ESSAYS ON MULTICHANNEL MARKETING

A Dissertation

by

TARUN LALBAHADUR KUSHWAHA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2007

Major Subject: Marketing

ESSAYS ON MULTICHANNEL MARKETING

A Dissertation

by

TARUN LALBAHADUR KUSHWAHA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Venkatesh Shankar
Committee Members,	David A. Bessler
	Arvind Rangaswamy
	Alina B. Sorescu
	P. Rajan Varadarajan
Head of Department,	Jeffrey S. Conant

August 2007

Major Subject: Marketing

ABSTRACT

Essays on Multichannel Marketing. (August 2007)

Tarun Lalbahadur Kushwaha, B.Sc., Gujarat University; M.A.S., Nirma University;

M.S., Texas A&M University

Chair of Advisory Committee: Dr. Venkatesh Shankar

Multichannel marketing is the practice of simultaneously offering information, goods, services, and support to customers through two or more synchronized channels. In this dissertation, I develop an integrated framework of multichannel marketing and develop models to assist managers in their marketing resource allocation decisions. In the first essay of the dissertation, I investigate the factors that drive customers multichannel shopping behavior and identify its consequences for retailers. In the second essay, I build on this work and develop a model that enables firms to optimize their allocation of marketing resources across different customer-channel segments.

In the first essay, I develop a framework comprising the factors that drive consumers' channel choice, the consequences of channel choice, and their implications for managing channel equity. The results show that customer-channel choice is driven in a nonlinear fashion by a customer demographic variable such as age and is also influenced by consumer shopping traits such as number of categories bought and the duration of relationship with a retailer. I show that by controlling for the moderating effects of channel-category associations, the influence of customers' demographics and

shopping traits on their channel choices can vary significantly across product categories. Importantly, the results show that multichannel shoppers buy more often, buy more items, and spend considerably more than single channel shoppers. The channel equity of multichannel customers is nearly twice that of the closest single channel customers (online or offline).

In the second essay, I propose a model for optimal allocation of marketing efforts across multiple customer-channel segments. I first develop a set of models for consumer response to marketing efforts for each channel-customer segment. This set comprises four models, the first for purchase frequency, the second for purchase quantity, the third for product return behavior, and the fourth for contribution margin of purchase. The results show that customers' responses to firm marketing efforts vary significantly across the customer-channel segments. They also suggest that marketing efforts influence purchase frequency, purchase quantity and monetary value in different ways. The resource allocation results show that profits can be substantially improved by reallocating marketing efforts across the different customer-channel segments.

DEDICATION

To my family

I dedicate this dissertation to my parents, wife, and sister. My parents have encouraged and supported me in all my professional and personal endeavors. They have been my inspiration in the pursuit of my doctoral degree. My family has made numerous personal sacrifices to ensure that I grow up in a nurturing environment where I can pursue my dreams. This dissertation is a tribute to their love and affection.

For the last three years my wife, Shipra, has been my closest friend and confidant. She has stood by my side during the ups and downs in my life. During this time without complaining even once, she has sacrificed our personal time together, tolerated my frustrations, and understood my disappointments. No words can express my gratitude for her unwavering support. This dissertation honors her dedication and love.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Venkatesh Shankar, and my committee members, Dr. David Bessler, Dr. Arvind Rangaswamy, Dr. Alina Sorescu, and Dr. Rajan Varadarajan for their guidance and support throughout the course of this research.

Dr. Shankar has been instrumental in providing me a sense of direction and has been a constant source of encouragement throughout my doctoral program. He has time and again helped me by providing his invaluable experience and methodological expertise to guide this, as well as numerous other research projects, towards fruitful outcomes. He has been an excellent mentor to me, who has taught me various nuances about the marketing academia. During my job search phase, he has been my biggest cheer leader, who has unflinchingly vouched for my research potential, tirelessly made dozens of phone calls to prospective employers, and wholeheartedly supported my decisions. I am deeply indebted for his dedication toward my well being.

Along with my advisor Dr. Shankar, Alina has been one of the biggest influences in evolution of my research thinking. As her research assistant, I learned about raising interesting research questions, linking research objectives to novel empirical contexts, and collecting secondary data through numerous sources for seeking answers to those interesting research questions. She has been extremely generous with her time in providing comments on drafts of manuscripts, feedback on job market talks, and guidance on several professional matters throughout my stay in the doctoral program.

I wish to thank Dr. Rajan for creating a nurturing environment for doctoral students in the department. During his tenure as Department Head, he ensured that doctoral students had adequate resources and were protected from any activities that might have posed competing demand on their time for research. His commitment is so strong towards doctoral student research that he gave up his generous personal computing budget to purchase a large computer server for exclusive use by doctoral students. This computer server was used for data analysis in this dissertation.

I also wish to thank Dr. Bessler and Dr. Rangaswamy for providing me valuable guidance during the dissertation research. As a domain expert in multichannel marketing and highly regarded marketing modeler, Dr. Rangaswamy has provided precious insights to the conceptual and empirical work in this research. His exceptional feedback has helped me fill several holes in this research. Dr. Bessler has helped me to improve this work by providing a fresh outside perspective to this research. His econometric expertise has tremendously helped me in addressing several methodological concerns.

I want to extend my gratitude to Dr. Leonard Berry for helping me obtain data from an anonymous company. I want to acknowledge this anonymous company, iBehavior, and Direct Marketing Education Foundation for providing critical data and doctoral support required for successful completion of this dissertation.

Last but not the least, I wish to thank professors, support staff, and doctoral colleagues in the department for their help and support. I especially wish to thank my doctoral colleague and close friend, Kartik Kalaignanam, for patiently listening to my problems and giving me useful advice at several junctures during the past seven years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

There's not a customer who wants to shop just by catalog or just by store. Everybody has different needs and those needs may even differ depending on the day or the time of day.

-Richard Thalheimer, CEO and Founder, Sharper Image

The above quote captures the essence of the complex decisions that managers face with respect to multichannel shopping and marketing, a topic that has generated tremendous interest among manufacturers, service providers, and retailers. A study by *DoubleClick* (2004) found that in 2003, 65% of consumers[1] were multichannel shoppers, an increase of more than 16% over the previous year. A similar study by Forrester research revealed that more than two thirds of the consumers search products online, but make a purchase offline (Johnson 2004). The *Marketing Science Institute* (2006) has included multichannel marketing in its top-tier research priorities for 2004-2006 for the Customer Management Community under the title "Managing and Maintaining Customers through Multiple Channels". Practitioners and researchers, therefore, need a deeper understanding of this phenomenon.

Multichannel marketing is the practice of simultaneously offering customers information, product, services, and support through two or more synchronized channels (Rangaswamy and Van Bruggen 2005). Organizations deploy marketing resources

This dissertation follows the style of *Journal of Marketing Research*.

[1] For expositional use, I use the terms, consumer, customer, and shopper interchangeably throughout the dissertation.

through multiple channels such as email, direct mail, the web, call center, direct sales force, and physical store (Levy and Weitz 2006; Neslin et al. 2006). In addition to offering customers more channels to shop from, multichannel marketing provides organizations with greater opportunities to interact with customers, promote other channels, use price differentiation tools, segment customers specific to a channel, and target product categories to specific customer segments. Thus, multichannel marketing capability allows organizations to build stronger relationships with the right customers. Therefore, a better understanding of multichannel purchase behavior, its drivers and consequences is important for marketing managers to formulate sound channel strategies.

