Retrieved from `https://www.optoro.com/2018-impact-report/` (Last Accessed: August 20, 2020)

Pagell, M. (2020). Replication without repeating ourselves: Addressing the replication crisis in operations and supply chain management research. *Journal of Operations Management*, *67*(1), 105–115.

Paolacci, G., & Chandler, J. (2014). Inside the turk: Understanding mechanical turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188.

Parasuraman, A. (1997). Reflections on gaining competitive advantage through customer value. *Journal of the Academy of Marketing Science*, *25*(2), 154.

Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991). Understanding customer expectations of service. *MIT Sloan Management Review*, *32*(3), 39–48.

Parasuraman, A., & Grewal, D. (2000). Serving customers and consumers effectively in the twenty-first century: A conceptual framework and overview. *Journal of the Academy of Marketing Science*, *28*(1), 9–16.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, *49*(4), 41–50.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12.

Pavlou, P. A., & Gefen, D. (2005). Psychological contract violation in online marketplaces:

Antecedents, consequences, and moderating role. *Information Systems Research*, *16*(4), 372–399.

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031.

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.

Petersen, J. A., & Kumar, V. (2009). Are product returns a necessary evil? Antecedents and consequences. *Journal of Marketing*, *73*(3), 35–51.

Petersen, J. A., & Kumar, V. (2010). Can product returns make you money? *MIT Sloan Management Review*, *51*(3), 85.

Picchi, A. (2018). Amazon bans customers with too many returns. *CBS News*. Retrieved from `https://www.cbsnews.com/news/amazon-bans-customers-with-too-many-returns/` (Last Accessed: December 15, 2018)

Posselt, T., Gerstner, E., & Radic, D. (2008). Rating e-tailers' money-back guarantees. *Journal of Service Research*, *10*(3), 207–219.

Powers, T. L., & Jack, E. P. (2013). The influence of cognitive dissonance on retail product returns. *Psychology & Marketing*, *30*(8), 724–735.

Price, E. (2018). L.L. Bean sued for ditching its lifetime return policy. *Fortune*. Retrieved from `http://fortune.com/2018/02/14/l-l-bean-sued-lifetime-return-policy/`

Rao, S., Lee, K. B., Connelly, B., & Iyengar, D. (2017). Return time leniency in online retail: A signaling theory perspective on buying outcomes. *Decision Sciences*.

Reinholtz, N., Bartels, D. M., & Parker, J. R. (2015). On the mental accounting of restricted-use funds: How gift cards change what people purchase. *Journal of Consumer Research*, *42*(4), 596–614.

Robinson, S. L. (1996). Trust and breach of the psychological contract. *Administrative Science*

Quarterly, 574–599.

Robinson, S. L., & Rousseau, D. M. (1994). Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior*, *15*(3), 245–259.

Rosato, D. (2019). Retailers with the best and worst policies. *Consumer Reports*. Retrieved from https://www.consumerreports.org/returns-refunds-exchanges/guide-to-returning-gifts (Last Accessed: November 10, 2020)

Rousseau, D. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements*. Sage publications.

Rungtusanatham, M., Miller, J., & Boyer, K. K. (2014). Theorizing, testing, and concluding for mediation in SCM research: Tutorial and procedural recommendations. *Journal of Operations Management*, *32*(3), 99–113.

Rungtusanatham, M., Wallin, C., & Eckerd, S. (2011). The vignette in a scenario-based role-playing experiment. *Journal of Supply Chain Management*, *47*(3), 9–16.

Safdar, K. (2018). How your returns are used against you at Best Buy, other retailers. *Wall Street Journal*. Retrieved from https://www.wsj.com/articles/how-your-returns-are-used-against-you-at-best-buy-other-retailers-1520933400

Samatli-Pac, G., Shen, W., & Hu, X. (2018). The impact of consumer loss aversion on returns policies and supply chain coordination. *International Journal of Operations Research and Information Systems (IJORIS)*, *9*(4), 1–20.

Samorani, M., Alptekinoglu, A., & Messinger, P. R. (2016). The analytics of product return episodes in retailing. *SSRN*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2813685

Seo, J. Y., Yoon, S., & Vangelova, M. (2016). Shopping plans, buying motivations, and return policies: impacts on product returns and purchase likelihoods. *Marketing Letters*, *27*(4), 645–659.

Shang, G., Ferguson, M. E., & Galbreth, M. R. (2018). Where should I focus my return reduction efforts? Empirical guidance for retailers. *Decision Sciences*, *50*(4), 877–909.

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.

Shang, G., Ghosh, B. P., & Galbreth, M. R. (2017). Optimal retail return policies with wardrobing. *Production and Operations Management*, *26*(7), 1315–1332.

Shang, G., Pekgün, P., Ferguson, M., & Galbreth, M. (2017). How much do online consumers really value free product returns? evidence from ebay. *Journal of Operations Management*, *53–56*(November), 45–62.

Shapiro, D. L. (1991). The effects of explanations on negative reactions to deceit. *Administrative Science Quarterly*, *36*(4), 614–630.

Shapiro, D. L., & Buttner, E. H. (1988). Adequate explanations: What are they, and do they enhance procedural justice under severe outcome circumstances. In *Annual meeting of the Academy of Management, Anaheim, CA.*

Shaw, J. C., Wild, E., & Colquitt, J. A. (2003). To justify or excuse?: A meta-analytic review of the effects of explanations. *Journal of Applied Psychology*, *88*(3), 444.

Shieh, S. (1996). Price and money-back guarantees as signals of product quality. *Journal of Economics & Management Strategy*, *5*(3), 361–377.

Shulman, J. D., Coughlan, A. T., & Savaskan, R. C. (2009). Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *Manufacturing & Service Operations Management*, *11*(4), 577–594.

Shulman, J. D., Coughlan, A. T., & Savaskan, R. C. (2010). Optimal reverse channel structure for consumer product returns. *Marketing Science*, *29*(6), 1071–1085.

Shulman, J. D., Coughlan, A. T., & Savaskan, R. C. (2011). Managing consumer returns in a competitive environment. *Management Science*, *57*(2), 347–362.

Simpson, D., Power, D., Riach, K., & Tsarenko, Y. (2019). Consumer motivation for product disposal and its role in acquiring products for reuse. *Journal of Operations Management*, *65*(7), 612–635.

Singhal, K., & Singhal, J. (2012). Imperatives of the science of operations and supply-chain management. *Journal of Operations Management*, *30*(3), 237–244.

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin*, *46*(2), 137.

Souza, G. C. (2013). Closed-loop supply chains: A critical review, and future research. *Decision Sciences*, *44*(1), 7–38.

Spector, P. E. (1981). *Research designs* (Vol. 23). Sage Publications Inc., Newbury Park, CA.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Spreng, R. A., Dixon, A. L., & Olshavsky, R. W. (1993). The impact of perceived value on consumer satisfaction. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, *6*(1), 50–55.

Su, X. (2009a). Consumer returns policies and supply chain performance. *Manufacturing & Service Operations Management*, *11*(4), 595–612.

Su, X. (2009b). A model of consumer inertia with applications to dynamic pricing. *Production and Operations Management*, *18*(4), 365–380.

Suwelack, T., Hogreve, J., & Hoyer, W. D. (2011). Understanding money-back guarantees: Cognitive, affective, and behavioral outcomes. *Journal of Retailing*, *87*(4), 462–478.

Swinney, R. (2011). Selling to strategic consumers when product value is uncertain: The value of matching supply and demand. *Management Science*, *57*(10), 1737–1751.

Teas, R. K., & Agarwal, S. (2000). The effects of extrinsic product cues on consumers' perceptions of quality, sacrifice, and value. *Journal of the Academy of Marketing Science*, *28*(2), 278–290.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*(3), 199–214.

Ülkü, M. A., Dailey, L. C., & Yayla-Küllü, H. M. (2013). Serving fraudulent consumers? the impact of return policies on retailer's profitability. *Service Science*, *5*(4), 296–309.

Ülkü, M. A., & Gürler, Ü. (2018). The impact of abusing return policies: A newsvendor model with opportunistic consumers. *International Journal of Production Economics*, *203*(September), 124–133.

UPS. (2017). UPS pulse of the online shopper. Retrieved from `https://pressroom.ups.com/mobile0c9a66/assets/pdf/pressroom/white\%20paper/UPS_2017_POTOS_media\%20executive\%20summary_FINAL.pdf` (Last Accessed: August 20, 2020)

Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest–posttest control group design. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 428–450.

Van den Poel, D., & Leunis, J. (1999). Consumer acceptance of the internet as a channel of distribution. *Journal of Business Research*, *45*(3), 249–256.

Vlachos, D., & Dekker, R. (2003). Return handling options and order quantities for single period products. *European Journal of Operational Research*, *151*(1), 38–52.

Wang, X. (2009). Retail return policy, endowment effect, and consumption propensity: an experimental study. *The BE Journal of Economic Analysis & Policy*, *9*(1).

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford Press, New York.

Welling, L. A. (1989). Satisfaction guaranteed or money (partially) refunded: Efficient refunds under asymmetric information. *Canadian Journal of Economics*, *22*(1), 62–78.

Williamson, O. E. (1989). Transaction cost economics. In *Handbook of industrial organization* (Vol. 1, pp. 135–182). Elsevier.

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.

Woodall, T. (2003). Conceptualising 'value for the customer': An attributional, structural and dispositional analysis. *Academy of Marketing Science Review*, *12*(1), 1–42.

Xiao, T., Shi, K., & Yang, D. (2010). Coordination of a supply chain with consumer return under demand uncertainty. *International Journal of Production Economics*, *124*(1), 171–180.

Xu, L., Li, Y., Govindan, K., & Xu, X. (2015). Consumer returns policies with endogenous deadline and supply chain coordination. *European Journal of Operational Research*, *242*(1), 88–99.

