2017

Sustainable, Operations-enabled Solutions for Reducing Product Waste

Erin C. McKie
University of South Carolina

Sustainable, Operations-enabled Solutions for Reducing Product Waste

by

Erin C. McKie

Bachelor of Business Administration
University of South Carolina, 2009

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Business Administration

Darla Moore School of Business

University of South Carolina

2017

Accepted by:

Mark Ferguson, Major Professor

Michael Galbreth, Major Professor

Sriram Venkataraman, Committee Member

David Crockett, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

Dedication

To my loving parents, Robert and Cassandra McKie.

I would not be where I am today without you. Thank you for giving me your all.

Acknowledgments

The completion of my dissertation would not have been possible without the support of the Moore School of Business faculty, department colleagues, and my dear friends and family.

Accordingly, I wish to express my utmost thanks to my dissertation committee co-chairs (Dr. Mark Ferguson and Dr. Michael Galbreth) for their enormous support throughout my tenure as a Ph.D. student. Dr. Ferguson, thank you for introducing me to the field of sustainable operations and allowing me to accompany you to several industry and academic conferences. These events provided the inspiration and foundation for my research. Dr. Galbreth, thank you for teaching me the fundamentals of developing a strong research paper, including how to effectively communicate my ideas. Your detailed instruction has been invaluable to my development. I also wish to recognize my dissertation committee members (Dr. Sriram Venkataraman and Dr. David Crockett) for their advisement on multiple research projects. Dr. Venkataraman, thank you for serving as an outstanding reference on a variety of analysis methods throughout the development of my dissertation. Dr. Crockett, thank you for your willingness to serve as an outside committee member and your input on my dissertation chapters.

I have been fortunate to interact with many other faculty members and students at the Moore School who I also owe special thanks. Dr. Manoj Malhotra, thank you for your mentorship over the last several years. I have learned so much from you and am privileged to have been an undergraduate and Ph.D. student in your department. Dr. Carrie Queenan, thank you for your willingness to help me prepare for numerous presentations, interviews, and teaching assignments. I have been touched by your generosity. Dr. Jim and Mrs. Virginia Ann Edwards, thank you for your professional and spiritual guidance. To Julia Witherspoon, our department administrator, thank you for being an incredible source of support and encouragement. To Scott Ranges, the director of Ph.D. programs, thank you for helping me navigate several administrative requirements. To my department colleagues, past and present (Dr. Cidgem Ataseven, Sanghoon Cho, Moonwon Chung, Justin Kistler, Dr. Ashley Metcalf, Dr. Mariana Nicholae, Olga Pak, Dr. Minseok Park and his wife, Aelim Kim, Dr. Soohoon Park, Yuqi Peng, Dr. Guangzhi Shang, Cherry Singhal, Dr. Deepa Wani, Dr. Övünç Yılmaz, and Zhihao Zhang) thank you all for being great neighbors, friends, and late-night study partners.

In addition to the faculty mentioned above, I would like to give a very special thanks to my career mentor, Dr. Sanjay Ahire. Dr. Ahire, thank you for inspiring and encouraging me to pursue a Ph.D. Ever since I was a student in your class nearly ten years ago, you have been committed to my professional and personal growth. For this, I am deeply grateful.

Finally, I wish to acknowledge my family who have and continue to be a tremendous source of support, love, and happiness. I am so thankful for all of you.

Abstract

The traditional linear production system where products are created, used, and then disposed of is no longer a viable business model for many firms. A combination of growing populations, increases in consumerism, and urbanization are placing unprecedented pressures on our world's natural resources. In addition to these motivations, strong demands from consumers and governments are requiring companies to reevaluate and prioritize their environmental strategies. Fortunately, there are several opportunities for firms to engage in more sustainable business practices throughout their entire supply chain, particularly at the end of their products' useful lives. However, moving from a linear model to a more closed-loop production system where products are recovered and reused brings a host of operational challenges, some of which remain unaddressed by the current literature.

In this dissertation, we examine a series of common, operations-related issues firms and government agencies face when pursuing sustainable waste management practices. In the first essay, we evaluate noted barriers operations managers face when entering the market for refurbished products. In the second study, we analyze the effectiveness of environmental legislation and consumer education efforts in promoting product reuse and recycling. In the last essay, we develop a robust consumer returns forecasting model to aid operations managers in their inventory, reverse logistics, and return recovery decisions. In addition to academic contributions, the results from these studies offer practitioners guidance needed to facilitate the transition to more circular production models and increase the number of sustainable, operations-enabled opportunities for reducing product waste.

Table of Contents

List of Tables

List of Figures

Chapter 1

Overview

The traditional linear production system where products are created, used, and then disposed of is no longer a viable business model for many firms. A combination of growing populations, increases in consumerism, and urbanization are placing unprecedented pressures on our world's natural resources. For example, since 1980 the amount of raw materials extracted and consumed globally has increased by 60% to over 62 billion metric tons (OECD, 2017). Furthermore, rapid increases in post-consumer waste are accelerating the threat of non-sustainable development and a potential global waste crisis. Although recycling rates have improved, the majority of waste in most developed countries is still sent to landfills (OECD, 2015). Last, in addition to the threats of resource scarcity and vast pollution, strong demands from consumers and governments are motivating firms to reevaluate their environmental strategies (Atasu, 2016; Macarthur-Foundation, 2013).

Fortunately, there are several opportunities for firms to engage in more sustainable business practices throughout their entire supply chain — particularly at the end of their products' useful lives (see Figure 1.1[1]). The EPA has identified and ranked various waste management strategies from most to least environmentally preferred, and the hierarchy places emphasis on reducing, reusing, and recycling as key to sustainable materials management (EPA.gov, 2016). However, moving from a linear model to a more closed loop production system where products are recovered and reused brings a host of operational challenges, some of which remain unaddressed by the current sus-

[1]Figure source: https://connect.innovateuk.org

tainable operations literature (Guide and Van Wassenhove, 2009). Examples include navigating environmental regulations, managing reverse supply chain complexity, and understanding market competition and consumer preferences for reused goods (Guide and Van Wassenhove, 2009; Atasu and Wassenhove, 2012; Agrawal et al., 2015). As noted by both industry and academic sources, additional research is needed in these areas to maximize firms' future economic and environmental performance (OECD, 2017; Guide and Van Wassenhove, 2009).

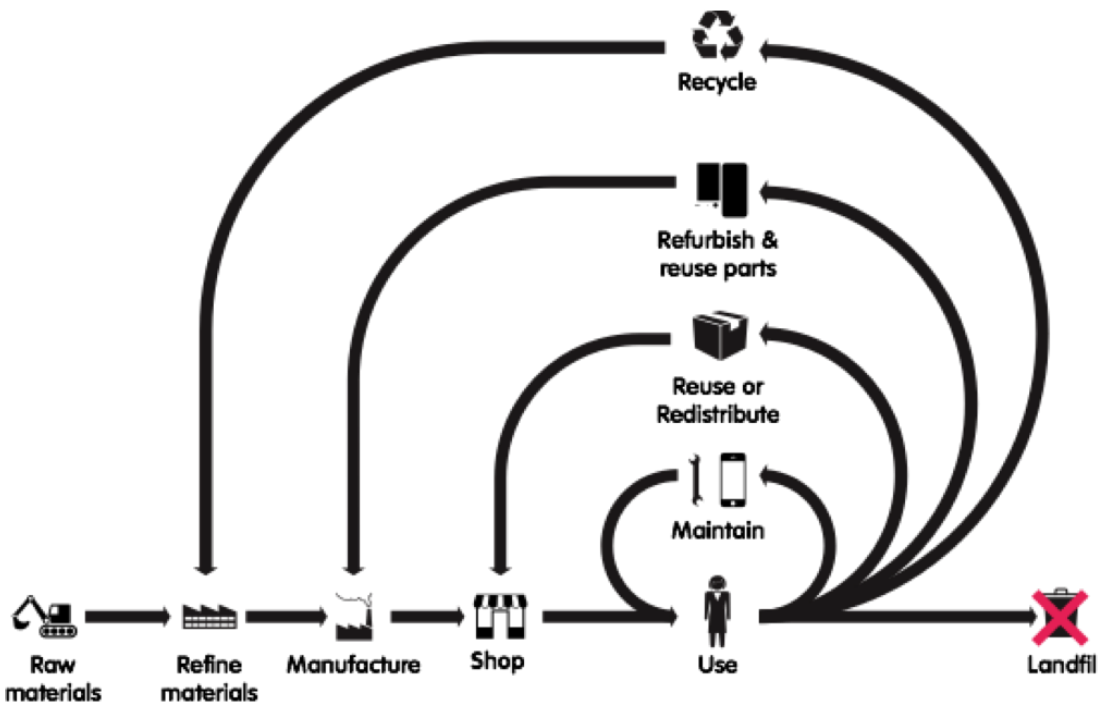


Figure 1.1: The Linear and Closed Loop Supply Chain

In response, in this dissertation we address a series of operations-related challenges firms and government agencies face in implementing sustainable waste management practices. As shown in Figure 2.3 and discussed below, we use the EPA's 3R hierarchy as a framework for our studies. In Chapter 2, we explore opportunities to promote product reuse through remanufacturing. We note that many companies are reluctant to enter the remanufacturing market because of concern with cannibalization of new sales, competition from current remanufacturers, and the willingness of consumers to purchase remanufactured products. What is often missing, however, is an in-depth understanding of how consumers make complex purchase decisions involving remanufactured items among numerous other options. Thus, this paper examines how consumers evaluate remanufactured products when there are multiple conditions and generations of the item available, and evaluates the risk that remanufactured products pose to new product sales. We leverage transaction data from eBay and structural estimation techniques developed in the industrial organization literature to conduct our analysis. We find that product generation, condition, and seller attributes are all highly influential in shaping consumers' purchasing decisions, and that the relationship between new and remanufactured products is much more nuanced and context-specific than previously thought. Counter to industry intuition, we find that remanufactured products pose the same amount of threat to new-condition goods as do used goods. Through these and other findings, we provide insights on how CLSC participants and those exploring entry into the remanufacturing business may achieve more profitable remanufacturing strategies.

	EPA Waste Management Hierarchy		
	Reduce	**Reuse**	**Recycle**
Study	**Chapter 4:** **Using Transaction Data to Improve Consumer Returns Forecasting**	**Chapter 2:** **How do Consumers Choose Between Multiple Product Conditions and Generations? An Empirical Study of iPad Sales on eBay**	**Chapter 3:** **Promoting E-Recycling: Effects of Legislation and Consumer Attributes on Recycling Outcomes**
Key Challenges Addressed	Variation in Consumer Product Evaluation Periods	Consumer Acceptance of Remanufactured Products, Firm Cannibalization Concerns	End-user participation in recycling programs, Legislation Configuration for Manufacturers
Research Domain	Closed Loop Supply Chains, Sustainable Operations	Closed Loop Supply Chains, Sustainable Operations	Environmental Regulations, Extended Producer Responsibility

Figure 1.2: Summary of Dissertation Essays

In the Chapter 3, we analyze the effectiveness of legislative tools and consumer education efforts in promoting product recycling. To protect public health and the environment, several states in the US have adopted various forms of legislation for electronic waste. It has been argued, however, that such legislation is ineffective and, in some cases, lays too heavy a responsibility on manufacturers when *consumer behavior* is actually the key to recycling success. Although both consumer attributes and legislation seem likely to impact e-recycling, a lack of empirical data has limited firms' and legislators' ability to assess the true impact of these factors. In this study, we quantify how recycling and reuse activities are shaped by consumer attributes and two popular forms of state-level e-waste legislation: extended producer responsibility (EPR) and landfill bans. To study our research questions, we leverage consumer survey data that details recycling activities in states with and without EPR policies and landfill bans for e-waste. We find, as expected, that in states with EPR legislation, consumers are more likely to recycle their electronic waste. Perhaps less intuitively, we find that consumers' knowledge of landfill bans increases their likelihood to *store* their electronics rather than recycle them. Lastly, we confirm that the biggest deterrents to consumer recycling are not knowing where to recycle electronic products, and inconvenient recycling locations. Through these and other observations, we provide insights on how administrators of recycling programs can best leverage e-waste legislation and structure consumer education efforts to maximize consumer e-waste recycling.

In the Chapter 4, we develop a robust consumer returns forecasting model to aid operations managers in inventory, reverse logistics, and return recovery decisions. In recent years, offering a generous return policy has become increasingly popular among U.S. retailers eager to win sales. Although lenient return policies have been shown to have marketing benefits such as a higher willingness to pay and a higher purchase frequency, counterbalancing these benefits with an increased volumes of returns presents operational challenges for both retailers and original equipment manufacturers (OEMs). To better manage consumer returns, operations managers need an accurate return forecast as an input into their strategic and tactical tools. We propose a consumer return forecasting framework and test our model on datasets provided by brick-and-mortar and online retailers. By more effectively utilizing transaction-level data such as purchase and return timestamps, our basic model demonstrates forecasting error reduction over benchmark models constructed from common industry practice and existing literature. We find that the reduction in forecasting error is likely more pronounced for product categories that have more variable return rates and less variability in average time-to-return durations. Such forecast accuracy improvement has broad implications for inventory, staffing, reverse logistics, and return recovery decisions.

Chapter 5 concludes with a summary of findings and future research directions. In addition to academic contributions, the results from these studies offer practitioners guidance needed to facilitate the transition to more circular production models and increase the number of sustainable, operations-enabled opportunities for reducing product waste.

Chapter 2

How do Consumers Choose Between Multiple Product Generations and Conditions? An Empirical Study of iPad Sales on eBay

2.1 Introduction and Motivation

Advances in technology, coupled with firms' needs for sustainable business practices, have resulted in the growth of industries that can deliver both economic and environmental value. One such industry, for which annual production exceeds $43 billion in the U.S. alone, is the remanufacturing sector (Ward, 2014; MSNBC, 2014).

Remanufacturing is an extensive, industrial process that restores goods to their original working condition (USITC, 2012a). The remanufacturing process is executed by participants of closed loop supply chains (CLSC) and offers wide-ranging benefits to the environment (reduced raw material and energy consumption), producers (average profit margins often exceed 20%), and consumers (lower prices and product failure rates) (Guide and Wassenhove, 2001).

Remanufactured products are sometimes referred to as "refurbished" in the marketplace and we will use these two terms interchangeably in this paper. Furthermore, remanufactured products are typically sold at a discount relative to the price of a new product but are often more expensive than used products. (Ferguson, 2010). We refer to *New, Remanufactured,* and *Used* as the product "conditions" of interest in this paper (subject to some additional sub-categorization, as explained in Section 3), with the understanding that the physical condition (although not the consumer perception) might be identical in some cases across these categories (e.g. Remanufactured vs. New).

Despite the demonstrated economic and subsequent environmental benefits, it is surprising that the remanufacturing industry remains underdeveloped with significant opportunities for growth (Hagerty and Glader, 2011). Remanufacturing intensity, measured as the ratio of remanufacturing to total manufacturing production, is "still small" (USITC, 2012b), p.1) and accounts for only 2% of the total $2.1 trillion manufacturing industry (USITC, 2012a). To better understand why remanufacturing is not more widespread, practitioners and academics have focused on isolating key issues and opportunities within the sector (see for example, Guide and Wassenhove (2001) Guide and Van Wassenhove (2009), and Souza (2013a)). As noted in this literature, both original equipment manufacturers (OEMs) and third-parties are reluctant to enter the remanufacturing market because of concern with cannibalization of new sales and competition from current remanufacturers. These threats, while theoretically viable, are based on speculation as opposed to empirical evidence. Still, in response to the threat of cannibalization and competition, OEMs have been reported to discourage remanufacturing through actions such as copyrighting repair manuals (Turner, 2016) and launching aggressive core destruction programs disguised as "recycling" tactics (USITC, 2012a).

Even beyond the fear of competition, firms are uncertain of how to promote remanufactured products, and fear that these goods will not be accepted by consumers in the marketplace. For example, the United States International Trade Commission (USITC) surveyed over 2,900 remanufacturers and reported that the (1) relative price of remanufactured goods to new goods and (2) consumers' perceptions remanufactured products were some of the top factors known to influence demand for remanufactured products (USITC, 2012a). However, the specific effects and consequences of price and consumers' perceptions were observed to be poorly understood by survey participants (USITC, 2012a).

Guide and Van Wassenhove (2009) summarize the effects of the discussed cannibalization, competition, and consumer acceptance concerns in their overview of the CLSC literature by noting that "if prices and markets are not fully understood, they become barriers to fully unleashing the value potential of CLSCs, no matter how well the operational system is designed." Therefore, in this paper we examine how consumers evaluate remanufactured products when there are multiple conditions and generations of the item available, and directly evaluate the risk that remanufactured products pose to new product sales. Specifically, we examine the relative effect of product (price, presence of warranty, return policy) and seller related (reputation score) attributes on consumers' choices for different conditions (New, Refurbished, Used) and different generations of a product.

Our study builds on a recent stream of work exploring how key factors such as brand equity (Abbey et al., 2015) and seller identity (Agrawal et al., 2015) influence consumers' attitudes and willingness to pay for remanufactured products; please see section 2 for a thorough review of this literature. Most papers in this stream compare consumers' behavior when faced with a simple binary choice: a new product or a refurbished version of that same product. In practice, however, consumers' preferences for refurbished products may be shaped by a dramatically wider selection of goods, including used-condition products. Further, the growth of online secondary markets has resulted in multiple *generations* of products being offered at the same time. For example, on Amazon.com, consumers shopping for iPads may choose between four generations (iPad 1 - iPad 4) and three conditions (New, Used, and Refurbished) of iPads. Additionally, consumers may choose between sellers with different consumer ratings and return policy offerings. Figure 2.1 shows iPad listings on WalMart.com, which similarly offers refurbished goods alongside multiple conditions and generations of the same products. Overall, our study includes a much broader choice set than those represented in previous studies, as shown in Figure 2.2. By analyzing the relative effects of noted product and seller related variables on consumers' choices, we extend the previous research by conducting a more direct examination of the cannibalization effect between product conditions and generations.

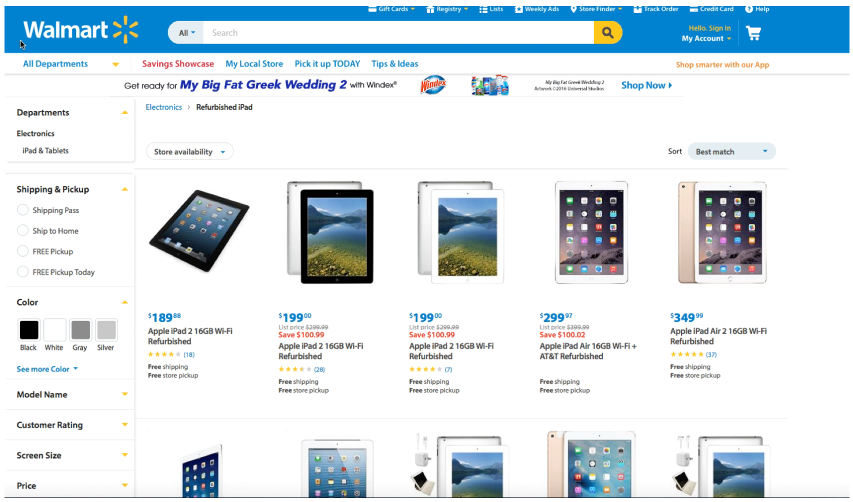


Figure 2.1: Example of Consumer Choices on WalMart.com

As described in detail in section 3, we focus our study on sales of iPads. Like many consumer electronics, iPads experience rapid growth and product design changes, leading to shorter life cycles and a growing number of products needing appropriate end-of-life management (EPA.gov, 2016). In addition, the presence of substances of concern in some electronics - such as lead, mercury, and chlorine - merit greater consideration for safe end-of-life management (Gayle, 2012). Remanufacturing provides a preferred alternative to these products being thrown away and consequently contributing to e-waste, and often provides a higher value recovery option than recycling (EPA.gov, 2016).

We leverage transaction data from eBay and structural estimation techniques developed in the empirical industrial organization literature to conduct our analysis (Berry, 1994). Our analysis method allows us to better understand why a consumer

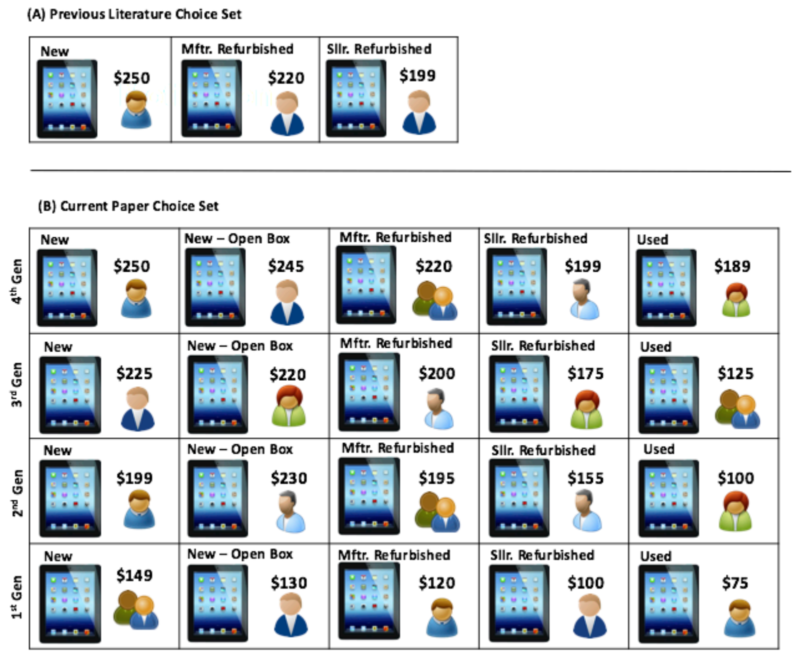


Figure 2.2: Illustration of Current Study Choice Set (prices shown are for illustrative purposes only)

makes a particular choice and how the consumer analyzes trade-offs among the attributes of the choices. We provide a preview of our main results here. We find that product generation, condition, and seller attributes are all highly influential in shaping consumers' purchasing decisions, and that the relationship between new and refurbished products is much more nuanced and context-specific than previously thought. Our results suggest that changes in Manufacturer 3P Certified products prices threaten the market share of New and New Open-Box approximately three times more than changes in Seller Refurbished products prices. Surprisingly, we find that remanufactured products pose the same amount of threat to New-condition goods as do Used goods. This finding runs counter to most existing industry intuition, where original equipment manufacturers fear cannibalization much more from remanufactured products than from used products. Through these and other observations, we provide insights on how CLSC participants and those exploring entry into the remanufacturing business may achieve more profitable remanufacturing strategies.

2.2 Literature Review

There are two streams of literature that are particularly relevant to our research. The first stream includes an emerging group of studies that explore market development issues of CLSCs. The second relates the growing application of our methodological approach, structural estimation modeling, to the operations literature.

2.2.1 How consumers choose between new and remanufactured products

Although companies that manage used electronics have existed for years, the market for refurbished goods has expanded rapidly as of late. The resulting concurrent availability of new, used, and refurbished goods, and the overlap of successive product generations, has raised several concerns about the potential for cannibalization and consumer acceptance of remanufactured products. In response, an emerging stream of literature has considered how consumers choose between new and remanufactured products, as shown in Figure 2.3. This set of papers can be classified into two sub-categories according to their key conclusions. In the first subset, the authors explore differences in consumers' willingness to pay for new and remanufactured products (Harms and Linton, 2015). Factors such as product category (Guide and Li, 2010), seller identity (Subramanian and Subramanyam, 2012; Agrawal et al., 2015), seller reputation (Subramanian and Subramanyam, 2012), and competitive intensity (Agrawal et al., 2015) are shown to influence the magnitude of these differences. A mixture of field and behavioral experiments, secondary market data, and surveys are leveraged in these studies.

In the second subset of studies, the authors identify factors that influence consumers' *preferences for remanufactured products* (Abbey et al., 2015), and the likelihood of new product cannibalization (Ovchinnikov, 2011). These papers show that factors such as brand equity (Abbey et al., 2015), product quality (Ovchinnikov, 2011; Abbey et al., 2015), and price discount (Ovchinnikov, 2011; Abbey et al., 2015) influence the utility of remanufactured products across product types and various consumer segments. Both surveys and behavioral experiments are leveraged in these studies. Overall, this stream of research provides insights on factors that influence 1) consumers' willingness to pay for products and 2) the attractiveness/utility of remanufactured products. A summary of the factors studied across all empirical studies is shown in Figure 2.4.

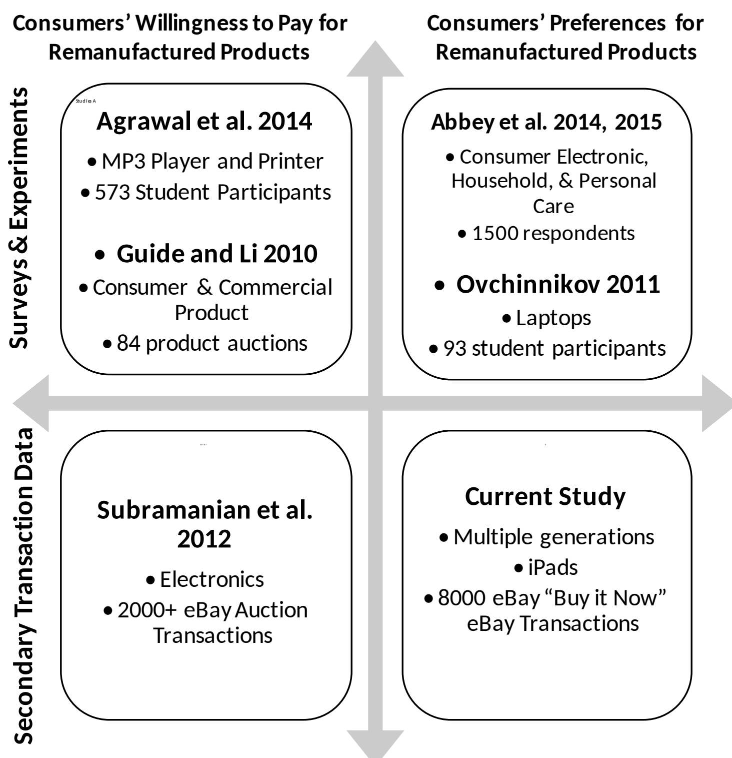


Figure 2.3: Empirical Studies in Closed Loop Supply Chains

The extant literature does not however, provide a clear understanding of how these factors influence consumers' preferences when choosing between multiple conditions (New, Refurbished, and Used) and generations of the same product (as noted earlier, a common choice set in practice). The effect of the remanufactured product's price on consumer purchasing behavior in particular is not well understood, even though

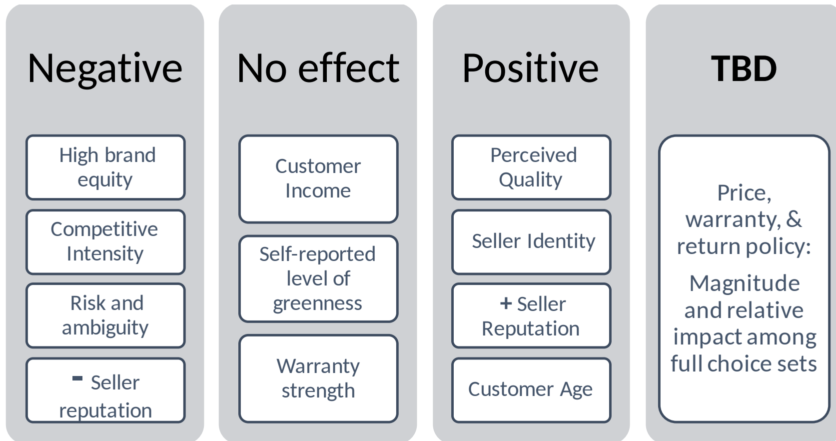


Figure 2.4: Effects of Noted Factors on Consumers' Willingness to Pay and Purchase Decisions

it is key to understanding the potential for cannibalization. As shown in Figure 2.3, we help fill this gap in the literature by being the first to use revealed preference data to calculate cross-price elasticities between product conditions and generations.

None of the studies mentioned above estimate the cross-price elasticities among the products considered in the customer's choice set. Cross-price elasticities provide a cleaner measure of how changes in the price of one condition (or generation) affects the demand of the other conditions (or generations). Further, this measure has been leveraged in previous literature to evaluate product-level competition between used and new-condition books (Ghose et al., 2006), CDs and DVDs (Smith and Telang, 2008), and successive generations of used-condition electronics (Elmaghraby et al., 2015). These papers collectively show that the degree of cannibalization varies widely by product type. For example, Smith and Telang (2008) find that the cross-price elasticity of New product demand with respect to used product prices (and consequently the potential for demand cannibalization) is far higher for CDs and DVDs than it is for books (Ghose et al., 2006).

2.2.2 Choice Modeling Techniques in Operations Management Literature

Increases in product options coupled with the accessibility of online marketplaces has dramatically expanded the number of options available to buyers (see Jagabathula, 2011, p.10). Consequently, capturing consumer choice behavior has become increasingly important to business managers (Garrow, 2016). Choice models, which allow researchers to understand how a consumer evaluates the attributes of alternatives within an offering, is one approach to studying such behavior. The discrete-choice demand model stemming from McFadden (1978) and Manski and McFadden (1981) random utility framework is, in particular, a widely-leveraged approach. The model assumes that when offered a set of alternatives, an individual assigns a utility to each attribute that influences their choice, then chooses the option that provides them their maximum utility (Garrow, 2016). Further, the model assumes that the purchase decisions of consumers are affected by the selection of products that a seller offers. In the Operations literature, choice modeling techniques have been leveraged extensively in a wide range of contexts. In retail operations, for example, researchers have used these methods to assist with assortment planning (Rusmevichientong et al., 2010; Kök and Xu, 2011; Li and Huh, 2011). In airline studies, they have been used to determine the allocation of aircraft seat capacity to multiple fare classes (Garrow and Koppelman, 2004; Adler et al., 2005; Vulcano et al., 2010). Last, in transportation analyses, these techniques have been leveraged to predict how consumers will respond to changes in existing and proposed transportation options (Lee and Waddell, 2010; Pinjari et al., 2007, 2011).

Given the ubiquity of consumers' choice behavior, several variants of discrete choice models exist (Jagabathula, 2011, see: p.21). To perform our analyses, we leverage the structural estimation technique introduced by Berry (1994) which allows for the development of models of demand and supply equations. The Berry (1994) and Berry et al. (1995) models were one of the first methods to estimate demand based on random utility maximization (RUM) models using *aggregate* market-level sales data. In the recent Operations literature, the Berry structural estimation model has been employed by Olivares et al. (2008) to study operating room staffing decisions, Allon et al. (2011) to estimate the value of reducing customer wait times in the drive-thru fast-food industry, and Guajardo et al. (2015) to examine how various product and service attributes affect demand for U.S. automobiles. Similarly, we leverage the model to understand how consumers' choices for Refurbished products are affected by seller and product related variables.

The Berry model distinctively assumes prices are endogenously determined by firms. This is in contrast to existing empirical studies on this topic that assume prices are exogenous. As the exogenous assumption has been recognized as an important shortcoming in the literature (Berry, 1994), we develop instrumental variables for price and nested market share using the sum of the other observations' characteristics (see section 5.1). Our findings indicate that when endogeneity is considered, the coefficient of price is significantly more negative. Thus, from a methodological perspective, we extend the previous research on this topic through explicitly dealing with the endogeneity of prices.

2.3 Data

We leverage transaction data on fixed price listings of a well-understood and ubiquitous product, Apple iPads, on eBay. It is estimated that 50 million iPads are sold annually and over 350 million iPads have been sold to date.

At the time of our data collection, three versions of iPads were currently available for sale on the site (and through retailers): iPad minis, iPad Airs, and classic iPads. Our dataset includes multiple generations (1 through 4) of classic iPads. We limit our sample to the classic iPads as this version had the largest number of transactions and successive versions available.

The eBay platform provides a useful data source for our research for multiple reasons. First, the site allows consumers to choose between various conditions (i.e. New, Used and Refurbished) and generations (1 through 4) of the classic iPad product. Second, eBay supplies historical information on all completed transactions, which can be readily collected by researchers. Last, the observed transactions exhibit sufficient variation on price, warranty offerings, seller reputation and other key variables of interest. Contrary to the auction format, with fixed price listings the sellers set prices for their products and consumers choose whether to purchase a given product at the stated price. Using this fixed price data, as opposed to auction results, allows for a larger sample size as nearly 90% of iPads sold on eBay are offered through this format. Additionally in practice, refurbished products are often sold by retailers (i.e. Best Buy and Amazon) and OEMs (i.e. HP, Dell, and Bose) through a fixed-priced format.