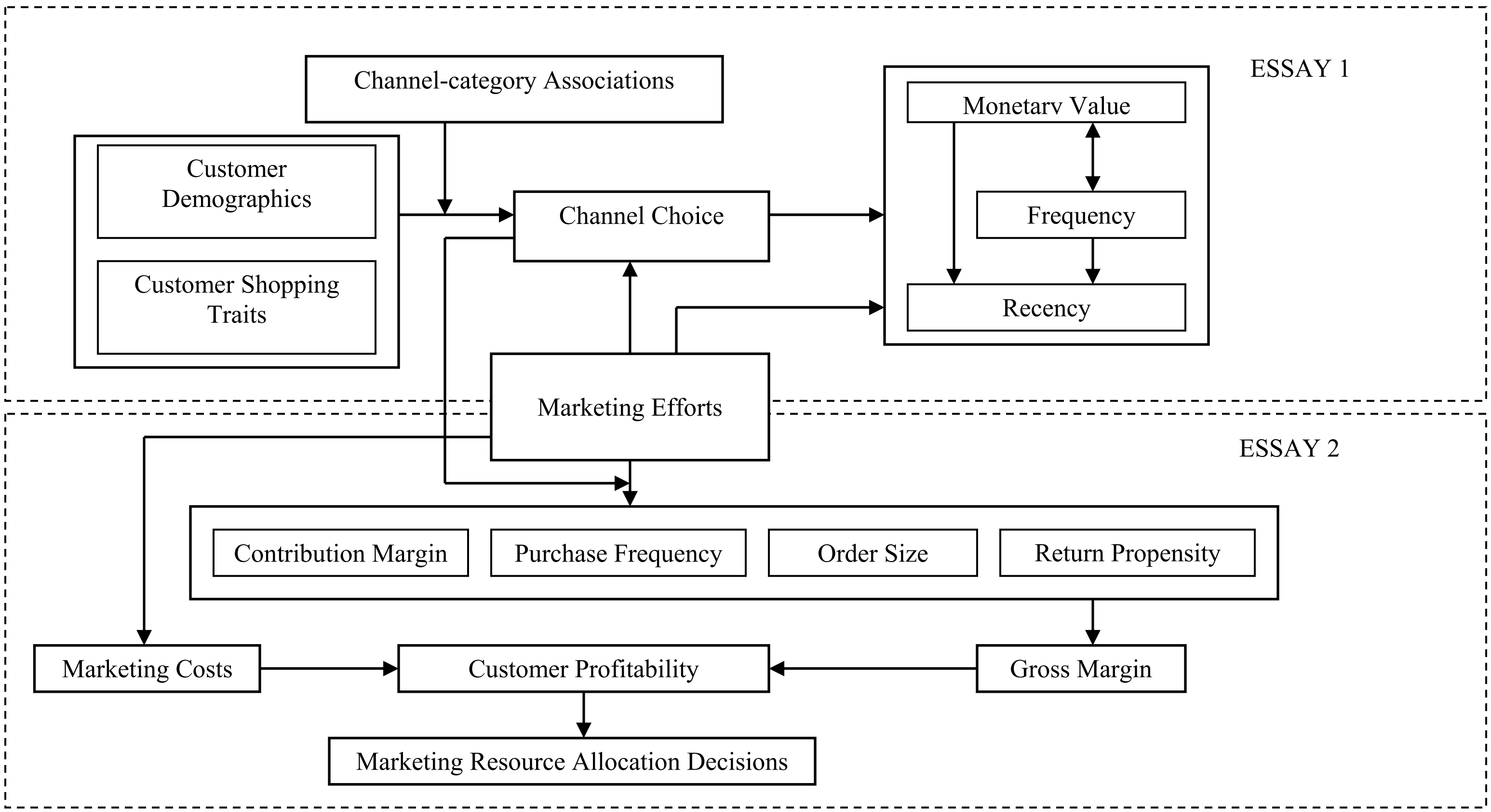
A few firms are noticing the dramatic impact of multichannel shopping behavior on their revenues. A study by J.C. Penney found that its customers who use all three channels (store, catalog and website) spent $887 per year compared to $150, $195 and $201 spent by customers who only use website, store or catalog, respectively (*Wall Street Journal* 2004). Another study by McKinsey & Company found that on average, retail customers using multiple channels spent about 20 to 30 percent more than those spent by customers using a single channel (Myers, Pickersgill, and Van Metre 2004).

Amid this emerging shopping behavior of consumers, retailers need to formulate a coordinated multichannel strategy to maximize their profits and make critical decisions on the allocation of resources across channels. However, there is a lack of adequate research that can provide a sound understanding of the differential characteristics of multichannel shoppers over single channel shoppers and of the implications of multichannel shopping on firm profitability.

In this dissertation, I develop an integrated framework of multichannel shopping to help managers identify and target profitable customers and make appropriate channel-specific resource allocation decisions based on the current and expected future profitability of its customers. Specifically, I address the following research questions in this dissertation through two essays. Which types of customers engage in multichannel shopping behavior? What drives customer channel choice? What are the effects of channel choice on the recency, frequency, and monetary value (RFM) of purchases? What are the implications of customer channel choice behavior for *channel equity* (the value of a channel to the firm)? How marketing resources should be optimally allocated across multiple channels? The overview of the integrated framework developed in the dissertation appears in Figure 1.1.

Despite the extensive research on channel management, these important research questions on multichannel shopping remain largely unexplored (Neslin et al. 2006). Rangaswamy and Van Bruggen (2005) note, "…..few academic studies have been devoted to systematically investigating the drivers and consequences of multichannel marketing."

Figure 1.1
Multichannel Marketing: An Integrated Framework



Previous work has looked at resource allocation decisions by identifying low and high value customer segments and by determining optimal allocation of marketing resources to those segments (Venkatesan and Kumar 2004). Some research has also examined allocation between customer acquisition and retention (Reinartz, Thomas, and Kumar 2005), and between retention and reacquisition of lost customers (Thomas, Blattberg, and Fox 2004). However, the problem of resource allocation has not been addressed at the customer-channel segment level. The primary advantage of developing future profitability based resource allocation metric at the customer-channel segment level is the ability to use channel as a segmenting tool and use marketing instruments in different channels with differing intensities. Previous work by Libai, Narayandas, and Humby (2002) has shown that a segment based approach to resource allocation can bring significant improvements in firm profitability. The topic's significance to mangers is also evident from its presence in the top-tier research priorities of *Marketing Science Institute* for the Marketing Productivity Community under the title "Using ROI to Allocate Resources across Functions, Marketing Vehicles, Geographies, and Over Time" (Marketing Science Institute 2006).

In the second essay, I seek to fill this gap by building on the work done in the first essay and by developing models to identify the responsiveness of different customer-channel segments to different marketing instruments. I develop a resource allocation model to optimize marketing resource allocation across customer-channel segments for firm profitability. I address two important research questions from a managerial decision making perspective. Do multichannel customers respond differently from single channel customers to a firm's marketing efforts? How should a firm allocate resources across its different customer-channel segments to maximize its total profits? I develop models that efficiently estimate and predict purchase frequency, purchase quantity, product return propensity, and contribution margin. These models capture the effects of marketing variables on customer behavior. The preliminary findings show that these models outperform pure stochastic models used in prior research. They extend prior research by differentiating the effects of order size and up selling and by accounting for product returns.

CHAPTER II

SINGLE CHANNEL VERSUS MULTICHANNEL CUSTOMERS: DETERMINANTS AND VALUE TO RETAILERS

What factors determine whether a customer purchases through only the online channel or only the offline channel or multiple channels? How do multichannel customers differ from single channel customers with regard to the recency, frequency, and monetary value (RFM) of purchases? What are the implications of shopping behaviors of single channel and multichannel customers for managing *channel equity* (the financial value to the firm of customers in a channel)? I address these important questions in the retailing context. I develop a conceptual model comprising factors that drive customer channel choice and the consequences of channel choice. I first develop a multinomial probit model of channel choice that captures the effects of customer demographics, customer shopping traits, marketing efforts, and product category. I then formulate a simultaneous system of Poisson, linear regression and negative binomial distribution (NBD) models for the RFM, respectively, of customer purchases as a function of channel choice. To provide empirically generalizable results, I estimate the models using a unique large scale dataset on the channel choice and purchase behavior of about one million customers, randomly selected from a cooperative database of 96 million US customers of 750 retailers, selling 24 product categories over a four year period. My models control for the interdependence and endogeneity of key variables. I estimate the multinomial probit model on the large scale data, using the marginal data augmentation algorithm based hierarchical Bayesian estimation approach. Some of the results are novel. They show that customer channel choice is driven in a nonlinear fashion by a demographic variable such as age and is also influenced by customer shopping traits such as number of categories bought and the length of shopping experience. The findings also show that the number of marketing mailers, not the time between such mailers, significantly influence the likelihood of multichannel shopping. Importantly, the results show that multichannel shoppers buy more often, buy more items, and spend considerably more than do single channel shoppers. They suggest that the equity of an average multichannel customer is nearly two (five and one half) times that of an offline (online) only customer.