Xu, L., Li, Y., Govindan, K., & Yue, X. (2018). Return policy and supply chain coordination with network-externality effect. *International Journal of Production Research*, *56*(10), 1–19.

Yalabik, B., Petruzzi, N. C., & Chhajed, D. (2005). An integrated product returns model with logistics and marketing coordination. *European Journal of Operational Research*, *161*(1), 162–182.

Yang, H., Chen, J., Chen, X., & Chen, B. (2017). The impact of customer returns in a supply chain with a common retailer. *European Journal of Operational Research*, *256*(1), 139–150.

Yoo, S. H. (2014). Product quality and return policy in a supply chain under risk aversion of a supplier. *International Journal of Production Economics*, *154*, 146–155.

Yoo, S. H., Kim, D., & Park, M.-S. (2015). Pricing and return policy under various supply contracts in a closed-loop supply chain. *International Journal of Production Research*, *53*(1), 106–126.

Zauberman, G. (2003). The intertemporal dynamics of consumer lock-in. *Journal of Consumer Research*, *30*(3), 405–419.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing*, *52*(3), 2–22.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, *60*(2), 31–46.

Zhang, J. (2013). Revenue maximizing with return policy when buyers have uncertain valuations. *International Journal of Industrial Organization*, *31*(5), 452–461.

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.

Zhou, W., & Hinz, O. (2016). Determining profit-optimizing return policies–a two-step approach on data from taobao. com. *Electronic Markets*, *26*(2), 103–114.

APPENDIX A

APPENDIX TO TAKING STOCK OF CONSUMER RETURNS: A REVIEW AND CLASSIFICATION OF THE LITERATURE

Table A.1: Classification of the Analytical Literature - Part A

Article	CB							RP							PE							RM		
	PV	CH	TC	TP	RT	DR	RA	G	M	T	F	S	X	O	PR	I	SC	FD	C	PC	PA	A	P	D
Akçay et al. (2013)	✓	✓							✓				✓		✓	✓								✓
Alptekinoğlu and Grasas (2014)	✓		✓					✓	✓												✓			✓
Altug and Aydinliyim (2016)	✓		✓						✓						✓	✓			✓					✓
Anderson, Hansen, and Simester (2009)	✓	✓	✓					✓							✓					✓				
Batarfi et al. (2017)						✓			✓						✓	✓		✓					✓	✓
Che (1996)	✓						✓	✓							✓									
B. Chen and Chen (2017a)	✓		✓					✓							✓				✓				✓	✓
B. Chen and Chen (2017b)	✓		✓					✓							✓			✓					✓	✓
J. Chen and Bell (2012)	✓							✓							✓	✓						✓		✓
J. Chen and Chen (2016)	✓												✓		✓				✓				✓	✓
J. Chen and Grewal (2013)	✓							✓							✓				✓					✓
J. Chen et al. (2018)	✓		✓					✓							✓				✓				✓	✓
Y.-J. Chen (2011)	✓	✓	✓						✓						✓									✓
Choi (2013)						✓			✓						✓					✓				✓
Choi et al. (2013)						✓		✓							✓					✓				
Chu et al. (1998)	✓		✓	✓					✓						✓									✓
Davis et al. (1995)	✓		✓	✓				✓							✓									✓
Davis et al. (1998)	✓		✓							✓					✓									✓
Difrancesco et al. (2018)					✓	✓			✓		✓											✓	✓	✓
Fruchter and Gerstner (1999)	✓	✓	✓					✓	✓						✓				✓					✓
Heal (1977)									✓						✓									
Heiman et al. (2002)			✓		✓				✓	✓					✓									
Hess et al. (1996)	✓		✓	✓					✓						✓									✓
Heydari et al. (2017)						✓		✓							✓	✓	✓						✓	✓
Hsiao and Chen (2012)	✓	✓	✓					✓	✓						✓					✓				✓
Hsiao and Chen (2014)	✓	✓	✓	✓				✓			✓				✓									✓
Hsiao and Chen (2015)	✓	✓		✓					✓						✓					✓		✓		✓
W. Hu et al. (2014)	✓					✓			✓						✓	✓	✓							✓
Huang et al. (2014)	✓								✓						✓	✓	✓							✓
Inderst and Tirosh (2015)	✓								✓						✓				✓					✓
Jalil and Shahzad (2013)					✓	✓				✓														
Ketzenberg and Zuidwijk (2009)	✓					✓							✓		✓	✓							✓	✓
C. H. Lee and Rhee (2018)	✓			✓					✓		✓				✓	✓						✓		✓

Table A.2: Classification of the Analytical Literature - Part B

Article	CB							RP							PE							RM		
	PV	CH	TC	TP	RT	DR	RA	G	M	T	F	S	X	O	PR	I	SC	FD	C	PC	PA	A	P	D
Letizia et al. (2018)	✓								✓						✓			✓						✓
Y. Li et al. (2013)					✓				✓						✓					✓				
G. Li et al. (2017)					✓			✓							✓			✓						✓
W. Li et al. (2018)	✓		✓					✓							✓			✓					✓	✓
Q. Liu and Xiao (2008)	✓	✓							✓						✓	✓								
N. Liu et al. (2012)					✓				✓						✓					✓				
J. Liu et al. (2014)	✓					✓			✓							✓	✓						✓	✓
Mann and Wissink (1988)	✓								✓						✓									✓
Matthews and Persico (2007)	✓		✓						✓						✓									✓
McWilliams (2012)	✓		✓					✓							✓				✓				✓	✓
Moorthy and Srinivasan (1995)	✓	✓	✓					✓							✓								✓	
Mukhopadhyay and Setaputra (2004)					✓				✓						✓									
Mukhopadhyay and Setaputra (2005)					✓				✓						✓					✓				
Mukhopadhyay and Setaputra (2007)					✓				✓						✓					✓				
Ofek et al. (2011)																								
Samatli-Pac et al. (2018)	✓					✓			✓						✓	✓	✓					✓	✓	✓
Shang, Ghosh, and Galbreth (2017)	✓	✓		✓					✓						✓									✓
Shieh (1996)	✓							✓	✓						✓									
Shulman et al. (2009)	✓	✓	✓						✓				✓		✓									✓
Shulman et al. (2010)	✓	✓							✓				✓		✓							✓		✓
Shulman et al. (2011)	✓	✓							✓				✓		✓	✓			✓					
Su (2009a)	✓	✓	✓					✓	✓						✓	✓	✓					✓	✓	✓
Swinney (2011)	✓		✓						✓														✓	
Ülkü et al. (2013)	✓		✓	✓						✓					✓									✓
Ülkü and Gürler (2018)	✓	✓		✓						✓					✓	✓								✓
Welling (1989)	✓					✓			✓						✓					✓				
Xiao et al. (2010)			✓			✓			✓						✓	✓	✓						✓	✓
Xu et al. (2015)	✓				✓				✓	✓					✓	✓	✓							✓
Xu et al. (2018)	✓								✓						✓	✓	✓							✓
Yalabik et al. (2005)	✓		✓	✓					✓						✓								✓	✓
Yang et al. (2017)	✓		✓					✓				✓			✓						✓			✓
Yoo (2014)	✓					✓			✓											✓				✓
Yoo et al. (2015)					✓				✓						✓		✓							✓
Zhang (2013)	✓		✓						✓						✓									✓

Table A.3: Classification of the Empirical Literature

Article	CB				RP							PE							RM		
	CR	AR	BI	BA	G	M	T	F	S	X	O	PR	I	SC	FD	C	PC	PA	A	P	D
Anderson, Hansen, and Simester (2009)			✓		✓							✓					✓				
Autry (2005)											✓									✓	
Bahn and Boyd (2014)		✓					✓											✓			
Bonifield et al. (2010)	✓		✓	✓							✓						✓				
Bower and Maxham III (2012)	✓	✓		✓		✓															
d'Astous and Guèvremont (2008)	✓						✓			✓											
Davis et al. (1998)											✓	✓						✓			✓
Gelbrich et al. (2017)			✓			✓															
Hawes and Lumpkin (1986)	✓				✓																
Heim and Field (2007)	✓					✓	✓		✓	✓											
Heiman et al. (2001)	✓				✓												✓				
Heiman et al. (2015)			✓				✓			✓											
Hjort and Lantz (2016)				✓		✓															
Hsieh (2013)	✓		✓								✓										
Janakiraman and Ordóñez (2012)	✓			✓			✓		✓												
Janakiraman et al. (2016)			✓	✓		✓	✓		✓	✓	✓				✓		✓	✓			
Jeng (2017)	✓		✓							✓											
Kim and Wansink (2012)	✓								✓												
Lantz and Hjort (2013)			✓			✓															
Mixon (1999)					✓							✓									
Mollenkopf et al. (2007)	✓		✓					✓													
Oghazi et al. (2018)	✓		✓							✓											
Pei et al. (2014)	✓		✓			✓															
Petersen and Kumar (2010)				✓						✓											
Posselt et al. (2008)		✓				✓	✓					✓									
Powers and Jack (2013)	✓	✓		✓						✓											
Rao et al. (2017)				✓			✓														
Seo et al. (2016)			✓		✓																
Shang, Pekgün, et al. (2017)				✓	✓											✓					✓
Suwelack et al. (2011)	✓	✓	✓				✓	✓									✓				
Van den Poel and Leunis (1999)			✓		✓							✓									
Wang (2009)				✓	✓		✓														
Wood (2001)	✓			✓		✓			✓												
Zhang et al. (2017)	✓		✓			✓	✓														
Zhou and Hinz (2016)				✓						✓											

APPENDIX B

APPENDIX TO HOW CONSUMERS VALUE RETAILER'S RETURN POLICY LENIENCY LEVERS: AN EMPIRICAL INVESTIGATION

B.1 Details for Pre-studies 1-5

B.1.1 Overview

We conducted a series of pre-studies that provided significant learning along the way toward designing a capstone main study to test our research hypotheses. Conducting a vignette-based experiment requires managing numerous tradeoffs, such as between maintaining the control and focus of the vignette in order to ensure internal validity of the results and increasing the degree of realism and immersion by creating a realistic scenario that enhances ecological and external validity (Rungtusanatham et al., 2011; Aguinis & Bradley, 2014). The learning from pre-studies helped us to better understand the design choices with respect to this and several other tradeoffs.