Our process for collecting data was as follows. On a monthly basis, our team identified all iPad listings that were sold, unsold (termed "completed" by eBay), and still active using eBay's advanced search tool. We extracted data from all sold product listings after applying two filters. First, we excluded all products offered from international sellers. Second, we excluded all products from our searches that were in non-working condition or that had severe defects/cosmetic issues. Listings that met the above criteria were then added to a master database. Subsequently, all viewable product and seller related variables were extracted from the identified listings.

2.3.1 Product and Transaction Related Variables

On the main product-listing page, we extracted the information about the product's price, the date and time at which the product was sold, the condition of the product, the return policy, shipping information (including shipping charges and shipping speed), and whether a warranty was offered with the product (Figure 2.5). Details about each of these variables are provided below.

Condition. Five conditions of iPad products are observed in our dataset: New, New Open-Box, Manufacturer Third-Party (3P) Certified Refurbished, Seller Refurbished, and Used. As defined by eBay, a New condition product is brand-new, unused, and in its original packaging. A New Open-Box product is similar to New, except the items' original packaging may not be present. A Manufacturer 3P Certified Refurbished product has been inspected, cleaned, and repaired and restored to a like—new condition (note the OEM does not actually do the refurbishing in the case of iPads, rather a seller or "third-party" who has been certified by Apple). A Seller Refurbished product is identical to a Manufacturer 3P Certified Refurbished product, except that a non-certified party has restored the product. Last, a Used condition product is fully operational, but may exhibit cosmetic signs of use and has not been inspected by a third party or the OEM. We thus model *CONDITION* as a categorical variable with a value set to 1 if the product is New, 2 if the product is New Open-Box, 3 if the product is Manufacturer 3P Certified Refurbished, 4 if the product is Seller Non-Certified Refurbished, and 5 if the product is Used.

Generation and Storage Size. Four generations of iPads with four different memory storage sizes are observed. We thus set our *iPad GENERATION* measure to 1 for iPad 1, 2 for iPad 2, 3 for iPad 3, and 4 for iPad 4. Our *iPad SIZE* measure is modeled according to the iPad's storage capacity: 1 for 16GB of storage, 2 for 32GB, 3 for 64GB, and 4 for 128GB.

Price and Shipping Fees. Contrary to the auction format, with fixed price listings sellers set prices for their products, and consumers choose whether to purchase a given product at the stated price. Using the fixed-price listings allows for a larger dataset but introduces endogeneity issues that arise from using set prices instead of auction prices. The final price a buyer pays for a product includes the price and the shipping fees, both of which are determined by the seller. For each transaction, we extract the values of these items as our *PRICE* and *SHIPPING* variables. Additionally we created a *SHIP* dummy variable and set the value to 0 if the seller offered free shipping and 1 if the seller charged for shipping the product. In our dataset the majority of the sellers offered free shipping with their listings, while shipping charges ranged from 0 to $41 dollars. For location-dependent shipping charges, estimates were obtained using a common address.

Return Policy and Warranty Length. In addition to price and shipping fees, sellers also decide whether to allow the consumer to return the product and, if so, whether they will refund the return shipping fees and/or charge a restocking fee. Return policy length ranged from 0 to 60 days, with the majority of listings featuring a 7-day policy. We set the *RETURN POLICY* variable to the length of the return policy with a value of 0 indicating that the seller does not accept returns. Similarly, *RESTOCK FEE* measures the value of any restocking fees charged by the seller with a value of 0 if the seller does not charge a fee. *RESHIP* indicates whether a seller pays for shipping fees and equals 0 if return shipping fees are paid by the buyer and 1 if the fees are paid by the seller. Additionally we created a *RETURN* dummy variable and set the value to 0 if the seller did not accept returns. A small percentage (approximately 3%) of products were sold with a seller's warranty, in addition to a return policy. Our *WARRANTY* measure indicates if a seller's warranty was offered with the product with 0 indicating that a warranty was not offered with the product. As the length of the warranty may vary from seller to seller, we also captured *WARRANTY LENGTH*, which measures the length of the warranty (in days) that is offered with the product with a value of 0 indicating that a warranty was not offered with the product.

Listing Time Details. To build our structural model, we captured the following temporal details for each listing. First, we captured the date and time at which the product was sold and extracted the *MONTH* out of each time stamp for aggregation purposes. Second, we captured the number of iPads of the same generation that were available at the time of purchase. We used the average number of iPad products for sale over a two-week period as a proxy for the number and types of categories included in the set of choices presented to a consumer at a given time of purchase. This information is represented by the *TOTAL ACTIVE* variable.

Other Controls. In addition to the aforementioned product-related variables, all text that appeared on the original product listing page, including all product description content, was captured as a reference to confirm the warranty length. We controlled for any product accessories that were sold with the iPad and created a dummy variable for each category of accessory: *HEADPHONES, EXTRACHGR, STYLUS, CASE, EARPHONE, DOCK, HEADSET, KEYBOARD, LIGHTCLIP, MOUNT*. In our analysis, we used sum of accessories as a single control variable rather than have 10 separate control variables since many of these variables had a value of 0 in our dataset. We also noted whether the product had 3G or 4G cellular capabilities and coded our 3G and 4G measures as 1 for products with these features and 0 for products with no cellular capabilities. Last, as a New condition product's packaging may or not be sealed, we created a dummy variable *SEALED* and set it to 1 if the iPad was sold in its original, unopened shrink wrap.

2.3.2 Seller Related Variables

Seller Reputation. When eBay members purchase or sell items, they are given the opportunity to leave feedback – often classified as positive, negative or neutral – about the transaction. Over time, eBay sellers develop a reputation based on this feedback. Abbreviated information about each seller is featured on each product–listing page in addition to a link to the seller's full feedback profile. From this profile page (Figure 2.6), we extracted several reputation measures noted in the previous literature to impact consumers' choices.

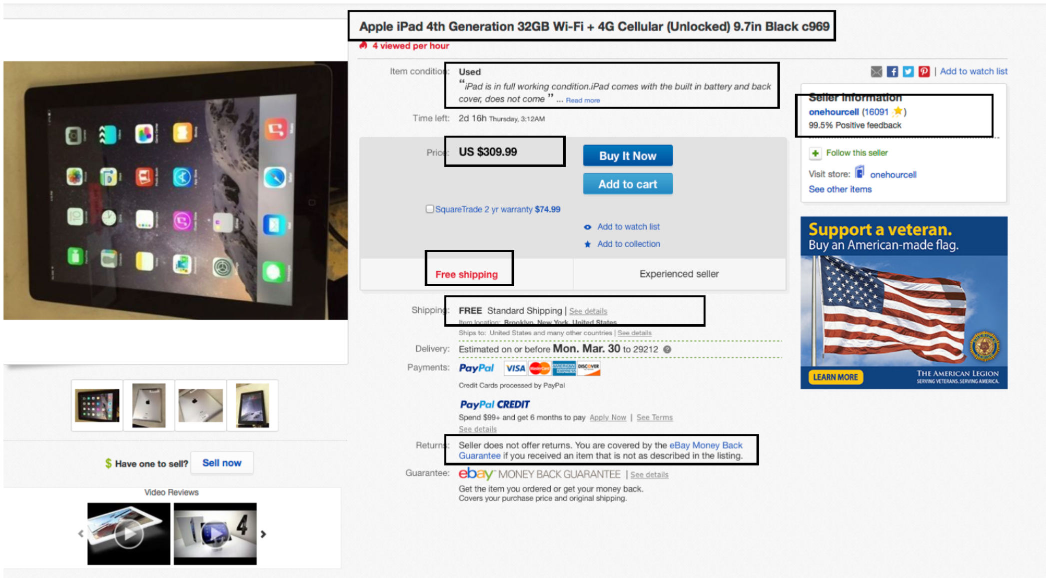


Figure 2.5: Example of Product Listing Page. Extracted Fields are Boxed.

Figure 2.6: Example of Seller Feedback Profile Page. Extracted Fields are Boxed.

Seller feedback score (provided directly below the seller name in Figure 2.6) is the ratio of positive to total reviews (i.e. positive, negative, and neutral) a seller has received. This variable is reported as a percentage and captured in our *FEEDBACKSCORE* measure.

A seller's *net* feedback score (located next to the seller name in Figure 2.6) represents the net number of positive, negative, and neutral Feedback ratings an eBay seller has received over their tenure. Sellers receive one point for each positive rating, no points for each neutral rating, and are penalized by one point for each negative rating. Given that the majority of feedback is positive, the feedback score can be viewed as a proxy for a seller's relative sales volume. We set this total to our *NET-FEEDBACK* variable. In accordance with Subramanian and Subramanyam (2012), we further separated this score into a measure of total neutral and negative feedback counts, represented by *NET NEGATIVE*, and the number of positive feedback counts as a measure of *NET POSITIVE*.

The detailed seller ratings represent the average score (out of five points) a seller has received in four specific performance categories: item as described, communication, shipping time, and shipping and handling charges. These ratings are independent of the seller's feedback and net feedback score. We thus include *ITEM DESCRIPTION SCORE, SHIPTIME SCORE, COMMUNICATION SCORE*, and *SHIPHAND SCORE* in our model as variables.

Last, *Seller incumbency* measures the length of time a seller has been a member of eBay. Similar to Subramanian and Subramanyam (2012), we measure *SELLER INCUMBENCY* as the number of days from the time of the sellers' initial registration with eBay to the start time of the product listing. A comparison of the data source and categorical and continuous variables included in the current and previous studies is shown in Figure 7.

2.3.3 Aggregation Technique and Descriptive Statistics

To prepare the data for analysis, we first segment the transactions by purchase date. Specifically, we separate the transactions by month (December 2014 to July 2015), thus defining multiple marketsİ. Since we use data from eBay, geographic conditions are less likely to impact sales. Thus, we segment markets by time so that factors like seasonality can be accurately captured. The total sales figures were extracted from eBay using the same technique leveraged to capture the primary dataset. We

	Current Study	Abbey et al 2015b	Abbey et al 2015c	Agrawal et al 2015	Guide and Li 2010	Ovchinnikov 2011	Subramanian & Subramanian 2012
Data Source							
eBay BIN Transactions	✔						✔
Apple iPads - Product Type	✔			✔		✔	
Continuous Variables							
Price	✔	✔	✔	✔	✔	✔	✔
Shipping Fees	✔						✔
Seller Incumbency	✔				✔		✔
Seller Avg. Feedback Score	✔		✔				✔
Seller Net Feedback Score	✔						✔
Seller Net Negative Score	✔						✔
Seller Net Positive Score	✔						✔
Warranty Length	✔						✔
Restock Fee	+						
Return Policy Length	+						
Seller Ship. & Hand. Score	+						
Seller Shiptime Score	+						
Seller Item Description Score	+						
Seller Communication Score	+						
Categorical Variables							
Five Product Conditions	+						
Four Generations	+						
Free Shipping Y/N	+						
Warranty Y/N	✔						✔
Return Policy Y/N	+						
Return Shipping Y/N	+						
Methodology							
Choice Model Methodology	+						
Corrections for Price Endogeneity	+						

Key:

✔ = similar or consistent with variables from previous paper

\+ = new variable or new technique employed

Figure 2.7: Comparison on Variables Included in Current and Previous Studies

aggregated the data in each market by seller. Thus, for each month, we calculated the average values for each seller that sells product condition, size and generation (see Table 2.1). Our 7,700 individual transactions representing over one million dollars of iPad products were transformed into 5,288 observations through this aggregation.

Descriptive statistics for the iPad conditions and generations are given in Tables 2.2, 2.3, and 2.6 and descriptive statistics of relevant continuous and categorical variables are given in Tables 2.5 and 2.4.

Table 2.1: Example of Aggregation Technique

month	sellerID	condtn.	size	4G	3G	gen.	ijklm	sold	avgprice	avgfdbk score
5	1010	5	1	0	0	1	1010,5,1,0,1	1	88	4391
5	1025	5	2	0	0	1	1025,5,2,0,1	1	127	66386
5	1160	5	3	0	0	1	1160,5,3,0,1	1	113.95	1088
6	1010	5	2	0	0	1	1010,5,2,0,1	1	95	4391
6	1025	5	2	0	0	1	1025,5,2,0,1	1	127	66386
6	1138	5	1	0	0	1	1138,5,1,0,1	1	109	91
7	1010	5	1	0	0	1	1010,5,1,0,1	1	88	4391
7	1010	5	2	0	0	1	1010,5,2,0,1	1	95	4391
7	1025	5	2	0	0	1	1025,5,2,0,1	1	127	66386
7	113	5	1	0	0	1	113,5,1,0,1	1	84.99	1233
7	1160	5	1	0	0	1	1160,5,1,0,1	1	85.75	1088

Table 2.2: Descriptive Statistics for iPad Conditions

Condition	Obsv.	Mean Price	Std. Dev.	Min	Max
New	209	\$343.97	89.76	134.00	694.87
New Open-Box	217	\$269.18	76.21	107.65	692.46
Manufacturer 3P Cert. Refurbished	194	\$230.92	70.40	89.99	535.00
Seller Refurbished	541	\$204.95	67.87	67.00	399.99
Used	4,127	\$180.94	71.56	60.21	499.99

Table 2.3: Descriptive Statistics for iPad Generations

iPad Generation	Obsv.	Mean Price	Std. Dev.	Min	Max
One (Oldest Model)	1,053	\$98.50	23.51	60.21	299.95
Two	1,792	\$164.05	32.35	104.94	369.99
Three	1,010	\$217.82	38.22	124.00	419.95
Four (Newest Model)	1,433	\$289.61	62.81	175.00	694.87

Table 2.4: Descriptive Statistics of Continuous Variables

Variable	Sample Size	Mean	Std. Dev.	Min.	Max.
Price	5,288	195.29	80.73	60.21	694.87
Shipping	5,288	4.18	5.58	0	41.19
Total Active	5,288	431.34	236.84	17	685
Restock Fee	5,276	0.01	0.04	0	0.20
Return Policy	5,277	10.75	11.52	0	60
Warranty Length	5,288	5.22	48.99	0	1,810.49
Seller Incumbency	4,669	3,075.84	1,198.70	14	42,208
Feedback Score	5,288	0.98	0.11	0	1
Net Feedback Score	5,288	15,791.15	58,788.47	0	638,792
Ship Hand Score	3,412	4.95	0.11	2.50	5
Shiptime Score	3,419	4.96	0.12	2.45	5
Communication Score	3,252	4.94	0.13	2.45	5
Item Description Score	3,420	4.88	0.13	2.40	5
Net Negative Score	2,762	162.26	517.13	1	3,686
Net Positive Score	4,614	2,927.55	11,360.40	1	176,132
Sealed	5,288	0.01	0.11	0	1
Sum of Accessories	5,274	0.23	0.70	0	5

Table 2.5: Descriptive Statistics of Categorical Variables

Variable	Categories	Count	Percentage
Condition	1: New	209	3.95
	2: New Open-Box	217	4.10
	3: Manufacturer 3P Cert. Refurbished	194	3.67
	4: Seller Refurbished	541	10.23
	5: Used	4,127	78.04
Generation	1	1,053	19.91
	2	1,792	33.89
	3	1,010	19.10
	4	1,433	27.10
Size	1: 16 GB	2,525	47.75
	2: 32 GB	1,666	31.51
	3: 64 GB	1,018	19.25
	4: 128 GB	79	1.49
4G	0: No 4G Available	4,858	91.87
	1: 4G Available	430	8.13
3G	0: No 3G Available	4,578	86.57
	1: 3G Available	710	13.43
Return Policy	0: No	2,313	43.74
	1: Yes	2,975	56.26
Restock Fee	0: No	4,956	93.72
	1: Yes	332	6.28
Shipping Fee	0: No	3,151	59.59
	1: Yes	2,137	40.41
Warranty	0: No	3,151	96.12
	1: Yes	175	3.88

Table 2.6: Descriptive Statistics for iPad Conditions and Generation

Generation	1			2			3			4		
Condition	Count	Percent	Avg Price	Count	Percent	Avg Price	Count	Percent	Avg Price	Count	Percent	Avg Price
1	12	1.1%	\$239.29	52	2.9%	\$261.49	19	1.9%	\$325.20	126	8.8%	\$390.81
2	22	2.1%	\$149.43	54	3.0%	\$223.88	46	4.6%	\$276.30	95	6.6%	\$319.21
3	10	0.9%	\$118.85	70	3.9%	\$179.72	40	4.0%	\$229.23	74	5.2%	\$295.42
4	72	6.8%	\$105.11	187	10.4%	\$168.60	109	10.8%	\$215.51	173	12.1%	\$279.14
5	937	89.0%	\$94.77	1429	79.7%	\$156.88	796	78.8%	\$211.61	965	67.3%	\$274.92
Total	1053			1792			1010			1433		

2.4 Analysis and Initial Results

Choice modeling is the most natural approach for determining how consumers choose between different conditions (i.e. New, Remanufactured, and Used) of the same product (Garrow, 2016). However, a direct application of choice models such as multinomial or nested logit is difficult as we do not know what other options each consumer was exposed to when they made their purchase decision. Thus, we leverage the Berry (1994) estimation method for demand estimation in differentiated markets. This model allows us to estimate the impact of price and other product and seller characteristics on demand using aggregate sales data. To develop the model, we first define the utility of an individual i purchasing j as

$$U_{ij} = \alpha p_j + \beta \mathbf{x}'_j + \xi_j + \epsilon_{ij}, \tag{2.1}$$

where $\mathbf{x}'_j$ is a vector of product and seller characteristics (return policy, product condition, seller reputation information, and shipping costs) observed by both researchers and consumers, p_j is the average price, ξ_j is a vector of product characteristics unobserved by the researchers but observed by consumers, and ϵ_{ij} is an error term representing consumer i's idiosyncratic preferences for product j. We express the aggregate utility for product j as

$$\delta_j = \alpha p_j + \beta \mathbf{x}'_j + \xi_j. \tag{2.2}$$

We use Berry's (1994) inversion method to derive the following non-nested model

$$\ln(s_j) - \ln(s_0) = \delta_j = \alpha p_j + \beta \mathbf{x}'_j + \xi_j, \tag{2.3}$$

where s_j is the market share of product j and s_0 is the market share of the outside option. We compute s_j as the total sales of classic iPads observed in our data set and s_0 as the total sales of other iPads – i.e. iPad Mini and iPad Air – not observed in our dataset (i.e. the outside option).

The non-nested model described above violates the independence of irrelevance alternatives (IIA) property. This is problematic because consumers who purchase a New condition product may be more likely to select another New condition product than a Used-condition product. A priori, we do not know what the best nesting structure is. It may be that customers are more likely to purchase within the same generation, or they may be more likely to purchase within the same condition (e.g. only buy New). Thus, we build two separate models where we nest our data both by condition in one model and by generation in the second model. A test comparing the coefficients of $log(s_{j|g})$ from both the models (see Tables 2.8 and 2.12) failed to reject the null hypothesis that both coefficients are different from each other. However, we present the nested by condition model in the main body of the paper and present the results obtained from the nested by generation model in the post-hoc section due to the higher significance levels of cross-price elasticities in the nested by condition case.

We created six nests (g= 0, 1,2,3,4,5) based on the product condition where $g = 1$ contains New product offerings, $g = 2$ contains New Open-Box offerings, $g = 3$ contains Manufacturer 3P Certified Refurbished offerings, $g = 4$ contains Seller Refurbished offerings, $g = 5$ contains Used offerings, and $g = 0$ contains the outside option only. Using the Berry inversion method, we derive the nested logit model as

$$\ln(s_j) - \ln(s_0) = \alpha p_j + \beta \mathbf{x}'_j + \sigma \log(s_{j|g}) + \xi_j, \tag{2.4}$$

where σ is a substitutability factor ($0 < \sigma < 1$), and $s_{j|g}$ is the market share of seller j within nest g.

2.4.1 Condition Nested Models

Table 2.7 presents the ordinary least squares (OLS) estimates of the condition nested model. For our return policy, shipping, and restocking fee variables, recall that we also created a binary variable with value = 0 if the original variable has value = 0 and 1 if the original variable has a value greater than 0. We add the interaction between the binary variable and the original continuous variable along with the binary variable in our estimation to account for the difference in behavior of consumers when the seller does not accept any returns, has free shipping, or has no restocking fee. As can be seen from Table 2.7, price has a negative ($\beta = -4 \times 10^{-3}$) and highly significant ($p < .001$) effect on consumers' choices. The estimate of the coefficient σ of nested market share is 0.51, which falls in the acceptable range between 0 and 1 indicating that it captures substitutability (Berry, 1994). Thus, the estimate of $\sigma = 0.51$ provides evidence that a nested model provides superior fit to a non-nested model.

It is likely that the product's price and nested market share are correlated with unobserved characteristics. For example, if there are unobserved factors (to the researcher) that may cause the demand for a particular condition or generation to be higher in a certain period, then a seller may set a higher price that period. To correct for this possible endogeneity, we use instrumental variables (IVs) for price and nested market share. The sum of the other observations' characteristics (i.e., total active, feedback score, net feedback score, and return policy) within a nest in a market are used as instruments for each observation's price and nested market share. Other observations' characteristics are appropriate instruments since they are excluded from the utility equation (u_{ij} or δ_j does not depend on product/seller characteristics of other observations) and they are correlated with prices via the markups in the first-order conditions (Berry, 1994; Berry et al., 1995). Similar set of instruments have been used in past operations management studies that have used aggregate choice models (e.g. Guajardo et al., 2015). We use the two stage least squares (2SLS) estimation method (Angrist et al., 1995) to get the IV estimates. We tested for endogeneity using both Durbin and Wu-Hausman scores. The Durbin χ^2 statistic was 691.09 ($p < .001$) and the Wu-Hausman F statistic was 512.28 ($p < .001$). These results indicate the presence of endogeneity of price and nested market share.

The results for the condition nested model with instruments are shown in Table

Table 2.7: Condition Nested Model: OLS Estimates [DV: $\ln(s_j) - \ln(s_0)$]

Variable	**Coefficient**	**(Std. Err.)**
Price	$-4.00 \times 10^{-3***}$	(3.38×10^{-4})
$\log(s_{j\|g})$	0.51^{***}	(0.01)
Total Active	$2.35 \times 10^{-3***}$	(8.17×10^{-5})
Feedback Score	3.64×10^{-3}	(1.24)
Net Feedback Score	$5.82 \times 10^{-7***}$	(1.52×10^{-7})
Warranty	-0.03	(0.06)
Communication Score	$0.34^{\wedge}$	(0.18)
Shiphand Score	0.25	(0.18)
Shiptime Score	-0.09	(0.17)
Item Description Score	-0.25	(0.19)
Warranty Length	-9.84×10^{-5}	(3.71×10^{-4})
Seller Incumbency	7.04×10^{-6}	(6.05×10^{-6})
Net Positive Score	$-2.30 \times 10^{-6**}$	(8.24×10^{-7})
Net Negative Score	-2.59×10^{-4}	(2.54×10^{-4})
Sealed	0.05	(0.10)
Accessories	-0.06^{***}	(0.02)
Return	0.02	(0.04)
Return $*$ Return Policy	$4.80 \times 10^{-3***}$	(1.17×10^{-3})
Shipping	0.23^{***}	(0.05)
Shipping $*$ Shipping Fee	-0.01^{**}	(5.23×10^{-3})
Restock	0.10	(0.12)
Restock $*$ Restock Fee	-0.50	(0.68)
3G	0.05	(0.03)
4G	0.05	(0.04)
iPad Size 32 GB^a	-0.04	(0.02)
iPad Size 64 GB^a	0.11^{***}	(0.03)
iPad Size 128 GB^a	0.61^{***}	(0.10)
iPad Generation 2^b	0.18^{***}	(0.04)
iPad Generation 3^b	0.98^{***}	(0.05)
iPad Generation 4^b	0.67^{***}	(0.07)
Intercept	-7.94^{***}	(1.22)

a: Holdout Group: IPad Size 16 GB, b: Holdout Group: IPad Generation 1, $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\wedge}p < .1$

2.8. The coefficient of price is negative ($\beta = -0.03$) and statistically significant ($p < .001$). As expected, the coefficient of price is more negative in the model with instruments than without instruments. This provides further support that price was endogenous with the nested market shares. The estimate of $\sigma = 0.41$ remains in the acceptable range to support nesting (Berry, 1994). The coefficient of net feedback score is positive and significant ($\beta = 1.24 \times 10^{-6}$, $p < .001$), while the coefficient of net negative score is negative and significant ($\beta = -1.78 \times 10^{-4}$, $p < .01$). The

Table 2.8: Condition Nested Model: IV Estimates [DV: $\ln(s_j) - \ln(s_0)$]

Variable	**Coefficient**	**(Std. Err.)**
Price	-0.03^{***}	(5.95×10^{-3})
$\log(s_{j\|g})$	0.41^{***}	(0.07)
Total Active	-5.00×10^{-5}	(2.72×10^{-4})
Feedback Score	2.93	(2.63)
Net Feedback Score	$1.24 \times 10^{-6***}$	(3.37×10^{-7})
Warranty	0.11	(0.12)
Communication Score	$-0.87^{\wedge}$	(0.47)
Shiphand Score	-0.65	(0.40)
Shiptime Score	0.27	(0.37)
Item Description Score	0.63	(0.44)
Warranty Length	-3.08×10^{-4}	(7.71×10^{-4})
Seller Incumbency	1.15×10^{-5}	(1.26×10^{-5})
Net Positive Score	-1.57×10^{-6}	(1.71×10^{-6})
Net Negative Score	$-1.78 \times 10^{-4**}$	(6.71×10^{-5})
Sealed	2.80^{***}	(0.59)
Accessories	$-0.07^{\wedge}$	(0.04)
Return	0.03	(0.08)
Return $*$ Return Policy	$4.63 \times 10^{-3\wedge}$	(2.46×10^{-3})
Shipping	-0.02	(0.16)
Shipping $*$ Shipping Fee	0.01	(0.01)
Restock	-0.23	(0.26)
Restock $*$ Restock Fee	0.84	(1.46)
3G	-0.10	(0.07)
4G	0.52^{***}	(0.14)
iPad Size 32 GB[a]	0.46^{***}	(0.11)
iPad Size 64 GB[a]	1.07^{***}	(0.23)
iPad Size 128 GB[a]	4.57^{***}	(0.94)
iPad Generation 2[b]	2.14^{***}	(0.42)
iPad Generation 3[b]	3.55^{***}	(0.66)
iPad Generation 4[b]	4.47^{***}	(0.89)
Intercept	-3.57	(2.73)

a: Holdout Group: IPad Size 16 GB, b: Holdout Group: IPad Generation 1, $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\wedge}p < .1$

coefficient of sealed is positive ($\beta = 2.80$) and highly significant ($p < .001$). Finally, the coefficients of sizes and generations are positive and significant, i.e. all else equal (including price), demand for larger memory and a newer generation is higher. We conducted further tests for the validity of our instruments. In the first stage of 2SLS, the tests for excluded instruments for both price and nested market share reject the null hypothesis of excluded instruments having no explanatory power. The F-statistic (p-value) for price, and nested market share are 22.74 ($p < .001$), and

334.09 ($p < .001$), respectively. These results support the strength of our instruments (Staiger and Stock, 1994). Next, we ran the test of underidentification of instruments and the null hypothesis that our instruments are underidentified is rejected (the Anderson Canonical Correlation LM Statistic is 28.45 ($p < .001$)). These tests provide further validity to our model specification and the use of instruments to address the endogeneity of price, and the nested market share variables (Guajardo et al., 2015).

2.5 Discussion of Main Results

Using the estimates in Table 2.8 and Equation 2.5, we compute the product's own and cross-price elasticities using conditions as nests, where $i, j \in 1, 2, ...5288$. For the own-price elasticities shown in Table 2.9, we estimate whether the own-price elasticities of product conditions g ($g = 2, 3, 4, 5$) are different than the own price elasticity of product condition 1 at the 95% significance level. We use the upper and lower confidence interval estimates of α and σ to estimate our significance level by checking to see if the elasticity estimates (calculated using either the upper or lower values from the 95% confidence intervals) have any overlaps. We estimate whether the cross-price elasticities of product pair (i, j) $[i \neq j, (i, j) \neq (2, 1)]$ are different from the cross price elasticity of product type pair $(2, 1)$ at the 95% significance level using the upper and lower confidence interval estimates of α and σ.

$$\frac{\partial s_i}{\partial p_j}\frac{p_j}{s_i} = \begin{cases} \alpha p_i \left[\frac{1}{1-\sigma} - \left(\frac{\sigma}{1-\sigma}\right) s_{i|g}\right] - \alpha p_i s_i & \text{if } i = j, \\ -\alpha p_j \left[\left(\frac{\sigma}{1-\sigma}\right) s_{j|g} + s_j\right] & \text{if } i \neq j,\ i \in g, j \notin g \\ -\alpha p_j s_j & \text{if } i \neq j,\ i, j \in g. \end{cases} \tag{2.5}$$

2.5.1 Own and Cross-Price Elasticity Calculations by Condition

In Table 2.9, the own and cross-price elasticities are averaged and reported at the condition level where we control for generation. The diagonal represents average own-price elasticities for each of the five conditions, and the off-diagonal represents average cross-price elasticity values for each condition pairing. The cross-price elasticity values show how price changes in the conditions listed in the columns affect market share for the conditions listed in the table rows. For example, a 1% increase in the New condition price (first column), increases market share for New Open-Box by 0.18%, Manufacturer 3P Certified Refurbished by 0.17%, Seller Refurbished by 0.18%, and Used by 0.18%. The own price elasticity values shows the resulting percent decrease in market share when the price of a single iPad increases by 1%. For example, a percent increase in the price of a single seller's Seller Refurbished condition iPad, decreases market share for that seller by 10.59% on average.

From the own-price elasticity results, we find that the New product condition category is the most sensitive to changes in its own price, followed by the New Open-Box-box category, Manufacturer 3P Certified Refurbished, Seller Refurbished and Used category. Further, we note significant differences in the own-price elasticity values across conditions. For example, a 1% increase in New-condition products' price decreases market share for New products by approximately 17.60%. In contrast, a 1% increase in the price of Used-condition products' price decreases market share for Used products by only 9.4%. A possible explanation for this result is that consumers are more sensitive to changes in prices for product conditions that are more standardized in nature i.e., the quality of a New condition product will not vary much across sellers whereas the quality of a Refurbished or Used condition product is more prone to vary. This finding is consistent with the previous literature - both Abbey et al. (2015) and Ovchinnikov (2011) discuss how consumers are often uncertain about the quality of Refurbished products.

Also from the own-price elasticity calculations, we find that for remanufactured products, the effect of changes in price differs slightly according to the seller's classification: a 1% increase in the price decreases market share for Manufacturer 3P

Certified Refurbished products by approximately 11.78% while a 1% increase in the price decreases demand for Seller Refurbished (non-certified) products by 10.59%. Although Subramanian and Subramanyam (2012) find that consumers are willing to pay slightly more on average for OEM-certified Refurbished products, our results suggest that consumers are also more sensitive to increases in OEM-certified prices (versus increases in non-certified Refurbished products' prices). This additional sensitivity to OEM-certified product prices could offset the benefits of higher prices if sellers of these products are not aware of the highlighted market effects.

From the cross-price elasticity values, we find that gains in market share of the other four conditions are higher when the price of New products increases as compared to similar increases in the price of Used products. Further, we are able to directly examine the threat of cannibalization from Refurbished products. Our results allow us to test the effects of (1) changes in New products' prices on Refurbished products' market share and (2) changes in Refurbished products' prices on New products' market share.

We find evidence that changes in New products' prices *do* impact the market share of Refurbished condition goods; however, our results suggest that the overall threat is moderate and is influenced by the sellers' classification: a 1% increase in the price of New-condition goods increases the market share of Refurbished condition products by only 0.17% to 0.18%. Further, we find that the degree to which Refurbished-condition products cannibalize New products' sales is no more than the amount of cannibalization posed by any of the other product conditions (i.e. New Open-Box and Used). Thus, Refurbished products do not appear to present a cannibalization threat to New product sales any more than do Open-Box or Used products.

Our results also suggest that changes in remanufactured products' price affect the market share of New and New Open-Box conditions differently: a 1% increase in the price of Manufacturer 3P Certified Refurbished products' price increases the market share of New and New Open-Box conditions by only 0.13%, while a 1% increase in the price of Seller Refurbished products' price increases the market share of New and New Open-Box conditions by 0.05% and 0.04% respectively. With regards to the two categories of remanufactured products, we find that Manufacturer 3P Certified Refurbished products are much closer substitutes to New-condition products than Seller Refurbished products (as shown in Figure 2.8). A 1% increase in the price of Manufacturer 3P Certified Refurbished products increases the market share of New-conditions products by 0.13% whereas a 1% increase in the price of Seller Refurbished products increases the market share of New-condition products by only 0.05%. This result supports and extends the findings of the Agrawal et al. (2015) study, which notes that the presence of Seller Refurbished products increases the perceived value of New products. We find in addition that Seller Refurbished products appear to be perceived as poor substitutes for New condition products and thus pose little threat of cannibalization to New condition goods.