INTRODUCTION

> *Some retailers are saying if our loyal customers like the store that much, how much more money would they spend if we can reach them X many more times through different channels?*
>
> \- Kate Delhagen, Research Director, Forrester Research

This quote captures the essence of the complex decisions that managers have to make with respect to multichannel shopping and marketing--a topic that has generated tremendous interest among manufacturers, service providers, and retailers. A study by the Direct Marketing Association (DMA) found that 40% of retailers sold through three or more channels, while another 42% sell through two channels (*The DMA* 2005). Practitioners and researchers, therefore, need a deeper understanding of this topic of emerging interest.

Multichannel marketing is the practice of simultaneously offering customers information, product, services, and support through two or more synchronized channels (Rangaswamy and Van Bruggen 2005). Organizations deploy marketing resources through multiple channels such as email, direct mail, the Web, call center, direct sales force, and physical store (Levy and Weitz 2006; Neslin et al. 2006). In addition to offering customers more channels to shop from, multichannel marketing provides organizations with greater opportunities to interact with customers, promote other channels, use price differentiation tools, segment customers specific to a channel, and target product categories to specific customer segments. Thus, multichannel marketing capability allows organizations to build stronger relationships with the right customers (Bolton and Tarasi 2006). Therefore, a better understanding of multichannel purchase behavior, its drivers and consequences is important for marketing managers to formulate sound channel strategies.

Multichannel marketing is particularly important to retailers as suggested by business press reports. J.C. Penney found that its customers who use all three channels (store, catalog and Web site) spent an average of $887 per year compared to averages of $150, $195 and $201 spent by customers who use only Web site, store and catalog, respectively (*Wall Street Journal* 2004). Another study by McKinsey & Company found that on average, retail customers who use multiple channels spent about 20 to 30 percent more than those by customers who use a single channel (Myers, Pickersgill, and Van Metre 2004).

Retailers need to formulate a coordinated multichannel strategy to maximize their profits and make critical decisions on targeting customers through the appropriate channels. However, there is inadequate research that can provide a solid understanding of the differential characteristics of multichannel shoppers over single channel shoppers and of the implications of multichannel shopping and their value to retailers (Neslin et al. 2006).

In this paper, I address the following important research questions. What demographic and behavioral factors determine whether a customer chooses only the online channel or only the offline channel or multiple channels?[2] Do the effects of the drivers of customer channel choice vary across product categories? What are the effects of channel choice on the recency, frequency, and monetary value (RFM) of purchases? What are the implications of the shopping behaviors of single channel and multichannel customers for managing *channel equity* (the financial value to the firm of customers in a channel)?

Despite the extensive research on channel management, these important research questions on multichannel shopping remain largely under explored (Neslin et al. 2006). In fact, Rangaswamy and Van Bruggen (2005) note, "…..few academic studies have been devoted to systematically investigating the drivers and consequences of multichannel marketing." In this paper, I seek to fill this gap and extend prior research in important ways. First, while we have some understanding of channel choice (e.g., Kumar

---

[2] I use the terms, determinant and driver to denote a *correlate* or a factor *associated* with channel choice or the consequences of channel choice such recency, frequency and monetary value of purchases. I do not imply a causal effect of such a factor on the dependent variable.

and Venkatesan 2005; Montoya-Weiss, Voss, and Grewal 2003) and channel migration behavior (e.g., Ansari, Mela, and Neslin 2007; Gensler, Dekimpe, and Skeira 2004; Knox 2006; Thomas and Sullivan 2005; Venkatesan, Kumar, and Ravishanker 2007; Verhoef, Neslin, and Vroomen 2007), we need a deeper understanding of a more comprehensive set of customer-level drivers of channel choice. To this end, in exploring the drivers of channel choice, I consider factors such as customer demographics and shopping traits, marketing efforts, and product categories in addition to their inherent preferences for certain channels. Furthermore, by accounting for the moderating role of product categories, I am able to more appropriately identify the main effects of customer variables on channel choice. Second, while customers' recency, frequency, and monetary value of purchases are interdependent, most studies in multichannel marketing treat them as independent and end up with debatable findings. I account for their interdependence and endogeneity in my models. Third, customer channel choice may be different for different product categories and different firms, so it can be best studied by examining a wide cross section of product categories and firms. Much prior research defines customers as online or offline or multichannel customers based on data from a single category or a single firm mainly due to data constraints although Du, Kamakura, and Mela (2007) note the need for augmenting a firm's customer transaction data with those from competing firms. I study single and multichannel purchase behavior across several firms and product categories. Fourth, to offer empirically generalizable insights, I analyze one of the largest and most representative samples of US customers -- a random sample of about one million customers selected from a cooperative database of

96 million customers of 750 retailers selling 24 product categories and hundreds of sub-categories over a four-year period. Finally, unlike prior studies, I develop a measure of customer-channel equity, offering managerial insights into the relative values of the customers shopping in different channels.

RELATED RESEARCH AND CONCEPTUAL DEVELOPMENT

Research on multichannel shopping has concentrated on a few specific aspects of multichannel shopping. Table 2.1 presents a review of some selected studies in multichannel marketing organized by sub-topics. Balasubramanian, Raghunathan, and Mahajan (2005) present a conceptual model of product and process utilities that determine channel choice for each shopping task. Kumar and Venkatesan (2005) model the impact of customer, firm and supplier characteristics on customer purchase behavior in a business-to-business (B2B) multichannel setting. They find that multichannel shoppers provide higher revenue and share of wallet than do single channel shoppers. Montoya-Weiss, Voss, and Grewal (2003) study the effects of channel and customer characteristics on service quality, satisfaction and channel usage. Their results show that overall satisfaction is driven by service quality delivered through all the channels. Verhoef, Neslin, and Vroomen (2007) use survey responses from consumers in six product categories to model the impact of search and purchase attributes, attractiveness, and crossover effects between channels on intended search and purchase behavior. They find significant within- and cross-channel effects within and across shopping tasks.

Some research has focused on the drivers and consequences of channel switching and channel utilization. Marketing efforts and customers' channel experience play significant roles in determining customers' channel selection (Ansari, Mela, and Neslin 2007), while the intensity of customers' channel usage affect their channel switching behaviors (Gensler, Dekimpe, and Skiera 2004). Additional research provides some insights into the roles of marketing efforts (Dholakia, Zhao, and Dholakia 2005; Knox 2006; Thomas and Sullivan 2005) and product characteristics (Thomas and Sullivan 2005) on multichannel shopping behavior. Venkatesan, Kumar, and Ravishanker (2007) study the timing of adoption of a new channel by a customer. The results from these studies suggest that customers' channel choice and channel switching behaviors are fairly complex and depend on factors such as product category and customer's shopping traits. The results also point out that customers can be segmented based on their channel usage and response behavior.

Other research in this area focuses on the role of channel category associations and the impact of acquisition channel on consumer behavior. A study by Inman, Shankar, and Ferraro (2004) found that shopper geodemographics influence the share of volume of purchases in a channel and that channel category associations moderate this relationship. The channel through which customers are acquired has a significant impact on their future purchase and channel switching behaviors (Verhoef and Donkers 2005; Villanueva, Yoo, and Hanssens 2006).

Table 2.1
Review of Selected Studies on Multichannel Marketing

Research Focus	Representative Studies	Drivers	Consequences	Research Context and Data	Methodology	Key Findings
Search and purchase behavior in a multichannel setting	Balasubramanian Raghunathan, and Mahajan (2005)	Utility from instrumental and non instrumental elements of process, Utility from product	Channel choice for a specific stage of purchase	Conceptual framework	NA	NA
	Kumar and Venkatesan (2005)	Customer, Supplier, and Firm characteristics	Revenues, Share of wallet, Probability of staying active	Sales data from a B2B computer hardware and software firm	Ordered logistic regression and MANOVA	Multichannel shoppers provide higher revenue, share of wallet, and are more active.
	Montoya-Weiss, Voss, and Grewal (2003)	Web site design assessment, Internet expertise	Service quality, Satisfaction, Channel use	Survey responses from customers in financial services industry and from a university	Structural equation model	Overall satisfaction with firm is determined by service quality provided through both channels.
	Verhoef, Neslin, and Vroomen (2007)	Search and purchase attributes and attractiveness, Between-channel cross over effect	Search and purchase channel choice	Survey response from panel of consumers in context of loans, holidays, books, computers, clothing, appliances	3SLS regression for search and purchase equations. Multivariate probit for choice	There are both within- and between-channel cross over effects across different shopping tasks.
Customer channel migration or choice in a multichannel setting	Gensler, Dekimpe, and Skiera (2004)	Tenure with firm, Channel usage, Product category	Channel switching behavior	Consumer semi durable from home shopping network	Colombo and Morrison brand switching model	Intrinsic preference for channel dominates channel choice. The effect is more pronounced for heavy users.
	Dholakia, Zhao, and Dholakia (2005)	Experience with firm, Similarity between channels	Choice of channel, Transaction size, Dollar value, Merchandise returns	Purchase history of customers of a retailer with store, catalog and Internet presence	Frequency cross tabulation and one way ANOVA	The use of multiple channels is greatest among customers who started on the Internet.