Our main research objective was to conduct the first comparative empirical assessment of all five practically-relevant leniency levers identified in the literature, predicated on a unified theoretical framework with multiple cognitive constructs. Thus, we also needed to 1) have a general idea of the expected effect sizes to understand the sample size and statistical power requirements, 2) identify relevant participant characteristics to select as covariates in the mediation analysis to reduce the residual variance (i.e., increase statistical power), 3) ensure the robustness of the scales and constructs in generating systematic variance as a result of manipulating different levers, and 4) lay a foundation for a vignette design whose scope can lend a broadly applicable context while providing internally valid experimental evidence (more on this in Appendix B below). Thus, the five pre-studies provided important conceptual knowledge and preliminary empirical insights that provided "building blocks" for the design of the confirmatory main study. Below, we describe these pre-studies in detail alongside with their contributions. Note that the pre-studies shared many aspects in terms of procedure and flow due to the "replication logic" that we follow (Pagell, 2020).

Thus, while we highlight key differences and unique characteristics as we progress discussing the studies, for brevity of presentation and to avoid repetition, we omit discussions of the common aspects. The vignettes used in the pre-studies are provided in Appendix C.

Each participant completed only one of the five pre-studies and no retake was allowed. To achieve this, we created a qualifier in MTurk that made subsequent studies not visible to a participant who participated in one of the prior studies and used Qualtrics options to prevent retaking. To ensure the quality of data collected through MTurk, we implemented a number of technical guards in alignment with Peer et al. (2014) and Y. S. Lee et al. (2018). In particular, we a) limited the location of participants to the U.S. using MTurk qualifiers (and cross-verified this with geolocation metadata collected via the Qualtrics interface), b) qualified only workers with a previous task approval rate greater than 95% to participate, c) restricted each participant to complete only one of the six studies, d) asked short text entry and attention check questions, and e) required participants to manually enter a randomly-generated code at the end of study in a designated area to receive the payment.

B.1.2 Pre-study 1

B.1.2.1 Motivation and Design

In Pre-study 1, we conducted a randomized vignette-based experiment with 2x2x2 factorial design with two most commonly studied leniency levers in the literature (Abdulla et al., 2019), namely, monetary and time, and a product category manipulation using apparel versus consumer electronics (two sectors with highest sales volume and return rates among consumer goods sectors (National Retail Federation, 2019)). We used the vignette in Figure B.1 in Appendix C, which only listed monetary and time levers and excluded the remaining three levers in order to narrow down participants' focus to the manipulations that allowed an initial assessment of participant attention. To operationalize high and low leniency levels in a realistic, practically-relevant way which would be applicable for both product categories, we chose 15% restocking fee vs full-refund for the monetary lever and 15 versus 60 days for the time lever. In Appendix B, we provide a more detailed rationale for the operationalizations of all leniency levers in pre-studies and ultimately in the main study.

B.1.2.2 Participants and Procedure

Four hundred seven participants on Amazon Mechanical Turk completed Pre-study 1 and received $1 in compensation. Several key sample characteristics for the pre-studies are provided in Table B.1. Participants were randomly assigned to one of the 8 experimental conditions, all of which were based on the vignette in Figure B.1. In addition to the experimental vignette part, participants answered several survey questions, which provided data regarding participants' general shopping behaviors and preferences, opinions about return policies and returning, as well as demographics (more on this later). At the end of the study, all participants answered two attention check questions (which also served as factual manipulation check questions (Kane & Barabas, 2019)), asking the restocking fee rate and time limit for returns mentioned in the experimental vignette (Abbey & Meloy, 2017). Of 407 participants, 270 participants who answered both attention check questions correctly, whose location was verified to be in the U.S. (using geolocation data provided by the Qualtrics interface), and who entered meaningful responses to the open text entry question that we asked toward the end of the study were included in the final sample. The overall data loss was in alignment with the rates reported in a survey of numerous studies in Abbey and Meloy (2017). No significant imbalances were observed in the data after eliminating the inattentive subjects, with the number of observations per each experimental condition ranging between 32 and 35.

B.1.2.3 Findings

After completing the measurement scale development and construct validation procedure detailed in the manuscript and in Appendix E, we ran a three-way ANOVA to examine potential main and interaction effects of the monetary and time leniency levers, as well as the product category factor, on purchase intentions – the focal outcome variable of our theoretical model. The analysis revealed a significant main effect of the monetary lever, nonsignificant main effects of time

Characteristics	Study 1		Study 2		Study 3		Study 4		Study 5		Row Sums	
Age	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
18–24	24	8.9	16	11.9	13	10.1	9	6.4	25	8.4	87	9.0
25–34	111	41.1	58	43	62	48.1	70	50	111	37.4	412	42.4
35–44	64	23.7	38	28.1	32	24.8	38	27.1	83	27.9	255	26.3
45–54	44	16.3	13	9.6	17	13.2	19	13.6	41	13.8	134	13.8
55 and above	27	10.0	10	7.4	5	3.9	4	2.9	37	12.5	83	8.5
Gender												
Male	132	48.9	84	62.2	75	58.1	73	52.1	141	47.5	505	52.0
Female	137	50.7	51	37.8	54	41.9	67	47.9	155	52.2	464	47.8
Other/NA	1	0.4	0	0.0	0	0.0	0	0.0	1	0.3	2	0.2
Income												
$< 30k$	80	29.6	34	25.2	50	38.8	56	40.0	86	29.0	306	31.5
$30 - 70k$	118	43.7	65	48.1	42	32.6	55	39.3	112	37.7	392	40.4
$70 - 100k$	44	16.3	18	13.3	18	14	22	15.7	58	19.5	160	16.5
$> 100k$	28	10.4	18	13.3	19	14.7	7	5.0	41	13.8	113	11.6
Purchase Channel												
Offline (Store)	52	19.3	20	14.8	21	16.3	27	19.3	48	16.2	168	17.3
Online	154	57.0	91	67.4	89	69.0	83	59.3	178	59.9	595	61.3
Either/No Preference	64	23.7	24	17.8	19	14.7	30	21.4	71	23.9	208	21.4
Return Channel												
Offline (Store)	144	53.3	60	44.4	70	54.3	89	63.6	166	55.9	529	54.5
Online (Mail)	71	26.3	46	34.1	38	29.5	32	22.9	79	26.6	266	27.4
Either/No Preference	55	20.4	29	21.5	21	16.3	19	13.6	52	17.5	176	18.1
# of Distinct Retailers												
< 5	50	18.5	51	37.8	31	24.0	35	25.0	66	22.2	233	24.0
5–10	126	46.7	59	43.7	64	49.6	66	47.1	128	43.1	443	45.6
> 10	94	34.8	25	18.5	34	26.4	39	27.9	103	34.7	295	30.4
Effective (% of Total)	270	67.5	135	67.5	129	64.5	140	70	297	74.3	971	100.0

Table B.1: Pre-study Sample Characteristics leniency and product category, and nonsignificant interaction effects (Table B.2). Observing no category contingency, we proceeded with consumer electronics as the context for the remainder of the pre-studies to reduce sample size requirements.

Predictor	df	Mean Square	F	p	partial η^2
Intercept	1	7231.665	5101.621	0.000	0.951
C	1	0.107	0.075	0.784	0.000
M	1	269.768	190.309	0.000	0.421
T	1	1.653	1.166	0.281	0.004
M x T	1	1.486	1.049	0.307	0.004
C x M	1	0.237	0.167	0.683	0.001
C x T	1	1.304	0.920	0.338	0.003
C x T x M	1	0.572	0.404	0.526	0.002
Error	262	1.418			

Table B.2: ANOVA Results – Pre-study 1

B.1.3 Pre-studies 2–4

B.1.3.1 Motivation and Design

Pre-studies 2–4 had a combined motivation, so we report them together. After examining monetary and time levers, we proceeded to test the remaining three levers, namely, effort, scope, and exchange levers by manipulating two of them in a given pre-study, resulting in the pre-studies that jointly tested each of these levers in two independent samples. In pre-studies 2–4, we fixed the monetary and time leniency levers to their high leniency levels that were observed more commonly in the industry (full-refund and 60-day return window as in Pre-study 1. Pre-studies 2–4 therefore had a 2x2 factorial design. Pre-study 2 manipulated effort and scope levers, Pre-study 3 tested effort and exchange, and Pre-study 4 tested scope and exchange levers.

B.1.3.2 Participants and Procedure

The procedure of Pre-studies 2–4 were the same as Pre-study 1, except for the experimental vignettes used in different treatment conditions. We used variants of the vignette in Figure B.2

in Appendix C for these pre-studies. All participants were randomly assigned to the experimental conditions. Participants answered the same survey questions and answered the same attention check questions (in order to compare the data loss rate across the samples). Two hundred participants completed each study to receive \$1 compensation. Using the same inclusion criteria as in Pre-study 1, of 200 participants, 135, 129, and 140 participants were included in the final sample of Pre-study 2, 3, and 4, respectively. The data loss rate was again comparable to Pre-study 1 and there were not significant imbalances in the number of observations across the experimental conditions in the final samples.