Last, from the cross-price elasticity calculations, we find evidence that competition between the two categories of remanufactured products does exist, but the effects are asymmetric. Manufacturer 3P Certified Refurbished products are much more vulnerable to this competition than Seller Refurbished products. A 1% increase in the price of Manufacturer 3P Certified Refurbished products increases demand for seller refurbished products by 0.12%, whereas a 1% increase in Seller Refurbished products increases demand for Manufacturer 3P Certified Refurbished products by on 0.04%.

2.5.2 Elasticities of Other Variables: Nested on Condition

Table 2.10 gives the elasticities of other significant variables for the overall sample as well as the average values for each condition. The elasticity values show how changes in the listed seller characteristics affect market share. These results add to the debate

Table 2.9: Own and Cross-Price Elasticities by Condition

Condition	New	New Open-Box	Mfr. 3P Cert. Refurbished	Seller Refurbished	Used
New	-17.60	0.15	$0.13^{\ddagger}$	$0.05^{\ddagger}$	$6.3 \times 10^{-3\ddagger}$
New Open-Box	0.18	-13.77	$0.13^{\ddagger}$	$0.04^{\ddagger}$	$5.6 \times 10^{-3\ddagger}$
Mfr. 3P Cert. Refurbished	0.17	$0.13^{\ddagger}$	$-11.78^{\dagger}$	$0.04^{\ddagger}$	$5.1 \times 10^{-3\ddagger}$
Seller Refurbished	0.18	0.13	$0.12^{\ddagger}$	$-10.59^{\dagger}$	$5 \times 10^{-3\ddagger}$
Used	0.18	$0.13^{\ddagger}$	$0.11^{\ddagger}$	$0.04^{\ddagger}$	$-9.40^{\dagger}$

† denotes own price elasticity of product type j ($j = 2, 3, 4, 5$) is different from own price elasticity of product type 1 at 95% significance level. ‡ denotes cross price elasticity of product type pair (i, j) $[i \neq j, (i, j) \neq (2, 1)]$ is different from cross price elasticity of product type pair $(2, 1)$ at 95% significance level.

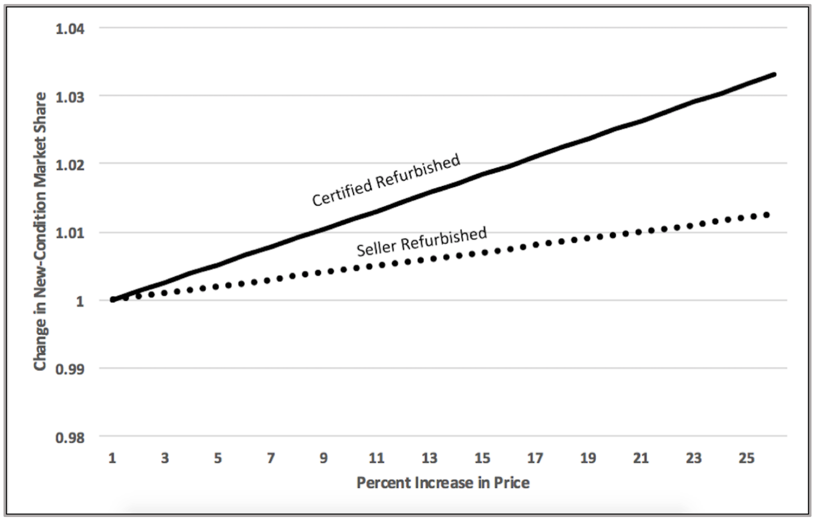


Figure 2.8: Change in Market Share (ms) of New Products from Price Increases of Refurbished Products

of whether price or other attributes dominate consumer choices (Ovchinnikov, 2011; Abbey et al., 2015). Our results indicate that price is by far the principal factor and, from a managerial perspective, the most important factor. For example, a 1% increase in the net feedback score and net negative score measure impact market share on average across all conditions by only 0.036% and 0.084% respectively, while a 1% increase price impacts market share on average by 12.61% (average of all own-price elasticities). Regarding seller attributes, we find that the two categories of Refurbished products are the most elastic to changes in net feedback scores followed by, New Open-Box, Used, and finally, New products. Although the two categories of Refurbished products react identically to changes in net feedback score, we find that there are significant differences to their changes in net negative scores. While the net

Table 2.10: Elasticities of Other Variables: Nested on Condition

Condition	Net Feedback Score Elasticity	Net Negative Score Elasticity
New	0.02	-0.05
New Open-Box	0.03	-0.04
Manufacturer 3P Cert. Refurbished	0.05	-0.19
Seller Refurbished	0.05	-0.11
Used	0.03	-0.03
Overall	0.036	-0.084

negative seller feedback score highly influences demand for Manufacturer 3P Certified Refurbished products, its influence on demand is for Seller Refurbished products is almost 50% less: a 1% increase in a sellers' net feedback score decreases market share by 0.19% for Manufacturer 3P Certified Refurbished products a 1% increase in a sellers' net feedback score decreases market share by 0.11% for Seller Refurbished products.

2.6 Post-Hoc and Robustness Tests

In this section, we consider a number of extensions and post-hoc and robustness tests to the previous analyses. To further explore the differences in elasticities between products of different generations, we estimate the interaction effects between generation and different variables. Table 2.11 gives the estimates of a model with interaction effects between generation and seller incumbency (Model 3), and generation and net feedback score (Model 4). As can be seen, the effect of seller incumbency on market share is greater for second-generation ($\beta = 5.81 \times 10^{-5}$, $p < .1$) and third-generation ($\beta = 9.18 \times 10^{-5}$, $p < .05$) products as compared to first-generation products. Similarly, the effect of net feedback score on market share is marginally higher for fourth-generation products (1.81×10^{-6}, $p < .1$) as compared to first-generation products.

Table 2.11: Condition Nested Model: IV Estimates with Interaction Effects with Generation [DV: $\ln(s_j) - \ln(s_0)$]

Variable	**Model 3**	**Model 4**
Price	-0.03^{***}	-0.03^{***}
$\log(s_{j\|g})$	0.41^{***}	0.41^{***}
Total Active	-3.91×10^{-5}	-5.06×10^{-5}
Feedback Score	2.95	2.75
Net Feedback Score	$1.24 \times 10^{-6***}$	3.32×10^{-7}
Warranty	0.08	0.12
Communication Score	$-0.86^{\wedge}$	$-0.83^{\wedge}$
Shiphand Score	-0.60	$-0.68^{\wedge}$
Shiptime Score	0.23	0.30
Item Description Score	0.62	0.58
Warranty Length	-2.33×10^{-4}	-2.80×10^{-4}
Seller Incumbency	-3.04×10^{-5}	(1.34×10^{-5})
Net Positive Score	-1.82×10^{-6}	$-3.13 \times 10^{-6\wedge}$
Net Negative Score	$-1.76 \times 10^{-4**}$	$-1.87 \times 10^{-4**}$
Sealed	2.79^{***}	2.79^{***}
Accessories	$-0.06^{\wedge}$	$-0.07^{\wedge}$
Return	0.03	0.03
Return $*$ Return Policy	$4.74 \times 10^{-3\wedge}$	$4.68 \times 10^{-3\wedge}$
Shipping	-0.02	-0.04
Shipping $*$ Shipping Fee	0.01	0.01
Restock	-0.24	-0.22
Restock $*$ Restock Fee	0.88	0.78
3G	-0.10	-0.09
4G	0.52^{***}	0.51^{***}
Seller Incumbency $*$ Generation 2	$5.81 \times 10^{-5\wedge}$	-
Seller Incumbency $*$ Generation 3	$9.18 \times 10^{-5*}$	-
Seller Incumbency $*$ Generation 4	1.32×10^{-5}	-
Net Feedback Score $*$ Generation 2	-	2.99×10^{-7}
Net Feedback Score $*$ Generation 3	-	1.32×10^{-6}
Net Feedback Score $*$ Generation 4	-	$1.81 \times 10^{-6\wedge}$
Intercept	-3.50	-3.34

a: Holdout Group: Interaction with Generation 1–New. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\wedge}p < .1$. Size, and Generation control estimates not shown for brevity.

Table 2.12: Generation Nested Model: IV Estimates with Interaction Effects [DV: $\ln(s_j) - \ln(s_0)$]

Variable	**Base Model**	**Model 1**	**Model 2**
Price	$-1.27 \times 10^{-3*}$	$-1.11 \times 10^{-3*}$	$-8.96 \times 10^{-4\wedge}$
$\log(s_{j\|g})$	0.33^{***}	0.33^{***}	0.33^{***}
Total Active	$3.62 \times 10^{-4*}$	$4.03 \times 10^{-4*}$	$4.92 \times 10^{-4**}$
Feedback Score	0.09	0.18	0.56
Net Feedback Score	$4.61 \times 10^{-7**}$	$4.56 \times 10^{-7**}$	$-1.80 \times 10^{-6\wedge}$
Warranty	0.05	0.04	0.06
Communication Score	0.59^{**}	0.56^{**}	0.55^{**}
Shiphand Score	0.09	0.09	0.14
Shiptime Score	$-0.29^{\wedge}$	-0.27	-0.34^{*}
Item Description Score	-0.37^{*}	$-0.35^{\wedge}$	$-0.32^{\wedge}$
Warranty Length	-1.55×10^{-4}	-1.51×10^{-4}	-1.88×10^{-4}
Seller Incumbency	$1.16 \times 10^{-5*}$	$-4.18 \times 10^{-5\wedge}$	$1.11 \times 10^{-5\wedge}$
Net Positive Score	$-2.14 \times 10^{-6**}$	$-2.18 \times 10^{-6**}$	-8.23×10^{-7}
Net Negative Score	3.65×10^{-5}	3.96×10^{-5}	$6.99 \times 10^{-5**}$
Sealed	0.10	0.07	0.05
Accessories	-0.09^{***}	-0.09^{***}	-0.09^{***}
Return	0.05	0.05	0.04
Return $*$ Return Policy	$5.60 \times 10^{-3***}$	$5.52 \times 10^{-3***}$	$5.72 \times 10^{-3***}$
Shipping	0.23^{***}	0.24^{***}	0.23^{***}
Shipping $*$ Shipping Fee	-0.02^{***}	-0.02^{***}	-0.02^{***}
Restock	0.16	0.16	0.15
Restock $*$ Restock Fee	-0.89	-0.89	-0.86
3G	0.16^{***}	0.16^{***}	0.15^{***}
4G	-0.11^{*}	-0.13^{**}	-0.14^{**}
Seller Incumbency $*$ Condition 2	-	4.09×10^{-5}	-
Seller Incumbency $*$ Condition 3	-	5.45×10^{-5}	-
Seller Incumbency $*$ Condition 4	-	$6.66 \times 10^{-5*}$	-
Seller Incumbency $*$ Condition 5	-	$5.51 \times 10^{-5*}$	-
Net Feedback Score $*$ Condition 2	-	-	3.13×10^{-7}
Net Feedback Score $*$ Condition 3	-	-	2.01×10^{-6}
Net Feedback Score $*$ Condition 4	-	-	9.96×10^{-7}
Net Feedback Score $*$ Condition 5	-	-	$2.30 \times 10^{-6*}$
Intercept	-7.22^{***}	-7.20^{***}	-7.75^{***}

a: Holdout Group: Interaction with Condition 1–New. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\wedge}p < .1$. Size, and Condition control estimates not shown for brevity.

2.6.1 Generation Nested Model

As an alternate model, we nested our observations based on the product's generation instead of its condition, i.e., we assume that consumers are more likely to substitute across generations than across conditions. Thus, a consumer who is considering a fourth-generation iPad is more likely to consider another fourth-generation iPad (in a different condition) than a third-generation iPad. This nested model also allows us to test for any interaction effects of different variables with condition. In this case, we estimate Equation 2.4 by nesting our observations based on generation. Columns 2, 3, and 4 of Table 2.12 give the estimates of the base model (without any interaction effects), the model with interaction effects between condition and seller incumbency (Model 1), and between condition and net feedback score (Model 2) respectively. As can be seen, the effect of seller incumbency on the market share is significantly higher for Seller Refurbished ($\beta = 6.66 \times 10^{-5}$, $p < .05$), and Used products ($\beta = 5.51 \times 10^{-5}$, $p < .05$) as compared to New products. Similarly, the effect of the net feedback score on market share is significantly higher for Used products ($\beta = 2.30 \times 10^{-6}$, $p < .05$) as compared to New products.

2.6.2 Own and Cross-Price Elasticity Calculations by Generation

Using the estimates in Table 2.13 and Equation 2.5, we compute the product's own and cross-price elasticities using *generations* as nests, where $i, j \in 1, 2, ...5288$. For the own-price elasticities shown in Table 2.13, we estimate whether the own-price elasticities of product generations g ($g = 2, 3, 4$) are different than the own price elasticity of product generation 1 at the 95% significance level. Similar to the process employed in section 5 we use the upper and lower confidence interval estimates of α and σ to estimate our significance level by checking to see if the elasticity estimates (calculated using either the upper or lower values from the 95% confidence intervals) have any overlaps. We estimate whether the cross-price elasticities of product pair (i, j) $[i \neq j, (i, j) \neq (2, 1)]$ are different from the cross price elasticity of product type pair $(2, 1)$ at the 95% significance level using the upper and lower confidence interval estimates of α and σ. In Table 2.13 we give the cross-price elasticities averaged and reported at the generation level where we control for product condition. The diagonal represents the average own-price elasticities for each of the four generations. From the own-price calculations, we find that the most recent model of iPads (i.e. generation

Table 2.13: Own and Cross-Price Elasticities by Generation

Generation	One	Two	Three	Four
One (Oldest Model)	-0.18	2.26×10^{-4}	5.10×10^{-4}	7.20×10^{-4}
Two	2.17×10^{-4}	-0.31	5.10×10^{-4}	7.20×10^{-4}
Three	2.18×10^{-4}	2.27×10^{-4}	-0.41	7.26×10^{-4}
Four (Newest Model)	2.42×10^{-4}	2.91×10^{-4}	5.85×10^{-4}	-0.54

4) are more sensitive to changes in price: a 1% increase in price for generation 4, decreases market share by 0.54%, whereas a 1% increase in price for generation 1 (the oldest iPad version) decreases market share by only 0.18%. Further, from the cross price elasticity calculations, we discover that when customers switch from generation 4 (because of a price increase) then they do so in order of generation (first 3rd, then 2nd then 1st). When there is a price increase in any of the other generations (1 - 3) however, the customers' choice is always to go to generation 4 first. Thus, a price increase in generation 1 causes customers to first switch to generation 4, then 3, then 2. Our results indicate that just as the price of New condition goods has a disproportionate influence on the demand for the other conditions, so does the price for the latest generation. We note however that none of our elasticity measures at the generation level are significant at the 95% level. This finding indicates that *the potential of cannibalization across conditions is greater than the threat of cannibalization across generations*.

2.6.3 Elasticities of Other Variables: Nested on Generation

Table 2.14 gives the elasticities of other significant variables for the overall sample as well as the average values for each generation. The elasticity values show how changes in the listed seller characteristics affect market share. Similar to the nested on condition results presented in section 5, our results indicate that price is by far the principal factor: a 1% increase in the net feedback score and seller incumbency score measure impact market share on average across all generations by 0.01% and

0.05% respectively, a 1% increase price impacts market share on average by 0.36% (average of all own-price elasticities from Table 2.13). Regarding seller attributes, we find that newer generation products are more sensitive to changes in seller attributes. For example, the impact of sellers' net feedback score on market share is nearly twice as large for fourth generation products compared to first generation products: a 1% increase in a sellers' net feedback score decreases market share by 0.01% for fourth generations products whereas a 1% increase in a sellers' net feedback score decreases market share by 0.005% for first generation products.

Table 2.14: Elasticities of Other Variables: Nested on Generation

Generation	Net Feedback Score Elasticity	Seller Incumbency Elasticity
1	5.52×10^{-3}	0.05
2	9.40×10^{-3}	0.05
3	0.01	0.05
4	0.01	0.06
Overall	0.01	0.05

2.6.4 Data Availability Robustness Check

Chintagunta and Dube (2005) and Bruno and Vilcassim (2008) point out that the Berry (1994) model, and its extensions yield biased estimates when some of the product types have zero sales during some of the time periods. As a robustness check, we used data from only those months in which all product types were sold. We have three months of such data. Table A.1 gives the estimates of the main nested model. As can be seen, the estimates are qualitatively similar to our estimates from the full dataset shown in Table 2.8. These results support the validity of our findings.

In addition to the above general observations, additional insights can be gleaned from a comparative look at Figures 2.9 and 2.10, which will be done in the next section to outline our managerial insights.

2.7 Conclusion

The refurbished product industry has experienced rapid growth, yet still remains a small fraction of the total market for new product sales. One of the main reasons it has not grown larger is that few OEMs understand how the sales of remanufactured products impact the sales of new products, thus causing them to try to avoid any competition from secondary markets (USITC, 2012a). Although a recent stream of empirical studies has explored how price and other key factors influence consumers' attitudes towards refurbished goods relative to new goods, most only compare consumers' purchasing behavior towards new and refurbished versions of the same products. However, the advent of consumer-to-consumer marketplaces such as eBay along with advances in product durability levels have facilitated the creation of used-product markets that feature a dramatically wider selection of product conditions and generations that consumers can choose from (Ghose, 2014). In response, we study how consumers evaluate refurbished products offered within a *product portfolio* of different generations and of different conditions of the same product, including Used, Open-Box and OEM Certified Refurbished. Among our main findings, as summarized in Figures 2.9 and 2.10, are that product generation, condition, and seller attributes are all highly influential in shaping consumers' purchasing decisions, and that the relationship between new and refurbished products is much more nuanced and context-specific than previously thought.

Through leveraging secondary transaction data on the sales of different generations and different conditions of Apple iPads on eBay, we calculate cross-price elasticity values for each product generation and condition pairing and provide empirical clarity to the cannibalization debate. Figure 2.11 summarizes the cross-price elasticity results between New and Refurbished condition products. We find that, when the consumers' choice set is explicitly considered, Refurbished products pose little threat to New products and to other product conditions. Specifically, the cross-price

	New	New Open-Box	Manufacturer 3P Cert Refurbished	Seller Refurbished	Used
Consumers' First Alternative Choice	New Open-Box, Seller Refurbished, and Used	New	New & New Open-Box	New	New
Own Price Elasticity	-17.60	-13.77	-11.78	-10.59	- 9.4
Net Feedback Score Elasticity	0.02	0.03	0.05	0.05	0.03
Net Negative Score Elasticity	-0.05	-0.04	-0.19	-0.11	-0.03

Figure 2.9: Summary of Condition Findings

	Generation One	Generation Two	Generation Three	Generation Four
Consumers' First Alternative Choice	Generation Four	Generation Four	Generation Four	Generation Four
Own Price Elasticity	-0.18	-0.31	-0.41	-0.54
Net Feedback Score Elasticity	5.52 E-3	9.4 E-3	0.01	0.01
Seller Incumbency Elasticity	0.05	0.05	0.05	0.06

Figure 2.10: Summary of Generation Findings

elasticity of New product condition demand with respect to Manufacturer 3P Certified Refurbished-condition prices is 0.13% (line A). Our results also indicate that Refurbished products pose a nearly equivalent threat to other product conditions as they do to New items. Managers of New products may be tempted to perceive Manufacturer 3P Certified and Seller Refurbished products as equivalent threats; however, our research suggests that Manufacturer 3P Certified products are much closer substitutes. Our cross-price elasticity calculations suggest that changes in Manufacturer 3P Certified products prices threaten the market share of New and New Open-Box approximately three times more than changes in Seller Refurbished products prices (lines A and B). However, we also find that the threat posed by Manufacturer 3P Certified Refurbished products to New goods is no different than the threat that New Open-Box products pose to New condition products (lines A and C). Last, we find ev-

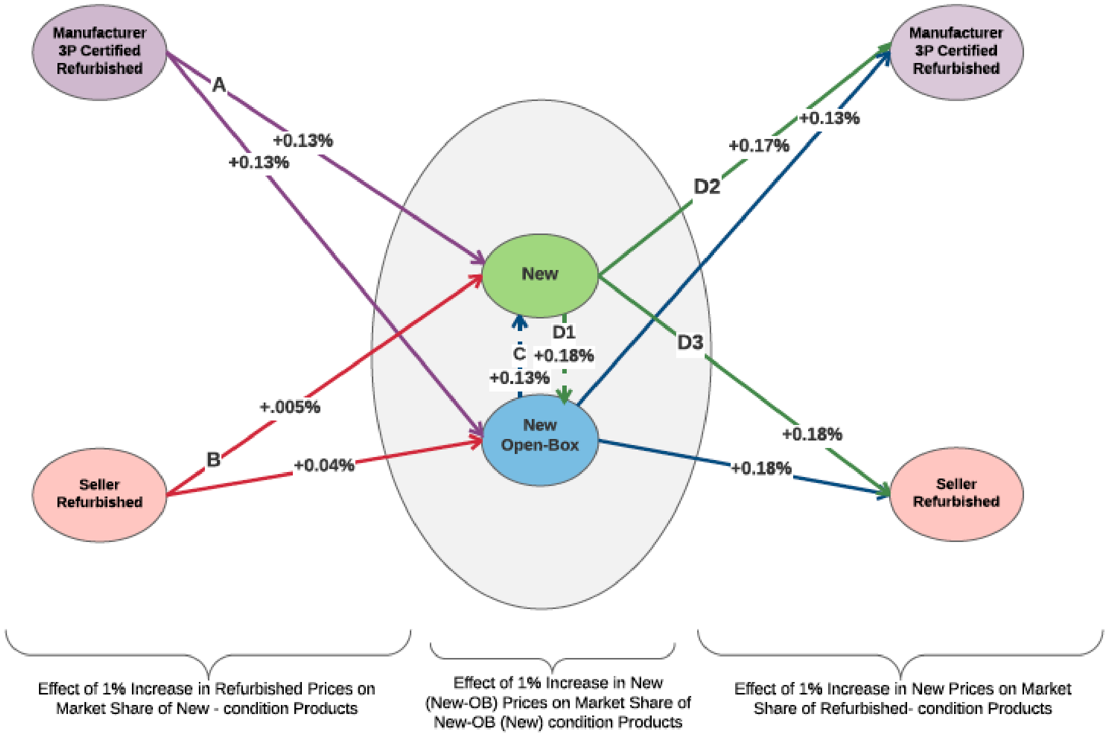


Figure 2.11: Summary of Cross-price Elasticities Between New and Refurbished Conditions

idence that changes in New products' prices impact the market share of Refurbished and other condition goods almost equivalently (lines D1, D2, and D3SS).

We find that evidence that competition between the two categories of remanufactured products does exist and that Manufacturer 3P Certified Refurbished products are three times more vulnerable to this competition than Seller Refurbished products. In addition, from our own-price elasticity results, we find that Manufacturer 3P Certified Refurbished products are more susceptible to within-category competition than Seller Refurbished products. However, overall, our results suggest that competition between the two refurbished categories and between new and refurbished conditions is modest, and thus should not strongly discourage firms from entering the remanufacturing industry.

Next, we find that the threat of cannibalization across conditions is much stronger than the threat of cannibalization between generations of products. Further, our own and cross-price elasticity calculations allow us to specifically determine those products for which sellers should be most attentive to changes in price (see Figures 2.9 and 2.10). For example, a seller of Used-condition products may be inclined to primarily monitor the prices of Seller Refurbished products; however, our results indicate that New condition product prices have the largest influence on, and thus are the biggest threat to, demand for Used-condition products.

Across all product generations and conditions, we find that the product's price and the seller's reputation measures (such as net feedback score) are significant predictors of consumer purchasing behavior. Product price is by far the dominant factor in shaping consumers' choices, but we find that the effect is not homogeneous across product conditions. We also find evidence that previously unstudied factors such as return policy length and the *communication* seller rating impact consumers' choices. Consistent with the previous literature, these results in sum support the need for all sellers in secondary markets to (1) build and communicate their reputations (2) understand the effects of changes of pricing on demand, and (3) strategically leverage return policies to stimulate sales (Ovchinnikov, 2011; Subramanian and Subramanyam, 2012; Abbey et al., 2015). In addition to these results, by including multiple generations and conditions of products in our study, we are able to provide further managerial insights specific to the studied products' categorization.

Our results suggest that sellers of newer-generation products should be conscious of their net feedback score measure, as consumers of more recent products focus on this attribute. Sellers of older generation products, however, should seek to build trust through highlighting their tenure as a seller. Our results also indicate that sellers of Refurbished products should be aware of differences in consumers' perceptions of remanufactured products, according to their sellers' classification. For example, our findings indicate that while changes in net feedback score influence market share for

Manufacturer 3P Certified and Seller Refurbished equally, the effect of changes in the net negative measure on market share are nearly two times larger for Manufacturer 3P Certified products than Seller Refurbished goods.

Both research and practice can benefit from related future research, and the following limitations from our study provide promising trajectories for further analyses. First, our study involved only one product (the Apple iPad) and product type (Electronics). Future research may consider replicating our analysis across other product categories and brands of interest, which may vary according to brand equity strength, product life cycle stage, and product durability levels. Second, we studied consumers' purchasing behavior on only one platform (eBay). As prior research suggests consumers' purchasing behavior may be influenced by shopping channel, it would be interesting to study how consumers' acceptance of remanufactured products vary across different secondary market platforms (i.e. retail managed sites such as Amazon, OEM managed sites such as HP or Dell, and third party logistic managed platforms such as Optoro or Newegg).

Chapter 3

Promoting E-Recycling: Effects of Electronic Waste Legislation and Consumer Attributes on Recycling Outcomes

3.1 Introduction

Rapid increases in the consumption and subsequent waste generation from electronic products (termed "e-waste") are growing international concerns (Atasu and Subramanian, 2012). In 2014 an estimated 92 billion pounds of e-waste was generated globally and, of this amount, only 15% was reported as recycled (Baldé et al., 2015).

The surge in e-waste generation presents an environmental risk as many electronic products contain toxic materials that can leach out of landfills into groundwater or emit dioxins when burned. E-waste is also subject to be dumped in places where the products are dismantled under crude and hazardous conditions (such as western Africa, China and other Asian nations), presenting a threat to people and the environment (ETBC, 2016).

To protect against these risks, several states in the US have adopted various legislative mechanisms to promote electronic reuse and recycling (see Figure 3.1[1] for a summary). One such mechanism is extended producer responsibility (EPR) legislation, which in addition to fostering e-waste recycling, seeks to (1) shift the costs of product recycling to manufacturers and (2) create incentives for manufacturers to de-

[1]Source: www.Deltainstute.org

sign products with fewer adverse impacts on the environment (Atasu and Wassenhove, 2012; Nash and Bosso, 2013). In the U.S., 25 states have adopted EPR legislation to address the e-waste problem, while the remaining 25 states favor voluntary efforts initiated by nonprofits, local governments and/or manufacturers. A second and complementary approach to addressing the risks of e-waste is landfill bans, which prohibit the disposal of electronic products in the trash. In the U.S, 20 states have adopted landfill bans; however, similar to EPR legislation, the scope of the products covered by this legislation vary widely by states. A third and final legislative option is advanced recycling fees. With this option, consumers pay a fee at the point of purchase on covered electronic products. These fees subsequently "go into a recycling fund administered by the State, reimbursing recyclers and collectors" (ETBC, 2016). Currently, California is the only state that leverages this model. In this study, we focus on the two most widely implemented e-waste legislative measures – landfill bans and EPR.

Although the motivations for these legislative measures are clear, there is much debate on the effectiveness of such policies in promoting consumer recycling - a key objective of e-waste regulation (Wharton, 2016; Nash and Bosso, 2013). A few papers have studied the effects of the presence of landfill bans in the U.S. on e-waste recycling. However, the conclusions on the impact of such legislation have been mixed: some studies have found that the presence of landfill bans does not explain past e-recycling behavior (Saphores et al., 2012), while others have concluded that landfill bans do influence consumers' disposal decisions. (Milovantseva and Saphores, 2013).

A lack of empirical data has limited firms' and legislators' ability to assess how e-waste recycling behavior is influenced by the presence of EPR programs. Indeed, we are unaware of any academic study that examines how the presence of EPR legislation affects e-waste consumer recycling behavior. Unlike legislation in other parts of the world, in the US consistent recycling performance measures are not tracked between

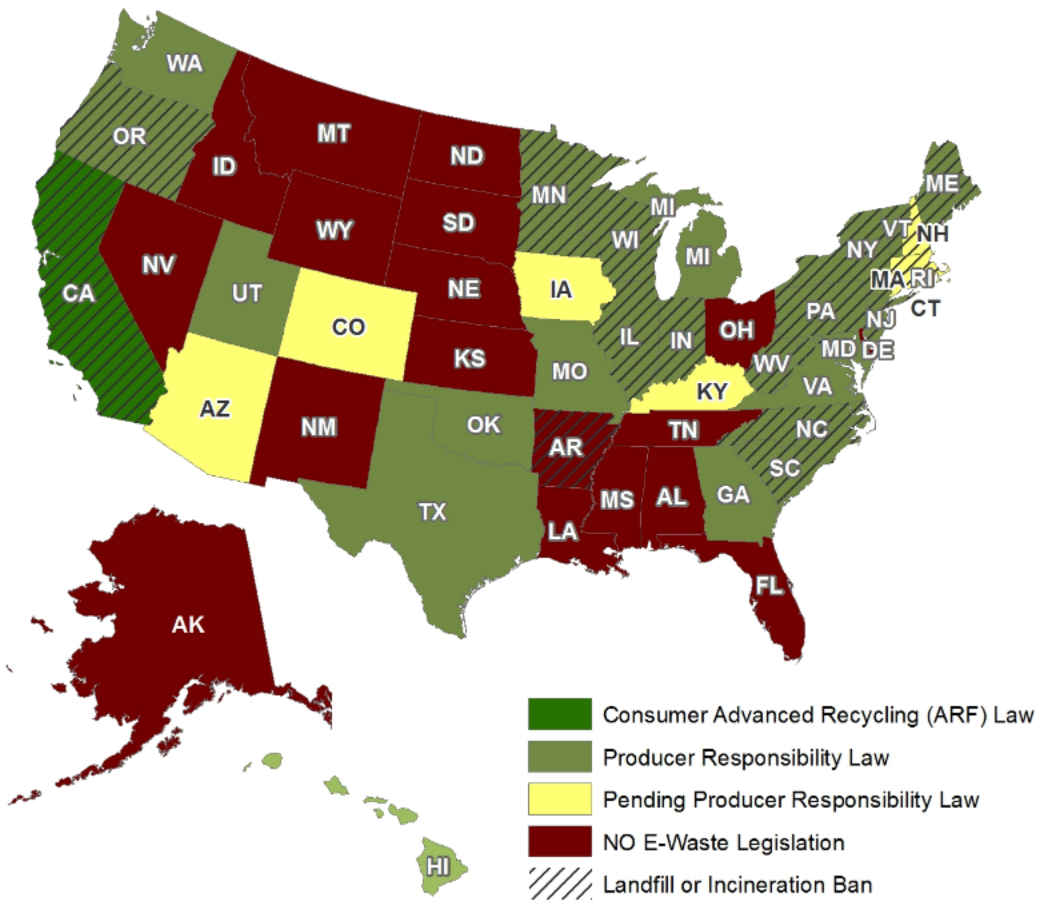


Figure 3.1: Summary of E-Waste Laws in the US

states with and without EPR programs (ETBC, 2016). For states with EPR legislation, per-capita collection rates are generally used to track performance. However, this standard metric has several flaws. First, the specific dimensions of EPR programs vary widely across states, making it difficult to compare performance. These dimensions include who pays for the recycling (producer or consumer), the collection method (municipalities, retailers, producers, etc.), performance goals and incentives (collection and/or pounds target), and the products included in scope of the legislation (televisions, tablets, cellphones, etc) (Atasu et al., 2013). Second, changes in the weight of the electronic products, which tend to become lighter over time, are not taken into account.

Some stakeholders argue that the legislation lays too heavy a responsibility for collection and recycling on manufacturers when *consumer behavior* is key to program success (Nash and Bosso, 2013). The Product Management Alliance (PMA), a coalition of firms from the carpet, electronics, packaging, and other sectors, note that, "EPR legislation adds costly and unnecessary mandates for consumers and local governments. Further, it damages vital American industries that produce and manufacture goods, which in turn leads to these industries moving overseas and the loss of American jobs" (PMA, 2016). As an alternative to EPR legislation, PMA and similar organizations focus on educating consumers on where and how to recycle their electronics.