Table 2.1 (Continued)

Research Focus	Representative Studies	Drivers	Consequences	Research Context and Data	Methodology	Key Findings
Customer channel migration/choice in a multichannel setting	Thomas and Sullivan (2005)	Product characteristics, Marketing communication, Intrinsic channel choice	Segmenting customers on their channel preferences for forming marketing communication strategies	Consumer goods retailer with multichannel presence	Multinomial logit model with heterogeneity and Markov chain to model predicted choices	Two distinct segments are: Catalog and Internet loyal, and bricks and mortar loyal.
	Knox (2006)	Firm's direct marketing efforts	Purchase decision, Channel choice decision, Quantity decision	US retailer selling through phone, fax, Internet and mail.	Partially observed Markov decision process to jointly model purchase decision, channel decision and purchase quantity	Customer behavior and response to marketing varies dramatically across segments.
	Ansari, Mela, and Neslin (2007)	Experience with channel, Communication	Channel selection, Amount purchased, Purchased volume	Consumer durable retailer with catalog and Internet presence	Tobit model of purchase volume, and Probit model of channel choice	Marketing and experience play a significant role in channel selection.
	Venkatesan, Kumar, and Ravishanker (2007)	Customer-firm interactions such as product attributes, purchase frequency, marketing frequency	Timing/Duration of next (second and third) channel adoption	Apparel data from discount and full price brick and mortar store and online store	Shared frailty hazard model	Purchase frequency, marketing communication frequency have the greatest influence on next channel adoption duration.
Drivers of channel choice, its consequences on RFM of purchases, implications for channel equity	My research	Customer characteristics, shopping traits, Product category, Marketing efforts	Recency, Frequency and Monetary value	Syndicated data covering several categories and firms with catalog and Internet presence over multiple years	Multinomial probit model of channel choice and simultaneous equation system for RFM, comprising Poisson, NBD, and linear regression models	Customer demographics, shopping traits, marketing efforts, and product category play important roles in channel selection, which in turn, affects RFM. Channel equity of multichannel customers is nearly two (five and one half) times that of offline (online) only customers.

I develop a conceptual model of the drivers and consequences of channel choice as shown in Figure 2.1. Although a customer's channel choice is determined by a host of factors, based on the literatures on brand choice, online shopping, and channel category association and on this empirical context, I focus on potential drivers such as customer demographics, customer shopping traits, marketing efforts, and product categories. Table 2.2 shows the expected directions of relationships between these variables and the likelihood of multichannel shopping and summarizes the rationale and supporting arguments for the hypotheses. The consequences of channel choice can be viewed in terms of the RFM of a customer. The expected directions of the relationships between channel choice and RFM appear in Table 2.3.

Drivers of Channel Choice

Customer Demographics

A customer's demographics may significantly influence his or her channel choice behavior. Different socio-economic classes may have different predispositions to buy different product categories from different types of channels and customer demographics play an important role in determining the choice and share of volume of a channel (Inman, Shankar, and Ferraro 2004). Age, income, family size, and education are key demographic variables influencing channel choice.[3]

---

[3] I do not include gender in the analysis because there are no strong theoretical reasons to expect differences in channel choice due to gender and because a subsequent empirical analysis involving gender showed that gender has an insignificant effect on channel choice, consistent with a similar finding that there are no significant differences between men and women with regard to online shopping (*Jupiter Research* 2006).

Figure 2.1
Conceptual Model of Drivers and Consequences of Channel Choice

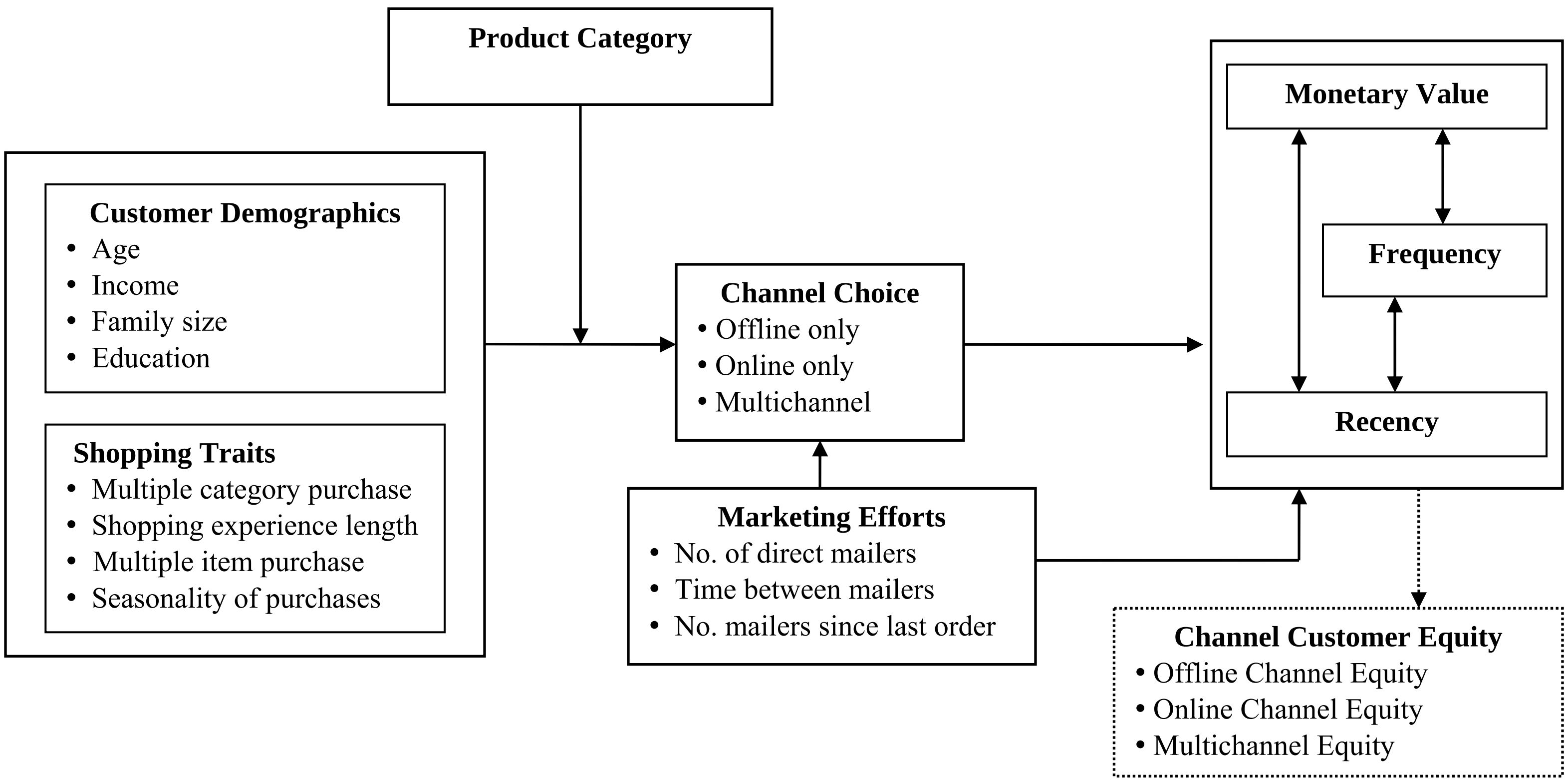


Table 2.2
Summary of Predicted Effects of Covariates on the Likelihood of Multichannel Shopping