B.1.3.3 Findings

After empirically (re)validating the measurement scales and constructs, we ran a two-way ANOVA with all three pre-study samples to examine the main and interaction effects of effort, scope, and exchange levers. The results are reported in Tables B.3, B.4, and B.5. We found a significant main effect of the exchange lever on purchase intentions in both Pre-study 3 and 4, whereas no significant main effect was observed for either effort or scope levers. Moreover, none of the two-way interaction effects turned to be significant. Combining with the evidence from Pre-study 1, we arrived to the preliminary conclusion that monetary and exchange leniency levers are likely to prove effective in influencing purchase intentions, whereas time (at least, in terms of a 15- versus 60-day window context), effort, and scope levers may not be as influential.

Predictor	df	Mean Square	F	p	partial η^2
Intercept	1	4494.703	3717.109	0.000	0.966
F	1	2.200	1.819	0.180	0.014
S	1	3.907	3.231	0.075	0.024
F x S	1	1.151	.951	0.331	0.007
Error	131	1.209			

Table B.3: ANOVA Results – Pre-study 2

Predictor	df	Mean Square	F	p	partial η^2
Intercept	1	4391.008	3865.058	0.000	0.969
F	1	0.983	0.865	0.354	0.007
X	1	16.254	14.307	0.000	0.103
F x X	1	0.015	0.014	0.907	0.000
Error	131	1.209			

Table B.4: ANOVA Results – Pre-study 3

Predictor	df	Mean Square	F	p	partial η^2
Intercept	1	4406.724	3028.709	.000	.957
S	1	0.007	0.005	0.944	0.000
X	1	30.560	21.003	0.000	0.134
S x X	1	0.012	0.008	0.928	0.000
Error	131	1.209			

Table B.5: ANOVA Results – Pre-study 4

B.1.4 Pre-study 5

B.1.4.1 Motivation and Design

In the last Pre-study, we jointly tested the monetary and exchange levers as two levers that showed significant effects in prior pre-studies. We also tested time lever in a second sample to replicate the finding from Pre-study 1. Thus, Pre-study 5 followed a 2x2x2 factorial design.

Participants and Procedure

In Pre-study 5, 410 participants completed the study on Amazon Mechanical Turk for a compensation of $1. The study procedure was in alignment with the earlier pre-studies. Again, participants were randomly assigned to one of the eight experimental conditions. Of 410 participants, 297 were included in the final sample as a result of applying the inclusion criteria.

B.1.4.2 Findings

After successfully (re)validating the measurement scales and constructs, we conducted a three-way ANOVA to test the main and interaction effects of monetary, exchange, and time leniency levers. The results of ANOVA are reported in Table B.6. Consistent with the earlier pre-studies, we again found a significant main effect for monetary and exchange levers (with monetary leniency larger in the effect size) and no main effect for the time lever, while none of the interaction effects were significant.

Predictor	df	Mean Square	F	p	partial η^2
Intercept	1	7145.475	3956.540	.000	.932
M	1	128.291	71.036	0.000	0.197
X	1	45.872	25.400	0.000	0.081
T	1	2.444	1.353	0.246	0.005
M x X	1	2.793	1.547	0.215	0.005
X x T	1	0.047	0.026	0.872	0.000
M x T	1	0.144	0.079	0.778	0.000
M x T x X	1	2.814	1.558	0.213	0.005
Error	289	1.806			

Table B.6: ANOVA Results – Pre-study 5

B.1.5 Triangulation with Survey Responses

We administered several survey questions within the pre-studies that provided data to triangulate with our theory and results from the experimental parts, which we discuss below.

B.1.5.1 Free Associates

Before participants entered the experimental part of a pre-study, we asked participants to indicate free associates (i.e., write down at least three words or phrases that comes to your mind when thinking of...) with “return policies" and “returning a product". In addition to ensuring data quality through meaningful responses to these free associates questions, analysis of these open responses provided supporting qualitative evidence that 1) participants could associate with different dimensions of return policies that we study and 2) participants thought of quality, cost, and value aspects of returns services. In particular, we observed multiple classes of responses that relate to different constructs in our empirical model that are antecedents of the outcome variable. In particular, analysis of the aggregated data showed that participants could identify different leniency levers, through phrases such as "refund," "restocking fee," "time limit," "time window," "hassle," "store credit," and "exchange". Second, participants frequently mentioned phrases such as "good customer service," "fast service," "friendly," "high quality," and "timely," which are attributed to returns service quality. Finally, through phrases such as "(in)convenient," "costly," "difficult," "complicated," "long," "time consuming," "frustrating," and "annoying," participants demonstrated their transaction cost perceptions when thinking of return policies and making returns.

B.1.5.2 Stated Importance of Leniency Across Different Levers

At the end of each pre-study, we also asked participants to indicate on a Likert scale of (1) Most Important to (7) Least Important, how important they find leniency across different return policy levers in choosing which retailers to shop (i.e., no restocking fees (monetary), a long return time window (time), easy and hassle-free return process (effort), not only store-credit/exchange, but also cash refund option (exchange), sale and clearance items allowed to be returned as regularly-priced items (scope)). The statements associated with the five leniency levers were provided in a *randomized order* to each participant. The purpose of this question was to get exploratory insights into whether participants possess systematically ordered perceptions of how important they find different forms of return policy leniency.

Interestingly, we found consistent evidence across all five pre-studies (which manipulated different subset of these five levers) that on average, participants have an order of importance of leniency across different levers. In particular, all five pre-study samples revealed that on average, monetary leniency was by far the most important, followed by exchange leniency, followed by effort, which was followed by scope and time leniency. Multiple t-tests revealed that monetary, exchange, and effort leniency were considered to be significantly more important relative to time and scope leniency. The difference in the importance of time versus scope leniency was not significant. Moreover, monetary leniency was considered as significantly more important than exchange leniency, which in turn was considered as more important than effort leniency. This consistent empirical evidence informed us about a potential systematic order of the effects of different leniency levers on purchase intentions, which we ultimately confirmed in the empirical testing of our cognitive process model. Further, this empirical evidence also built confidence that the particular ordering of return policy levers in the vignettes was not driving the assessment of different levers and this allowed us to design a main study vignette that followed a clear, realistic, and practically-relevant ordering of the levers instead of randomizing the order for all participants. For example, note that even in Pre-study 1 where we did not include effort, scope, and exchange levers in the vignette, the order of importance for different leniency aspects was consistent with the other pre-studies. Table B.7 reports the corresponding means and standard deviations.

Pre-Study #	1	2	3	4	5
Monetary	2.544 (1.795)	3.100 (1.921)	3.350 (1.903)	3.057 (1.806)	2.910 (1.857)
Exchange	3.674 (1.195)	3.570 (1.191)	3.609 (1.282)	3.589 (1.266)	3.495 (1.224)
Effort	4.074 (2.017)	3.700 (2.048)	3.910 (2.069)	4.000 (2.060)	3.930 (2.070)
Scope	4.726 (1.739)	4.700 (1.706)	4.580 (1.810)	4.629 (1.698)	4.820 (1.736)
Time	4.915 (1.868)	4.950 (2.038)	4.810 (1.995)	5.050 (1.879)	5.110 (1.897)

Note: Lower mean value means greater importance.

Table B.7: Mean and standard deviations for importance ratings of levers in pre-studies.

B.1.5.3 Complexity and Cognitive Demand of Return Policy Framing

In the exit survey of Pre-Study 5, before we asked the attention check questions, we asked participants to evaluate on a 7-point Likert scale (1–Strongly Disagree to 7–Strongly Agree) to what extent participants think that the return policy statement presented to them were complex, cumbersome to ready, hard to comprehend. The analysis of these data, which had a significant right-skewness (i.e., responses largely concentrated on the lower end), showed that participants did not find the return policy statement presented in the vignettes complex (median = 2.00, 75th percentile = 3.00), cumbersome to read (median = 2.00, 75th percentile = 4.00), or hard to comprehend (median = 2.00, 75th percentile = 3.00). This built confidence that the participants did not experience significant challenge in reading and understanding the return policies presented to them.

B.1.6 Preliminary Insights into Mediation Effects

In addition to analyzing the main effects of the five leniency levers on the focal outcome of our research—consumer purchase intention—we also examined the theorized parallel-serial mediation model by running a regression based mediation analysis with non-parametric bootstrapping on each pre-study sample. This analysis was motivated by the following.

First, recent developments in the area of mediation analysis have established that even when the direct effects of a treatment variable on an outcome variable are not statistically significant, it is worthwhile to analyze the indirect (i.e., mediation) effects and such effects may turn out to be significant (Rungtusanatham et al., 2014; M. K. Malhotra et al., 2014; Hayes, 2018), which is a reversal from the long-standing Baron and Kenny (1986) approach. Thus, we were interested in observing if a similar phenomenon could be observed in the pre-studies where we found statistically insignificant main (i.e., direct) effects for time, effort, and scope levers.

Second, we wanted to understand how the number of bootstrap re-samples used in the estimation of the indirect effects influences the precision of the estimates and confidence intervals (theoretically, as the number of resamples increase, the statistical power and the precision of the estimates increase) and what would be an appropriate choice of the number of bootstrap re-samples. Finally, we wanted to have an overall understanding of the potential indirect effect sizes and the potential for a full mediation of the direct effect of a lever on purchase intentions, captured by our theory-driven empirical model.

To this end, we ran mediation analysis on all five pre-study samples with the parallel-serial mediation model that we theorized on, with the number of bootstrap resamples set as 1,000, 5,000, 10,000, and 20,000. We also included age, gender, income, number of distinct retailers shopped during a year, return rate, channel preference, and tendency to view a return policy before purchasing as covariates, given their correlations with one or several of the empirical constructs in our model. This aimed to improve the statistical power.