Given these competing perspectives on the importance of consumer attributes versus legislative actions, the purpose of this study is to examine how recycling and reuse behaviors are shaped by both consumer characteristics (attitudes and awareness) and the presence of e-waste legislation. Specifically, we address the following research question: How are consumers' e-waste disposal choices influenced by their (i) awareness of landfill bans, (ii) attitudes towards recycling, (iii) awareness of recycling locations, and the (iv) presence of EPR legislation?

Our research leverages survey data provided by the Electronics Recycling Coordination and Clearinghouse (ERCC), an organization that collects detailed information on consumers' recycling behaviors, attitudes towards recovering electronics, and awareness of EPR programs. As part of the questionnaire, respondents specified how they dispose of their electronics once the products have reached their end-of-use. We group consumers' responses into three categories — trash, store, and recycle — and develop a multinomial logistic model to study our research questions.

Our main results can be summarized as follows. We find, as expected, that in states with EPR legislation, consumers are more likely to recycle their electronic waste. Perhaps less intuitively, we find that consumers' knowledge of landfill bans increases their likelihood to *store* their electronics rather than recycle them. Finally, we confirm that the biggest deterrents to consumer recycling are not knowing where to recycle electronic products and inconvenient recycling locations. Through these and other observations, we provide insights on how administrators of recycling programs can best leverage e-waste legislation and structure consumer education efforts to maximize consumer recycling participation.

The remainder of this study is organized as follows. In section 3.2, we review the relevant literature. In sections 4.4 and 4, we describe our data and methodology, respectively. In section 5, we present our results and discuss the implications of these results for practice. We conclude in section 6 with a discussion of limitations of the study and directions for future research.

3.2 Literature Review

Factors that affect consumers' likelihood to participate in pro-environmental behaviors (such as e-waste recycling) have been studied extensively within the economic (Viscusi et al., 2011), environmental psychology (Bagozzi and Dabholkar, 1994), and consumer behavior literatures (Iyer and Kashyap, 2007). Specifically, studies have investigated and debated that consumers' pro-environmental behaviors are influenced by factors such as social norms, religious and cultural beliefs, habits, demographics, degree of environmental knowledge/concern, incentives, and legislative policy (Sabbaghi et al. (2015), see Iyer and Kashyap (2007) for a review). This literature, however, has often disputed the generalizability of consumers' recycling behaviors to other materials or environmentally responsible behaviors (Domina and Koch, 2002). Thus, the relatively new challenges posed by e-waste have initiated a new stream of empirical research that focuses specifically on factors that influence e-waste recycling.

In practice, and as noted by Milovantseva and Saphores (2013), e-waste recycling can be enhanced by a number of complementary measures, which include (1) educating consumers about the dangers of dumping e-waste and the social benefits of recycling it properly, (2) creating economic incentives to foster recycling, (3) requiring producers to be responsible for the end-of-life of their products (i.e. EPR), and (4) banning the improper disposal of e-waste.

While the effects of legislative measures (such as disposal bans and EPR for e-waste) on consumer recycling have not been studied extensively in the e-waste literature, several studies have explored the effects of consumer attributes on recycling outcomes. For example, through an extensive survey of US households, Saphores et al. (2012) analyzed factors associated with (i) consumers' who had previously recycled and (ii) consumers' willingness to recycle e-waste using curb-side services. The authors found that consumers who *had previously recycled* were older (i.e. over 60 years old), aware of the toxicity of electronics materials, and had strong moral norms (Saphores et al., 2012). The authors also found that consumers' *willingness to recycle* was strongly associated with factors such as recycling convenience, knowledge of toxicity of e-waste, and prior e-waste experience (Saphores et al., 2012). Last, sociodemographic factors such as education, gender, and ethnicity were found to be weakly associated with consumers' willingness to recycle, but not associated with prior e-waste drop off behavior (Saphores et al., 2012). In a related study, Milovantseva and Saphores (2013) examined the characteristics of consumers who *store* electronics that are no longer in use. The authors found that pro-environmental attitudes, age, marital status, gender, and geographic location are all significant in explaining the number of broken or obsolete TVs stored by US households (Milovantseva and Saphores, 2013).

In addition to these papers, several studies have examined consumers' e-waste behaviors in locations outside of the US. For example, Echegaray and Hansstein (2017) through a survey of Brazilian citizens, finds that there is a significant gap between consumers' intention to recycle and actual recycling behavior, and this gap is particularly salient among the higher income echelons and Southeast regions of Brazilian society. Dwivedy examines recycling preferences for Indian consumers, while Yin et al. (2014) and Wang et al. (2011) explore consumers' attitudes toward electronic recycling in China. The authors find that recycling convenience, recycling habits, economic benefits, and residential conditions are all associated with consumers' willingness to participate in recycling (Yin et al., 2014; Wang et al., 2011).

In sum, this body of literature indicates that e-waste recycling can be stimulated by (i) promoting moral norms, (ii) educating the public about the benefits of recycling e-waste, and (iii) making e-waste recycling more convenient (Saphores et al., 2012). Factors such as awareness of recycling locations, beliefs about recycling, and knowledge of the potential environmental hazards caused by e-waste were generally consistent in explaining consumers' past recycling behaviors. However, apart from age and gender, most other demographic variables (such as income, marital status, family size, etc.) were inconsistently significant across studies.

As mentioned earlier, research examining the effects of legislative measures on consumer e-waste recycling is in its early stages. A few papers have studied the effects of landfill bans on e-waste recycling. However, the conclusions on the impact of such legislation have been mixed. For example, Saphores et al. (2012) find that consumers' awareness of e-waste laws does not impact their willingness to recycle or explain previous recycling behavior. However, Milovantseva and Saphores (2013) found that the presence of state e-waste bans does promote the recycling of cell phones but not consumers' intentions to recycle televisions.

Further, although end consumer participation has been acknowledged as a critical success factor of EPR programs (Lai et al., 2014), to our knowledge there have been no academic studies of how/if the presence of EPR legislation affects e-waste consumer recycling behavior. In general, most research on e-waste EPR focuses on how producers can achieve operational efficiency and minimize compliance costs (Atasu and Wassenhove, 2012; Esenduran and Kemahlıoğlu-Ziya, 2015), or on understanding the effects that e-waste EPR has on firms' product designs and prices (Plambeck and Wang, 2009; Favot and Marini, 2013). In addition, a few papers provide anecdotal information on features of e-waste EPR programs in various geographical locations (Mayers et al., 2011).

We address this gap in the literature by examining how consumer e-waste recycling behavior is influenced by two forms of e-waste legislation: EPR and landfill bans. In addition, we extend previous literature, by assessing how consumers' *awareness* of (versus the mere *presence* of) landfill bans affects their e-waste disposal choices. Further, we leverage an expanded choice set, which allows us to uniquely study how consumer attributes and legislative mechanisms affect consumers' use of a wide range of e-waste disposal options.

3.3 Data

The data for our study were provided by the Electronics Recycling Coordination and Clearinghouse organization (ERCC). Founded in 2010, this non-profit organization serves as an information forum for state governments that have, or are considering implementing, electronics recycling laws. It also serves as an information source for those that are complying with state e-waste legislation. Members of ERCC meet bi-monthly to coordinate various research initiatives and formalize joint responses on key implementation issues. Two levels of membership exist: voting members (includes the environmental agencies of states with EPR legislation) and affiliate members (includes industry representatives, non-profits, and other key stakeholders). A list of ERCC legislative and affiliate members is provided in the appendix in Table B.1.

In March of 2014, ERCC members developed a survey to measure consumers' awareness of e-waste recycling locations and legislation (i.e. landfill bans). The motivation for the survey stemmed from the need to generate a comparable e-waste recycling performance measure across states. Prior to this study, it was difficult to evaluate EPR program performance using the standard "pounds of e-waste collected per resident" measure as states differ in the types of products (i.e. laptops, desktops, cellphones) and entities (i.e. consumers, businesses, nonprofits) that are covered by their EPR programs. However, in measuring consumer awareness rates — defined by ERCC as, *"the level of awareness of electronics recycling programs among consumers for whom the services are available"* — a comparable and key recycling performance metric could be obtained.

The survey script and questions were developed independently by ERCC members and affiliates (please see Figure B.1 in the appendix). In total, ten questions with pre-specified answers were included in the survey and measured consumers' (i) disposal choices, (ii) awareness of recycling location and legislation (iii) interest in recycling activities and (iv) demographic information. Respondents were also able to provide comments for each question, which our team reviewed as a part of our analysis. In the following sections and tables we discuss how we translate the survey data into categorical variables for our analysis.

3.3.1 Consumers' E-Waste Disposal Behavior

We define our dependent variable *DISPOSAL* as consumers' answer to the survey question, "What do you typically do with electronics, such as televisions, computers, monitors and printers that you no longer want?" For our initial analysis, we group consumers' responses into three categories and set the *DISPOSAL* variable to 0 if the consumer stated that they throw their electronics in the trash, 1 if the consumer stated that they store their electronics, and 2 if the consumer stated that they recycle or reuse their electronics. In subsequent analysis, we use the ungrouped *DISPOSAL* variable (see Table 3.1) as our dependent variable. As shown in Table 3.1, nearly 80% of respondents stated that they generally recycle or reuse their electronics. Specifically, donating electronics to a nonprofit organization was the most popular disposal choice (approximately 26% of respondents), followed by recycling at a town or county center (approximately 21% of respondents), and recycling at a store (approximately 15% of respondents).

Table 3.1: Descriptive Statistics for Dependent Variables

Variable	Description	Categories	Count	Percentage
Disposal (grouped)	Consumers' e-waste disposal choice	1: Trash	174	8.93
		2: Store	252	12.94
		3: Recycle or Reuse	1524	78.23
Disposal (ungrouped)	Consumers' e-waste disposal choice	1: Trash	174	8.93
		2: Store	252	12.94
		3: Give to Family Member	300	15.4
		4: Donate	512	26.28
		5: Recycle at Store	296	15.2
		6: Recycle at Town/County Center	416	21.36

3.3.2 Consumer Awareness-Related Variables

Three questions were used to gauge consumers' awareness of e-waste recycling locations and legislation.

First, each survey participant was asked if they were aware of electronic landfill bans in their resident state. We model the *LEGALAWARENESS* variable as 1 if the consumer responded "yes" and 0 if the consumer responded "no."

Location Awareness. In addition to awareness of landfill bans, survey participants were asked if they knew where they could recycle their used, unwanted electronic products. We model the variable *LOCATIONAWARENESS* as 0 if the consumer responded "no," 1 if the consumer indicated that he/she was somewhat aware of e-waste recycling locations, and 2 if the consumer indicated that he/she was certain of recycling locations.

Last, consumers were asked to identify where they go to find information on where to recycle their used electronics. We model the variable *INFOSOURCE* according to the following grouping scheme: 0 if the consumer indicated that he/she does not gather information on e-waste recycling, 1 if the consumer indicated general sources

- such as the internet, phone book, etc., 2 for information from manufacturers or retailers, 3 for non-profits (i.e. greenergadgets.org), and 4 for government sources.

Table 3.2: Descriptive Statistics for Predictor Variables

Variable	Description	Categories	Count	Percentage
Age	Age of survey participant	1*: 18-35	676	34.7
		2: 36-45	306	15.71
		3: 46-55	358	18.38
		4: 56-65	372	19.1
		5: 66+	236	12.11
EPR	Indicates if EPR legislation is present in consumers' resident state	0*: No	441	22.64
		1: Yes	1507	77.36
Distance	Distance that consumer is willing to travel to recycle electronics	1: 1-5 miles	559	28.7
		2: 6-10 miles	787	40.4
		3: 11-20 miles	444	22.79
		4*: 21 miles +	158	8.11
Legalawareness	Indicates if consumer is aware of landfill bans in resident state	0*: No	1245	63.91
		1: Yes	703	36.09
Locationawareness	Indicates if consumer is aware of recycling locations in resident state	0*: No	565	29
		1: Somewhat	601	30.85
		2: Yes	782	40.14
Importance	Importance of electronic recycling to consumer	1*: Not	23	1.18
		2: Not Very	2	2.67
		3: Neutral	168	8.62
		4: Somewhat	539	27.67
		5: Very	1166	59.86
Infosource	Primary source of consumers' recycling information	0*: None	112	5.75
		1: General	434	22.28
		2: Government	736	37.78
		3: RetailManu	243	12.47
		4: NonProfit	398	20.43
		5: Garbage	25	1.28
Prevent	Stated reasons for not recycling	0*: None	108	5.54
		1: Too Expensive	350	17.97
		2: Inconvenient Location	487	25
		3: Unable to Transport Items	434	22.28
		4: Unaware of Recycling Locations	569	29.21
Rural	Indicates if consumer lives in area with less than 50,000 residents	0*: No	1246	63.96
		1: Yes	702	36.04

*denotes reference category

3.3.3 Consumers' Attitudes Towards Recycling Activities

The second section of the survey measured consumers' interest in recycling activities.

Survey participants were asked to identify the top factor that would prevent them from recycling. We modeled the variable *PREVENT* as 0 if the consumer indicated that there were no limiting factors, 1 if the consumer selected "not knowing where to recycle", 2 if the consumer selected "unable to transport items", 3 if the consumer selected "inconvenient recycling location", and 4 if the consumer selected "too expensive."

We model the variable *DISTANCE* according to consumers' response to the question "what is the farthest you would be willing to travel to deliver your unwanted device?": 0 if the consumer selected "1-5 miles away," 1 if the consumer selected "6-10 miles away," 2 if the consumer selected "11-20 miles away," and 3 if the consumer selected "21 miles away or more."

Consumers were asked, "In general, how important do you feel it is to recycle electronic devices that you no longer need or use?" As responses were captured using a Likert scale, we accordingly modeled the variable *IMPORTANCE* from 1 if the consumer selected "very unlikely" to 5 if the consumer selected "very likely."

3.3.4 EPR Legislation

In addition to the survey-derived variables, we also include a binary indicator variable *EPR* in our analysis. We set *EPR* to 1 if a state had EPR legislation in place in at the time of the survey and 0 if the state did not. As shown in Table 3.3, we observed slight differences in the grouped disposal choices of consumers between states with and without EPR legislation.

Table 3.3: Distribution of Disposal by EPR

	Disposal Choice		
Legislation	Trash	Store	Recycle
EPR = 0	58	60	323
	13.15%	13.61%	73.24%
EPR = 1	116	192	1,199
	7.70%	12.74%	79.56%
Average	8.93%	12.94%	78.13%

3.3.5 Demographic Information

Each survey participant was asked if they live in a rural area (defined by ERCC as a location with less than 50,000 residents). We model the $RURAL$ variable as 1 if the consumer responded "yes" and 0 if the consumer responded "no".

Survey participants were asked to identify their age range. We modeled the variable AGE according to the choices selected by the respondents: 1 if the consumer selected "18-35 years old", 2 if the consumer selected "36-45", 3 if the consumer selected "46-55", 4 if the consumer selected "56-65", and 5 if the consumer selected "66 years or older."

The survey was administered between August of 2014 and March of 2015 by *Service 800* (http://www.service800inc.com), a market research company. Prior to launching the data collection effort, participating state environmental agencies determined the number of consumers to be surveyed within a state, according to their desired level of response confidence. Service 800 randomly contacted consumers within a state until the desired number of complete responses were secured. Twelve states participated in the survey in total – including six with EPR legislation and five with landfill bans (see Table 3.4). However, we exclude the results from the state of Hawaii from our analysis, as a different set of questions was used to survey consumers in this state. Of the 2,709 consumers contacted, 1,948 participated in the survey for a response rate of 72%.

Table 3.4: Survey Respondents per State

State	EPR Legislation	Landfill Bans	Complete Responses	Percent of Total
Arizona	No	No	70	4%
Florida	No	No	72	4%
Maine	No	Yes	76	4%
Ohio	No	No	86	4%
Tennessee	No	No	72	4%
Wyoming	No	No	65	3%
Conneticut	Yes	Yes	281	14%
Michigan	Yes	No	398	20%
New York	Yes	Yes	369	19%
Oregon	Yes	Yes	369	19%
Texas	Yes	No	90	5%
	Total		**1948**	**100%**

3.4 Empirical Model

To determine the key drivers of consumers' e-waste disposal choices, we use a multinomial choice (MNL) model. This model is based on random utility theory (McFadden, 1978) and is used to predict the probability that an alternative is chosen, given a set of independent variables. The utility function for a MNL model is defined as

$$\begin{aligned} U_{ni} &= \beta_i' \mathbf{x}_{ni} + \varepsilon_{ni} \\ &= V_{ni} + \varepsilon_{ni} \end{aligned} \tag{3.1}$$

where U_{ni} is the total utility derived from alternative i for individual n; V_{ni} is the observed portion of the utility; and β' is the vector of parameters associated with attributes x. Utility is a linear function of the x attributes which vary across individuals and alternatives, and ε_{ni} is the unobserved portion of the utility function.

In our model, a consumer makes a disposal choice among the three alternatives in the choice set $C = \{$Trash, Store, and Recycle$\}$. The utility associated with each of the alternatives for respondent n is estimated as:

$$V_{ni} = \hat{\beta}_i \mathbf{x}_{ni} + \varepsilon_{ni} \tag{3.2}$$

where i=0,1,2 represents the alternatives Trash, Store, and Recycle respectively and $\hat{\beta}_i$ represents a vector of parameter estimates of β_i. In equation 3.2, $\mathbf{x}_{ni}$ is a vector that of attributes which vary across alternatives i and individuals n. Using the ERCC survey data, $\mathbf{x}_{ni}$ includes information on the presence of EPR legislation, consumers' awareness of recycling legislation, consumers' attitudes towards recycling activities, and demographic information.

We assume that the error terms are independently and identically Gumbel distributed. Accordingly, the probability that consumer n chooses alternative i among j alternatives is as follows:

$$P_{ni} = \frac{e^{\hat{\beta}_i \mathbf{x}_{ni}}}{\sum_j e^{\hat{\beta}'_j \mathbf{x}_{nj}}} \tag{3.3}$$

In addition, the odds that consumer n will choose alternative i over alternative j can be expressed as

$$\frac{P_{ni}}{P_{nj}} = \frac{exp(V_{ni})}{exp(V_{nj})} = exp(X_i \beta_{i|j}) \tag{3.4}$$

We estimated our model using the `mlogit` command in Stata 14. Recall that prior to collecting the survey data, participating state environmental agencies determined the number of consumers to be surveyed within a state, according to their desired level of response confidence. Since our data samples were collected using state-level strata, the usual standard errors associated with the model above are incorrect as they do not adjust for the lack of independence (Long and Freese, 2006; Korn and Graubard, 1990). To address this issue, we employ the `svy` command as suggested by Long and Freese (2006) and obtain model estimates that account for the stratification of our data.

Last, to assess model fit, we performed a series of tests to confirm the independence of variable values (Wald), test whether the disposal options should be combined (Likelihood Ratio Test), and confirm the independence of irrelevant alternatives

(Small-Hsiao) as suggested by Long and Freese (2006). The results from these tests validate our model assumptions.

3.5 Results

The results from our MNL model that describe how e-waste legislation and consumer attributes affect consumers' disposal choices are presented in Table 3.5. To provide general, high level insights, we first provide model output for the three grouped disposal options — Recycle, Store, and Trash — as follows:

1. Columns I - III and VII - VIIII show the model output with *Trash* as the reference category. Note that in a multinomial model the reference category is the category used to compare differences.

2. Columns IV - VI show the model output with *Store* as the reference category.

3. Columns I, IV, and VII present the coefficient values. These values indicate the rates at which the predicted log odds increase or decrease with each successive unit of an independent variable. The significance level and standard error for each coefficient are also shown in the columns to the right.

4. Columns III, VI, and IX list the odds ratios for each variable level. In MNL models, odds ratios (OR) are used to explain the dynamics among the outcomes. Further, the ORs for a variable represent how the odds change with a unit increase in that variable, holding all other variables constant. In our model, the ORs describe the odds that a consumer will choose disposal choice i over disposal choice j for a given variable value.

In addition to understanding the dynamics between the outcomes, we are interested in understanding the marginal effects of the variables in our model. Further, a potential limitation of odds ratios is that the odds can be large in magnitude, even when the underlying probabilities of occurrence are low. Thus, to complement the OR calculations, we compute the average marginal effects for all the variables in our model. Note that for this analysis, we compute the marginal effects for the six ungrouped disposal choices, and provide the corresponding OR calculations in Table B.2 in the appendix. Marginal effects show the change in probability of an outcome, for a unit change in x, while holding all other variables constant at a specified value. In our model, and as as shown in Table 3.6, the marginal effects describe the change in probability that a consumer will choose a disposal choice, while holding all other variables at their respective mean values.

3.5.1 Effects of Consumer Attributes: Awareness

Our results show that consumers' awareness of e-waste landfill bans does impact their e-waste disposal choices. Specifically, awareness decreases the odds that consumer will recycle versus store their electronics by a factor of 0.71. Stated differently, consumers' awareness of landfill bans increases the odds that they will store instead of recycle their electronics by a factor of 1.4. Surprisingly, we do not find evidence that consumers' awareness of landfill bans increases their odds to recycle electronic products compared with the odds of the other disposal choices.

We find that consumers' awareness of e-waste recycling locations also plays a significant role in consumers' disposal choices. The odds of consumers recycling versus trashing their e-waste is almost four times higher for consumers who are somewhat aware of recycling locations. The odds of consumers recycling versus trashing their e-waste is over five times higher for consumers who are aware of recycling locations. More specifically, the probability that consumers recycle in-store or using the town/county recycling center increases by 22 and 19 percentage points, respectively.

Table 3.5: MNL Model with Grouped Disposal as DV

	Recycle vs. Trash			**Recycle vs. Store**			**Store vs Trash**		
	I Coef.	II Std. Error	III Odds ratio	IV Coef.	V Std. Error	VI Odds ratio	VII Coef.	VIII Std. Error	IX Odds ratio
1. EPR: Yes	0.44*	0.21	1.55	0.06	0.18	1.06	0.38	0.24	1.46
1. LEGALAWARENESS: Yes	0.07	0.22	1.07	-0.34*	0.16	0.71	0.41	0.25	1.50
1. LOCTAWARENESS: Maybe	1.37***	0.23	3.92	1.17	0.19	3.92	0.20	0.28	1.22
2. LOCTAWARENESS: Yes	1.67***	0.26	5.30	1.48***	0.20	4.41	0.18	0.30	1.20
2. IMPORTANCE: Not very	-0.08	0.66	0.95	-0.65	0.95	0.52	0.56	0.88	1.76
3. IMPORTANCE: Neutral	1.03^	0.58	2.80	0.25	0.87	1.28	0.78	0.82	2.18
4. IMPORTANCE: Somewhat Important	1.95**	0.57	7.06	-0.32	0.84	0.73	2.28**	0.80	9.74
5. IMPORTANCE: Very Important	2.34***	0.57	10.41	0.02	0.85	10.41	2.32**	0.81	10.17
1. INFOSOURCE: General	0.41	0.34	1.51	1.19***	0.31	3.30	-0.78*	0.40	0.46
2. INFOSOURCE: Government	0.76*	0.36	2.14	1.08***	0.30	2.94	-0.32	0.40	0.73
3. INFOSOURCE: Retailer Manufacturer	0.55	0.39	1.74	0.77**	0.34	2.66	-0.42	0.46	0.65
4. INFOSOURCE: Non Profit	0.77*	0.38	2.16	1.09***	0.32	2.98	-0.32	0.44	0.73
1. PREVENT: Too Expensive	-1.19	0.77	0.31	-0.60	0.47	0.55	-0.59	0.89	0.55
2. PREVENT: Inconvenient Location	-1.73*	0.74	0.18	-1.06*	0.45	0.35	-0.67	0.86	0.51
3. PREVENT: Unable to Transport	-1.13	0.75	0.32	-0.22	0.47	0.81	-0.91	0.88	0.40
4. PREVENT: Unaware of Location	-1.57*	0.73	0.21	-0.87*	0.46	0.42	-0.70	0.86	0.50
1. RURAL: Yes	-0.23	0.19	0.78	-0.37**	0.15	0.69	0.14	0.22	1.16
2. AGE: 36-45	0.35	0.29	1.43	0.05	0.20	2.95	0.31	0.32	0.58
3. AGE: 46-55	0.00	0.25	0.98	0.94***	0.23	2.55	-0.94**	0.31	0.39
4. AGE: 56-65	0.44	0.28	1.49	1.09***	0.24	2.95	-0.65*	0.33	0.52
5. AGE: 66+	0.60^	0.34	1.82	1.14***	0.29	3.13	-0.54	0.41	0.58
2. DISTANCE: 1-5 miles	0.61**	0.23	1.83	0.45	0.28	1.57	-0.47**	0.41	0.62
3. DISTANCE: 6-10 miles	0.38	0.25	1.46	0.33	0.27	1.39	0.25	0.40	1.29
4. DISTANCE: 11-20 miles	0.02^	0.34	1.02	0.54	0.29	1.42	-0.19	0.42	0.83

*******p < .001, ** p <.01, *p<.05, ^ p<.1*

Table 3.6: Marginal Effects from Ungrouped Disposal MNL Model

Variables	Trash	(Std. Error)	Store	(Std. Error)	Family	(Std. Error)	Donate	(Std. Error)	ReStore	(Std. Error)	ReTown	(Std. Error)
1. EPR: Yes	-0.03^{**}	(-0.02)	-0.01	(-0.02)	-0.03	(-0.02)	0.02	(-0.03)	0	(-0.02)	0.05^{***}	(-0.02)
2. IMPORTANCE: Not very	0	(-0.14)	0.07	(-0.08)	0.01	(-0.1)	0.14	(-0.1)	-0.15	(-0.13)	-0.06	(-0.09)
3. IMPORTANCE: Neutral	-0.18	(-0.13)	0	(-0.06)	0.08	(-0.09)	0.26^{***}	(-0.08)	-0.15	(-0.12)	-0.02	(-0.09)
4. IMPORTANCE: Somewhat Important	-0.28^{**}	(-0.13)	0.07	(-0.06)	0.02	(-0.09)	0.21^{***}	(-0.08)	-0.08	(-0.12)	0.06	(-0.08)
5. IMPORTANCE: Very Important	-0.30^{**}	(-0.13)	0.03	(-0.06)	-0.02	(-0.09)	0.27^{***}	(-0.07)	-0.04	(-0.12)	0.06	(-0.08)
1. INFOSOURCE: General	-0.01	(-0.03)	-0.16^{***}	(-0.05)	-0.02	(-0.05)	0.18^{***}	(-0.05)	0	(-0.05)	0.01	(-0.05)
2. INFOSOURCE: Government	-0.04	(-0.03)	-0.14^{***}	(-0.05)	-0.03	(-0.04)	0.09^{*}	(-0.05)	-0.04	(-0.05)	0.15^{***}	(-0.05)
3. INFOSOURCE: Retailer Manufacturer	-0.02	(-0.03)	-0.13^{**}	(-0.06)	-0.01	(-0.05)	0.13^{**}	(-0.06)	0.12^{**}	(-0.06)	-0.08^{*}	(-0.05)
4. INFOSOURCE: Non Profit	-0.03	(-0.03)	-0.14^{**}	(-0.05)	0.05	(-0.05)	0.15^{***}	(-0.05)	0.03	(-0.05)	-0.07	(-0.05)
1. LEGALAWARENESS: Yes	-0.01	(-0.01)	0.04^{**}	(-0.02)	0.01	(-0.02)	-0.05^{*}	(-0.03)	-0.01	(-0.02)	0.01	(-0.02)
1. LOCTAWARENESS: Maybe	-0.09^{***}	(-0.02)	-0.14^{***}	(-0.03)	-0.04	(-0.03)	0.06^{**}	(-0.03)	0.12^{***}	(-0.02)	0.08^{***}	(-0.02)
2. LOCTAWARENESS: Yes	-0.10^{***}	(-0.02)	-0.16^{***}	(-0.02)	-0.10^{***}	(-0.03)	-0.04	(-0.03)	0.22^{***}	(-0.02)	0.19^{***}	(-0.02)
1. PREVENT: Too Expensive	0.04^{*}	(-0.02)	0.04	(-0.04)	0.03	(-0.06)	0.02	(-0.06)	-0.03	(-0.04)	-0.09^{*}	(-0.05)
2. PREVENT: Inconvenient Location	0.07^{***}	(-0.02)	0.09^{**}	(-0.04)	-0.03	(-0.06)	0.03	(-0.06)	-0.03	(-0.04)	-0.13^{***}	(-0.05)
3. PREVENT: Unable to Transport	0.03^{*}	(-0.02)	0	(-0.04)	-0.03	(-0.06)	0.07	(-0.06)	-0.03	(-0.04)	-0.06	(-0.05)
4. PREVENT: Unaware of Location	0.06^{***}	(-0.02)	0.06^{*}	(-0.04)	-0.03	(-0.05)	-0.01	(-0.06)	-0.03	(-0.04)	-0.05	(-0.05)
1. RURAL: Yes	0.01	(-0.01)	0.04^{**}	(-0.02)	-0.01	(-0.02)	-0.04^{*}	(-0.02)	-0.03	(-0.02)	0.03^{*}	(-0.02)
2. AGE: 36-45	-0.02	(-0.02)	0	(-0.03)	-0.03	(-0.03)	0.03	(-0.04)	-0.01	(-0.02)	0.04^{*}	(-0.02)
3. AGE: 46-55	0	(-0.02)	-0.10^{***}	(-0.02)	-0.06^{**}	(-0.03)	-0.05	(-0.03)	0	(-0.02)	0.21^{***}	(-0.03)
4. AGE: 56-65	-0.02	(-0.01)	-0.11^{***}	(-0.02)	-0.12^{***}	(-0.03)	0	(-0.03)	0.06^{**}	(-0.03)	0.19^{***}	(-0.03)
5. AGE: 66+	-0.03^{*}	(-0.02)	-0.12^{***}	(-0.02)	-0.09^{***}	(-0.03)	-0.06	(-0.04)	0.06^{**}	(-0.03)	0.25^{***}	(-0.04)
2. DISTANCE: 1-5 miles	0.01	(-0.03)	-0.06	(-0.04)	0.04	(-0.04)	-0.01	(-0.05)	0	(-0.04)	0.02	(-0.03)
3. DISTANCE: 6-10 miles	-0.03	(-0.02)	-0.04	(-0.04)	0.06^{*}	(-0.04)	0.03	(-0.05)	-0.04	(-0.03)	0.02	(-0.03)
4. DISTANCE: 11-20 miles	-0.02	(-0.02)	-0.06	(-0.04)	0.05	(-0.04)	0.04	(-0.05)	-0.02	(-0.04)	0.02	(-0.03)

****p < .001, ** p <.01, *p<.05, ^ p<.1*

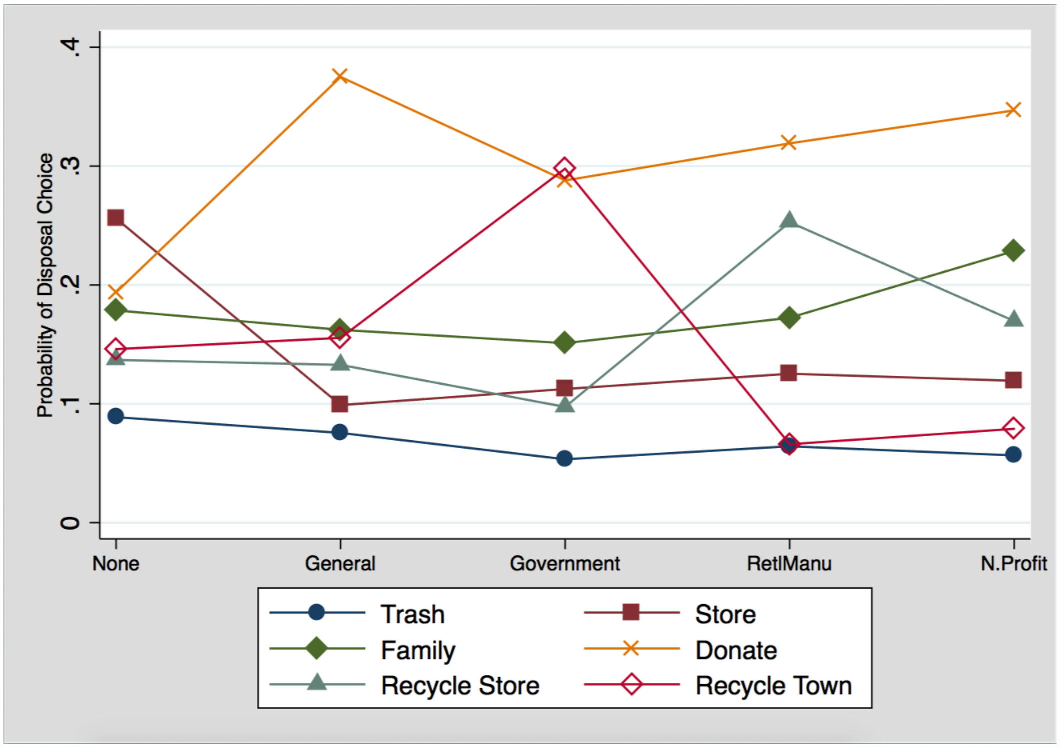
The probability that consumers reuse their electronic waste, in the form of giving unwanted items to family members, decreases by 10 percentage points. Finally, the probability that consumers choose a non-sustainable disposal option such as storing or trashing their electronics decreases by 16 and 10 percentage points, respectively.