Independent Variable	Expected Relationship		Rationale
	Offline Only*	Online Only*	
Age	Positive	Inverted 'U'	Older customers have lower Internet usage and lower propensity to shop online and are more likely to be offline only shoppers. At the same time, younger customers have lower disposable incomes that adversely influence their online shopping incidences.
Income	Negative	ND	Internet usage and propensity to shop online increases with increase in disposable income. Thus, affluent customers are more likely to be multichannel customers (using both online and offline channels) than they would be offline only customers.
Family Size	Negative	Positive	Larger families, due to time pressure, are likely to value convenience highly and may select the online channel that generally offers greater convenience than the offline channel.
Education	Negative	ND	Internet usage and propensity to shop online increase with number of years of education. Thus, multichannel and online only customers are likely to be more educated than offline only customers.
Multiple Category Purchase	Negative	Negative	The preferred channel of purchase depends on the product category. The greater the number of product categories bought, the higher the likelihood of using multiple channels.
Shopping experience	Negative	Negative	Customers with longer shopping experience are more familiar with selling practices and channels, have greater trust and lower perceived risk in making purchases across multiple channels.
Multiple Item Purchase	Negative	Negative	All SKUs are unlikely to be available in any one channel. Firms spread their portfolio of SKUs across multiple channels to reduce inventory carrying costs. Thus, customers buying more items are more likely to be multichannel customers.
Seasonality (Holiday season) of Purchase	Negative	Positive	During the busy holiday season, timely product availability and utility of time are among customers' top priorities. Since the online channel is fast and convenient for purchase, it is generally preferred to the offline channel during the holiday season.

Notes: * Likelihood of channel choice with multichannel being the base channel. A negative estimate indicates likelihood in favor of multichannel customers, while a positive estimate indicates vice-versa. ND–Not Different.

Table 2.3
Summary of Predicted Effects of Channel Choice on RFM

Dependent Variable	Expected Relationship	Rationale
Frequency	Multichannel > Single Channel	An extra channel provides firms an additional opportunity to interact with customers. This enables a firm to cross-promote through other channels, sell more frequently, upgrade customers, cross-sell, and accelerate customers' purchase cycles.
Monetary Value	Multichannel > Single Channel	
Recency	Multichannel > Single Channel	

Age is a likely determinant of channel choice. There is a significant relationship between the age of the US population and the likelihood of Internet usage (Bart et al. 2005). I expect an inverted U-shaped relationship between the age of the shopper and the likelihood of a shopper being an online only shopper relative to being a multichannel shopper. Internet-savy increases with age until late forties, beyond which it is likely to be lower (*Jupiter Research* 2006), so I anticipate that middle-aged customers to be more likely an online (relative to multichannel) shopper than those at lower and higher age groups. I also expect age to be positively related to the likelihood of a shopper being an offline only shopper relative to being a multichannel shopper. The older the shopper, the more likely he/she shops only in an offline channel than in multiple channels.

Income may play an important role in channel choice, particularly in the choice of the online channel. In 2004, Internet usage was only 31% for households with annual income of less than $15,000, but was as much as 83% for households making $75,000 or more a year (*National Telecommunication and Information Administration* 2004). Thus, higher income individuals are more likely to be online or multichannel shoppers than they are offline shoppers.

Family size is also likely to affect channel choice. Customers with larger families are more pressed for time and effort, so they are likely to prefer the online channel over the offline channel. In contrast, customers with small to moderate size families have more time to search and shop. Such customers will likely be offline only or multichannel shoppers, depending on their time constraints.

Education will also likely drive channel choice. The likelihood of Internet usage increases with education as only 15.5% of the US population with an education of high school or less use the Internet, compared to 88% of those with undergraduate degree or higher (*National Telecommunication and Information Administration* 2004). If we assume that the higher the likelihood of Internet usage, the higher the likelihood of online buying, we can expect that more educated and affluent users are online only or multichannel shoppers than they are offline shoppers.

Customer Shopping Traits

In addition to customer demographics, customer shopping traits such as the number of items and categories of purchase, the length of shopping experience, and seasonality of purchases will likely influence a customer's channel choice. Kumar and Venkatesan (2005) consider similar determinants of channel choice, calling them "customer characteristics" as they examine a B2B context. I label them as customer shopping traits because I am examining a consumer retailing context.

A customer's choice of a channel will likely depend on the number categories and items that she is able to purchase. The store choice literature suggests that assortment ranks right behind location and price as the major drivers of customer store choice (Hoch, Bradlow, and Wansink 1999). Store choice is also driven by variety seeking behavior (Popkowski and Timmermans 1997). However, in the context of channel choice, variety seeking could be reflected by the purchase of different assortment of product categories and SKUs available in different channels. Customers who seek to buy a variety of products or SKUs may have to use multiple channels to fulfill their requirements (Kumar and Venkatesan 2005). Because of these reasons, I expect customers with wider product assortment needs to be more likely multichannel customers than they would be single channel customers.

The length of a customer's shopping experience will also likely decide the channel he/she would use to make purchases. Typically, customers with longer shopping experience are more knowledgeable about selling practices and channels and are likely to have a greater shopping involvement. Customers who have long tentures with a firm buy from multiple channels (Kumar and Venkatesan 2005). These customers are more likely to buy from multiple channels than those with shorter experiences. Therefore, I expect that customers with longer shopping experience to be more likely multichannel customers than they would be single channel customers.

The seasonality of a customer's purchases has an important bearing on that customer's channel selection. During the holiday season, customers value their time and effort significantly more than they do during other times of the year. This period is generally marked by crowded stores, long check out lines, busy call centers, and uncertain availability of desired products. The online channel allows a customer to draw from a wider selection and avoid lines and stockouts associated with the offline channel. Thus, I expect the likelihood of shopping online shopping (relative to multichannel and offline shopping) to be greater for customers who make most of their annual purchases during the holiday season.

Marketing Efforts

The effects of marketing efforts on channel choice are directly relevant to marketers. In the interactive marketing context, two marketing variables of interest to marketers, the number of marketing mailers sent to a customer and the time duration between two successive mailers.[4] There is a positive synergy toward multichannel shopping when customers are contacted more through marketing communications (Kumar and Venkatesan 2005). Therefore, the number of marketing mailers may be positively related to the likelihood of multichannel shopping. Marketing communications are also positively associated with customer utility (Ansari, Mela, and Neslin 2007), which typically is higher when products are available from multiple channels. Number of communications is positively associated with the likelihood of the choice of an offline channel over the online channel (Thomas and Sullivan 2005). Based on these arguments, I expect that with greater number of marketing mailers, the

[4] Price could be a determinant of channel choice and its consequences such as recency. However, at a broad level of analysis cross multiple product categories over a long time period, it is not reasonably straightforward to assign a price level to a channel. Even so, in subsequent empirical analyses, I tried two alternative proxies of channel price level per customer, namely, the average price per order and the average price per item. None of these measures were significant in any of the models, so I do not discuss it.

likelihood of a customer being a multichannel shopper is the highest, followed by that of being an offline only shopper, and that of being an online only shopper.

Inter-mailer time's effect on channel choice will likely be opposite to that of the number of marketing mailers. The longer the time between two marketing communication pieces, the less likely the customer is a multichannel shopper. Marketing mailers highlight products to buy for the customer and also serve as a reminder to shop by thinking about multiple channels or contact points (Venkatesan, Kumar, and Ravishanker 2007). As the time between mailers increases, the less a customer is exposed to products and the less is her retrieval of various shopping alternatives. Therefore, I expect the likelihood of being an online only or an offline shopper to higher than that of a multichannel shopper as the inter-mailer time increases.

Product Categories

I expect the effects of demographic variables and shopping traits on channel choice to be moderated by the product categories purchased. When buying a product, a customer will likely choose a channel based on the perceived associations of that product category with different channels. Some products are strongly associated with certain channels, prompting people to buy them more from those channels than other channels (Inman, Shankar, and Ferraro 2004). The choice of offline vs. online channel is significantly influenced by product categories (Thomas and Sullivan 2005).