The analyses revealed important insights. First, we found that increasing the number of bootstrap resamples beyond 10,000 did not result in improved precision in the mediation effect estimates (measured in terms of the width of the 95% bootstrapped confidence intervals for the estimates). Hence, we choose 10,000 as the number of resamples to be used in the mediation analysis of the main study data.

Second, we found statistically significant (i.e., 95% bootstrapped confidence intervals did not include zero) total indirect effects for the monetary (average total indirect effect across two studies was 1.606) and exchange (average total indirect effect across three studies was 0.793) leniency, aligned with the expectations based on our ANOVA results presented earlier. Interestingly, we found that for all cases of significant mediation effect, the analyses suggested a full mediation.

For the remaining three levers that did not show significant main effects (i.e., direct effects) in ANOVA, we found significant indirect effect of effort leniency in one of the two samples (with effect size of 0.385 in Pre-study 3), whereas in Pre-study 2, the effect was positive (0.118) but not statistically significant. Time and scope levers did not generate statistically significant indirect effects in both samples tested, though the point estimates for the indirect effects were positive as expected (average total indirect effects across two samples being 0.146 and 0.149, respectively).

The estimates in the total effect models (i.e., a regression model regressing purchase intention on treatment variables and covariates) as part of the mediation analyses revealed that among the covariates included in the analysis, only tendency to review return policies before purchasing could explain statistically significant variability in purchase intentions beyond what was already explained by the treatment variables (i.e., manipulation of leniency levers). This was observed in Pre-study 1 and 5 samples, in which we found a negative association. Nevertheless, due to the correlations of these covariates with the mediators in the model, we decided to keep all covariates in the (confirmatory) mediation analysis of the main study sample.

Thus, we learned from the preliminary mediation analysis of the pre-study samples that 1) our theory-driven parallel-serial mediation model would be able to fully explain the direct causal effect, if any, of a leniency lever on purchase intentions, 2) monetary, followed by exchange lever, has significantly higher effects on purchase intentions relative to the remaining levers, 3) though not conclusive in terms of statistical significance, the effects of effort, scope, and time levers were practically small and similar to each other.

B.1.7 Conclusion

To conclude, five pre-studies that we conducted provided us significant conceptual knowledge and preliminary empirical insights to inform the design, execution, and analysis of the confirmatory main study that tested the research hypotheses. Next, we discuss how in light of findings from the pre-studies and through triangulation with observations from practice, we made key design decisions for the main study.

B.2 Main Study Vignette Design Choices

In designing vignette for the main study, we made several choices based on what we learned from Pre-studies 1–5, as well as additional theoretical and practical rationale that we discuss below.

First, we decided to increase the level of immersion and realism for the vignette used (Aguinis & Bradley, 2014). To this end, we designed a vignette that is based on websites of typical retailers that included several key features (i.e., product search box, product category tabs, etc.) in addition to the return policy statement. By doing this, we also shifted from asking participants to review the return policy of a retailer as presented by the researchers (i.e., the retailer provides the following return policy), to a return policy statement presented from a retailer's perspective. This was done also to increase the engagement with the vignette scenario and to stimulate realism and ecological validity.

Second, we chose a multiple-category retailer based in the U.S. for the main study. From a theoretical standpoint, this allowed estimating the average treatment effects of different return policy leniency levers in a broad context (i.e., without priming participants on a narrow set of product types or price levels). Indeed, understanding the effect of different leniency levers averaged across all potential product categories and/or price levels would be of practical significance to the large retailers (i.e., big-box stores, department stores) who may have negative connotations about offering complex return policies with many category-based exclusions. To further rationalize the choice of a multiple-category retailer, we surveyed the top 20 U.S-based retailers by sales revenue. We found that 16 of them were general merchandise stores (i.e., department stores, big-box stores) and e-tailers that carried multiple-category assortments sold through online and offline channels (the remainder was supermarkets specialized in grocery). Considering that the sales volume of these retailers constitute a significant percentage of all U.S. retail sales (i.e., $1.4 of $5.5 trillion as of 2019), at an average return rate of 10 percent, these retailers alone would account for approximately $140B out of $369B of annual returns in the U.S. (National Retail Federation, 2018, 2019). Therefore, among a myriad of other options, choosing a multiple-category retailer selling through both online and offline channels for the vignette served the purpose of a broader practical relevance.

Pre-studies revealed no statistically significant difference in purchase intentions between 15 versus 60 days operationalizations. In the main study, we decided to increase the contrast between the low and high time leniency levels in order to provide a more conservative assessment of the time lever relative to other levers. In order to decide on an actual operationalization that is practically relevant, we looked at the return policy time windows of the 16 largest multiple-category retailers mentioned above. At the time of the main study (we used Wayback Machine web archive[*] to extract this information), we found that these 16 retailers were offering return policies ranging from 15 days (2 retailers) to 180 days (3 retailers), with 5 retailers offering 30 days and 6 retailers offering 90 days as their predominant return policy applicable to most items. Thus, we chose the contrasted high and low leniency for the time lever as 180 days and 15 days, respectively.

As for monetary leniency, we analyzed top 100 Google search engine results for "restocking fee". Overall, we observed that a 15% restocking fee was most frequently mentioned, such as in consumer forums (e.g., Reddit[†]), marketplace seller forums (Amazon.com[‡]), popular press articles on consumer returns (Safdar, 2018; Rosato, 2019). A majority of the remaining entries mentioned restocking fee rates above 15%, up to 35%. Thus, we manipulated monetary leniency in terms of offering a full refund versus a 15% restocking fee—based on which we found a highly dominant

[*]https://archive.org/web/
[†]https://www.reddit.com/search/?q=restocking%20fee
[‡]https://sellercentral.amazon.com/forums

effect of monetary leniency in the pre-studies, relative to the remaining levers.

In determining the specific operationalizations of high versus low scope, effort, and exchange levers in both pre-studies and the main study, we gave a consideration to 1) be consistent with the existing literature, 2) provide realistic and/or actionable levels, and 3) significantly differentiate between low and high levels of leniency. For scope leniency, our low leniency operationalization (i.e., all sales are final for clearance and sale items and no returns allowed) was based on the seminal and often-cited contribution by Wood (2001) and is employed by numerous retailers (e.g., Michael Kors, Tommy Hilfiger, Best Buy) to prevent significantly discounted items from being returned. Exchange leniency, by definition, is uniformly operationalized in the existing literature as whether a retailer offers only store credit/exchange or offers a cash refund, so the choice here was relatively straightforward (Abdulla et al., 2019).

We surveyed the literature, retail practice, consumer forums to take note of different ways that retailers impose “hassle" to the return process and designed the low effort leniency condition accordingly. We then designed a similarly-sized framing for the high effort leniency condition, which involved different measures of convenience provided to shoppers in returning a product. Though the effort leniency statements may appear to be lengthier than the remaining levers, this was intentional, since it was critical to provide significant details for the process of returning to be able to decently signal the effort (low) leniency or convenience (high) through salience and minimize ambiguity and guesswork from the participants in this regard (that may increase the unsystematic variance in responses and make the estimates inefficient). Note that we found from Pre-study 5 that participants did not consider the return policy presented to them as complex, cumbersome to read, or hard to comprehend.

B.3 Experimental Vignettes used in Pre-studies and Main Study

Below are the snapshots of the experimental vignettes provided to the participants of Pre-study 1 (Figure B.1), Pre-studies 2–5 (Figure B.2), and the Main Study (Figures B.3 and B.4). The low leniency condition for each lever is given in the square brackets. Participants in the Main Study read the same background information in the pre-study vignettes (i.e., “We are interested in knowing your perceptions about the return policy of one of the largest retailers in the country...") prior to viewing the respective vignette.

> We are interested in knowing your perceptions about the return policy of one of the largest [consumer electronics or apparel] retailers in the country. This retailer offers free shipping for all sales. The retailer is interested in knowing how you perceive their return policy that is the same for both online and brick & mortar purchases. Note that online returns also have free return shipping.
>
> The highlights of the retailer's return policy:
>
> <Monetary> A full refund (0% restocking fee) is available for all returns. [A partial refund (15% restocking fee) is available for all returns.]
>
> <Time> All returns must be made within 60 days of the original purchase. [All returns must be made within 15 days of the original purchase.]

Figure B.1: Pre-study 1 Vignette

We are interested in knowing your perceptions about the return policy of one of the largest consumer electronics retailers in the country. This retailer offers free shipping for all sales. The retailer is interested in knowing how you perceive their return policy that is the same for both online and brick & mortar purchases. Note that online returns also have free return shipping.

The highlights of the retailer's return policy:

<Scope> All products, including clearance or sale products, have the standard return policy outlined below. [All regularly priced products have the standard return policy outlined below. However, all sales are final for clearance and sale items (no returns for clearance or sale items are allowed).]

<Monetary> A full refund (0% restocking fee) is available for all returns. [A partial refund (15% restocking fee) is available for all returns.]

<Exchange> Customers have the option to receive a refund either as cash or store credit. [All refunds are store credit only (no cash refunds available).]

<Time> All returns must be made within 60 days of the original purchase. [All returns must be made within 15 days of the original purchase.]

<Effort> For customers' convenience, the retailer provides both a pre-paid return shipping label and return receipt with every online order. Customers can place the return receipt in the box, apply the shipping label, and send any product(s) back. Alternatively, customers can bring the product(s) back to any of the retailer's stores at the return counter and provide the receipt for return processing. [The retailer requires all returns to be pre-authorized either by calling customer service or completing an online return authorization form. If the return is authorized, the customer will be given a pre-paid return shipping label and return authorization form to be packaged with the product(s) for return processing. Alternatively, customers can bring the product(s)back to any of their stores and provide the original order information, receipt, and reason for return at the return counter.]