From the odds ratios of our MNL model, we see that nonprofit and government sources are two of the most effective e-waste information sources. Compared to consumers who have not received information on e-waste recycling locations, consumers who receive information through nonprofit and government sources are over 2 times more likely to recycle rather than trash their electronics. These information sources also increase the odds that consumers will recycle rather than store their electronics by a factor of nearly 3. In general, the other information sources also result in positive e-waste disposal behavior. Information from retailers and manufacturers increases the odds that consumers will recycle rather than store their electronics by a factor of 2.6 and information from general sources (such as newspapers, the internet, etc.) increases the odds that consumers will recycle rather than store their electronic by a factor of 3.3. From our marginal effects calculations, and, as one might suspect, we find that the information sources correspond with consumers' ultimate disposal choices (see Figure 3.2). For example, the probability that consumers recycle using government and town services increases by 15 percentage points for those who receive information on e-waste recycling from the government.

3.5.2 Effects of Consumer Attributes: Attitudes

Our model results show that, consistent with previous literature, consumers' interest in e-waste activities strongly affects their disposal choices. The odds of consumers recycling versus trashing their e-waste is 7.06 times higher for consumers who stated that recycling was somewhat important to them, and 10.41 times higher for consumers who stated that recycling was very important to them. Additionally, the odds of con-

Figure 3.2: Predictions of E-Waste Disposal Choice Based on Consumers' InfoSource



sumers storing rather than trashing their e-waste is 9.74 times higher for consumers who stated that recycling was somewhat important to them and 10.17 times higher for consumers who stated that recycling was very important to them. We did not find evidence that the odds of consumers' relative disposal choices were effected if consumers stated that they were "neutral" or that recycling was "not very" important to them. Results from our marginal calculations complemented these findings: the probability that consumers trash their electronics decreases by 30 percentage points for those who stated that recycling was "very important" to them (see Figure 3.3).

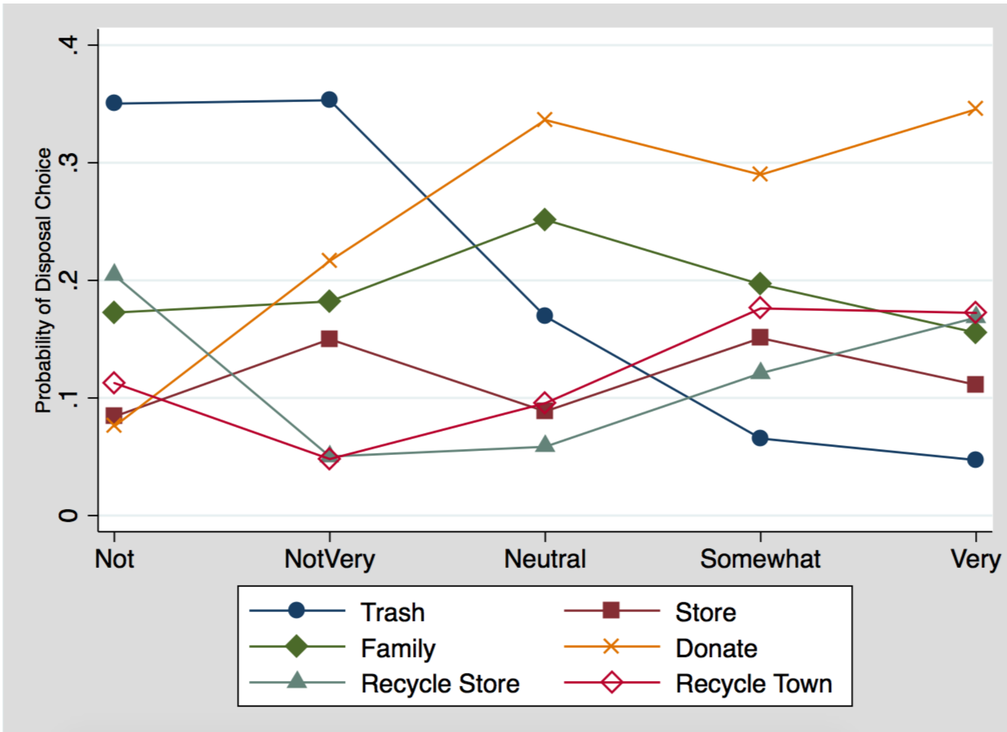
Our results suggest that consumers' perceived barriers to recycling significantly increase the probability that they will trash or store their electronics. From our odds ratio calculations, we find that the odds that consumers recycle versus trash or store their electronics decreases when e-waste recycling location is unknown (OR = 0.21) or inconvenient (OR = 0.18). Indeed our marginal effect calculations confirm that, of the options tested, the probability that consumers trash their electronics is most influenced by these location factors. The probability that consumers trash their electronics increases by 7 percentage points for those who indicated that "inconvenient locations" were their top barrier to recycling, and by 6 percentage points for those who indicated that "unawareness of location" was a top preventer. The other factors, expense and transportation, also increase the probability that consumers trash their electronics, but to a lesser extent. The probability that consumers trash their electronics increases by 4 percentage points for those who indicated that expense was their top barrier to recycling, and by 3 percentage points for those who indicated that transportation issues was a top concern.

We found limited insights on effects of distance on consumers' disposal choices. From the odds ratio calculations, we find that the odds that consumers recycle versus trash increases by a factor of 1.83 when recycling locations are between the 1-5 mile range versus the 21+ mile range. Additionally, the results from our marginal analysis suggest that the probability that consumers give their electronics to their family members increases by 6 percentage points when recycling locations are between the 6-10 mile range.

3.5.3 Effects of E-Waste Legislation

From our model calculations, we find that the odds of consumers choosing to recycle versus throwing items in the trash are approximately 1.6 times higher for states with EPR legislation for e-waste versus states without EPR legislation for e-waste. This finding provides support for the argument that EPR promotes e-waste recycling; however, it also sheds light on the relative impact of legislation versus consumer attributes. For example, our marginal effect calculations show that for states with EPR legislation in place, the probability that consumers use the town and county

Figure 3.3: Predictions of E-Waste Disposal Choice Based on Importance of Recycling to Consumer

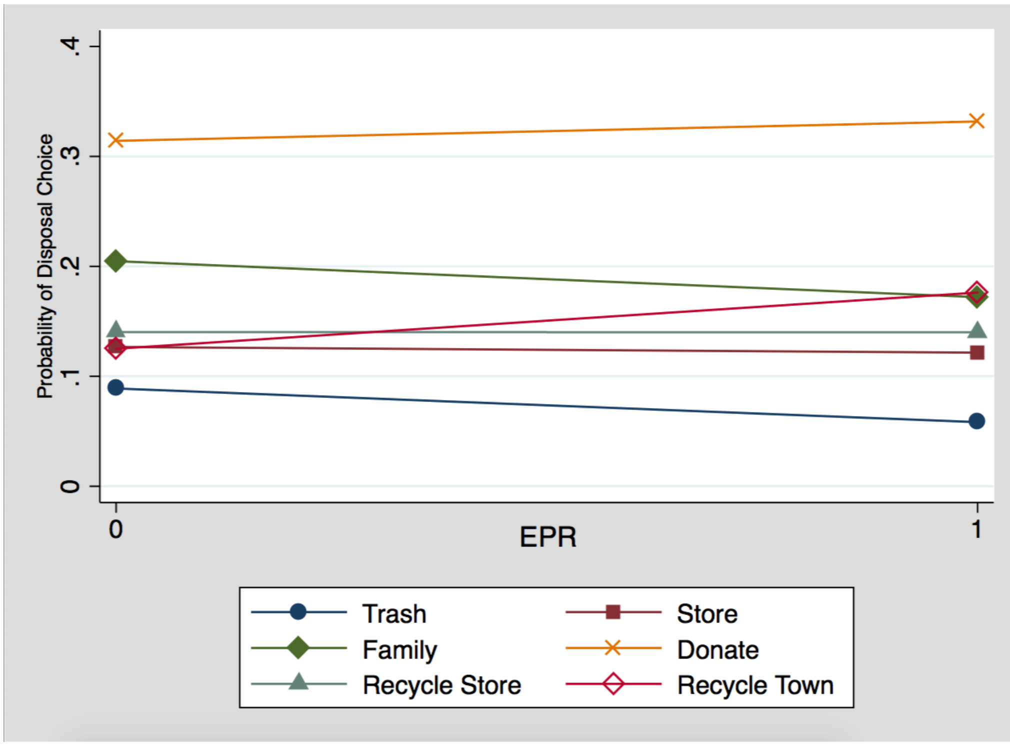


recycling options increases by 5 percentage points (see Figure 3.4). Additionally, the probability that consumers trash their electronics decreases by 3 percentage points.

3.5.4 Consumer Demographics

Consistent with previous literature, we find that older consumers are more likely to recycle their electronics versus throw them in the trash: compared to consumers who are between the ages of 18 and 35, the odds of consumers recycling rather than storing their e-waste are 2.5 times higher for consumers who are between the ages of 36 and 45, 2.95 times higher for consumers between the ages of 56 and 65, and and 3.3 times higher for consumers who are 66 or older. Finally our results suggest that consumers living in rural areas are more likely to store their electronics rather than recycle them: the probability that consumers choose the store disposal option

Figure 3.4: Predictions of E-Waste Disposal Choice Based on Presence of EPR Legislation



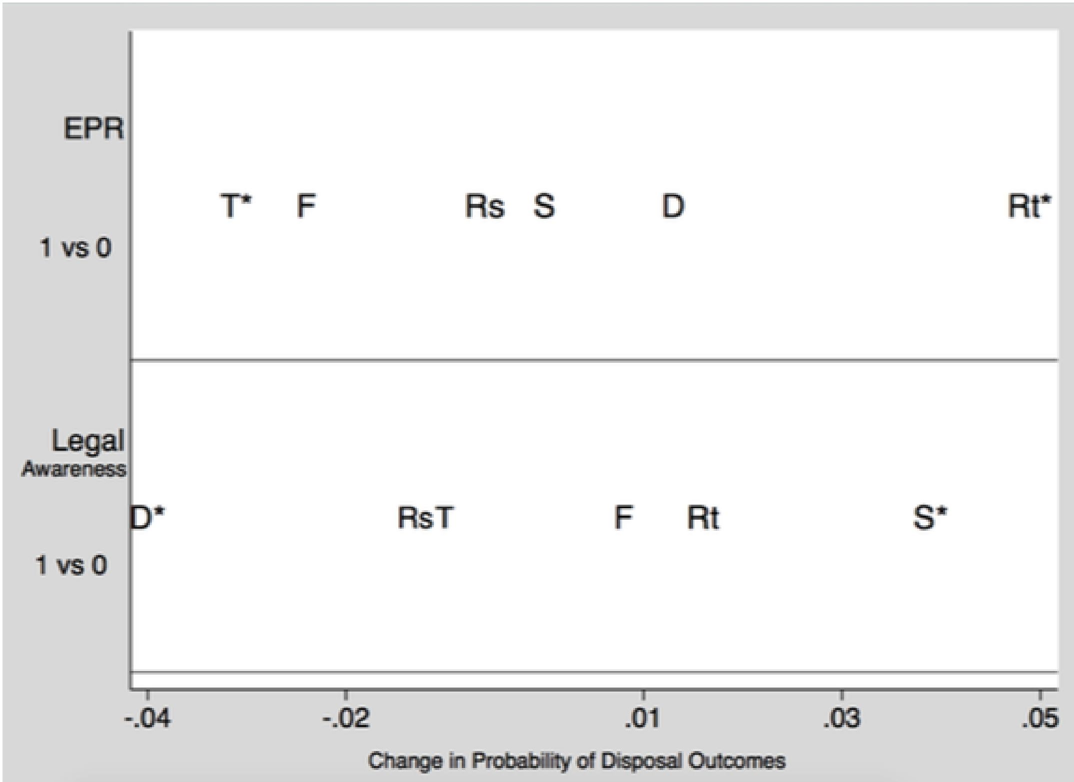
increases by 4 percentage points for consumers living in these locations.

3.6 Conclusion and Future Research Directions

The recent surge in e-waste generation poses vexing problems for regulators seeking to protect human health and the environment, and for manufacturers seeking to comply with various educational and legislative policies. Although several states in the US have adopted such approaches, the effectiveness of these policies (in particular, EPR for e-waste) in promoting consumer recycling continues to be questioned. In response, and through use of consumer survey data, we investigate the effects of e-waste legislation and consumer attributes on consumers' disposal choices to provide clarity to this debate.

The marginal effects of e-waste legislation on consumer e-waste disposal choices from our analysis are summarized in Figure 3.5. We find that that both legislative options (EPR and landfill bans) affect consumers' e-waste disposal behavior. The presence of EPR legislation, for example, increases the likelihood that consumers will choose to recycle and decreases the likelihood that consumers will trash their e-waste. However, in the case of landfill bans, we find that the legislation's impact on consumers' disposal choices may be different from the envisioned outcomes. Specifically, our results show that consumers' awareness of landfill bans increases the likelihood that they will store their electronics. While storing e-waste is preferred to consumers' throwing their used electronics in the trash, ideally landfill bans would also promote recycling. Even if consumers are temporarily storing their electronics, as noted by Sabbaghi et al. (2015) time delays can increase the obsolescence rate of potentially still-functional products and lead to fewer recovery opportunities such as reuse, upgrade, and refurbishment. Thus, our results suggest that efforts that focus on informing consumers about the dangers of e-waste and where they can recycle their electronic products may need to be coupled with e-waste legislation and to significantly promote e-waste recycling.

For instance, we find that e-waste location awareness plays a significant role in consumers' disposal decisions. The odds of consumers recycling rather than throwing the items away in the trash or storing them significantly increases with awareness levels. In addition, of the options explored, consumers named inconvenient recycling locations and not knowing where to go as the biggest deterrents to recycling. Last, as regulators seek to improve awareness levels, our results show that consumers often seek information from a variety of sources and that all are effective in reducing the probability that consumers throw their electronics away in the trash. We note that consumers ultimate disposal choice is highly influenced by their original information source. In states such as New York and Indiana where manufacturers are required and sometimes struggle to meet annual collection targets, these results suggest that



Key: T = Trash, S = Store, F = Family, D = Donate, Rs= Recycle Store, and Rt = Recycle Town and County Center. *=$p<.05$

Figure 3.5: Effect of Legislation on Changes in Probability of E-Waste Disposal Choices

consumers' disposal behavior may be very responsive to e-waste location information.

Beyond improving consumer knowledge of recycling locations, our results suggest that administrators of recycling programs should also focus their efforts on educating consumers about the importance of recycling. Of all the factors tested, appreciation of recycling was most influential in reducing the likelihood that consumers threw their electronics items in the trash. Our results also indicate that younger consumers and those living in rural areas may benefit most from such educational efforts.

There are several limitations to our study which may provide interesting avenues for future research. First, we focus on consumers' general e-waste disposal behaviors. In practice, consumers' e-waste behaviors may vary across product types. For example, cell phones may be easier to dispose of versus products that are in general more expensive to recycle and heavier to transport, such as televisions. Second, because of data limitations, we were not able to explore how specific differences in the structure of state legislation may affect consumers' choices. Finally, again due to data limitations, we did not explore the effects of Advance Recycling Fees, which are currently implemented in the state of California. Also due to data limitations, we were unable to explore interaction effects between the different factors, which may provide additional insights. Although not exhaustive, the above limitations provide promising trajectories for further analyses.

Chapter 4

Using Transactions Data to Improve Consumer Returns Forecasting

4.1 Introduction

In recent years, offering a generous return policy has become increasingly popular among U.S. retailers. Eager to win sales, big brick-and-mortar chains such as Walmart and Target promise a full money back guarantee upon return if customers are not satisfied with their purchases. Responding to this competitive pressure, Best Buy famously eliminated its 15% restocking fee for consumer electronics in 2011, aligning itself with the majority of the U.S. retail industry (Reisinger, 2010).[1] In the online sector, lenient return practices have also become popular, as several retailers even inlcude free return shipping in their policies (examples include Zappos, Nordstrom, Gap, and Urban Outfitters).

Behind the trend of lenient return practices the belief that consumers highly value the option to return products penalty free, which in turn generates higher demand and better customer satisfaction for the retailer (Mollenkopf et al., 2011). The recent consumer returns studies offer more specific evidence: Anderson et al. (2009) and Heiman et al. (2015) estimate that consumers value a full refund policy for an apparel item purchased through a catalog or physical channel at 10% to 25% of the product's price. It has also been shown that frequent returners are associated

[1] There are retailers that do not accept returns, such as baby products retailer Zulily (Ng, 2015) and some small-to-medium sized sellers on eBay. To our knowledge, these cases are rare in the whole retail sector.

with more frequent, and higher-dollar future purchases (Griffis et al., 2012), stronger brand loyalty (Ramanathan, 2011), and even larger customer lifetime value (Petersen and Kumar, 2015). For online retailers, the importance of a generous return policy can be even more critical to sales. Since consumers are not able to physically inspect products, a liberal returns policy can help consumers overcome any reluctance to make a purchase (WSJ, 2017).

While enjoying the demand-side benefits a full refund policy entails, retailers and OEMs also find the management of returns a challenging task – often ranked among the top managerial concerns (Brill, 2015). At the macro level, industry research has estimated the total value of returns that U.S. consumers make annually to be around \$260 billion (Kerr, 2013; Ng, 2015), which makes up roughly 7% of sales for physical stores and over twice as much for online retailers (approximately 15%). Further, the National Retail Federation (2008, 2016) has observed a steady increase in average return rates over the past decade, from 7% to 12%. Adding to the sheer volume of returns are the operational costs that must occur to ensure their flow along the reverse supply chain and value recovery through various options. A recent study by Accenture (Douthit et al., 2011) suggests that return-related operations such as inspection, reverse logistics, and refurbishment or disposal accounts for 5% of a typical OEM's revenue and 4% of a typical retailer's sales. Overall, just a 1% return rate has been estimated to cost a large retail chain \$17 million (Douthit et al., 2011) and the whole U.S. economy \$32 billion (National Retail Federation, 2016).

Perhaps the most straightforward "solution" to reducing returns is to charge a restocking fee that limits consumers' incentive to return. However, given the prevalent adherence to full refund policies in practice and the strong belief in their marketing benefits, much of the return management burden falls instead on the operational levers that reduce costs by optimizing operational decisions along the reverse supply chain, while treating the return volume as exogenous. Examples include better inventory management, reverse logistics, staffing, and product recovery. As described in detail below, a common crucial input for the effectiveness of these operational decisions is an accurate return forecast.

From a retailer's perspective, the influx of returns requires adjustments to inventory policies, since the current level of inventory replenishment should consider the volume of future returns (Ketzenberg and Zuidwijk, 2009). In such cases, forecasting returns becomes a prerequisite for determining the order quantity. Similarly, many other inventory-related tactical decisions along the reverse supply chain also involve return forecasting, including determining the optimal number of parts for refurbishing and repairs, the number and location of stock points, and the allocation of inventories across them (Fleischmann et al., 1997).

Planning for reverse logistics is generally more difficult than coordinating the forward supply chain due to the increased uncertainty in quantity, time, and quality of returned products (Agrawal et al., 2015). Moreover, returned products often accumulate in a warehouse and are eventually routed to a low payback salvage channel because their value has depreciated past the point of being profitable to remarket. One of the main reasons for this wasteful practice is that return volumes are not built into the the logistics network, since there is no reliable forecast for them(Phillips, 2015). It is reported that the retail sector's annual spending on reverse logistics for processing and disposition of consumer returns is more than $40 billion (Enright, 2003), and OEMs spend 2 to 3 times more on reverse logistics costs than on the original outbound logistics of the same product (Stock et al., 2006). The lack of an accurate returns forecast and the resultant inability to plan effectively are contributing to these high costs.

As the prominence of returns management increases, how to best staff the return counter using a reliable returns forecast becomes part of a retailer's labor planning process, complementing the existing traffic-based sales force staffing (Chuang et al., 2016). Furthermore, with returns moving upstream in the reverse supply chain, distribution centers and refurbishment centers also face staffing challenges of similar nature, yet technically more complex. For example, Wal-Mart sends returns to one of its six regional return centers across the U.S., where employees sort the returned goods into four tiers. The return flow is highly variable: although 45 million pallets of returns are processed annually, over 40% of this volume occurs in January and February after the holiday time period (Souza, 2013b). As a result, Wal-Mart staffs with seasonal employees. The performance of this approach is highly dependent on the accuracy of returns forecast. Our conversation with managers at the Bose refurbishment center in South Carolina[2] reveals a similar staffing problem for their part-time labor force.

The salvage value of a returned merchandise largely depends on the recovery channel. For consumer electronics, it ranges from as high as 70% value recovery by selling through an online resell channel to as low as 20% by selling to liquidators who buy in pallets (Ng, 2015). Other recovery options include reusing for parts, refurbishing, and recycling. For example, Optoro – a third-party reverse logistics provider specializing in swift value recovery of electronic products – “recovers” returns through these various channels based on expected return volume, consumer preference for open-box items, and discount store demand (Tabuchi, 2015). An accurate return forecast is again a crucial input parameter. According to Shorewood Liquidators, who processes returns for Groupon, better planning on the recovery strategy can lift salvage value from 20% to 50% (Ng and Stevens, 2015). For an OEM who accepts returns from retailers[3], the recovery decision often involves allocating returns between two options, restocking for open-box sale and earmarking for future warranty demand.

[2]Bose, a major OEM of high-end consumer audio equipment, operated its only returns processing center for the North America market in South Carolina where consumer returns are sorted, refurbished, and tested.

[3]Different from the dominance of full refund policies between retailers and consumers, the policy between OEMs and retailers has a much higher degree of variation (Crocker and Letizia, 2014).

Pince et al. (2016) derive optimal strategies under such a setting where the future return forecast is one of the critical inputs in their disposition decision model.

In general, just as more accurate sales forecasts enable retailers and OEMs to improve their operational performance, more accurate returns forecasts can also increase the effectiveness of return-related decisions along various stages of the reverse supply chain (Agrawal et al., 2015). Given that 95% of the consumer returns are "no-problems-found," the cost saving potential appears promising (Douthit et al., 2011). Additionally, the operations literature has proposed many decision support tools for managing returns such as retail store inventory (Ketzenberg and Zuidwijk, 2009), disposition decisions (Pince et al., 2016), and reverse logistics network design (Guide et al., 2006). All of these models share the common trait that an accurate return forecast is critical for their successful implementation.

So what methods do companies commonly apply for forecasting consumer returns? Toktay et al. (2003) find that many rely on simple heuristics such as multiplying projected sales by the historical return rate of a product or time-series methods such as ARIMA models. Our conversations with over 15 retailers and OEMs during the 2013, 2014, and 2015 Consumer Returns Annual Conferences confirm that these are still the most common choices[4]. Furthermore, although there is a rich stream of studies on *end-of-use or end-of-life returns* forecasting, the type of data available in that context is very different from our context of *consumer returns.* Specifically, consumer returns are typically processed by a retailer through Point-Of-Sale (POS) technologies, which yields a rich transaction-level data set that includes purchase and return timestamps of each individual transaction. In contrast, historical period-level (e.g. weekly, monthly) sales and return quantities are the most common data inputs for end-of-use and end-of-life returns.

[4]The Consumer Returns Annual Conferences is the largest industry meeting for practitioners who focus on managing consumer returns. It is annually held in September or October. More details could be found at `http://consumerreturns.wbresearch.com/`.

In response to the need of more accurate returns forecasts and the lack of advanced methods that effectively utilize the rich transaction-level data available, we propose a "causal model" approach that predicts future return quantities in two steps. First, we fit an econometric model that simultaneously estimates the likelihood of a purchase being returned and how long it takes the consumer to make the keep-versus-return decision. Based on the purchase timestamp and the estimated model, we compute a probability for each purchase to be returned in a specific time window (e.g. next week, month, etc). Second, we aggregate the probabilities across the past purchases to compute the *expected* number of returns in that time window. We label this as the "predict-aggregate" approach. Our econometric model in the first step (predict) has the flexibility to accommodate two unique features of the consumer returns context. The first is that consumers differ in how certain they are about their valuations of a product when they make a purchase. During the trial period, each consumer experiences the product and adjusts her valuation accordingly; the scale of this adjustment is likely to be larger for those who are *ex ante* more uncertain. This heterogeneous nature of trial uncertainty requires the modeling of a non-constant error variance. The second is that consumers also differ in how long it takes them to sufficiently experience a given product. Since the retailer cannot observe this duration for products that are not returned, but only for products that are actually returned, our model accounts for this sample selection bias.

We test the performance of our model on two large Point-Of-Sale (POS) transactions data sets provided by a major consumer electronics retail chain (Ni et al., 2012) and an online fashion-jewelry store. Testing on these two datasets has several advantages and allows us to compare our model's performance on samples with different return rates, average time-to-return durations, and other key metrics. Further, managing consumer returns are particularly important for the product types and retail channels represented in our data. For example, the National Retail Federation

(2016) estimates that return rates for "hard goods," including consumer electronics, are 50% more than the average across product categories. Early identification and processing of returned products is especially important for short life cycle products such as consumer electronics, which can lose value at a rate greater than 1% per week (Guide et al., 2006). Thus, better return forecasts can be particularly valuable for this product category. In addition, consumer electronics have relatively low purchase frequency compared to other categories such as apparel, which makes the task of forecasting more difficult. Generating an accurate return forecasts for online sales (as in our second dataset) can be similarly challenging as e-retailers generally observe higher return rates and return variability than brick-and-mortar stores.

To assess the performance of our model, we draw from industry practice and the existing literature to construct benchmark models, which as we discuss in §3 can be summarized as the "aggregate-predict" approaches. The performance comparison between our proposed predict-aggregate (P-A) and the existing aggregate-predict (A-P) approaches is presented in a sequential manner. We start with a "bare-bones" version of our approach, which already shows significant accuracy gains over common practice benchmarks. We then add heteroskedasticity and sample selection components to our model, showing how each addition improves the forecast accuracy. In addition, since the time-to-return distribution might exhibit shapes quite different from what we observe in our data, we make a methodological generalization to our econometric model in an extension. In light of the spike of same-day return observed in the return lag data, we also consider a "two-part" model that explicitly accounts for the high volume of same day returns in another extension.

4.2 Literature Review

An emerging stream of literature in the closed-loop supply chain area studies the forecasting problem of *end-of-use* and *end-of-life* returns. For example, Toktay et al.

(2000) examine the case of Kodak's single-use cameras and Clottey et al. (2012) expand on this work by considering alternative distributions for the return lag. Li et al. (2011) apply count regression techniques to forecast trade-in returns in a business-to-business context. The consumer returns context is different in that the relevant data is most often collected by retailer POS technologies, resulting in rich transaction-level information such as purchase and return timestamps. Naturally, how to effectively utilize the granularity of this transaction-level data becomes an important empirical question. To our knowledge, Daim et al. (2012) is the only study that forecasts consumer returns. Their approach is to first aggregate the transactions data into weekly sales and returns, and then predict future returns using a moving average of return percentages[5], which as we discuss in the next section belongs to the family of our A-P benchmark approach. In contrast, our P-A approach better utilizes POS data by predicting the probability of a return for each individual purchase through an econometric model, and then aggregates these probabilities into period-level return quantities. In addition, our econometric model captures two transaction-level outcome variables, return probability and experience duration, through a very general setup. While they do not forecast the quantity of returns, Hess and Mayhew (1997) suggest the use of a logistic regression for predicting the return probability and a five-parameter non-negative survival distribution for return lag. Our general model nests their suggested setup as a special case, which allows us to test how their setup fares for consumer returns forecasting (discussed in detail in Appendix C.4).

The process flow of our P-A forecasting approach is in spirit analogous to the bottom-up method often found in the context of time-series sales forecasting. Con-

[5]The authors apply this approach to three categories of camcorder returns based on return reason code. Despite the theoretical appeal of using reason codes, it is well known in the consumer returns industry that these codes are rarely reliable. First, consumers often claim that the product is broken when making a return even though this is rarely the case. Second, return counter staff tend to select the most convenient reason code (often the first one on a drop down list) when processing consumer returns.

sider a television manufacturer who needs to predict demand for its assortment of TV models. Based on historical sales for each SKU, the bottom-up method first creates individual time-series forecasts at the SKU-level and then sums them up to generate the sales forecast for the whole assortment of TVs. The efficiency of top-down and bottom-up methods has been examined empirically in multiple contexts including telephone demand (Dunn et al., 1971), multiple-competition (Dangerfield and Morris, 1992), and market earnings (Kinney, 1971). More recently, Kremer et al. (2015) consider the effects of behavioral factors on the accuracy of these models. The accuracy and efficiency of the bottom-up method is often compared with the alternative top-down method, where the assortment-level forecast is created before being partitioned into individual SKU-level forecasts (e.g. Dangerfield and Morris, 1992; Kremer et al., 2015). A common finding from these studies is that the bottom-up method often outperforms the top-down. The comparison between our P-A approach and the benchmark A-P approaches for forecasting consumer returns resembles the top-down vs. bottom-up comparison in the sales forecasting literature and yields a similar finding: the P-A approach, which conducts predictions at the bottom level, outperforms the A-P approach. As we explain in later sections, however, the P-A approach requires a much more sophisticated econometric treatment than do most bottom-up forecasting methods for new product sales.

4.3 Forecasting Approach

Consider an empirical setting where a retailer collects transaction-level data through POS technologies. While the collected data will likely encompass product, transaction, and consumer characteristics, we focus only on the timestamps of purchase and return to demonstrate the performance of our forecasting approach relative to benchmarks. Note that additional predictors can be easily added to our approach, which we discuss in §4.6.4. Before moving into model details, we contextualize the forecasting problem, discuss the conceptual difference between two potential approaches, and provide intuition on why the P-A approach exhibits superior performance.

4.3.1 A Conceptual Comparison between Two Approaches

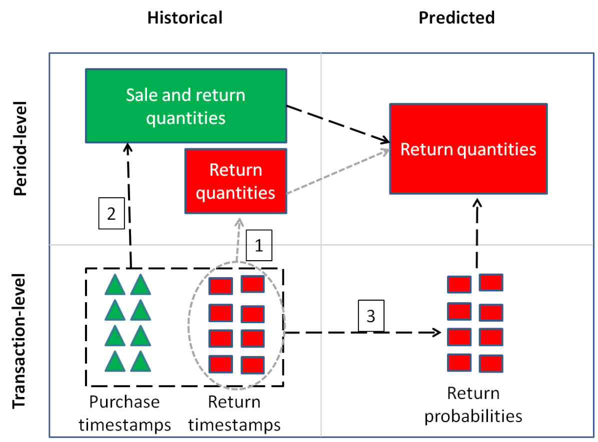


Figure 4.1: Existing and Proposed Return Forecasting Approaches

Figure 4.1 presents different paths for using POS data to forecast return quantities. The starting point is the raw data of purchase and return timestamps (left bottom quadrant). To achieve the goal of predicting the number of returns in a given period (upper right quadrant), such as a month, the forecaster can take two paths. The existing forecasting methods require the forecaster to first aggregate the transaction-level timestamps into period-level sales and return quantities and then predict future returns based on these historical quantities. Thus, with the A-P approach the forecasting model is applied at the period level. In doing so, the forecaster can choose to ignore the aggregated sales quantities and rely solely on historical return quantities to fit a time-series model such as an ARIMA model (Path 1), as is often done in practice. Alternatively, she can exploit the fact that future returns come from past sales and predict returns through a regression model (Path 2), which is the typical approach used in forecasting *end-of-life* returns. We propose a forecasting approach that deviates substantially from these existing methods – we predict return probabilities at the transaction-level and then aggregate these individual probabilities into return quantities (Path 3). In other words, our proposed method follows a P-A path, which applies a forecasting model at the transaction-level. Next, we provide some intuition on why the P-A path might better utilize POS data and thus deliver superior performance.