I expect the drivers of customer channel choice to vary across different product categories. Consider for example, the moderating effect of technology products such as computers and electronics on the effect of income on channel choice. These technology products draw higher income shoppers and may be more associated with the online channel than they are with the other channels, so the odds of being an online only shopper relative to being an offline only or multichannel shopper of such products will likely be greater at higher income levels than at lower income levels. Such moderating effects of product category are empirical issues and can be best determined by including them in the channel choice model and estimating the model on choice data.

Consequences of Channel Choice

A customer's choice of a channel will likely result in a specific pattern of recency, frequency, and monetary value (RFM) of purchases for that customer. I expect the three outcomes, RFM, to be interrelated. In addition, frequency will likely be determined by channel choice, the number of categories and items purchased and by the marketing efforts. Monetary value will likely depend on channel choice, frequency, the number of categories and items purchased, and marketing efforts. Recency will also likely be driven by channel choice, frequency, monetary value, and marketing efforts since the last order.

The RFM may be significantly different across online only, offline only, and multichannel customers. An extra channel provides firms with an additional opportunity to interact, build, and strengthen relationships with customers. More frequent communications from a firm to its customers helps it to improve the relationship commitment of its customers to the firm (Morgan and Hunt 1994). This improved frequency of interaction and higher commitment to relationship by customers permit firms to cross promote other channels, sell more frequently, upgrade customers, cross-sell, and accelerate customers' purchase cycles. Customers with higher recency, frequency and monetary value have higher customer life time values (CLVs) (Fader, Hardie, and Lee 2005b). In non-contractual settings, customers with higher CLVs are more profitable than those with lower CLVs (Reinartz and Kumar 2000). Thus, I expect that customers who interact through multiple marketing channels will likely have greater RFM and be more valuable to the firm than those who use only one channel.

DATA

To test my conceptual model and predictions, I examine an empirical context, comprising a carefully compiled unique and very large cross-sectional database obtained from about one million US customers, who were randomly selected from a cooperative database of 96 million customers of 750 retail firms covering 24 product categories and several sub-categories during a four year period (2001-2004). The data were provided by *i-Behavior*, a syndicated data aggregator firm. Firms in the cooperative database have only Internet and catalog channels, but no physical stores, so catalog is their offline channel. The cooperative dataset has information on the demographic characteristics of customers, their shopping traits, the channel used for purchases, order summary, and product categories summarized over a four year period. This time period adequately captures the initial phase of growth of the Internet as a distribution channel.

A advantage of this dataset is that it is not firm and industry specific and that it captures customer's purchase behavior across a comprehensive set of product categories and competing retail firms. Datasets used by prior research primarily have data from a single firm across one or a few product categories. Thus, research based on prior data would classify a customer as single channel customer even if that customer may be using different channels when transacting with different firms. For example, a customer may be a catalog only customer for J.C. Penney, but an online only customer for Nordstrom. Similarly, another customer may prefer to buy electronic items only through the online channel, but purchase holiday gifts only through the catalog channel. Because this database covers a wide range of product categories from apparel to accessories, gifts to hobby items, and electronics to music for 750 multichannel retailers, I can develop a richer understanding of customer multichannel purchase behavior. The operationalization of the variables in this data appears in Table 2.4.

Table 2.5 provides summary statistics for the key variables in my model. Of the usable sample (those with data on every variable in the database), 71.8% purchased only through the catalog channel, 5.3% purchased only through the Internet, and the remaining 22.9% purchased through both the channels. Although purchases of online only shoppers are much smaller than those of offline only and multichannel shoppers, online retail sales is growing at a high compounded annual growth rate of 12% through 2010 (*Jupiter Research* 2006). The summary snapshot suggests that multichannel customers spend more than one and one half times as much as offline only customers and about five and one half times as much as online only customers. Similarly, multichannel customers buy more often (higher frequency) than do single channel customers. The probability that multichannel customers would have bought more recently also appears to be greater than that for single channel customers.

Table 2.4
Operationalization of Variables

Variable	Operationalization
Channel Choice	Based on a four-year purchase history, customers are classified as online (Internet) only, offline (Catalog) only, and multichannel (both online and offline)
Consequences	
Frequency	Number of orders by the customer in the four-year window
Monetary Value	Total dollars spent by the customer in the four-year window
Recency	Number of weeks elapsed between the last order date and the end of the data period
Demographics	
Age	The mid point of the age range to which the customer belongs (7 intervals). For the last age range, which is open- ended (75 years +), the lower bound of the range is taken as the measure.
Income	The mid point of the household income range to which the customer belongs. For the first and last intervals, which were open ended (i.e., less than \$15K, more than \$150K), the customer's household income was assumed to be the upper and lower limits, respectively.
Family Size	Number of adults and children in the customer's household
Education	Number of years of education of the customer
Shopping Traits	
No. of Categories	Number of different product categories the customer bought in the four-year window
Shopping Experience	Number of weeks since the customer placed the first order
No. of Items	Number of items (SKUs) the customer bought in the four-year window
Seasonality of Purchases	Number of purchases made in November and December as a percentage of total purchase in the calendar year
Lifetime Highest Purchase	Dollar value of the highest value of purchase made during the customer's life
No. of High-end Orders	Number of orders of high-end items by the customer during the four-year window
Last Order Amount	Dollar value of the customer's last order before the end of the data period
Marketing Efforts	
No. of Direct Mailers	Number of catalogs mailed to the customer in the last 24 months
Time Between Mailers	Average months elapsed between two mailings to the customer
No. of Mailers since Last Order	Number of times catalogs were mailed to the customer since the customer's last order

Table 2.5
Summary Statistics of Key Variables in the Data

	Offline	Online	Multichannel
Sample Size (n)	296,550	21,810	94,719
Channel Choice (%)	71.79	5.28	22.93
Frequency	8.73	1.95	12.01
Monetary Value ($)	640.95	206.81	1,144.33
Recency (weeks)	25.55	30.90	24.33
Age (years)	57.22	45.99	48.99
Income ($ per annum)	71,950.50	86,686.57	89,331.50
Family Size	2.42	2.63	2.67
Education (years)	13.31	13.96	13.78
No. of Categories Bought	3.54	1.49	4.57
Shopping Experience (weeks)	157.44	79.77	166.95
No. of Items Bought	25.09	6.26	40.27
Seasonality of Purchases (%)	24.18	28.31	25.41
No. of Mailers	9.79	1.61	11.75
Inter- Mailer Time	9.49	17.64	8.66
No. of Mailers since Last Order	3.97	1.32	4.38

The summary statistics also suggest that multichannel customers appear to have different demographics and shopping traits than do single channel customers. Multichannel customers, on average, buy more items, buy more product categories, have longer shopping experience, are more affluent and more educated than do/are single channel customers. Online only customers are the youngest among the three channel customer groups, and are also more educated than offline only customers. Although online only customers spend the least and have the least number of orders, the average size of their orders is the largest among the three channel customer groups. These summary statistics suggest that multichannel customers may be more valuable than single channel customers and may have a different profile than those of single channel customers. The purpose of this research is to rigorously investigate this issue through my models.

In Table 2.6, I provide a summary of the purchases of the product categories across the three channels to get a view of the association of product categories with the channels. I classify the 24 product categories in the data set into five broader product categories based on their characteristics, consistent with Bart et al. (2005). While most of the product categories are associated with the offline channel, music and general merchandise have the highest share of purchases in the offline only channel. In the online only channel, among the product categories, photography and video has the highest share. Among the product categories bought by multichannel customers, photography and video, antiques and electronics have the highest share.

Table 2.7 presents the correlation matrix among the key variables in this data. An analysis of the correlations suggest that the variance inflation factors are not high (from 1.1 to 2.7), suggesting that multicollinearity is not an issue with this data. Although the correlation between no. of mailers and mailers since last order is high, these variables are not used as independent variables in the same model.