Figure B.2: Pre-study 2–5 Vignette

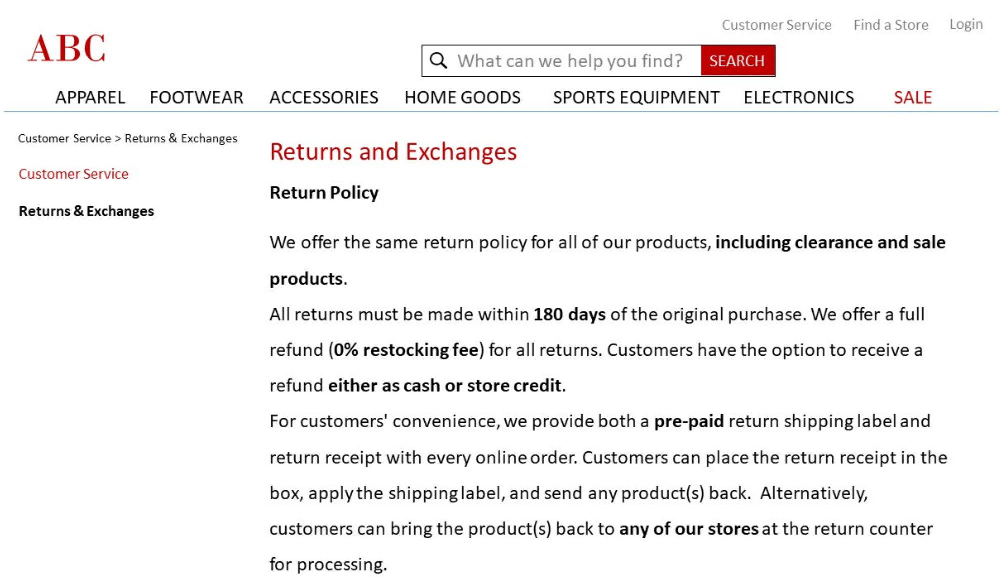

Figure B.3: Main Study Vignette – Highest Leniency Condition

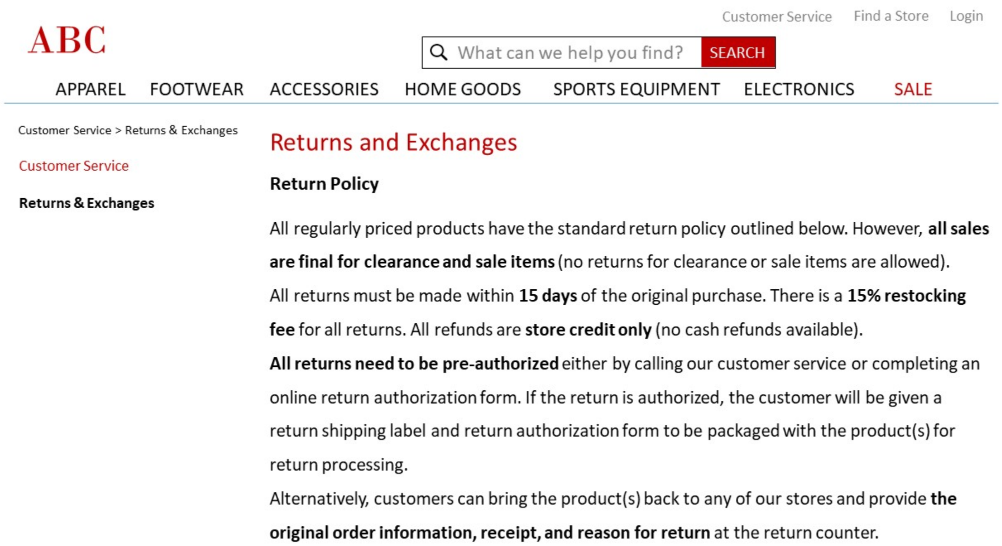

Figure B.4: Main Study Vignette – Lowest Leniency Condition

Appendix D: Measurement Scale Development and Construct Validation

The full set of the measurement items used in the experimental part of the studies is provided in Table B.8.

Associated Construct	Measurement Item	Based on/Adapted from
Perceived Service Quality	I would expect this retailer to have a reliable returns service. I would expect this retailer to promptly process my returns and exchanges. I would expect this retailer to willingly handle my returns and exchanges. I would expect a positive experience when making a return with this retailer.	Dabholkar et al. (1996)
Perceived Transaction Costs	I would expect it to be costly for me to make a return to this retailer. I would need to be more careful about selecting the right product for my needs because of this return policy. I would expect to spend too much time completing a return with this retailer. I would feel annoyed if I have to make a return to this retailer. It would be a burden to return a product to this retailer. Getting a refund from this retailer would be difficult.	Zeithaml (1988) Dodds et al. (1991)
Perceived Service Value	This retailer has a valuable return policy. I consider this retailer's return policy to be beneficial for me. This retailer offers a valuable returns service. This retailer would provide a valuable returns service for the time I may need to sacrifice. It would be worth my effort to return a product to this retailer.	Cronin et al. (2000)
Purchase Intention	I would buy from this retailer. I would purchase from this retailer if their prices are reasonable. I would buy from this retailer if they carry the products I need. My likelihood of buying from this retailer would be high. This retailer would be a reasonable choice for my shopping needs.	Baker et al. (2002)

Table B.8: Measurement Items for Scale Development

We perform Exploratory Factor Analysis (EFA) using principal components analysis with Varimax rotation for dimension reduction to establish unidimensionality. On average, items loaded significantly on their respective latent factors, with across-study factor loadings higher than 0.7. We also conduct an EFA using principal axis factoring as the extraction and oblique Promax as the rotation method (Fabrigar et al., 1999) and ensure robustness with respect to methods. We report the results of these analyses in Online Appendix E. In doing so, we apply EFA to the items developed to measure Perceived Service Quality and Perceived Transaction Costs – two parallel mediators in our empirical model. Next, we apply the EFA framework to the items that are theoretically related to Perceived Service Value and Purchase Intention.

Next, we perform Confirmatory Factor Analysis (CFA) on the four latent factors to further assess the unidimensionality of the constructs and test the goodness-of-fit for the overall measurement model suggested by the EFA results through cross-validation across the samples. Table B.9 reports values for several popular fit indexes across the five study samples. Overall, the CFA results suggest “very good" to “good" fit for the measurement model across the samples (L.-t. Hu & Bentler, 1999).

Construct reliability is assessed by the corresponding Cronbach's α and Composite Reliability (CR) measures reported in Table B.10. For all constructs in the empirical model across all five samples, the corresponding measurement scales have reliability scores greater than 0.7, indicating high reliability (Nunnally & Bernstein, 1994).

Fit Index	Pre-study 1	Pre-study 2	Pre-study 3	Pre-study 4	Pre-study 5
CFI	0.977	0.955	0.957	0.943	0.968
TLI	0.971	0.943	0.945	0.927	0.959
RMSEA	0.072	0.086	0.077	0.101	0.079
SRMR	0.031	0.044	0.048	0.043	0.033

Notes: Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square of Error Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR).

Table B.9: Values for Fit Indexes from CFA Results

	Pre-study 1	Pre-study 2	Pre-study 3	Pre-study 4	Pre-study 5
PSQ	0.856 [0.874]	0.877 [0.902]	0.840 [0.859]	0.844 [0.843]	0.813 [0.866]
PTC	0.857 [0.900]	0.885 [0.837]	0.864 [0.849]	0.852 [0.863]	0.873 [0.918]
PSV	0.935 [0.902]	0.906 [0.770]	0.889 [0.787]	0.925 [0.854]	0.937 [0.858]
PI	0.946 [0.798]	0.837 [0.754]	0.906 [0.814]	0.945 [0.823]	0.931 [0.768]

Notes: Cronbach's α [Composite Reliability]

Table B.10: Construct Reliabilities Across the Pre-studies

We assess convergent and discriminant validity of the constructs following the Fornell and Larcker (1981) criterion. According to this criterion, discriminant validity is established if the square root of the amount of variance captured by the construct, as measured by Average Variance Extracted (AVE), is greater than the shared variance with other constructs, as measured by inter-construct correlations. Using inter-variable correlations given in Table B.11 and AVE values from Tables B.12 and B.13 we establish the discriminant validity of the constructs across five studies by applying the criterion. To see this, for example, notice that the highest inter-variable correlation is between PI and PSQ (0.735), while the minimum value for the square root of AVE for these two constructs are 0.801 (in Study 2 sample) and 0.788 (in Study 4 sample), respectively. Finally, AVEs and CRs for all constructs are significantly greater than 0.5 and 0.7, respectively, in all study samples. This indicates good convergent validity.