4.3.2 Advantages of the Predict-Aggregate Approach

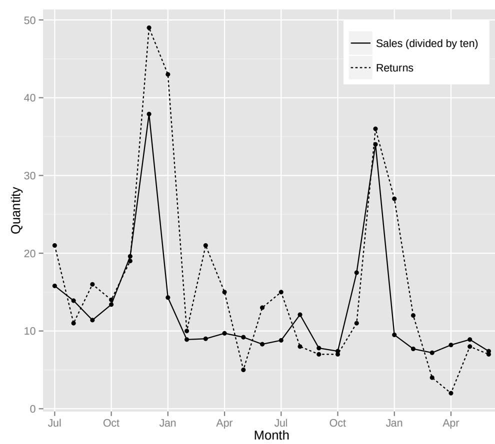


Figure 4.2: An Example of Time Series of Sales and Returns (generated from the audio speaker category in our data set)

Figure 4.2 depicts a typical series of monthly returns data, generated from the audio speaker category in our data set. To illustrate the relationship between sales and returns, we overlay the returns with the series of monthly sales data. We make several observations from Figure 4.2 to help explain our intuition. First, the return-series lags behind the sales series, which is especially obvious during the period when sales fall back to normal levels after the holiday season. For example, while the sales volume decreases sharply from December to January, the decrease in the volume of returns appears to lag behind by at least a month. The time that consumers take to experience their purchases contributes to the lag of return for individual purchases, which in turn aggregates to the phenomenon of monthly returns lagging behind monthly sales. Second, when sales quantities are relatively stable, such as between February and October, the return quantities appear to be more volatile. This suggests the number of returns is much noisier than what can be captured by a simple product of the average return rate and the sales volume. One major driver of this noisiness is the purchase timestamps within each month, which might vary from month to month. For example, a sale in the latter part of a month is more likely to contribute to the next month's returns than to the current month's. Apart from the timing of purchases, the heterogeneity of the return probability and the return lag across individual transactions also gives rise to the fluctuation in return quantities. When a retailer or an OEM forecasts the quantity of returns for a given period of time, both the lagging behavior of returns and its noisiness are better captured through a transaction-level model than an aggregate model. Therefore, while our proposed P-A and the existing A-P methods both start with purchase and return timestamps, the former utilizes the available data more effectively.

4.3.3 Model Building

While both A-P and P-A approaches can predict the number of returns for any arbitrary period of time, we will compare their performance on monthly returns since this is a common planning time segment for both OEMs and retailers. The total number of returns t in a given month j, (R_j^t) can be decomposed into two parts: those attributed to the sales completed in the previous months, (R_j), and those attributed to the current month's sales, (R_j^c). At the end of month $j-1$, the existing purchases contribute to R_j but not to R_j^c. In order for our proposed approach to predict R_j^c, we need a predicted purchase pattern for month j, which adds additional noise to the return forecasts. Therefore, to minimize this additional noise introduced by the sales forecast and ensure a clean comparison among different methods, we focus our main analysis on the prediction of R_j. In §4.6.5, we extend our analysis to predict R_j^t by employing a simple sales forecast.

In the following, we build our forecasting model in three steps, starting with a basic version, which enables us to outline its procedures in a straightforward manner. Then, we address two important econometric challenges that arise in a returns context – heterogeneous return probability and its correlation with experience duration. Our performance testing in §4.5.2 is synchronized with these steps. Presenting our results in this manner has two benefits. First, it helps demonstrate that superior forecasting performance can be achieved even with a reasonably simple setup under the P-A approach. Second, the predictive value of each additional model component becomes more evident.

Basic Model.

For a given purchase i, denote its probability of being returned in a future month as P_{ij}, which is the joint probability of two events happening together – the consumer has had sufficient time to experience the purchased product i by month j and the consumer decides to return product i. Label the probabilities for the two events as P^r_{ij} for the return and P^d_{ij} for the duration probabilities, respectively. If we assume a consumer's decision to return is independent from the required time for her to experience the product, we have $P_{ij} = P^r_{ij} \times P^d_{ij}$. Furthermore, if we assume a homogeneous return probability across purchases, P^r_{ij} is essentially the return rate in the estimation sample. These two assumptions are relaxed in following sections. Denote the time it takes the buyer of i to fully experience this product as d_i. Let the timestamp for purchase i be t_i and that for the start and end dates of month j be t^{start}_j and t^{end}_j. It follows that P^d_{ij} is the probability that d_i is between $t^{start}_j - t_i$ and $t^{end}_j - t_i$, both of which are always positive since we are looking at future months. Let the CDF of d_i be F_d so that P^d_{ij} is simply $F_d(t^{end}_j - t_i) - F_d(t^{start}_j - t_i)$. Grounded in the empirical characteristics of our data (see §4.4), we assume d_i follows a log-normal distribution such that $\log(d_i)$ is normally distributed, $\mathcal{N}(\mu_d, \sigma_d)$. We later generalize the distributional choice in §4.6.2. The parameters μ_d and σ_d can be estimated through the linear regression: $\log(d_i) = \mu_d + \epsilon_d$, where the error term ϵ_d is a zero mean normal distribution with variance σ_d^2 and $\log(d_i)$ is the logged return lag. We estimate the log-normal duration through the `fitdist` command in R. The above constitutes the *predict* step in our proposed P-A approach.

Next, we turn to the *aggregate* step. Because each purchase that happened before month j has a non-zero probability of being returned in month j and these probabilities are not equal, the total number of returns in month j follows a Poisson-Binomial distribution. Thus, the sum of these Bernoulli probabilities is $\hat{R}_j = \sum_{i=1}^{N_j} P_{ij}$, where N_j is the total number of purchases made before month j.

4.4 Data

To test the performance of our proposed approach, we apply our model on data from two different retailers (Table 4.1). The first, is the Ni et al. (2012) data set of consumer electronics transactions published in *Manufacturing & Service Operations Management*. In an effort to promote quality empirical research, the Institute for Operations Research and the Management Sciences (INFORMS) encourages scholars to publish unique and comprehensive data sets. The outcomes of this effort includes three data sets published in *Manufacturing & Service Operations Management* (Bodea et al., 2009; Willems, 2008; Mumbower and Garrow, 2014) and four in *Marketing Science* (Ni et al., 2012; Bronnenberg et al., 2008; Wang et al., 2014; Goldenberg et al., 2010). We make use of the Ni et al. (2012) data since the publishers suggest that it is by far the most comprehensive data set of consumer electronics transactions available, and encourage researchers to specifically leverage it for product return issues (p. 1012). The original data contains $173,262$ transactions provided by a major U.S. electronics chain, the identity of which, for confidentiality reasons, is not

Table 4.1: Description of the Datasets

	Dataset 1	Dataset 2
Product Type	Consumer Electronics	Fashion Jewerly
Retail Channel	Brick-and-Mortar	Online Only
Number of Categories	9: Audio, Auto Parts, Cable, Computer Imaging, Mble Phone, Phone, TV, and TV Box	5: Accessories, Bracelets, Earrings, Necklaces, and Rings
Return Policy	No Specified Limit	30 Days and Free Return Shipping
Date of Transactions	December 1998 - October 2001	December 2015 - April 2017
Number of Transactions	85,725	616,127

provided. All returns at this retailer are accepted open-box and given a full refund. The transactions in this data set took place between December 1998 and November 2004. We limit our attention to a subset of the original data that is applicable to our analysis (the specific subsetting procedure is described in Appendix C.1). The remaining applicable data for our analysis contains $85,725$ transactions spanning across nine product categories.

The second data set leveraged in our analysis was provided by an online jewelry retailer. The identity of this retailer is again masked for confidentiality reasons. We make use of this additional data to test the robustness of our model and gain insights on how differences in return lag, return rates, and other key metrics impact our model's performance. The retailers' return policy states that merchandise can be returned within 30 days; however, this return policy is not heavily enforced. In addition, the retailer does not charge a restocking fee and return shipping is free. The original data set contains 620,470 transactions that took place between December 2015 and April 2017. We exclude transactions in which the product categories could not be identified (approximately .03% of our observations). The remaining applicable data for our analysis contains $616,127$ transactions spanning across five product categories. In the following we discuss key variable definitions and descriptive statistics of both datasets.

Since each purchase may or may not be returned, each observation takes either

Table 4.2: Purchase and Return Timestamps Example

Category	Purchase Date	Return Date
Audio	1-Dec-2001	.
TV	1-Dec-2001	10-Jan-2002

one of the two forms shown in Table 4.2. Return lag is calculated by differencing the purchase and return timestamps. Tables 4.3 and 4.4 provide descriptive statistics for the return rate and lag of each category. The average return rate for our electronics data matches closely with the 12% average for consumer electronics reported by the National Retail Federation (2016), and the average return lag ranges from 12 days to 30 days. The average return rate for our jewelry data is 7% and the average return lag for this data is between 23 and 24 days (the National Retail Federation (2016) does not report a return rate for this category of merchandise). To gain some exploratory insights into the distribution of the return lag, we present a histogram of both data sets in Figure 4.3. The shape of these histograms resembles a lognormal distribution with a small mean and a fairly large variance, except that the spike at same-day return is substantial in the electronics data. Therefore, while we use the lognormal distribution to model return lag in §4.3.3, we generalize our approach to allow for other distributions as an extension in §4.6.2. In light of the spike at $return\ lag = 1$, we also consider a "two-part" duration model that explicitly accounts for the high volume of same day returns in another extension in §4.6.3.

Table 4.3: Return Lag and Return Rate Descriptive Statistics for Electronics Dataset

Category	Return Rate			Return Lag			Total Observations
	Average	Std. Dev.	Coeff. of Var.	Average	Std. Dev.	Coeff. of Var.	
Audio	12%	0.32	2.76	12.9	20.0	1.6	11,367
Auto Parts	14%	0.35	2.44	21.9	54.1	2.5	7,211
Cable	11%	0.31	2.91	10.9	19.8	1.8	9,398
Computer	11%	0.31	2.90	11.1	23.0	2.1	21,811
Imaging	11%	0.31	2.92	11.8	19.9	1.7	7,823
Mobile Phone	16%	0.37	2.27	28.9	74.4	2.6	3,126
Phone	14%	0.35	2.47	16.1	30.6	1.9	5,777
TV	8%	0.28	3.28	11.3	21.5	1.9	9,413
TV Box	14%	0.34	2.52	14.7	28.5	1.9	9,799
Sample Average	**12%**	**0.33**	**2.72**	**15.5**	**32.4**	**2.0**	**1,201.05**

Table 4.4: Return Lag and Return Rate Descriptive Statistics for Jewelry Dataset

Category	Return Rate			Return Lag			Total Observations
	Average	Std. Dev.	Coeff. of Var.	Average	Std. Dev.	Coeff. of Var.	
Accessories	4%	0.19	5.18	23.28	19.69	0.85	20,583
Bracelets	7%	0.26	3.64	22.70	14.71	0.65	61,997
Earrings	8%	0.27	3.46	24.40	17.30	0.71	308,385
Necklaces	7%	0.25	3.67	23.74	18.12	0.76	255,470
Rings	8%	0.27	3.43	24.94	23.01	0.92	51,121
Sample Average	**7%**	**0.25**	**3.87**	**23.81**	**18.57**	**0.78**	**616,127**

Table 4.5: Monthly Sales and Returns (Electronics Dataset)

		Sales		Returns	
Category	Months	Average	Std. Dev.	Average	Std. Dev.
Audio	71	158.23	182.47	6.44	8.39
Auto Parts	71	100.08	90.36	4.99	4.19
Cable	71	130.97	202.46	3.46	4.52
Computer	71	302.99	421.42	8.73	8.69
Imaging	71	107.30	102.46	3.92	3.74
Mobile Phone	70	44.64	54.73	2.67	2.88
Phone	71	80.80	82.08	4.13	3.29
TV	71	131.10	206.03	3.38	3.76
TV Box	71	136.63	195.98	6.69	9.58

Note: Returns in this table do not including same months returns.

Table 4.6: Monthly Sales and Returns (Jewelry Dataset)

		Sales		Returns	
Category	Months	Average	Std. Dev.	Average	Std. Dev.
Accessories	16	1,102.39	893.31	27.83	28.45
Bracelets	16	3,211.00	1,285.31	152.11	80.20
Earrings	16	15,425.94	7,111.96	835.22	475.11
Necklaces	16	13,593.67	4,874.56	648.28	291.82
Rings	16	2,761.11	1,293.78	155.56	105.94

Note: Returns in this table do not including same months returns.

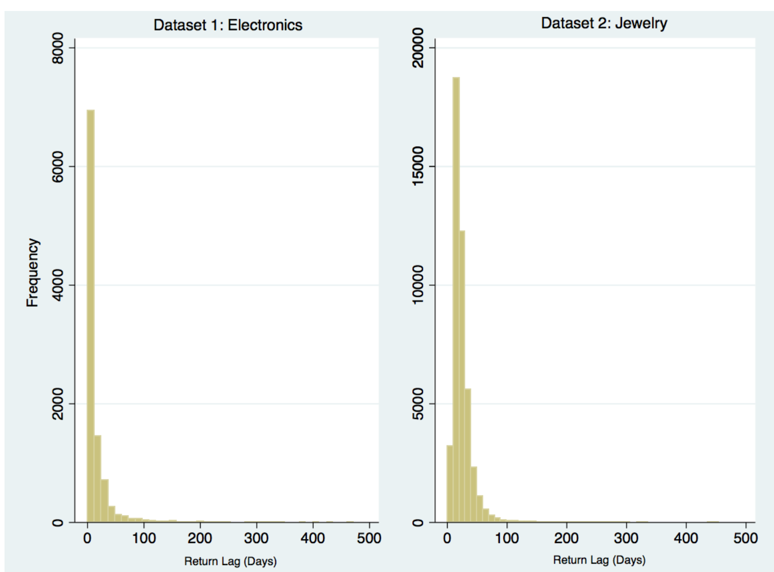


Figure 4.3: Histogram of Return Lags

4.5 Forecasting Performance

We demonstrate the performance of our proposed P-A forecasting approach for consumer returns in two steps. First, we describe two commonly used benchmark models and evaluate their performance against the basic version of our approach from §4.3.3. Second, we evaluate the forecasting performance after adding heteroskedasticity and sample selection, one at a time, against the basic model to show the value of these additional model components.

4.5.1 Benchmark Models

As discussed in §4.3.1, the existing forecasting approaches follow the A-P work flow, where prediction is carried out at the period-level. Two representative prediction models are ARIMA (often used in practice) and Lagged Sales (often discussed in the *end-of-use* returns forecasting literature). Both A-P approaches first use the purchase and return timestamps to compute the monthly sales and return quantities (the aggregate step), and then apply either a time-series model on the return quantities or a lagged sales regression model on both quantities. The fitted model is in turn used to predict future return quantities (the predict step). Next, we outline these two steps in more detail.

The monthly sales, S_j, is a simple count of purchases within each month. R_j is a simple count of returns from purchases made in previous months. Table 4.5 provides a by-category view of the period-level variables. This completes the aggregate step.

Moving on to the prediction step, in the first method we fit an ARIMA(p,d,q) model to the historical monthly return data. The need for differencing (i.e. $d = 1$ or $d = 0$) is determined by a unit root test. Since none of our return time-series show an obvious trend, the test confirms our intuition that no differencing is needed and thus $d = 0$. To determine the appropriate autoregressive parameter, p, and the moving average parameter, q, we use a stepwise model selection procedure as documented in

Hyndman and Khandakar (2008), which is implemented in the `auto.arima` command in R. The most frequently chosen model is ARIMA(1,0,0) such that $R_j = \alpha_0 + \alpha_1 R_{j-1} + e_j$.

In the second method, we regress monthly returns on past sales, which is referred to as the Lagged Sales model in the literature (Toktay et al., 2000; De Brito and Dekker, 2003; Clottey et al., 2012). The rationale for this model is that the returns in a given month are attributed to sales made in previous months, where the central question is how far back to look. The 99th percentile of return lag in our data is 150 days for our electronics data and 91 for our jewelry data set. As a result, we include the past five months' sales in our model for the electronics forecast and three months for the jewelry model. The Lagged Sales model is given by $R_j = \alpha_0 + \sum_{k=1}^{5} \alpha_k S_{j-k} + e_j$ and implemented with the `lm` command in R. Both the ARIMA and the Lagged Sales models are "trained" on the first 35 months' of data for the electronic transactions and 9 months for the jewelry data set. The remaining periods are reserved as the prediction sample (i.e. a half-half division for the training and prediction samples).

We employ two prediction accuracy measures: the mean absolute error (MAE) and the root mean squared error (RMSE). While the former reflects the average bias of forecasts, the later penalizes the large errors more severely. Define MAE $= \frac{1}{J-35} \sum_{36}^{J} |\hat{R}_j - R_j|$ and RMSE $= \frac{1}{J-35} \sum_{36}^{J} \sqrt{(\hat{R}_j - R_j)^2}$, where J is the total number of months as shown in Table 4.5. The results for the benchmark models are presented in Table 4.7. In general, the Lagged Sales model performs better than the ARIMA model, which is expected since the former uses both sales and returns to calibrate the model and also has a better theoretical motivation. Next, we compare the performance of these A-P approaches to our P-A approach.

Table 4.7: Comparison of the Aggregate-Predict and Basic Predict-Aggregate Approaches (Electronics Dataset)

	ARIMA		Lagged Sales		Basic P-A		Basic P-A vs ARIMA		Basic P-A vs L. S.	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Audio	4.11	5.97	3.95	5.68	3.42	5.55	-16.9%	-7.0%	-13.4%	-2.2%
Auto Parts	3.34	4.50	2.97	3.84	2.77	3.57	-17.2%	-20.7%	-6.8%	-7.2%
Cable	3.95	5.55	2.60	3.74	2.29	3.27	-42.0%	-41.1%	-12.0%	-12.8%
Computer	5.28	8.09	5.01	7.59	4.36	6.78	-17.4%	-16.2%	-12.8%	-10.7%
Imaging	3.32	4.25	2.44	2.92	2.14	2.80	-35.3%	-34.0%	-11.9%	-4.1%
Mobile Phone	2.61	3.61	2.66	3.69	2.55	3.73	-2.2%	3.5%	-4.0%	1.2%
Phone	2.50	3.60	2.44	3.49	2.15	3.32	-14.0%	-7.9%	-12.0%	-5.0%
TV	1.95	3.28	1.87	3.20	2.00	3.76	2.5%	14.6%	7.0%	17.3%
TV Box	5.13	11.19	4.17	10.07	4.01	9.61	-21.9%	-14.1%	-3.9%	-4.6%
Simple Avg.							**-18.3%**	**-13.7%**	**-7.8%**	**-3.1%**
Weighted Avg.							**-17.7%**	**-13.3%**	**-7.9%**	**-3.5%**

Abbreviations: L.S. = Lagged Sales; Avg. = Average. The percentages are calculated as P-A measure divided by A-P measure minus one. Weighted average is calculated as percentage measures weighted by categorical return rates.

Table 4.8: Comparison of the Aggregate-Predict and Basic Predict-Aggregate Approaches (Jewelry Dataset)

	ARIMA		Lagged Sales		Basic P-A		Basic P-A vs ARIMA		Basic P-A vs L. S.	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Accessories	23.81	27.13	6.24	6.85	5.16	6.10	-78.3%	-77.5%	-17.3%	-10.9%
Bracelets	93.24	103.08	75.03	88.27	25.08	30.77	-73.1%	-70.2%	-66.6%	-65.1%
Earrings	466.59	643.32	253.25	389.66	164.90	224.61	-64.7%	-65.1%	-34.9%	-42.4%
Necklaces	324.05	337.90	230.32	248.69	98.90	131.97	-69.5%	-60.9%	-57.1%	-46.9%
Rings	54.43	68.04	47.93	57.29	21.69	28.38	-60.1%	-58.3%	-54.7%	-50.5%
Simple Average							**-69.1%**	**-66.4%**	**-46.1%**	**-43.2%**
Weighted Average							**-67.9%**	**-65.1%**	**-48.9%**	**-46.5%**

Abbreviations: L.S. = Lagged Sales; Avg. = Average. The percentages are calculated as P-A measure divided by A-P measure minus one. Weighted average is calculated as percentage measures weighted by categorical return rates.

4.5.2 Comparison between P-A and A-P Approaches

Tables 4.7 and 4.8 contains the performance comparison between the basic P-A from §4.3.3 approach and the two A-P approaches (ARIMA and Lagged Sales). As shown in Figure 4.4, our basic P-A approach outperforms the baseline models in most product categories. Two exceptions are the phone and TV categories from the electronics dataset, where the Lagged Sales Model performs better. However, as we introduce more sophistication into our approach, these exceptions disappear.

In general, we observe that the reductions in forecast error from the basic P-A approach are more significant for the jewelry data set than the electronics data set. There are several potential causes for the differences in performance as summarized in Figure 4.5. First, we observe more volatility in the monthly return rate for the jewelry dataset than the electronic data. The average coefficient of variation for return rate is 3.87 for the jewelry data and 2.72 for the electronic dataset. Recall that the baseline ARIMA model only uses past returns to predict future returns, while the lagged sales model leverages both sales and returns. As a result, both models may be prone to error when there is significant variability in the sales-to-returns patterns and return rates. However, our P-A forecasting approach explicitly controls for differences in individual consumers' return probabilities, a key source of return variability not captured by the ARIMA and Lagged Sales models. As a result, we observe significant reductions in error from the ARIMA and Lagged Sales model for the jewelry product categories. Second, we observe a longer, less variable return lag in the online jewelry dataset than the electronic dataset. The average return lag for the jewelry retailer is approximately 24 days, CV = 0.78 and 15 days, CV = 2.0 for the electronic retailer. We suspect that the jewelry data's higher and more consistent average return lag (due in part, to the presence of a return policy) decreases the likelihood that products will be returned across multiple months and allows for easier prediction. Last, the performance gain from the P-A model is likely to be affected

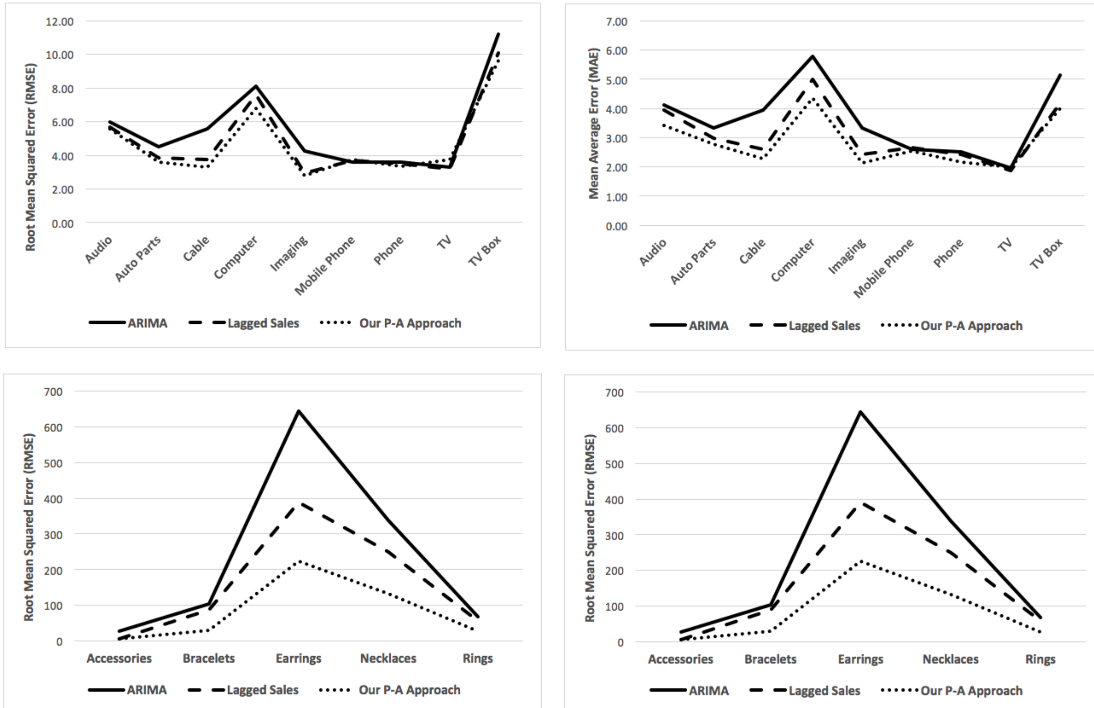


Figure 4.4: Graphical Comparison of the A-P and Basic P-A Approaches

by our data sets' respective sample sizes. Specifically, the higher purchase frequency observed in the jewelry data set likely allows for further reductions in forecasting error.

To summarize, our basic P-A approach demonstrates an accuracy improvement over the baseline models when applied to both the electronic and jewelry data sets. Specifically, the basic P-A approach reduces forecast error by 3% to 18% on average for the electronic data, and by 43% to 69% for the jewelry data set, depending on the benchmark model and error estimated used. We suspect that the improvement in forecasting error is likely more pronounced for product categories with higher variability in return rate and less variability in return lag.

Metric	Electronics Data	Jewelry Data	Effect on P-A Model Accuracy
Return Rate Volatility	Less Variable (Return Rate CV = 2.72)	More Variable (Return Rate CV = 3.87)	Higher reduction in forecast error compared to ARIMA model and potentially Lagged Sales model for jewelry dataset
Return Lag	Shorter (Mean = 15.5 days) More Variable (CV = 2.0)	Longer (Mean = 23.81 days) Less Variable (CV = 0.78)	Number of returns per month more consistent. Higher reduction in forecast error from ARIMA model and Lagged Sales model for jewelry dataset
Purchase Frequency	Lower (Mean = 1,200 purchases per month)	Higher (Mean = 38,507 purchases per month)	Larger sample size likely increases accuracy of P-A forecasting model. Higher reduction in forecast error compared to ARIMA model and Lagged Sales model for jewelry dataset

Figure 4.5: Summary of Key Metrics Effects' on P-A Model Accuracy

4.6 Extensions

In this section, we consider a number of extensions to the previous analyses. These extensions accomplish two goals: generalizing our proposed forecasting approach, and examining its performance over the benchmark models in different applications. To facilitate a succinct exposition, we discuss our motivation for and the main insights from each extension in the following, while providing the empirical details in the appendices. We note that the extensions are only performed on the electronics data set.

4.6.1 Heteroskedasticity of Return Probability and Sample Selection

To relax the assumption that the return probability, P_{ij}^r, is homogeneous across all purchases, we use a Probit regression such that $P_{ij}^r = \Phi(\dfrac{\mu_r}{\sigma_r})$, where $\Phi(\cdot)$ is the CDF of a standard normal distribution. Again, this distributional restriction is relaxed in §4.6.2. In this case, consumers return when $\mu_r + \epsilon_r > 0$ and keep when $\mu_r + \epsilon_r < 0$, where the error term follows a zero mean normal distribution $\epsilon_r \sim \mathcal{N}(0, \sigma_r)$. Recall that the only input information we have are the purchase and return timestamps. Therefore, we provide a time-related motivation for this heterogeneous return probability. Since products purchased during the holiday season (November and December) are more likely to be gifts, and evaluating what others might like is more difficult than choosing the best product for oneself, these holiday season purchases should have a wider experience variation, which is captured by σ_r. Identification of the Probit regression requires σ_r be set to a fixed number (usually 1) if it is a constant. We set σ_r for the non-holiday purchases to 1 and allow it to be estimated for the holiday purchases. Essentially, we estimate a heteroskedastic Probit model, which is implemented through the `hetglm` command in R. The probability P_{ij}^d is estimated in the same manner as in the basic model.

Adding Sample Selection.

In our context, the experience duration is manifested only in a selected sample – as the return lag of the returned purchases. Under the assumption behind the basic model that a consumer's decision to return a product is independent of the length of her product trial, estimating experience duration from this selected sample will not introduce bias because the mechanism that "selects" the sample (i.e. keep versus return decision) is uncorrelated with the outcome to be estimated from the selected sample (i.e. experience duration). However, there are reasonable arguments for why these two events may be correlated, in both the negative and positive directions. For example, consumers who face "buyer's remorse" might be more likely to return the product and also be more likely to end the product trial early – a negative correlation. On the other hand, a perfectionist spends a long time testing the product, while her perfectionism may also cause even small mismatches between her expectations and the product's functionality to result in a product return – a positive correlation. This correlation will bias the estimation of experience duration and hence P_{ij}^d. Further, for forecasting purposes, the implication of a non-zero correlation between the return decision and experience duration is that $P_{ij} = P_{ij}^r \times P_{ij}^d$ no longer holds. As a result, we must specify a bivariate distribution for P_{ij}, which incorporates the correlation between the two events. Recall that the error terms for the return regression and the duration regression, ϵ_r and ϵ_d, are both normally distributed. Therefore, their correlation can be incorporated through a bivariate normal distribution such that $\epsilon_r, \epsilon_d \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \left(\begin{smallmatrix} \sigma_r^2 & \rho\sigma_r\sigma_d \\ \rho\sigma_d\sigma_r & \sigma_d^2 \end{smallmatrix}\right)$. This setup is commonly referred to as a sample selection model (Heckman, 1979) in econometrics. Note that a sample selection model is not the same as the conventional censored regression model and the latter does not fit our problem (see Appendix C.3 for a detailed explanation).

To estimate this model, we proceed with a Maximum Likelihood (ML) method.

Let I_i be the return indicator. If purchase i is not returned, the likelihood is simply

$$L_i|_{I_i=0} = \Pr(\epsilon_r < -\mu_r) = \Phi(-\frac{\mu_r}{\sigma_r}) \tag{4.1}$$

If purchase i is returned, we know two things: the return lag is d_i and $\mu_r + \epsilon_r > 0$. The likelihood in this case is therefore

$$L_i|_{I_i=1} = \Pr(\epsilon_r > -\mu_r \cup \epsilon_d = \log(d_i) - \mu_d) = \frac{1}{\sigma_d}\phi(\frac{\log(d_i) - \mu_d}{\sigma_d})\Phi(\frac{\frac{\mu_r}{\sigma_r} + \frac{\rho}{\sigma_d}(\log(d_i) - \mu_d)}{\sqrt{1-\rho^2}}) \tag{4.2}$$

Derivations are provided in Appendix C.5. The log-likelihood for the whole sample is

$$LL = \sum_{i=1}^{N}\{(1 - I_i) \times L_i|_{I_i=0} + I_i \times L_i|_{I_i=1}\} \tag{4.3}$$

Again, σ_r is estimated for the holiday purchases and set to 1 for the non-holiday purchases, which incorporates heteroskedasticity in the return probability regression. The above model is implemented through the `bbmle` package in R.

After estimating the parameters (μ_r, μ_d, σ_r, σ_d, and ρ), the probability for each purchase to be returned in month j is given by

$$P_{ij} = \Pr(\epsilon_r > -\mu_r, t_j^{end} - t_i > e^{\mu_d + \epsilon_d} > t_j^{start} - t_i) = \int_{\log(t_j^{start} - t_i) - \mu_d}^{\log(t_j^{end} - t_i) - \mu_d} \int_{-\mu_r}^{+\infty} F_{r,d}(\epsilon_r, \epsilon_d) d\epsilon_r d\epsilon_d \tag{4.4}$$

where $F_{r,d}(\cdot)$ is the bivariate normal distribution for the two error terms (calculation implemented in R's `pmvnorm` command). The aggregate monthly returns are then obtained by summing these P_{ij}'s. In the next section, we describe the data set we use to test the performance of our proposed approach.

We present the additional accuracy gain caused by the inclusion of heteroskedasticity (denoted by Het.) over the basic P-A model in Appendix C.2. The average accuracy gain is around 6% (Table C.4). While we observe an improvement in accuracy in all the categories, it is more pronounced for products that are likely to be purchased as gifts. For example, the accuracy gain for the audio category, a common gift choice, is between 16% and 19%. On the other hand, auto parts and cables might be less common choices for gifts, and the accuracy gains for these two categories are much smaller. Our primary motivation for the addition of heteroskedastic error in §4.6 is the change in consumers' return behavior due to gifting in the holiday season. The above categorical differences in the accuracy improvement is aligned with this motivation.

4.6.2 Arbitrary Distributions

Throughout §4.3.3, we assumed that experience duration follows a log-normal distribution, which simplifies our analysis in two ways. First, when sample selection is not incorporated, the parameters related to experience duration can be estimated through a linear regression. Second, when sample selection is present, it allows us to use a bivariate normal distribution and follow the standard (Heckman, 1979) setup to estimate model parameters. Although the log-normal distribution fits our data set well (see Appendix C.4), there is no guarantee that it will do so for other product returns data. To expand the flexibility of our proposed approach, we introduce a more general setup of our econometric model in Appendix C.4 that accommodates an arbitrary distributional choice for both the return probability and experience duration. After comparing the forecasting accuracy obtained from using an alternative distribution with that using the normal distribution, we find no evidence that the former dominates the latter. As a cautionary note, we stress that the "optimal distribution" will likely vary by the data at hand. This is exactly the reason why we generalize our model to allow for an arbitrary distributional choice.