Table 2.6
Product Categories (% Breakout of Total Category Purchases by Channel)

	Offline (%)	Online (%)	Multichannel (%)	n
Technology				
Automotive	71.56	3.86	24.58	58,856
Computing	71.31	4.36	24.33	23,996
Electronics	59.11	5.46	35.43	49,840
Telecommunication	68.55	.49	30.96	2,175
Leisure				
Books & Magazines	71.67	.96	27.37	65,461
Music	83.03	.23	16.74	17,676
Sports	62.66	1.52	35.82	81,542
Travel	68.70	.69	30.61	30,304
Videos/DVDs	74.56	.63	24.81	23,175
Family and Personal				
Apparel	68.40	2.09	29.51	346,748
Children's Products	65.32	1.61	33.07	107,752
Health & Beauty	75.74	1.56	22.70	181,010
Home & Garden	70.45	2.18	27.37	251,347
Home Furnishings	69.13	2.48	28.39	335,802
Jewelry	70.68	1.06	28.26	89,362
Pets	65.95	.79	33.26	49,637
Hobby				
Antiques	61.42	1.23	37.35	13,686
Collectibles	72.93	2.08	24.99	122,493
Crafts/Hobbies	67.04	2.23	30.73	20,083
Photography & Video	54.29	5.07	40.64	7,544
Miscellaneous				
Food/Beverages	62.45	2.43	35.12	39,522
General Merchandise	79.43	3.42	17.15	91,399
Gifts	67.53	2.27	30.20	241,233
Other	62.25	1.44	36.31	104,705

Table 2.7
Correlation Matrix of Key Variables

	Offline only Channel Dummy	Online only Channel Dummy	Frequency	Monetary Value	Recency	Age	Income	Family Size	Education
Offline only Channel Dummy	1								
Online only Channel Dummy	-.38	1							
Frequency	-.05	-.13	1						
Monetary Value	-.12	-.10	.66	1					
Recency	.06	.05	-.22	-.19	1				
Age	.24	-.12	.14	.03	-.05	1			
Income	-.16	-.02	-.02	.12	-.02	-.16	1		
Family Size	-.07	.01	-.04	-.02	.02	-.11	.21	1	
Education	-.13	.03	-.01	.10	-.02	-.06	.55	.02	1
Categories	-.07	-.18	.58	.39	-.16	.15	.01	-.02	.01
Experience	-.10	-.26	.46	.38	-.10	.15	.06	-.01	.05
Items	-.07	-.08	.57	.53	-.16	.06	.02	-.02	.02
Seasonality	-.03	.04	-.07	-.04	-.11	-.05	.07	.03	.06
No. of Mailers	-.03	-.05	.56	.69	-.16	.07	.08	-.04	.09
IMT*	.02	.11	-.30	-.32	.12	-.09	-.06	.04	-.06
MSLO**	-.02	-.05	.39	.61	-.03	.08	.11	-.05	.13

Table 2.7 (Continued)

	Categories	Experience	Items	Seasonality	Mailers	IMT*	MSLO**
Categories	1						
Experience	.55	1					
Items	.39	.32	1				
Seasonality	-.01	-.08	-.05	1			
No. of Mailers	.45	.23	.42	-.05	1		
IMT	-.32	-.29	-.22	.07	-.03	1	
MSLO	.38	.19	.31	-.01	.80	-.38	1

Notes: * Inter Mailer Time, ** No. of Mailers since Last Order. n = 413,080

MODEL FORMULATION AND ESTIMATION

I first develop a model of channel choice with covariates comprising eight customer characteristics (demographics and shopping traits), marketing efforts, and dummy variables for broad product categories and their interactions with customer characteristics. I then develop the RFM model, by analyzing the effects of channel choice on recency, frequency and monetary value. I divide the sample into three equal parts. I use the first one-third of the sample to estimate the channel choice model. I use the second one-third of the sample to determine the predicted channel choices based on the parameter estimates obtained in the first one-third sample. I also use this sample to estimate the RFM model. I use the last one-third of the sample for predicting the RFM model.

I develop my model in two stages. In the first stage, I identify the drivers of channel choice through an independent multinomial probit model. I chose the probit model over the logit model for the following reasons. First, unlike the logit model, the probit model does not make the unrealistic assumption of independence of irrelevant alternatives (IIA). Second, the multinomial probit model assumes that the errors follow the more commonly observed normal distribution (vis-à-vis a Type 1 extreme value distribution for the multinomial logit model) and thus its estimates are asymptotically normal, consistent, and efficient (Currim 1982). Using the parameter estimates, I predict the channel choice probability for each individual.

If a customer i is faced with J alternative channels, the utility U derived by the customer for each channel j can be expressed as:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{2.1}$$

where V is the deterministic component of the utility and ε is the random component. The deterministic component of the utility is determined by the attributes of the customer and the channel. The deterministic component as a function of these attributes can be expressed as:

$$V_{ij} = X_i \beta_j = \sum_{k=o}^{K} \beta_{jk} x_{ki} \tag{2.2}$$

Where, X_i is a vector of k characteristics or covariates for customer i, β_j is the response parameter vector for j[th] channel, and K is the number of covariates in the model. The covariates include customer demographics, customer shopping traits, product category dummies, and interaction of effects of product categories and demographics and product categories and shopping traits, and customer-specific marketing efforts.

The probability of customer i choosing channel j is given by (Jones 2000):

$$\mathrm{P_{ij}} = \int_{-\infty}^{\mathrm{V_{i1}}} \dots \int_{-\infty}^{\mathrm{V_{ij-1}}} \phi(\varepsilon_{i1} \dots, \varepsilon_{ij-1}) d\varepsilon_{i1} \dots, d\varepsilon_{ij-1} \tag{2.3}$$

where ϕ is the probability density function of normal distribution. Note that the independent multinomial probit assumes that ε are independent and identically distributed with multivariate normal distribution $\varepsilon \sim N(0, \Sigma)$. Σ is the error variance covariance matrix.

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & . & \sigma_{1j} \\ . & . & . \\ \sigma_{j1} & . & \sigma_{jj}^2 \end{bmatrix} \tag{2.4}$$

In this data, customers can be online only users, offline only users or multichannel users, that is, a customer has three choice alternatives. Let U_{i1}, U_{i2}, and U_{i3} be the utilities derived by customer i by using online only, offline only, and both the channels, respectively. The first choice alternative will be chosen by customer i if $U_{i1} > U_{i2}$ and $U_{i1} > U_{i3}$. The probability that the first alternative is preferred to the second and the third alternatives can also be expressed as (Maddala 1983):

$$\Pr ob(\varepsilon_{i2} - \varepsilon_{i1} < V_{i1} - V_{i2}, \varepsilon_{i3} - \varepsilon_{i1} < V_{i1} - V_{i3}) \tag{2.5}$$

This probability is given by:

$$P_{i1} = \int_{-\infty}^{V_{i1}-V_{2i}} \int_{-\infty}^{V_{i1}-V_{i3}} \phi(\varepsilon_{i2} - \varepsilon_{i1}, \varepsilon_{i3} - \varepsilon_{i1}) d(\varepsilon_{i2} - \varepsilon_{i1}) d(\varepsilon_{i3} - \varepsilon_{i1}) \tag{2.6}$$

where $\varepsilon_{i2} - \varepsilon_{i1}$ and $\varepsilon_{i3} - \varepsilon_{i1}$ will have bivariate normal distributions because the first alternative is being evaluated against the second and third alternatives. The covariance matrix of $\varepsilon_{i2} - \varepsilon_{i1}$ and $\varepsilon_{i3} - \varepsilon_{i1}$ is given by

$$\Omega_1 = \begin{bmatrix} \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} & \sigma_1^2 - \sigma_{13} - \sigma_{12} + \sigma_{23} \\ \sigma_1^2 - \sigma_{13} - \sigma_{12} + \sigma_{23} & \sigma_1^2 + \sigma_3^2 - 2\sigma_{13} \end{bmatrix} \tag{2.7}$$

Similarly, the probability that the second and the third alternatives are preferred over the other two by customer i is given by:

$$P_{i2} = \int_{-\infty}^{V_{i2}-V_{i1}} \int_{-\infty}^{V_{i2}-V_{i3}} \phi(\varepsilon_{i1} - \varepsilon_{i2}, \varepsilon_{i3} - \varepsilon_{i2}) d(\varepsilon_{i1} - \varepsilon_{i2}) d(\varepsilon_{i3} - \varepsilon_{i2}) \tag{2.8}$$

$$P_{i3} = \int_{-\infty}^{V_{i3}-V_{i1}} \int_{-\infty}^{V_{i3}-V_{i2}} \phi(\varepsilon_{i1} - \varepsilon_{i3}, \varepsilon_{i2} - \varepsilon_{i3}) d(\varepsilon_{i1} - \varepsilon_{i3}) d(\varepsilon_{i2} - \varepsilon_{i3}) \tag{2.9}$$

where $(\varepsilon_{i1} - \varepsilon_{i2}, \varepsilon_{i3} - \varepsilon_{i2})$ and $(\varepsilon_{i1} - \varepsilon_{i3}, \varepsilon_{i2} - \varepsilon_{i3})$ each have a bivariate normal distribution with covariance matrix Ω_2 and Ω_3, respectively. Both Ω_2 and Ω_3 are identical to Ω_1.