Variable	PSQ	PTC	PSV	PI
PSQ	1.000 [0.000]			
PTC	-0.444 [0.068]	1.000 [0.000]		
PSV	0.691 [0.102]	-0.565 [0.165]	1.000 [0.000]	
PI	0.735 [0.072]	-0.531 [0.147]	0.719 [0.014]	1.000 [0.000]

Table B.11: Mean and Standard Deviation of Inter-variable Correlations Across the Pre-studies

Tables B.12 and B.13 report results of a dimension reduction procedure using principal components analysis with a Varimax rotation, which has become a norm in the field (O'Leary-Kelly &

Latent Variables	Measurement Items	Study 1 (77.73%)	Study 2 (78.55%)	Study 3 (76.85%)	Study 4 (74.85%)	Study 5 (74.77%)
Perceived Service Quality (PSQ)	1. I would expect this retailer to have a reliable returns service.	0.811 (-0.360)	0.895 (-0.244)	0.798 (-0.250)	0.801 (-0.342)	0.811 (-0.275)
	2. I would expect this retailer to promptly process my returns and exchanges.	0.896 (-0.215)	0.876 (-0.123)	0.845 (-0.266)	0.807 (-0.168)	0.862 (-0.102)
	3. I would expect this retailer to willingly handle my returns and exchanges.	0.798 (-0.222)	0.835 (-0.330)	0.811 (-0.231)	0.796 (-0.331)	0.804 (-0.306)
	Average Variance Extracted	0.699	0.755	0.669	0.642	0.682
Perceived Transaction Costs (PTC)	1. I would need to be more careful about selecting the right product for my needs because of this return policy.	0.847 (-0.179)	0.878 (-0.175)	0.844 (-0.149)	0.877 (-0.159)	0.861 (-0.145)
	2. I would feel annoyed if I have to make a return to this retailer.	0.838 (-0.287)	0.675 (-0.469)	0.859 (-0.214)	0.804 (-0.335)	0.894 (-0.178)
	3. It would be a burden to return a product to this retailer.	0.824 (-0.342)	0.703 (-0.379)	0.701 (-0.331)	0.764 (-0.416)	0.846 (-0.258)
	4. I would expect it to be costly for me to make a return to this retailer.	0.824 (-0.318)	0.602 (-0.469)	0.643 (-0.353)	0.674 (-0.333)	0.829 (-0.242)
	Average Variance Extracted	0.694	0.567	0.589	0.613	0.736

Notes: Cross-loadings in the parentheses under factor loadings. Total variance explained in the parentheses under the study titles. Using principal components analysis with a Varimax rotation.

Table B.12: EFA Results Across the Pre-studies - Part A

Latent Variables	Measurement Items	Study 1 (90.65%)	Study 2 (79.09%)	Study 3 (81.61%)	Study 4 (88.86%)	Study 5 (88.19%)
Perceived Service Value (PSV)	1. This retailer has a valuable return policy.	0.830 (0.481)	0.734 (0.337)	0.897 (0.236)	0.865 (0.413)	0.855 (0.435)
	2. This retailer offers a valuable returns service.	0.858 (0.423)	0.753 (0.537)	0.781 (0.499)	0.827 (0.429)	0.831 (0.454)
	3. This retailer would provide a valuable returns service for the time I may need to sacrifice.	0.819 (0.420)	0.897 (0.249)	0.658 (0.425)	0.783 (0.441)	0.773 (0.399)
	4. I consider this retailer’s return policy as beneficial.	0.834 (0.479)	0.664 (0.399)	0.754 (0.521)	0.827 (0.464)	0.855 (0.418)
	Average Variance Extracted	0.698	0.588	0.604	0.682	0.688
Purchase Intention (PI)	1. I would buy from this retailer.	0.732 (0.302)	0.757 (0.404)	0.746 (0.511)	0.765 (0.544)	0.663 (0.359)
	2. I would purchase from this retailer if their prices are reasonable.	0.878 (0.391)	0.791 (0.309)	0.879 (0.312)	0.849 (0.447)	0.860 (0.409)
	3. I would buy from this retailer if they carry the products I need.	0.828 (0.477)	0.817 (0.279)	0.845 (0.371)	0.876 (0.397)	0.852 (0.435)
	Average Variance Extracted	0.664	0.622	0.681	0.691	0.635

Notes: Cross-loadings in the parentheses under factor loadings. Total variance explained in the parentheses under the study titles. Using principal components analysis with a Varimax rotation.

Table B.13: EFA Results Across the Pre-studies - Part B

Vokurka, 1998). We also report EFA results for pre-study samples based on principal axis factoring and Promax rotation that some scholars (e.g., Fabrigar et al. (1999)) argue to be more appropriate as a dimension reduction method while dealing with potentially correlated constructs. Tables B.14 and B.15 report total variances extracted, factor loadings, and cross-loadings. Again, the reflective indicators included in the measurement of the empirical constructs of our model show sufficiently high factor loadings and low cross-loadings using alternative factor extraction and rotation methods.

Latent Variables	Measurement Items	Study 1 (71.23%)	Study 2 (67.33%)	Study 3 (59.73%)	Study 4 (61.29%)	Study 5 (67.93%)
Perceived Service Quality (PSQ)	1. I would expect this retailer to have a reliable returns service.	0.829 (0.000)	0.932 (0.001)	0.700 (-0.037)	0.801 (-0.342)	0.756 (-0.062)
	2. I would expect this retailer to promptly process my returns and exchanges.	0.978 (0.130)	0.872 (0.098)	0.915 (0.051)	0.584 (-0.092)	0.799 (0.105)
	3. I would expect this retailer to willingly handle my returns and exchanges.	0.672 (-0.147)	0.765 (-0.132)	0.721 (-0.021	0.845 (0.015)	0.761 (-0.079)
Perceived Transaction Costs (PTC)	1. I would need to be more careful about selecting the right product for my needs because of this return policy.	0.903 (0.099)	0.831 (0.147)	0.894 (0.116)	0.877 (-0.159)	0.865 (0.053)
	2. I would feel annoyed if I have to make a return to this retailer.	0.846 (-0.028)	0.668 (-0.222)	0.823 (0.008)	0.789 (-0.025)	0.944 (0.059)
	3. It would be a burden to return a product to this retailer.	0.707 (-0.145)	0.905 (-0.013)	0.715 (-0.291)	0.782 (-0.097)	0.785 (-0.081)
	4. I would expect it to be costly for me to make a return to this retailer.	0.849 (-0.048)	0.634 (-0.248)	0.626 (-0.242)	0.705 (-0.326)	0.798 (-0.061)

Notes: Cross-loadings in the parentheses under factor loadings. Total variance explained in the parentheses under the study titles. Using principal axis factoring with a Promax rotation.

Table B.14: EFA Results Across the Pre-studies - Part A

Latent Variables	Measurement Items	Study 1 (86.84%)	Study 2 (70.99%)	Study 3 (74.82%)	Study 4 (84.85%)	Study 5 (83.97%)
Perceived Service Value (PSV)	1. This retailer has a valuable return policy.	0.855 (0.120)	0.879 (0.359)	0.937 (-0.115)	0.980 (-0.023)	0.921 (0.028)
	2. This retailer offers a valuable returns service.	0.923 (0.028)	0.915 (0.119)	0.790 (0.163)	0.842 (0.117)	0.854 (0.102)
	3. This retailer would provide a valuable returns service for the time I may need to sacrifice.	0.733 (0.171)	0.706 (0.221)	0.690 (0.285)	0.672 (0.216)	0.617 (0.228)
	4. I consider this retailer's return policy as beneficial.	0.861 (0.116)	0.918 (-0.056)	0.733 (0.211)	0.797 (0.127)	0.921 (0.028)
Purchase Intention (PI)	1. I would buy from this retailer.	0.642 (0.338)	0.778 (0.370)	0.624 (0.296)	0.661 (0.296)	0.686 (0.484)
	2. I would purchase from this retailer if their prices are reasonable.	0.871 (0.043)	0.722 (0.394)	0.913 (-0.039)	0.875 (0.080)	0.780 (0.129)
	3. I would buy from this retailer if they carry the products I need.	0.876 (0.081)	0.849 (-0.020)	0.862 (0.032)	0.928 (0.008)	0.945 (0.006)

Notes: Cross-loadings in the parentheses under factor loadings. Total variance explained in the parentheses under the study titles. Using principal axis factoring with a Promax rotation.

Table B.15: EFA Results Across the Pre-studies - Part B

B.4 Construct Validation in the Main Study Sample

We conduct EFA and CFA with the main study sample to (re)validate the measurement scales and constructs in our empirical model, since the main study ultimately tested the research hypotheses. In doing so, we conducted EFA using both principal components analysis and principal axis factoring methods for factor extraction, and both Varimax and Promax rotations to establish robustness with respect to extraction and rotation methods (O'Leary-Kelly & Vokurka, 1998; Fabrigar et al., 1999; MacKenzie et al., 2011). Tables B.16 and B.17 report factor loadings, total variances extracted, and AVEs under both methods. As can be seen from these tables, the reflective measurement items load highly on their respective latent factors and sufficiently high AVEs under both extraction and rotation method with minor differences. Applying Fornell and Larcker (1981) criteria, convergent and discriminant validity are also confirmed in the main study sample. The results of a CFA, again, suggested a “good" to “very good" fit of for the overall factor structure (CFI = 0.988 , TLI = 0.985, RMSEA = 0.050, SRMR = 0.022). The inter-variable correlations in the main study sample is provided in Table B.18.

Latent Variables	Measurement Items	Method 1 (70.39%)	Method 2 (78.71%)
Perceived Service Quality (PSQ)	1. I would expect this retailer to have a reliable returns service.	0.873	0.855
	2. I would expect this retailer to promptly process my returns and exchanges.	0.804	0.846
	3. I would expect this retailer to willingly handle my returns and exchanges.	0.800	0.832
	Average Variance Extracted	0.683	0.713
Perceived Transaction Costs (PTC)	1. I would need to be more careful about selecting the right product for my needs because of this return policy.	0.862	0.859
	2. I would feel annoyed if I have to make a return to this retailer.	0.920	0.862
	3. It would be a burden to return a product to this retailer.	0.702	0.767
	4. I would expect it to be costly for me to make a return to this retailer.	0.744	0.804
	Average Variance Extracted	0.659	0.679

Notes: Method 1: Principal Axis Factoring with Promax Rotation. Method 2: Principal Components Analysis with Varimax Rotation. Total variance explained in the parentheses under the method titles.