4.6.3 Inflated Same-Day Returns

Our data exhibits a high volume of same-day returns, as shown in Figure 4.3. This could occur for several reasons: consumers might have simply purchased the wrong item, they might test the product immediately after purchase, or it may simply be a case of immediate consumer remorse without an actual product trial. From a modeling standpoint, the spike of same-day returns leads to a poor fit at the left corner for most distributions. In Appendix C.6, we explore whether explicitly accounting for these "inflated same-day returns" improves the forecast accuracy. The basic idea is to modify the distribution of experience duration into a mixture of continuous and discrete components, where the discrete component helps to better capture the spike. Our analysis reveals that paying separate attention to the same-day returns provides only a marginal forecasting improvement. The performance similarity between using a continuous and a mixture distribution might be explained as follows. For each purchase that took place before month j, we predict its probability of being returned in the time interval $[t_j^{start} - t_i, t_j^{end} - t_i]$. A mixture distribution should predict this probability more accurately when t_j^{start} is only several days ahead of t_i. However, the proportion of such purchases is very small when using our monthly buckets. Thus, if our goal had been to estimate the probability that a purchase will be returned in two or three days, a mixture distribution would be more likely to outperform its continuous counterpart.

4.6.4 Flexibility of Adding Transaction-Level Predictors

In §4.3.3, we only used the purchase and return timestamps in building our econometric model, which helps to cleanly demonstrate the superiority of our proposed P-A approach over the benchmark A-P approaches. In practice, one additional advantage of using our proposed approach is its flexibility of incorporating additional transaction-level predictors of return probability and experience duration. The granularity of these variables is hard to capture in period-level models. In Appendix C.7, we include *purchase price* and a *holiday* dummy variable into our model and observe a 1% to 4% reduction in forecast error, depending on the econometric specification.

4.6.5 Total Number of Returns

To ensure a clean and direct comparison between the benchmarks and our approach, we focused on predicting monthly return quantities attributed to previous months' sales (i.e. R_j) in §4.5.2. Now, we extend our analysis to predict the total monthly return quantities[6], R_j^t. In practice, a retailer or OEM may be interested in all three quantities. While R_j^t provides a holistic view of the return flow, R_j^c and R_j provide additional information on the age of the returns. The retailer or OEM may recapture the value of newer returns in a different way than older returns. On average, the total monthly returns are around

[6]Recall that R_j^c denote the same month returns and R_j^t denote the total number of returns in month j such that $R_j^t = R_j^c + R_j$.

Table 4.9: Comparison of MAE of Aggregate-Predict and Predict-Aggregate Approaches for Forecasting Total Returns

	A-P Approach		P-A Approach		
	ARIMA	Lagged Sales	Het. & Sel.	vs ARIMA	vs Lagged Sales
Audio	9.13	9.60	6.24	-31.6%	-35.0%
Auto Parts	9.07	7.26	8.18	-9.8%	12.7%
Cable	8.49	8.12	7.13	-16.0%	-12.2%
Computer	10.62	10.62	9.24	-12.9%	-13.0%
Imaging	6.47	6.36	5.55	-14.3%	-12.8%
Mobile Phone	5.37	4.83	4.85	-9.7%	0.5%
Phone	4.73	4.83	4.12	-12.9%	-14.7%
TV	5.19	5.22	4.80	-7.7%	-8.2%
TV Box	9.91	9.75	8.28	-16.5%	-15.0%
Simple Average				**-14.6%**	**-10.9%**
Weight Average				**-14.4%**	**-9.9%**

three times the number of returns attributed to the previous months' sales in our data (see Appendix C.8 for a detailed breakdown across categories).

To predict R_j^c, we use the same historical timestamps for calibration, but in the prediction stage, we need a forecast of the purchase pattern in month j. Since producing the sales forecast of new purchases is outside the scope of this research, we proceed with a naïve approach – using the purchase pattern in month $j-12$ as the forecast of the purchase pattern in month j. In other words, we assume consumers' purchasing behavior repeats from the previous year. In practice, a forecaster will most likely substitute this simple method with their preferred sales forecasting method. Appendix C.8 describes how a forecast of total monthly returns is computed by our proposed and the benchmark approaches. Table 4.9 presents the MAE comparison between them. The RMSE results are similar and available upon request.

When predicting R_j, the MAE improvement of our full P-A model over the ARIMA model is 24.5%. In the case of total returns, R_j^t, the performance gap between these two approaches narrows to 14.6%. The direction of this change is expected due to two reasons. First, the predictive ability of the ARIMA model should be relatively stable across R_j and R_j^t, since each case uses one series of actual returns data. Second, the predictive ability of our econometric model should be slightly lower when predicting R_j^t, because a proportion of it (i.e. the same month returns R_j^c) is based not on actual purchase data but on forecasts. The performance gap between the Lagged Sales model and our econometric model also narrows, but the change is at a smaller rate. When predicting R_j, the total MAE improvement is 14.8%. In the case of R_j^t, it is 10.9%. In summary, despite the fact that applying our proposed approach to forecast total monthly returns uses a simplistic approach for predicting purchase patterns, its performance is still substantially better than the benchmark approaches.

4.7 Conclusion

In recent years, retailers have often considered the design of a return policy to be as important as product pricing in generating demand (Brill, 2015). The prevailing practice adopted in the industry is to provide the most lenient return policy – refunding the full purchase price upon return. The consequence of such a policy is the large volume of consumer returns that parties along the reverse supply chain must process, with recent return rates at around 12% and on an upward trend (National Retail Federation, 2016).

The above industry status and trend increase the urgency and importance of optimizing return-related operational activities such as inventory management, staffing of retailers' return counters and OEMs' refurbishing centers, reverse logistics planning, and allocation of returns for reselling and warranty stocks. For all of these activities, an accurate returns forecast is a critical input parameter. Although cost reduction potentials resulting from these activities are believed to be quite promising, the topic of forecasting consumer returns has received little attention in the academic literature.

In this paper, we present a new return forecasting framework that enables retailers and OEMs to better leverage their transaction-level POS data. Our proposed approach follows a predict-aggregate sequence, while existing ones are based on an aggregate-predict sequence. With the same data input, we demonstrate that our approach substantially outperforms existing approaches, even with a very basic setup – so basic in fact that it could be implemented in a simple spreadsheet. In addition, by incrementally building our forecasting model, we show the predictive value of each model component. By more effectively utilizing transaction-level data such as purchase and return timestamps, the basic P-A approach reduces forecast error by 3% to 18% on average for the electronic data, and by 43% to 69% for the jewelry data set, depending on the benchmark model and error estimated used. Such forecast accuracy improvement presents broad implications for inventory, staffing, reverse logistics, and return recovery disposition decisions. We also extend our forecasting model in several directions to make it more generally applicable and show the robustness of its superior performance in a number of empirical settings.

Our study uncovers several interesting areas for future research. First, with our data set, we are unable to quantify the cost savings from better return forecasts, which will depend on the current structure of the reverse supply chain. Thus, estimating these savings under different reverse supply chain structures would be a valuable contribution. Second, the current literature on closed loop supply chains has provided many decision support tools for better managing consumer returns, such as Pince et al. (2016) for the OEM's optimal disposition between which returns to return to the market and which ones to save to meet future warranty demand. Our study provides a methodology for improving the accuracy of one of the key input parameters of these decision support tools – the return volume forecast. Third, a typical idea for forecasting new product sales is to calibrate model parameters on a sample of existing products and then apply the model to new products. Future studies could borrow from new product sales forecasting and adapt our approach to forecast returns of items freshly added to a retailer's assortment.

Chapter 5

Conclusion

The main objective of this dissertation was to advance strategies that can not only provide economic opportunities for firms, but also reduce or eliminate the introduction of waste into the environment. To achieve this, we empirically investigated operational challenges associated with three preferred waste management strategies.

In Chapter 2, we evaluated the potential for demand cannibalization of refurbished products in a multi-generation, multi-condition setting. Fears of product cannibalization have prevented firms from fully engaging in the remanufacturing industry, and despite noted economic and environmental benefits, the sector remains largely underdeveloped. Using structural estimation modeling techniques and data from over 8,000 online transactions we showed that product generation, condition, and seller attributes are highly influential in shaping consumers' purchasing decisions, and that the relationship between new and remanufactured products is much more nuanced and context-specific than previously thought. Counter to industry intuition, we found that remanufactured products pose the same level of threat to new-condition goods as do used goods. Future studies may consider incorporating expanded product types in consumers' choice sets to draw additional insights.

In Chapter 3, we analyzed the effectiveness of legislative tools and consumer education efforts in promoting product reuse and recycling. Although both consumer attributes and legislation seem likely to impact e-recycling, a lack of empirical data has limited firms' and legislators' ability to assess the true impact of these factors. Using survey data from 11 state environmental agencies, we found that in states with EPR legislation, consumers are more likely to recycle their electronic waste. Perhaps less intuitively, we found that consumers' knowledge of landfill bans increases their likelihood to store their electronics rather than recycle them. Our results suggest that efforts that focus on informing consumers about the dangers of e-waste and where they can recycle their electronic products may need to be coupled with e-waste legislation and to significantly promote e-waste recycling. In future studies, understanding how consumers' e-waste behaviors may vary across product types may be an interesting area of inquiry.

In Chapter 4, we developed a robust consumer returns forecasting model to aid operations managers in inventory, reverse logistics, and return recovery decisions. Although lenient return policies have been shown to have marketing benefits such as a higher willingness to pay and a higher purchase frequency, counterbalancing these benefits with an increased volumes of returns presents operational challenges for both retailers and original equipment manufacturers (OEMs). To better manage consumer returns, operations managers need an accurate return forecast as an input into their strategic and tactical tools. We proposed a consumer return forecasting framework and tested our model on datasets provided by brick-and-mortar and online retailers. Finally, we illustrated how our proposed approach demonstrates significant error reduction over benchmark models found in the literature and in practice.

Bibliography

Abbey, J. D., J. D. Blackburn, and V. D. R. Guide (2015). Optimal Pricing for New and Remanufactured Products. *Journal of Operations Management 36*, 130–146.

Abbey, J. D., M. G. Meloy, V. D. R. Guide, and S. Atalay (2015). Remanufactured Products in Closed-Loop Supply Chains for Consumer Goods. *Production and Operations Management 24*(3), 488–503.

Adler, T., C. Falzarano, and G. Spitz (2005). Modeling Service Trade-offs in Air Itinerary Choices. *Transportation Research Record: Journal of the Transportation Research Board* (1915), 20–26.

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

Agrawal, V. V., A. Atasu, and K. Van Ittersum (2015). Remanufacturing, third-party competition, and consumers' perceived value of new products. *Management Science 61*(1), 60–72.

Allon, G., A. Federgruen, and M. Pierson (2011). How Much is a Reduction of Your Customers' Wait Worth? An Empirical Study of the Fast-Food Drive-Thru Industry Based on Structural Estimation Methods. *Manufacturing & Service Operations Management 13*(4), 489–507.

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

Angrist, J., G. Imbens, and A. B. Krueger (1995). Jackknife Instrumental Variables Estimation.

Atasu, A. (2016). *Environmentally Responsible Supply Chains*, Volume 3. Springer.

Atasu, A., Ö. Özdemir, and L. N. Van Wassenhove (2013). Stakeholder perspectives on e-waste take-back legislation. *Production and Operations Management 22*(2), 382–396.

Atasu, A. and R. Subramanian (2012). Extended producer responsibility for e-waste: Individual or collective producer responsibility? *Production and Operations Management 21*(6), 1042–1059.

Atasu, A. and L. N. Wassenhove (2012). An operations perspective on product take-back legislation for e-waste: Theory, practice, and research needs. *Production and Operations Management 21*(3), 407–422.

Bagozzi, R. P. and P. A. Dabholkar (1994). Consumer recycling goals and their effect on decisions to recycle: A means-end chain analysis. *Psychology and Marketing 11*(4), 313–340.

Baldé, C., F. Wang, R. Kuehr, and J. Huisman (2015). The global e-waste monitor 2014: Quantities, flows, and resources. United Nations University.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994). Estimating Discrete-choice Models of Product Differentiation. *The RAND Journal of Economics*, 242–262.

Bodea, T., M. Ferguson, and L. Garrow (2009). Data set – Choice-based revenue management: Data from a major hotel chain. *Manufacturing & Service Operations Management 11*(2), 356–361.

Brill, J. (2015). Rethinking online returns. *UPS White Paper*. Available at https://pressroom.ups.com/pressroom/.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008). Database paper – the IRI marketing data set. *Marketing Science 27*(4), 745–748.

Bruno, H. A. and N. J. Vilcassim (2008). Research Note-structural Demand Estimation with Varying Product Availability. *Marketing Science 27*(6), 1126–1131.

Chintagunta, P. K. and J.-P. Dube (2005). Estimating a Stockkeeping-unit-Level Brand Choice Model that Combines Household Panel Data and Store Data. *Journal of Marketing Research 42*(3), 368–379.

Chuang, H. H.-C., R. Oliva, and O. Perdikaki (2016). Traffic-based labor planning in retail stores. *Production and Operations Management 25*(1), 96–113.

Clottey, T., W. Benton, and R. Srivastava (2012). Forecasting product returns for remanufacturing operations. *Decision Sciences 43*(4), 589–614.

Crocker, K. J. and P. Letizia (2014). Optimal policies for recovering the value of consumer returns. *Production and Operations Management 23*(10), 1667–1680.

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

Dangerfield, B. J. and J. S. Morris (1992). Top-down or bottom-up: Aggregate versus disaggregate extrapolations. *International Journal of Forecasting 8*(2), 233–241.

De Brito, M. P. and R. Dekker (2003). Modelling product returns in inventory control: Exploring the validity of general assumptions. *International Journal of Production Economics 81-82*, 225–241.

Domina, T. and K. Koch (2002). Convenience and frequency of recycling: implications for including textiles in curbside recycling programs. *Environment and behavior 34*(2), 216–238.

Douthit, D., M. Flach, and V. Agarwal (2011). Reducing the quantity and cost of product returns in consumer electronics. *Accenture*.

Dunn, D. M., W. H. Williams, and W. A. Spivey (1971). Analysis and prediction of telephone demand in local geographical areas. *The Bell Journal of Economics and Management Science*, 561–576.

Echegaray, F. and F. V. Hansstein (2017). Assessing the intention-behavior gap in electronic waste recycling: the case of brazil. *Journal of Cleaner Production 142*, 180–190.

Elmaghraby, W., A. Gopal, and A. Pilehvar (2015). Cannibalization and Reference Prices in Secondary Market Auctions for IT Equipment: Evidence from Field Experiments. *Working Paper*.

Enright, T. (2003). Post-holiday logistics. *Traffic World* (January 6), 20.

EPA.gov (2016, February). Promoting and Practicing Environmental Stewardship for Electronic Products. Online. https://archive.epa.gov/wastes/conserve/tools/stewardship/web/html/electronics.html.

Esenduran, G. and E. Kemahlıoğlu-Ziya (2015). A comparison of product take-back compliance schemes. *Production and Operations Management 24*(1), 71–88.

ETBC (2016, January). E-Waste Problem Overview. Electronics Take Back Coalition. Report. http://www.electronicstakeback.com/.

Favot, M. and A. Marini (2013). A statistical analysis of prices of electrical and electronic equipment after the introduction of the weee directive. *Journal of Industrial Ecology 17*(6), 827–834.

Ferguson, M. (2010). *Strategic and Tactical Aspects of Closed-loop Supply Chains*, Volume 8. Now Publishers Inc.

Fleischmann, M., J. M. Bloemhof-Ruwaard, R. Dekker, E. Van der Laan, J. A. Van Nunen, and L. N. Van Wassenhove (1997). Quantitative Models for Reverse Logistics: A Review. *European journal of operational research 103*(1), 1–17.

Garrow, L. A. (2016). *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*. Routledge.

Garrow, L. A. and F. S. Koppelman (2004). Multinomial and Nested Logit Models of Airline Passengers' No-show and Standby Behavior. *Journal of Revenue and Pricing Management 3*(3), 237–253.

Gayle, D. (2012, October). Chemical Breakdown: The Toxic Substances Inside Your Mobile Phone. Online. Read more: http://www.dailymail.co.uk.

Ghose, A., M. D. Smith, and R. Telang (2006). Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact. *Information Systems Research 17*(1), 3–19.

Goldenberg, J., B. Libai, E. Muller, and S. Stremersch (2010). Database submission-the evolving social network of marketing scholars. *Marketing Science 29*(3), 561–567.

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

Guajardo, J. A., M. A. Cohen, and S. Netessine (2015). Service Competition and Product Quality in the US Automobile Industry. *Management Science*.

Guide, V., G. Souza, L. Van Wassenhove, and J. Blackburn (2006). Time value of commercial product returns. *Management Science 52*(8), 1200–1214.

Guide, V. D. R. and J. Li (2010). The Potential for Cannibalization of New Products Sales by Remanufactured Products. *Decision Sciences 41*(3), 547–572.

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

Guide, V. D. R. and L. N. Wassenhove (2001). Managing Product Returns for Remanufacturing. *Production and operations management 10*(2), 142–155.

Hagerty, J. and P. Glader (2011). US News: From Trash Heap to Store Shelf–Refurbished Goods Industry Seeks US Support for Freer Global Trade, more R&D. *Wall Street Journal 24*.

Harms, R. and J. D. Linton (2015). Willingness to Pay for Eco-Certified Refurbished Products: The Effects of Environmental Attitudes and Knowledge. *Journal of industrial ecology*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society 47*(1), 153–161.

Heiman, A., D. R. Just, B. P. McWilliams, and D. Zilberman (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*(February), 105–117.

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

Hyndman, R. J. and Y. Khandakar (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software 26*(3), 1–22.

Iyer, E. S. and R. K. Kashyap (2007). Consumer recycling: Role of incentives, information, and social class. *Journal of Consumer Behaviour 6*(1), 32–47.

Jagabathula, S. (2011). *Nonparametric Choice Modeling: Applications to Operations Management.* Ph. D. thesis, MIT.

Kerr, J. (2013, August 12). Serial returners, beware: Retailers are tracking you. *Today.com*. Available at http://www.today.com/money/serial-returners-beware-retailers-are-tracking-you-6C10900265.

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

Kinney, W. R. (1971). Predicting earnings: entity versus subentity data. *Journal of Accounting Research*, 127–136.

Kök, A. G. and Y. Xu (2011). Optimal and Competitive Assortments with Endogenous Pricing under Hierarchical Consumer Choice Models. *Management Science 57*(9), 1546–1563.

Korn, E. L. and B. I. Graubard (1990). Simultaneous testing of regression coefficients with complex survey data: Use of bonferroni t statistics. *The American Statistician 44*(4), 270–276.

Kremer, M., E. Siemsen, and D. J. Thomas (2015). The sum and its parts: Judgmental hierarchical forecasting. *Management Science forthcoming*.

Lai, K.-h., C. W. Wong, and Y. V. Lun (2014). The role of customer integration in extended producer responsibility: a study of chinese export manufacturers. *International Journal of Production Economics 147*, 284–293.

Lee, B. H. and P. Waddell (2010). Residential Mobility and Location Choice: A Nested Logit Model with Sampling of Alternatives. *Transportation 37*(4), 587–601.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica: Journal of the Econometric Society 51*(2), 507–12.

Li, H. and W. T. Huh (2011). Pricing Multiple Products with the Multinomial Logit and Nested Logit Models: Concavity and Implications. *Manufacturing & Service Operations Management 13*(4), 549–563.

Li, K. J., D. K. Fong, and S. H. Xu (2011). Managing trade-in programs based on product characteristics and customer heterogeneity in business-to-business markets. *Manufacturing & Service Operations Management 13*(1), 108–123.

Long, J. S. and J. Freese (2006). *Regression models for categorical dependent variables using Stata*. Stata press.

Macarthur-Foundation, T. E. (2013, January). Towards the Circular Economy: Economic and business rationale for an accelerated transition. Online. https://www.ellenmacarthurfoundation.org/.

Manski, C. F. and D. McFadden (1981). Alternative Estimators and Sample Designs for Discrete Choice Analysis. *Structural analysis of discrete data with econometric applications*, 2–50.

Mayers, K., R. Peagam, C. France, L. Basson, and R. Clift (2011). Redesigning the camel. *Journal of Industrial Ecology 15*(1), 4–8.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record* (673).

Milovantseva, N. and J.-D. Saphores (2013). E-waste bans and US households' preferences for disposing of their e-waste. *Journal of Environmental Management 124*, 8–16.

Mollenkopf, D., R. Frankel, and I. Russo (2011). Creating value through returns management: Exploring the marketing–operations interface. *Journal of Operations Management 29*(5), 391–403.

MSNBC (2014, December). How remanufacturing is booming across the us. Online Video. http://www.msnbc.com.

Mumbower, S. and L. A. Garrow (2014). Data set-online pricing data for multiple us carriers. *Manufacturing & Service Operations Management 16*(2), 198–203.

Nash, J. and C. Bosso (2013). Extended producer responsibility in the united states. *Journal of Industrial Ecology 17*(2), 175–185.

National Retail Federation (2008, January). Consumer returns in the retail industry. http://www.nrf.com.

National Retail Federation (2016, January). Consumer returns in the retail industry. http://www.nrf.com.

Ng, S. (2015). Zulily tests online returns program. *Wall Street Journal* (June 23, 2016).

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

Ni, J., S. A. Neslin, and B. Sun (2012). Database submission - the ISMS durable goods data sets. *Marketing Science 31*(6), 1008–1013.

OECD (2015, January). Environment at a Glance: OECD Indicators. Online. http://www.oecd-ilibrary.org/environment/.

OECD (2017, January). Material Resources, Productivity, and the Environment: Key Findings. Online. http://www.oecd.org/greengrowth/.

Olivares, M., C. Terwiesch, and L. Cassorla (2008). Structural Estimation of the Newsvendor Model: An Application to Reserving Operating Room Time. *Management Science 54*(1), 41–55.

Ovchinnikov, A. (2011). Revenue and Cost Management for Remanufactured Products. *Production and Operations Management 20*(6), 824–840.

Petersen, J. A. and V. Kumar (2015). Perceived risk, product returns, and optimal resource allocation: Evidence from a field experiment. *Journal of Marketing Research 52*(2), 268–285.

Phillips, E. (2015). Do customers have a "constitutional right" to return stuff ordered online? *Wall Street Journal* (May 1, 2015).

Pince, C., M. Ferguson, and B. Toktay (2016). Extracting maximum value from consumer returns: Allocating between selling refurbished product and meeting warranty demand. *Manufacturing & Service Operations Management forthcoming*.

Pinjari, A. R., R. M. Pendyala, C. R. Bhat, and P. A. Waddell (2007). Modeling Residential Sorting Effects to Understand the Impact of the Built Environment on Commute Mode Choice. *Transportation 34*(5), 557–573.

Pinjari, A. R., R. M. Pendyala, C. R. Bhat, and P. A. Waddell (2011). Modeling the Choice Continuum: An Integrated Model of Residential Location, Auto ownership, Bicycle ownership, and Commute tour Mode Choice Decisions. *Transportation 38*(6), 933–958.

Plambeck, E. and Q. Wang (2009). Effects of e-waste regulation on new product introduction. *Management Science 55*(3), 333–347.

PMA (2016). Product management alliance: Did you know. Online. http://www.productmanagementalliance.org.

Prentice, R. L. (1974). A log gamma model and its maximum likelihood estimation. *Biometrika 61*(3), 539–544.

Prieger, J. E. (2002). A flexible parametric selection model for non-normal data with application to health care usage. *Journal of Applied Econometrics 17*(4), 367.

Ramanathan, R. (2011). An empirical analysis on the influence of risk on relationships between handling of product returns and customer loyalty in e-commerce. *International Journal of Production Economics 130*(2), 255–261.

Reisinger, D. (2010). Best Buy ends (most) restocking fees. *CNET* (December 10, 2010).

Rusmevichientong, P., Z.-J. M. Shen, and D. B. Shmoys (2010). Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint. *Operations research 58*(6), 1666–1680.

Sabbaghi, M., B. Esmaeilian, A. R. Mashhadi, S. Behdad, and W. Cade (2015). An investigation of used electronics return flows: A data-driven approach to capture and predict consumers storage and utilization behavior. *Waste Management 36*, 305–315.

Saphores, J.-D. M., O. A. Ogunseitan, and A. A. Shapiro (2012). Willingness to engage in a pro-environmental behavior: An analysis of e-waste recycling based on a national survey of us households. *Resources, conservation and recycling 60*, 49–63.

Smith, M. D. (2003). Modelling sample selection using archimedean copulas. *The Econometrics Journal 6*(1), 99–123.

Smith, M. D. and R. Telang (2008). Internet Exchanges for Used Digital Goods. *Working Paper*.

Souza, G. C. (2013a). Closed-Loop Supply Chains: A Critical Review, and Future Research. *Decision Sciences 44*(1), 7–38.

Souza, K. (2013b). Retail returns, reverse logistics ramp up for the holidays. *TalkBusiness.net*. Available at http://talkbusiness.net/2013/10/retail-returns-reverse-logistics-ramp-up-for-the-holidays.

Staiger, D. O. and J. H. Stock (1994). Instrumental Variables Regression with Weak Instruments.

Stock, J., T. Speh, and H. Shear (2006). Managing product returns for competitive advantage. *MIT Sloan Management Review 48*(1), 57–62.

Subramanian, R. and R. Subramanyam (2012). Key Factors in the Market for Remanufactured Products. *Manufacturing & Service Operations Management 14*(2), 315–326.

Tabuchi, H. (2015). In season of returning, a start-up tries to find homes for the rejects. *The New York Times* (December 28, 2015).

Toktay, B., E. A. van der Laan, and M. P. Brito (2003). Managing product returns: The role of forecasting. Technical report, Erasmus Research Institute of Management (ERIM).

Toktay, L., L. Wein, and S. Zenios (2000). Inventory management of remanufacturable products. *Management Science 46*(11), 1412–1426.

Turner, K. (2016, September). Apple Made It Cheaper to Repair Cracked iPhone Screens for Some Owners. *The Washington Post*.

USITC (2012a, October). Remanufactured Goods: An Overview of the U.S. and Global Industries, Markets, and Trade. Report. https://www.usitc.gov/publications/332/pub4356.pdf.

USITC (2012b, October). Summary of Remanufactured Goods: An Overview of the U.S. and Global Industries, Markets, and Trade. Online. https://www.usitc.gov/publications.

Viscusi, W. K., J. Huber, and J. Bell (2011). Promoting recycling: Private values, social norms, and economic incentives. *The American Economic Review 101*(3), 65–70.

Vulcano, G., G. Van Ryzin, and W. Chaar (2010). OM Practice Choice-based Revenue Management: An Empirical Study of Estimation and Optimization. *Manufacturing & Service Operations Management 12*(3), 371–392.

Wang, X., F. Mai, and R. H. Chiang (2014). Database submission-market dynamics and user-generated content about tablet computers. *Marketing Science 33*(3), 449–458.

Wang, Z., B. Zhang, J. Yin, and X. Zhang (2011). Willingness and behavior towards e-waste recycling for residents in beijing city, china. *Journal of Cleaner Production 19*(9), 977–984.

Ward, A. (2014, December). Remanufacturing Benefits Both Business and Earth. Online Article. Http://www.aol.com/article/2014/12/15/remanufacturing-benefits-both-business-and- earth/21115507/?gen=1.

Wharton (2016, April). How U.S. Laws Do (and Don't) Support E-Recycling and Reuse.

Willems, S. P. (2008). Data set: Real-world multiechelon supply chains used for inventory optimization. *Manufacturing & service operations management 10*(1), 19–23.

WSJ (2017). Strategies for managing omnichannel returns. *Wall Street Journal*, http://partners.wsj.com/ups/strategies–managing–omnichannel–returns/.

Yin, J., Y. Gao, and H. Xu (2014). Survey and analysis of consumers' behaviour of waste mobile phone recycling in china. *Journal of Cleaner Production 65*, 517–525.

Appendix A

Chapter 2 Appendices

Table A.1: Robustness Check: Nested Model: IV Estimates using 3 months of data [DV: $\ln(s_j) - \ln(s_0)$]

Variable	**Coefficient**	**(Std. Err.)**
Price	-0.01^{***}	(1.76×10^{-3})
$\log(s_{j\|g})$	0.37^{***}	(0.04)
Total Active	$1.01 \times 10^{-3***}$	(1.58×10^{-4})
Feedback Score	1.35	(1.68)
Net Feedback Score	$7.92 \times 10^{-7***}$	(1.86×10^{-7})
Warranty	0.06	(0.08)
Communication Score	-0.02	(0.24)
Shiphand Score	5.09×10^{-4}	(0.26)
Shiptime Score	0.01	(0.22)
Item Description Score	0.09	(0.25)
Warranty Length	-6.37×10^{-4}	(6.05×10^{-4})
Seller Incumbency	1.16×10^{-5}	(7.52×10^{-6})
Net Positive Score	-1.63×10^{-6}	(1.08×10^{-6})
Net Negative Score	$-5.71 \times 10^{-5\wedge}$	(3.35×10^{-5})
Sealed	1.06^{***}	(0.19)
Accessories	-0.03	(0.02)
Return	0.05	(0.05)
Return $*$ Return Policy	$4.61 \times 10^{-3**}$	(1.69×10^{-3})
Shipping	0.18^{*}	(0.07)
Shipping $*$ Shipping Fee	-0.01	(0.01)
Restock	0.03	(0.15)
Restock $*$ Restock Fee	-0.39	(0.84)
3G	-0.05	(0.04)
4G	0.24^{**}	(0.07)
iPad Size 32 GB^a	0.17^{***}	(0.04)
iPad Size 64 GB^a	0.41^{***}	(0.07)
iPad Size 128 GB^a	1.77^{***}	(0.29)
iPad Generation 2^b	0.89^{***}	(0.12)
iPad Generation 3^b	1.75^{***}	(0.21)
iPad Generation 4^b	2.08^{***}	(0.28)
Intercept	-7.84^{***}	(1.64)

a: Holdout Group: IPad Size 16 GB, b: Holdout Group: IPad Generation 1, $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\wedge}p < .1$

Appendix B

Chapter 3 Appendices

Table B.1: List of ERCC Members

2016-2017 ERCC Members:	Membership Type
California Recycle Agency	Voting
Connecticut Department of Energy and Environmental Protection	Voting
Hawaii State Department of Health	Voting
Maine Department of Environmental Protection	Voting
Maryland Department of Education	Voting
Michigan Department of Environmental Quality	Voting
Minnesota Pollution Control Agency	Voting
New Jersey Department of Environmental Protection	Voting
New York State Department of Environmental Conservation	Voting
NC Department of Environmental Quality	Voting
Oregon Department of Environmental Quality	Voting
Pennsylvania Department of Environmental Protection	Voting
Rhode Island DEM	Voting
South Carolina Department of Health and Environmental Control	Voting
Vermont Department of Environmental Conservation	Voting
Washington DC Department of Energy and EnvironmentÊ	Voting
Wisconsin Dept. of Natural Resources.	Voting
Best Buy	Non-voting
Barnes and Noble	Non-voting
Brother International	Non-voting
Consumer Technology Association	Non-voting
Dell	Non-voting
ECS Refining	Non-voting
ERI, Inc.	Non-voting
Funai	Non-voting
IMS Electronics	Non-voting
LG	Non-voting
Microsoft	Non-voting
MRM	Non-voting
Novotec	Non-voting
Panasonic	Non-voting
Pennsylvania Recycling Markets Center	Non-voting
Rhode Island Resource Recovery Corporation	Non-voting
Ricoh	Non-voting
RLGA	Non-voting
Samsung	Non-voting
Sustainable Electronics Recycling International	Non-voting
Tongfang Global	Non-voting
URT	Non-voting
Vintage Tech	Non-voting

ERCC Survey Questions		Response Options
Section 1 - Consumer Recycling Behaviors		
1	What do you typically do with electronics, such as	I put it in storage I gave it to family and friends Donated it Threw it out with the trash Used town or county recycling program or household hazardous waste program Recycled at store
2	In the future, how likely is it that you would recycle electronics such as computers, televisions, printers, and monitors, when you no longer have a use for them?	Very Likely Likely Not Sure Unlikely Very Unlikely
Section 2 - Interest in Recycling Activities		
3	In general, how important or unimportant do you feel it is to recycle electronic devices that you no longer need or use?	Very Important Somewhat Important Neutral Not Very Important Not at all Important
4	Which of the following would most likely prevent you from recycling your electronics?	Too Expensive Inconvenient Recycling Location Unable to transport items Not knowing where to recycle items
5	Thinking about possible locations for dropping off your old electronics, what is the farthest you would be willing to travel to deliver your unwanted device?	1-5 miles away 6-10 miles away 11-20 miles away 21 miles away or more
6	Where do you go to find information on where to recycle your used electronics?	State government agency website National websites, such as greenergadgets.org, earth911.org Manufacturer or retailer website Phone book Other
7	Do you know if it's legal in [your state] to place electronic items into your trash?	Yes No
8	Do you know where you can recycle used, unwanted electronics products such as computers, TVs, monitors and printers?	Yes - I'm certain of where to take them I think I know where to take them No, I don't know where to take them
Section 4 - Demographic Questions		
9	Do you currently live on a farm or in an area with fewer than 50,000 people in your town?	Yes No
10	In which of the following ranges is your current age?	18-35 years old 36-45 years old 46-55 years old 56-65 years old 66 years or older