Table B.16: EFA Results of Main Study - Part A

Latent Variables	Measurement Items	Method 1 (86.36%)	Method 2 (90.24%)
Perceived Service Value (PSV)	1. This retailer has a valuable return policy.	0.893	0.842
	2. This retailer offers a valuable returns service.	0.870	0.836
	3. This retailer would provide a valuable returns service for the time I may need to sacrifice.	0.804	0.839
	4. I consider this retailer's return policy as beneficial.	0.834	0.847
	Average Variance Extracted	0.745	0.710
Purchase Intention (PI)	1. I would buy from this retailer.	0.744	0.796
	2. I would purchase from this retailer if their prices are reasonable.	0.901	0.877
	3. I would buy from this retailer if they carry the products I need.	0.899	0.853
	Average Variance Extracted	0.725	0.710

Notes: Method 1: Principal Axis Factoring with Promax Rotation. Method 2: Principal Components Analysis with Varimax Rotation. Total variance explained in the parentheses under the method titles.

Table B.17: EFA Results of Main Study - Part B

B.5 Details for the Study on the Moderating Role of Price on the Effects of Monetary and Exchange Leniency

In a randomized online experiment, we explored the moderating role of price in determining the effect of monetary and exchange leniency on consumers' purchase intentions. The experiment

Variables	PSQ	PTC	PSV	PI	RETURN RATE	POLICY READ	NRETAILERS	CHANNEL	GENDER	AGE	INCOME
PSQ	1.000	-0.523**	0.717**	0.674**	-0.072*	0.063	0.024	0.046	-0.035	0.057	0.029
PTC	-0.523**	1.000	-0.642**	-0.553**	0.003	0.089**	-0.022	-0.061	0.034	-0.043	-0.048
PSV	0.717**	-0.642**	1.000	0.719**	-0.162**	0.129**	-0.014	-0.034	-0.075*	-0.083*	0.022
PI	0.674**	-0.553**	0.719**	1.000	-0.115**	-0.011	0.041	-0.004	-0.094**	-0.095**	-0.001
RETURN RATE	-0.072*	0.003	-0.162**	-0.115**	1.000	-0.090*	0.033	0.078*	0.008	0.092**	0.025
POLICY READ	0.063	0.089**	0.129**	-0.011	-0.090*	1.000	0.055	-0.003	0.011	0.011	0.053
N RETAILERS	0.024	-0.022	-0.014	0.041	0.033	0.055	1.000	0.115**	0.006	-0.046	0.150**
CHANNEL	0.046	-0.061	-0.034	-0.004	0.078*	-0.003	0.115**	1.000	0.068*	0.070*	0.040
GENDER	-0.035	0.034	-0.075*	-0.094**	0.008	0.011	0.006	0.068*	1.000	0.106**	-0.037
AGE	0.057	-0.043	-0.083*	-0.095**	0.092**	0.011	-0.046	0.070*	0.106**	1.000	0.033
INCOME	0.029	-0.048	0.022	-0.001	0.025	0.053	0.150**	0.040	-0.037	0.033	1.000

* $p < 0.05$, ** $p < 0.01$

Table B.18: Inter-Variable Correlations in the Main Study Sample had a six-cell (partial factorial) design, price-level (low = \$50, high = \$500), monetary leniency (low = 15% restocking fee, high = no restocking fee), and exchange leniency (low = no cash refund/store credit only, high = both cash refund and store credit as options) as the manipulated factors. We excluded low monetary and exchange leniency conditions from this study because they are unrealistic from a practical standpoint (i.e., charging a restocking fee on a store credit). We chose desk chairs for the purchase scenario, since they have considerably high price variability in the market and significant product fit/quality uncertainty. We used experimental vignettes that are consistent with our earlier studies. In particular, we created new vignettes for the product listings (Figures B.5 and B.6) and used the vignettes for the return policies from the main study.

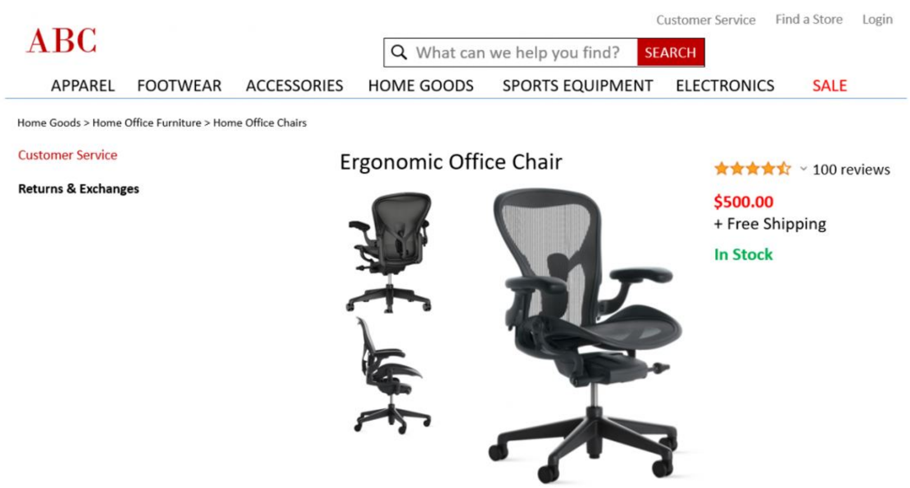

Figure B.5: Product Vignette - High Price Condition

Four hundred fifty participants (49% female) recruited through Prolific Academic completed the study for a compensation of \$0.70 (\$8.5 on an hourly basis). Participants were asked to imagine themselves buying a desk chair. Then, participants were told that they found a desk that they like from retailer ABC. Background information about ABC in line with the main study was provided. After viewing the page for the desk chair, participants were instructed to review the return policy of

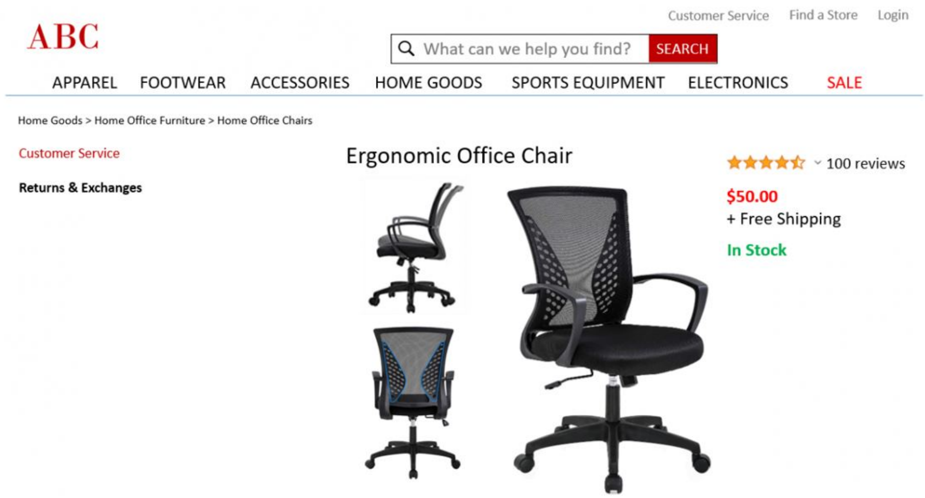

Figure B.6: Product Vignette - Low Price Condition

ABC, knowing that they may end up returning the chair. The rest of the experiment was in line with the main study. At the end of the study, participants were asked three attention check questions based on price, monetary, and exchange leniency manipulations (i.e., factual manipulation checks). Of 450 participants, 383 who passed three attention check questions based on manipulations were included into the final analysis.

A (fractional factorial) three-way ANOVA with purchase intention as the dependent variable showed that the main effects of price ($F(1, 377) = 8.49, p = 0.003$), monetary leniency ($F(1, 377) = 82.58, p = 0.000$), and exchange leniency ($F(1, 377) = 49.23, p = 0.000$) were statistically significant. The interaction effects between price and monetary leniency ($F(1, 377) = 1.59, p = 0.208$) and between price and exchange leniency ($F(1, 377) = 2.31, p = 0.129$) were not statistically significant. Thus, we concluded that price-level does not significant impact the effects of monetary and exchange leniency on purchase intentions.

APPENDIX C

APPENDIX TO RESTRICTIVE CHANGES TO LONG-ESTABLISHED LENIENT RETURN POLICIES AND CONSUMER REACTIONS TO THEM

C.1 Measurements

Measurement Item	Mean (S.D.)
1. I end up returning items quite often	2.544 (1.420)
2. Lenient return policies are important for me while choosing a retailer I shop with	4.812 (1.564)
3. In the U.S., lenient return policies are abused a lot	4.920 (1.470)
4. Overall, retailers I usually shop with well respond to consumer needs during the pandemic	5.100 (1.112)

Notes: On a 7-point Likert scale from 1 – Strongly Disagree to 7 – Strongly Agree

Table C.1: Descriptive Statistics for Study 1 Covariates

Latent Variable	Measurement Item
General Tendency to Trust Others ($\alpha = 0.84$)	1. I generally believe that others can be counted on to do what they say they will 2. I usually trust others until they give me a reason not to trust them. 3. My tendency to trust others is high. 4. It is usually hard for me to lose my trust for others.

Notes: On a 7-point Likert scale from 1 – Strongly Disagree to 7 – Strongly Agree

Table C.2: Measurement Scale for General Tendency to Trust Others

Latent Variable	Measurement Item
Consumer Trust toward Retailer ($\alpha = 0.92$)	1. Bank Name seems to have high integrity in doing business 2. Bank Name would be strongly committed to what its customers value. 3. Bank Name would protect its customers' best interests while doing business. 4. Bank Name is a caring bank. 5. Bank Name has customers in mind while making business decisions.
Favorable Behavioral Intentions ($\alpha = 0.78$)	1. I would sign up for another credit card from Bank Name. 2. I would recommend Bank Name to others. 3. I would use the services of ACME in the future.

Notes: On a 7-point Likert scale from 1 – Strongly Disagree to 7 – Strongly Agree

Table C.3: Measurement Scales for Study 3