Figure B.1: Copy of ERCC Consumer Awareness Survey

Table B.2: Odds Ratios from Ungrouped Disposal MNL Model with Trash as Reference Category

Variables	Store	(Std. Error)	Family	(Std. Error)	Donate	(Std. Error)	ReStore	(Std. Error)	ReTown	(Std. Error)
1. EPR: Yes	1.46	(-0.35)	1.28	(-0.3)	1.61**	(-0.36)	1.53*	(-0.38)	2.15***	(-0.54)
2. IMPORTANCE: Not very	1.77	(-1.55)	1.05	(-0.77)	2.83	(-3.4)	-0.28	(-0.28)	0.42	(-0.59)
3. IMPORTANCE: Neutral	2.18	(-1.79)	3.01*	(-1.9)	9.17*	(-10.48)	0.59	(-0.57)	1.75	(-1.79)
4. IMPORTANCE: Somewhat Important	9.58***	(-7.69)	6.08***	(-3.79)	20.42***	(-23.2)	3.16	(-2.78)	8.35**	(-8.17)
5. IMPORTANCE: Very Important	9.80***	(-7.96)	6.69***	(-4.18)	33.82***	(-38.31)	6.12**	(-5.31)	11.35**	(-11.05)
1. INFOSOURCE: General	0.45**	(-0.18)	1.07	(-0.43)	2.28**	(-0.94)	1.14	(-0.6)	1.25	(-0.6)
2. INFOSOURCE: Government	0.73	(-0.3)	1.41	(-0.59)	2.47**	(-1.07)	1.18	(-0.64)	3.40**	(-1.66)
3. INFOSOURCE: Retailer Manufacturer	0.68	(-0.31)	1.33	(-0.61)	2.28*	(-1.05)	2.55*	(-1.42)	0.62	(-0.34)
4. INFOSOURCE: Non Profit	0.73	(-0.32)	2.01	(-0.89)	2.81**	(-1.27)	1.94	(-1.07)	0.85	(-0.45)
1. LEGALAWARENESS: Yes	1.53*	(-0.39)	1.18	(-0.29)	0.95	(-0.22)	1	(-0.26)	1.21	(-0.3)
1. LOCTAWARENESS: Maybe	1.23	(-0.34)	2.44***	(-0.64)	3.53***	(-0.89)	11.76***	(-3.97)	6.15***	(-1.83)
2. LOCTAWARENESS: Yes	1.13	(-0.34)	2.02**	(-0.58)	3.27***	(-0.9)	24.55***	(-8.76)	13.55***	(-4.23)
1. PREVENT: Too Expensive	0.5	(-0.45)	0.4	(-0.32)	0.37	(-0.3)	0.27	(-0.23)	0.21*	(-0.17)
2. PREVENT: Inconvenient Location	0.48	(-0.42)	0.19**	(-0.15)	0.25*	(-0.19)	0.19**	(-0.15)	0.11***	(-0.09)
3. PREVENT: Unable to Transport	0.37	(-0.33)	0.31	(-0.25)	0.45	(-0.35)	0.3	(-0.25)	0.28	(-0.22)
4. PREVENT: Unaware of Location	0.47	(-0.41)	0.22*	(-0.17)	0.25*	(-0.19)	0.22*	(-0.17)	0.20**	(-0.16)
1. RURAL: Yes	1.16	(-0.26)	0.78	(-0.17)	0.74	(-0.15)	0.7	(-0.16)	1.02	(-0.23)
2. AGE: 36-45	1.37	(-0.44)	1.19	(-0.37)	1.54	(-0.46)	1.24	(-0.42)	2.09**	(-0.74)
3. AGE: 46-55	0.42***	(-0.13)	0.68	(-0.2)	0.8	(-0.22)	0.9	(-0.28)	3.62***	(-1.1)
4. AGE: 56-65	0.57*	(-0.19)	0.69	(-0.22)	1.48	(-0.44)	2.17**	(-0.72)	5.30***	(-1.77)
5. AGE: 66+	0.59	(-0.24)	1.01	(-0.39)	1.35	(-0.49)	2.45**	(-0.98)	7.31***	(-2.86)
2. DISTANCE: 1-5 miles	0.62	(-0.25)	1.22	(-0.52)	0.92	(-0.35)	0.96	(-0.39)	1.04	(-0.42)
3. DISTANCE: 6-10 miles	1.27	(-0.51)	2.40**	(-1.02)	1.78	(-0.67)	1.2	(-0.49)	1.82	(-0.73)
4. DISTANCE: 11-20 miles	0.83	(-0.35)	1.71	(-0.76)	1.43	(-0.56)	1.09	(-0.46)	1.41	(-0.58)

****p < .001, ** p <.01, *p<.05, ^ p<.1*

Appendix C

Chapter 4 Appendices

C.1 Data Set Construction

The original Ni et al. (2012) data set contains six transaction types, namely: product purchase, product return, service contract purchase, service contract return, sales discount, and miscellaneous transaction (Table 1A in Ni et al., 2012). As our proposed forecasting framework requires only purchase and return timestamps for each transaction as data input, we are able to use most of the transactions in Ni et al. (2012). We create the data set used in the current study through a few screening procedures. Our goal is to subset the products that are suitable for our analysis and make sure each transaction has the correct purchase and return timestamps. The amount of data reduction for each screening step is presented in Table C.1. We elaborate the screening process as follows.

Table C.1: Data Screening Process

Data Screening Step	Data Reduction	Sample Size
Ni et al. (2012) data set		173,262
Remove service, miscellaneous, and on-line transactions	20,035	153,227
Matching returns and discounts with purchases	27,800	125,427
Product categorization	27,009	98,418
Remove appliances and information goods	12,693	**85,725**

First, we remove the two service transaction types as well as the miscellaneous ones, since it is unclear what the return policy is for the service contracts. In addition, many service contracts were returned along with the product they were attached to. We exclude the “miscellaneous” transactions, because they do not have product descriptions. On-line transactions were a very small fraction (1.9%) of this retailer’s sales when the data was gathered. As a result, we are not in a position to investigate differences between on-line and in-store return behavior, and hence the on-line transactions are eliminated as well.

Second, the remaining three types of observations in Ni et al. (2012) are product purchase, return, and discounts. We match the 14,707 product returns with their corresponding purchase observations, which is necessary because a purchase contains information about whether it was returned but does not show when the return happened. In other words, if a consumer purchases an item and then later returns it, the purchase and return timestamps are entered in two separate observations, which need to be linked for our analysis. Similarly, if a purchased item has a discount, the amount of discount is entered as a separate observation. To obtain the actual price paid for the discounted items, we matched the purchases with the discounts. Next, we describe how the matching is carried out. The purchase-return matching is not straightforward since the original data does not have an identifier variable which directly links a purchase observation with its return observation. We resolve this by sorting the data on certain existing variables that will make the purchase return link. Specifically, we sorted the data by three variables – original ticket number, return indicator, and transaction type, in descending priority. See Ni et al. (2012) for variable definitions. Table C.2 contains an example of the outcomes of this sorting method. Note that if a purchase is returned, both the purchase and the return observation have the return indicator set to "Y". The rationale for this sorting is straightforward – if a purchase is returned, its return observation is now right below the purchase observation. We make two additional checks for each matched purchase-return pair: return comes after purchase, and they have the same price. This procedure is able to match 12,783 (87%) return observations with their purchase counterparts. We also matched the 1,447 observations of purchase discounts with their associated purchases. The matching strategy for the purchase-discount pairs is similar to the purchase-return case, using the original ticket number, product ID, and transaction type as the sorting variables. Note that there is no indicator variable to tell whether a purchase observation is associated with a discount observation. Therefore, we resort to product ID for matching. Again, an example is provided in Table C.2. Although the product ID variable is missing in 24% of the observations, this set of sorting variables matched 1,221 (84%) of the discount observations with their associated purchases. To sum-

Table C.2: Examples of Matched Purchase-Return and Purchase-Discount Pairs

	Original Ticket Number	Product ID	Return Indicator	Transaction Type	Unit Price	Date
			(purchase-return matching example)			
Purchase	86702346281	538012	Y	1	39.99	Feb 16 2000
Return	86702346281	538012	Y	2	-39.99	Feb 17 2000
			(purchase-discount matching example)			
Purchase	373201712537	808789	N	1	179.99	Dec 15 2012
Discount	373201712537	808789	N	5	-4.99	Dec 15 2012

marize, the reduction of 27,800 observations in this screening step is due to three reasons: merging of returns and discounts into purchases, missing product IDs, and missing return indicators.

Third, we categorize products so that a separate analysis could be conducted for each category. Note that this data set is not appropriate for product-specific analysis, since a product on average has one return per 23 days. Our categorization is a fine-tuning on the original categorization in Ni et al. (2012), which blends the accessories for a product along with the product itself into the same category. For example, many products in the TV category are in fact TV stands and cables. Table C.3 contains details of the categorization process. Some transactions have either no or uninformative product descriptions. Along with the transactions that do not fit into any of the categories in Table C.3, we have left 27,009 transactions uncategorized, which are removed in this screening step.

Last, we ensure that each category has a fair number of returns. The "appliances" category has 58 returns in total, which makes it inappropriate for analysis. The "information goods" category has a very low return rate (4.5%) and a very short return lag (10 days). In addition, the common return policy for information goods is to not accept open-box items for refund, which is distinct from other products. Therefore, we exclude appliances and information goods from our final sample for the forecasting analysis.

C.2 Heteroskedasticity and Sample Selection Models

The forecasting accuracy of the full model, after further adding in sample selection, is also presented in Table C.4 (denoted by Het. & Sel.). The accuracy gain in this case shows a steady increase from the previous case where dependence between the return probability

Table C.3: Product Categorization

Category	Observations	Percentage	Subcategory No. in Appendix I in Ni et al. (2012)
Appliances	1,674	1.33%	128, 129, 200, 210, 215, 219, 220, 221, 230, 232, 239, 240, 253, 254, 305, 306, 312, 315, 318, 319, 321, 329, 355, 600
Audio	11,367	9.06%	132, 272, 273, 393, 605, 609
Auto Parts	7,211	5.75%	108, 110, 111, 123, 124, 142, 143, 170
Cable	9,398	7.49%	179, 181, 182, 185, 186, 334, 389, 390, 391
Computer	21,811	17.39%	400, 401, 402, 405, 410, 415, 420, 425, 427, 430, 435, 440, 475, 480, 525
Imaging	7,823	6.24%	274, 275, 276, 277, 279, 282, 283, 288, 289, 290, 291, 298
Information Goods	11,019	8.79%	1, 9, 11, 13, 15, 17, 19, 21, 38, 41, 43, 45, 47, 56, 65, 74, 77, 376, 377, 380, 384, 392, 502, 503, 504, 505, 506, 507, 508
Mobile Phone	3,126	2.49%	323, 338, 366, 369, 370, 371, 372, 373, 375, 383, 387, 395, 397, 398, 591
Phone	5,777	4.61%	320, 360, 362, 363, 365, 382, 386
TV	9,413	7.50%	100, 102, 103, 104, 117, 120, 121, 125, 126, 127, 130, 131, 172, 174, 263, 347
TV Box	9,799	7.81%	292, 294, 295, 296, 304
Not Categorized	27,009	21.53%	
Total	125,427	100.00%	

Description for each subcategory are in Appendix I in Ni et al. (2012). "Auto parts" are off-market speaker, video, and GPS systems for cars. "TV box" are recorder, DVR, VCR, DVD player, and other box-shaped items that connect to TV.

and experience duration is not incorporated. Unlike the heteroskedasticity-only model, however, the dependence between the two events appears to apply similarly to all product categories.

Table C.4: Comparison within the Predict-Aggregate Approach

	Het.		Het. vs Basic		Het. & Sel.		Het. & Sel. vs Basic	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Audio	2.84	4.54	-16.9%	-18.3%	2.48	3.96	-27.4%	-28.7%
Auto Parts	2.72	3.52	-1.8%	-1.3%	2.70	3.50	-2.5%	-1.9%
Cable	2.28	3.25	-0.3%	-0.3%	2.30	3.29	0.7%	0.8%
Computer	4.03	6.32	-7.7%	-6.8%	4.07	6.17	-6.6%	-9.0%
Imaging	1.96	2.53	-8.4%	-9.7%	1.95	2.55	-9.0%	-9.1%
Mobile Phone	2.40	3.49	-6.0%	-6.5%	2.38	3.45	-7.0%	-7.5%
Phone	2.13	3.25	-1.0%	-2.2%	2.16	3.32	0.5%	0.2%
TV	1.90	3.38	-5.1%	-10.1%	1.82	3.17	-8.7%	-15.5%
TV Box	3.81	9.06	-5.0%	-5.6%	3.69	8.52	-7.9%	-11.3%
Simple Avg.			**-5.8%**	**-6.8%**			**-7.6%**	**-9.1%**
Weighted Avg.			**-5.6%**	**-6.4%**			**-7.3%**	**-8.6%**

Abbreviations: Het. = Heteroskedasticity; Sel. = Selection; Avg. = Average. The percentages are calculated as het. (or het. & sel.) measures divided by basic measures minus one. Weighted average is calculated as percentage measures weighted by categorical return rates (see Table 4.3).

C.3 Sample Selection versus Censored Regression (Tobit) Models

Although both models deal with some form of missing information in the data, they are fundamentally different, the key of which is the variable that determines the censoring rule. Suppose the data generating process (i.e. population regression) is $y = \beta x + \epsilon$. The Tobit model deals with the situation where the dependent variable y is "cut off" at a specific value, e.g. 0. In this case, we observe y for $y > 0$ and observe 0 for $y \leq 0$. In other words, the censoring rule of y is defined by y itself. However, the sample selection model deals with a situation where y is unobserved because of a "third variable" – a variable outside the x-y relationship. Moreover, when y is unobserved, we do not have any information on the value of y (e.g. whether it is greater or smaller than zero). In our context, experience duration is not censored by itself but by the decision to return. When an item is returned, we observe its experience duration. Otherwise, we have no information on experience duration. Thus, we employ the sample selection model, but not the Tobit model.

C.4 Predict-Aggregate Approach with Arbitrary Distributions

In the following, we describe a copula-based approach to generalize the distributional choice in the prediction model of our P-A approach. This relaxation is most important for the duration part, since the return probability is modeled through a binary choice and the common distributional choices, normal and logistic, produce highly similar results. Therefore, we demonstrate the general setup by choosing an alternative distribution for experience duration, while staying with the normal distribution for return probability. Specifically, the generalized gamma distribution is chosen from a set of commonly used distributions for time data based on likelihood ratio tests. A side benefit of generalizing the distributional choice is that we are able to directly test how well the distributions suggested by Hess and Mayhew (1997) work in our data. Specifically, they use a logit model for return probability and a five-parameter survival distribution for experience duration.

Let r_i be the latent variable that determines the return-versus-keep decision, such that $r_i > 0$ for returns and $r_i < 0$ for keeps. Its PDF and CDF are denoted by f_r and F_r.

The PDF and CDF for experience duration, d_i, are f_d and F_d, respectively. If dependence between return probability and experience duration is ignored (i.e. $corr(r_i, d_i) = 0$), r_i and d_i are estimated separately, which is straightforward. To account for the dependence, we resort to the copula approach that creates the bivariate CDF, $F_{r,d}$, from the two marginal CDFs, F_r and F_d. The copula approach is commonly used in the econometric literature to accommodate sample selection models with non-normal marginal distributions (e.g. Lee, 1983; Prieger, 2002; Smith, 2003). Specifically, $F_{r,d} = C(F_r, F_d; \theta)$, where $C(\cdot, \cdot; \theta)$ is called the copula function and the θ parameter measures the correlation between r_i and d_i. We use Frank's copula, $C(F_r, F_d; \theta) = -\theta^{-1} \log(1 + \frac{(e^{-\theta F_r}-1)(e^{-\theta F_d}-1)}{e^{-\theta}-1})$, which is both simple to estimate and stable in our data. See Smith (2003) for a review of different copula functions for the sample selection model.

Next, we are ready to construct the ML estimator, using the same flow as in §4.6.1. For the kept purchases, the likelihood is $L_i|_{I_i=0} = F_r(0)$. For the returned purchases, the likelihood is given by

$$L_i|_{I_i=1} = \Pr(r_i > 0, d_i) = \frac{e^{\theta F_d}(e^{\theta F_r} - e^{\theta})}{e^{\theta(F_r+F_d)} + e^{\theta}(1 - e^{\theta F_r} - e^{\theta F_d})} f_d$$

where $f_d(d_i)$, $F_d(d_i)$, and $F_r(0)$ are abbreviated as f_d, F_d, and F_r in the final expression. See Appendix C.5 for a detailed derivation. The log-likelihood for the whole sample is therefore: $LL = \sum_{i=1}^{N}\{(1 - I_i)L_i|_{I_i=0} + I_i L_i|_{I_i=1}\}$. Note that this likelihood function does not involve pre-specified distributions and hence allows general distribution choices.

Before selecting a specific distributional choice for experience duration, we test a number of potential candidates on the return lag data. Table C.5 shows the log-likelihood of these fitted distributions. The increment in log-likelihood from the one-parameter exponential distribution to the two-parameter Weibull is fairly large. In addition, while both involves two parameters, the log-normal distribution clearly fits our data better than Weibull, signified by the uniformly higher log-likelihoods across categories. Therefore, the log-normal distribution appears to be a good choice if we use a duration distribution that involves two or fewer parameters. To explore whether adding more "flexibility" to the duration distribution is beneficial, we also fitted the three-parameter generalized gamma distribution

Table C.5: Log-Likelihood of Different Distributions Fitted to Return Lags

	One Parameter	Two Parameters		Three Parameters
	Exponential	Weibull	Log-Normal	Generalized Gamma
Audio	-4779	-4737	-4662	-4661
Auto Parts	-4291	-4003	-3886	-3879*
Cable	-3461	-3410	-3297	-3281*
Computer	-8116	-7961	-7697	-7650*
Imaging	-2920	-2894	-2839	-2838
Mobile Phone	-2238	-2057	-1999	-1995*
Phone	-3127	-3067	-2999	-2996
TV	-2807	-2749	-2678	-2664*
TV Box	-4997	-4896	-4773	-4766*

* Likelihood ratio test between generalized gamma and log-normal distributions is significant at $p = 0.01$ level.

(the PDF and CDF for this distribution are given in Prentice (1974); accessible through the `flexsurv` package in R). The increase in log-likelihood caused by this additional parameter appears to be more pronounced in certain categories. Since generalized gamma nests log-normal as a special case, we implement a likelihood ratio test to examine whether the additional goodness-of-fit provided by the former is statistically significant (significance indicated by stars in Table C.5). Indeed, significance is found in certain categories.

In the following, we proceed with the generalized gamma distribution to illustrate the P-A approach with arbitrary distributions. For the return decision, we still use the normal distribution. The probability of transaction i being returned during month j is given by

$$P_{ij} = \Pr(r_i > 0, t_j^{end} - t_i > d_i > t_j^{start} - t_i) = \int_{t_j^{start}-t_i}^{t_j^{end}-t_i} \int_0^{+\infty} C(F_r, F_d; \theta) dr_i dd_i$$

The estimation of the likelihood function and calculation of P_{ij} are implemented through R's `bbmle` and `copula` packages, respectively.

We compare the forecasting accuracy of using the generalized gamma distribution for experience duration against using the log-normal distribution. We do this comparison for the three scenarios shown in §4.3.3 to §4.6.1. The results are presented in Table C.6. Since MAE and RMSE results are similar, we present the former, with the latter are available upon request. We observe that increasing the "distributional flexibility" in experience duration does not yield significantly better forecasts in our data. Specifically, while the generalized

Table C.6: Comparison of MAE of Predict-Aggregate Approaches with Log-Normal and Generalized Gamma Distributions for Experience Duration

	Panel A: Basic			Panel B: Het.			Panel C: Het. & Sel.		
	L. N.	G. G.	Diff.	L. N.	G. G.	Diff.	L. N.	G. G.	Diff.
Audio	3.42	3.41	-0.1%	2.84	2.81	-1.0%	2.48	2.82	13.6%
Auto Parts	2.77	2.71	-2.0%	2.72	2.69	-1.0%	2.70	2.69	-0.4%
Cable	2.29	2.28	-0.6%	2.28	2.27	-0.6%	2.30	2.27	-1.7%
Computer	4.36	4.33	-0.8%	4.03	4.10	1.9%	4.07	4.12	1.2%
Imaging	2.14	2.14	-0.1%	1.96	1.96	-0.1%	1.95	1.97	0.7%
Mobile Phone	2.55	2.50	-2.2%	2.40	2.35	-2.1%	2.38	2.35	-1.1%
Phone	2.15	2.15	0.3%	2.13	2.13	0.1%	2.16	2.13	-1.3%
TV	2.00	1.95	-2.3%	1.90	1.87	-1.7%	1.82	1.87	2.3%
TV Box	4.01	3.98	-0.7%	3.81	3.79	-0.5%	3.69	3.79	2.7%
Simple Average			**-1.0%**			**-0.6%**			**1.8%**
Weighted Average			**-1.0%**			**-0.6%**			**1.6%**

gamma distribution performs better in the basic and heteroskedastic scenarios (Panels A and B), its performance is lower in the last scenario (Panel C). In addition, the performance difference is very small across all scenarios. Therefore, the log-normal distribution appears to be a good choice for our data.

Hess and Mayhew (1997) suggest the use of a logit regression for estimating the return probability, $P^r_{ij} = \frac{e^{\mu_r}}{1+e^{\mu_r}}$, and a five-parameter non-negative distribution for estimating the experience duration, whose CDF and PDF are as follows. $F_d = 1 - e^{-\alpha_1\left(-\text{erf}(\alpha_3)+\text{erf}(\alpha_2 t+\alpha_3)-e^{\alpha_4}\right)e^{\beta}}$ and $f_d = (1-F_d)\left(\frac{2\alpha_1\alpha_2}{\sqrt{\pi}}e^{-(\alpha_2 t+\alpha_3)^2} + e^{\alpha_4}\right)e^{\beta}$, where $\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}$ is the error function. As discussed above, using the three-parameter generalized gamma distribution for experience duration does not clearly outperform the more parsimonious two-parameter log-normal distribution in forecasting returns with our data. In addition, the predictive ability of probit and logit regressions are often similar. Therefore, adding further flexibility into the duration distribution, as in Hess and Mayhew (1997), is not expected to improve forecasting accuracy with our data. Since they do not consider heteroskedasticity or sample selection, we explore the forecasting performance of their distributions with the basic model in §4.3.3. Table C.7 contains a comparison between the forecasting accuracy of "probit & log-normal distributions" (those used in §4.3.3) and "Hess and Mayhew (1997) distributions." We removed the mobile phone category because the ML estimation of the Hess and Mayhew (1997) duration distribution did not converge in this case. Performance between the two distribution sets is very similar across most of the categories. However, the five-parameter duration distribution appears to overfit our data for the auto parts and TV categories, which is evident by the higher MAE and RMSE measures.

Table C.7: Comparison of Predict-Aggregate Approaches with Different Distributions

	Probit & Log-Normal		Hess and Mayhew (1997)		Comparison	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Audio	3.42	5.55	3.38	5.61	-1.2%	1.0%
Auto Parts	2.77	3.57	2.86	3.79	3.4%	6.2%
Cable	2.29	3.27	2.29	3.26	0.0%	-0.2%
Computer	4.36	6.78	4.43	6.85	1.5%	1.1%
Imaging	2.14	2.80	2.16	2.86	1.0%	2.1%
Phone	2.15	3.32	2.18	3.42	1.5%	3.1%
TV	2.00	3.76	2.07	4.06	3.4%	8.0%
TV Box	4.01	9.61	3.98	9.49	-0.7%	-1.2%
Simple Average					**1.1%**	**2.5%**
Weighted Average					**1.1%**	**2.4%**

C.5 Derivations

Derivations for §4.6.1. $\Pr(\epsilon_r > -\mu_r \cup \epsilon_d = \log(d_i) - \mu_d) = f_d(\log(d_i - \mu_d)) \Pr(\epsilon_r > -\mu_r | \epsilon_d = \log(d_i - \mu_d))$. The first term is simply $\frac{1}{\sigma_d}\phi(\frac{\log(d_i)-\mu_d}{\sigma_d})$, while evaluating the second terms requires invoking the property of the conditional bivariate normal distribution. Conditional on $\epsilon_d = \log(d_i - \mu_d)$, ϵ_r is normally distributed as $\mathcal{N}\big(\frac{\rho\sigma_r}{\sigma_d}(\log(d_i) - \mu_d), \sigma_r^2(1-\rho^2)\big)$. Therefore, the second term is given by $1 - \Phi(\frac{-\mu_r - \frac{\rho\sigma_r}{\sigma_d}(\log(d_i)-\mu_d)}{\sigma_r\sqrt{1-\rho^2}}) = \Phi(\frac{\frac{\mu_r}{\sigma_r} + \frac{\rho}{\sigma_d}(\log(d_i)-\mu_d)}{\sqrt{1-\rho^2}})$. Assembling the two terms yields the expression in (4.2).

Derivations for Appendix C.4. $\Pr(r_i > 0, d_i) = \Pr(r_i > 0) f_{d|r}(d_i | r_i > 0)$, where the latter term is the conditional PDF of d_i given $r_i > 0$. $f_{d|r}(d_i|r_i > 0) = \frac{\partial}{\partial d_i} F_{d|r}(d_i | r_i > 0) = \frac{\partial}{\partial d_i}\big(\frac{F_d - C(F_r, F_d; \theta)}{1 - F_r}\big) = \frac{f_d - \frac{\partial C(F_r, F_d;\theta)}{\partial d_i}}{1-F_r}$. Applying the chain rule to $\frac{\partial C(F_r,F_d;\theta)}{\partial d_i}$, we have $\frac{\partial C(F_r,F_d;\theta)}{\partial F_d}\frac{\partial F_d}{\partial d_i} = \frac{\partial C(F_r,F_d;\theta)}{\partial F_d} f_d$. Therefore, $\Pr(r_i > 0, d_i) = f_d - \frac{\partial C(F_r,F_d;\theta)}{\partial F_d}$. Plugging in the expression of $C(F_r, F_d; \theta)$ and simplifying the partial derivative, we obtain $\Pr(r_i > 0, d_i) = \frac{e^{\theta F_d}(e^{\theta F_r} - e^{\theta})}{e^{\theta(F_r + F_d)} + e^{\theta}(1 - e^{\theta F_r} - e^{\theta F_d})} f_d$.

C.6 Predict-Aggregate Approach with Inflated Same-Day Returns

In §4.3.3, we assumed that consumers' experience duration, d_i, follows a log-normal distribution, regardless of how short it is. To give the same-day returns a special treatment, we modify this distribution to allow for a discrete point at $d_i = 1$ with probability ϕ. For $d_i > 1$, we still assume it is log-normally distributed. Essentially, we have modified the distribution of experience duration into a mixture of continuous and discrete components. Since incorporating the dependence between return probability and experience duration requires a continuous distribution for the latter, we

Table C.8: Comparison of MAE of Predict-Aggregate Approaches with Continuous and Mixture Distributions for Experience Duration

	Panel A: Basic			Panel B: Heteroskedastic		
	Continuous	Mixture	Difference	Continuous	Mixture	Difference
Audio	3.42	3.43	0.4%	2.84	2.83	-0.3%
Auto Parts	2.77	2.72	-1.7%	2.72	2.67	-1.7%
Cable	2.29	2.28	-0.3%	2.28	2.28	-0.3%
Computer	4.36	4.30	-1.4%	4.03	3.96	-1.8%
Imaging	2.14	2.14	-0.2%	1.96	1.96	-0.3%
Mobile Phone	2.55	2.57	0.7%	2.40	2.42	0.7%
Phone	2.15	2.14	-0.4%	2.13	2.12	-0.5%
TV	2.00	1.99	-0.6%	1.90	1.88	-0.7%
TV Box	4.01	4.02	0.2%	3.81	3.81	0.2%
Simple Average			**-0.4%**			**-0.5%**
Weight Average			**-0.3%**			**-0.5%**

restrict the above modification to the basic and heteroskedastic models. Table C.8 presents the results. It appears that paying separate attention to the same-day returns provides only a marginal improvement in forecasting accuracy; the MAE reduction is less than 1% under both the basic and heteroskedastic models.

C.7 Predict-Aggregate Approach with Additional Predictors

We linearly parameterize the mean of the return utility, μ_r, and the mean of the logged experience duration, μ_d. Let $\mu_r = X\beta$ and $\mu_d = Z\gamma$, where X and Z are exogenous variables, and β and γ are coefficient vectors. We include a purchase price variable in X and Z, and also a holiday dummy variable in Z. We do not include the holiday dummy in X because the holiday effect in return probability is already captured by the heteroskedasticity and entering it again in X results in over fitting. Results are presented in Table C.9. As shown by the percentage differences, the additional information contained in the exogenous variables increases the accuracy of returns forecasting. In addition, the forecast error reduction is most substantive with the basic Predict-Aggregate approach (e.g. the simple average is 3.9%) and decreases as the model specification becomes more sophisticated. Therefore, the accuracy gain increases at a diminishing rate as the overall model sophistication increases.

Table C.9: Comparison of MAE of Predict-Aggregate Approaches with and without Additional Predictors

	Without Exogenous Vars			With Exogenous Vars			Difference		
	Basic	Het.	Het. & Sel.	Basic	Het.	Het. & Sel.	Basic	Het.	Het. & Sel.
Audio	3.42	2.84	2.48	3.08	2.52	2.52	-9.9%	-11.3%	1.7%
Auto Parts	2.77	2.72	2.70	2.72	2.70	2.70	-1.9%	-0.8%	0.0%
Cable	2.29	2.28	2.30	2.11	2.12	2.12	-7.7%	-7.0%	-7.9%
Computer	4.36	4.03	4.07	4.24	4.14	4.14	-2.8%	2.7%	1.6%
Imaging	2.14	1.96	1.95	2.10	1.95	1.95	-2.0%	-0.9%	-0.2%
Mobile Phone	2.55	2.40	2.38	2.52	2.36	2.35	-1.4%	-1.7%	-1.1%
Phone	2.15	2.13	2.16	2.14	2.13	2.13	-0.3%	0.2%	-1.2%
TV	2.00	1.90	1.82	1.91	1.82	1.81	-4.7%	-4.1%	-0.6%
TV Box	4.01	3.81	3.69	3.84	3.72	3.72	-4.2%	-2.3%	0.8%
Simple Average							**-3.9%**	**-2.8%**	**-0.8%**
Weighted Average							**-3.6%**	**-2.6%**	**-0.7%**

Table C.10: Breakdown of Monthly Returns

	R_j^t	R_j^c	R_j	R_j/R_j^t
Audio	18.44	12.00	6.44	34.9%
Auto Parts	14.51	9.52	4.99	34.4%
Cable	13.93	10.46	3.46	24.9%
Computer	32.41	23.68	8.73	26.9%
Imaging	11.35	7.44	3.92	34.5%
Mobile Phone	7.27	4.60	2.67	36.7%
Phone	11.44	7.31	4.13	36.1%
TV	11.17	7.79	3.38	30.3%
TV Box	18.63	11.94	6.69	35.9%
Average	**15.46**	**10.53**	**4.93**	**32.7%**

C.8 Forecasting Total Monthly Returns

Table C.10 contains descriptive statistics for the three return quantities. We use the same econometric setup as in §4.6.1 for our P-A approach. It follows that the probability of each purchase made in month j being returned in the same month is obtained by modifying (4.4) such that $P_{ij} = \Pr(\epsilon_r > -\mu_r, e^{\mu_d+\epsilon_d} < t_j^{end} - t_i) = \int_{-\infty}^{\log(t_j^{end}-t_i)-\mu_d} \int_{-\mu_r}^{+\infty} F_{r,d}(\epsilon_r, \epsilon_d) d\epsilon_r d\epsilon_d$. Then, R_j^c is simply the sum of all these probabilities. R_j is predicted the same way as in §4.6.1. Thus, our P-A forecast of R_j^t is obtained by adding up R_j^c and R_j. Forecasting R_j^t with the aggregate models (ARIMA and Lagged Sales) are rather straightforward. We replace the series of R_j with R_j^t and then proceed in the same manner as in §4.5.1.