In estimating the multinomial probit model on the massive data set, I introduce a new estimation method to marketing. The standard maximum likelihood estimation method (MLE) for a multinomial probit model on a large dataset with multiple alternatives is computationally demanding, involving numerical evaluation of multiple integrals and leading to convergence problems. Therefore, I need a feasible estimation technique that performs at least as well as the MLE method. I estimate the independent multinomial probit model, using the marginal data augmentation algorithm based hierarchical Bayesian estimation approach suggested by Imai and Dyk (2005). This approach is not only faster than the MLE method but also produces more interpretable estimates which are consistent in directionality and effect sizes across multiple estimation sequences and starting values. Imai and Dyk (2005) show that this estimation algorithm also outperforms existing hierarchical Bayesian estimation algorithms on speed, simplicity and prior specification and that it is simpler than the algorithm suggested by McCulloch and Rossi (1994), faster than the hybrid Markov chain based approach suggested by Nobile (1998), and superior in prior specification to the approach suggested by McCulloch, Polson, and Rossi (2000). The algorithm uses improper prior distribution on β and diffused prior distribution on Σ. I report the results of my model estimated with 50,000 draws of the priors and 45,000 used for the burn-in period.

In the second stage, I model the consequences of channel choice through a simultaneous equation system, comprising models for recency, frequency, and monetary value of purchase. In this model, I use the predicted channel choice from the first stage of the model. To test the predictive accuracy of my model, I compare the predicted values of purchases generated by each channel with the actual values of purchases in each channel in a holdout sample.

While modeling the consequences of multichannel shopping behavior, I encounter the endogeneity problem in two ways. First, the monetary value generated by a customer and her frequency of purchase are interdependent. The two equations, one each for frequency and monetary value cannot be estimated using the ordinary least squares (OLS) method since each dependent variable appears as an independent variable in the other two equations and are thus, endogenous to the system. Estimating these equations by OLS would lead to biased and inconsistent parameter estimates (Zellner and Theil 1962). Similarly, the recency of a customer's purchase depends on the frequency and monetary value of her purchases in the data window. Second, channel choice is likely to be an endogenous decision with respect to the consequences modeled. To account for the first endogeneity problem, I use the instrumental variable approach to estimate the first system of two simultaneous equations. Using the predicted values of frequency and monetary value in the third equation, I account for their endogeneity. To account for the second endogeneity problem, I use the predicted channel choice from the independent multinomial probit model in the simultaneous equation system. The system of simultaneous equations used to model the consequences of channel choice is given by:

$$E(ORDR_i \mid X_i) = \exp[\alpha_0 + \alpha_1 OFF_i + \alpha_2 ON_i + \alpha_3 DLR_i + \alpha_4 CAT_i + \alpha_5 ITEM_i + \alpha_6 MAIL_i + \alpha_7 IMT_i] + \psi_i \tag{2.10}$$

$$DLR_i = \beta_0 + \beta_1 OFF_i + \beta_2 ON_i + \beta_3 ORDR_i + \beta_4 CAT_i + \beta_5 ITEM_i + \beta_6 HIGH_i + \beta_7 LHPA_i + \beta_8 MAIL_i + \xi_i \tag{2.11}$$

$$P(LOT_i \mid \mu, k) = \left(1 + \frac{\mu}{k}\right)^{-k} \frac{\Gamma(k+y)}{y!\,\Gamma(k)} \left(\frac{\mu}{\mu + k}\right)^{y} \tag{2.12}$$

$$\mu_i = \exp[\gamma_0 + \gamma_1 OFF_i + \gamma_2 ON_i + \gamma_3 ORDR_i + \gamma_4 DLR_i + \gamma_5 LOAM_i + \gamma_6 MSLO_i + \zeta_i] \tag{2.13}$$

where ORDR, DLR and LOT are frequency of ordering (frequency), dollar spending (monetary value), and last order time (recency), respectively and X is the set of covariates presented in Equation (10). OFF (offline only) and ON (online only) are dummy variables for the channel chosen and multichannel is the base channel. CAT, ITEM, HIGH, LHPA, and LOAM are number of categories bought, number of items bought, number of high end purchases, life time highest purchase amount, and last order amount, respectively. MAIL, IMT and MSLO are catalogs mailed, average time between two mailers, and catalogs mailed since last order was placed, respectively. α, β, and γ are parameter vectors and Equation 12 shows a negative binomial distribution with mean μ and dispersion parameter k. ψ, ξ, and ζ are error terms.

In each of the first two equations (Equations 2.10 and 2.11), I have one left hand side endogenous variable (ORDR and DLR), two or three exogenous variables (MAIL and IMT; HIGH, LHPA, MAIL), and one right hand side endogenous variable (DLR and ORDR). Since there are two endogenous variables and two equations, to uniquely identify the system of two equations, I have the one instrument needed for each equation. At the same time, I need to account for the effects of marketing efforts on the

RFM of customer purchases. I choose the inter mailing time as the instrument for Equation (2.10) and the life time highest purchase of the customer and the number of high-end categories bought as the instruments for Equation (2.11).

In the above system, while ORDR and LOT are count variables, DLR is a continuous variable. Equation (2.10) and (2.11) form a two equation system in which a count variable and a continuous variable are simultaneously determined. The 2SLS or 3SLS estimates of a linear system of equations would be inconsistent in such a situation. I use the Generalized Method of Moment estimation for a system, comprising count and continuous dependent variables, consistent with Windmeijer and Santos Silva (1997). For a given set of exogenous instruments for the endogenous variables, the GMM estimation method would produce consistent estimates of the parameters. Finally, I estimate Equation (2.12) by accounting for the endogeneity of frequency and monetary value through the predicted values of ORDR and DLR. LOT as measured in this data is a count variable given by the weeks elapsed since the last order placed by a customer. Thus, it is appropriate to model the equation using a Poisson or a negative binomial distribution (NBD) regression, which is appropriate for modeling count data.[5]

Because I am dealing with time series data summarized over four year window and the analysis is essentially a cross-sectional analysis, identifying a good instrument poses significant challenge. Theoretically, a good instrument should be correlated with the left hand side endogenous variable, but uncorrelated with the independent variables.

[5] The choice of Poisson or NBD model is based on a test of overdispersion in the data. Poisson distribution assumes equality of mean and variance, so I first perform a Lagrange multiplier test of overdispersion (Greene 2003). If the overdispersion test is rejected, then I use a Poisson regression model. Otherwise, I use the NBD model.

The choice of weak instrumental variables can lead to poor estimates (Bound, Jaeger, and Baker 1995). I test for the quality of instruments using the approach suggested by Staiger and Stock (1997). As suggested by Staiger and Stock (1997), I test the first stage F-statistic for each equation with the instrumental variables. The bias introduced by the weak instruments is of the order of the inverse of the F- statistic. As a rule of thumb, Stock and Watson (2003) suggest that an F-statistic of greater than 10 is acceptable as it corresponds to a bias of less than 10% in the estimates.

RESULTS AND DISCUSSION

In Table 2.8, I present the results of channel choice model with multichannel as the base channel. Models I, II, and III are multinomial probit models estimated using a Markov Chain Monte Carlo Bayesian approach algorithm. In Model 1, I include only demographic variables, shopping traits, and marketing efforts in explaining channel choice. In Model II, I include the direct effect of product categories on channel choice decisions. In Model III, I also include the moderating role of product categories on the relationships between customer demographics and channel choice and between customer shopping traits and channel choice.

The results of Model II and III are interesting. In the absence of the moderating role of product categories, the effects of income, family size, number of items bought, and seasonality are significant on customer channel choice ($p < .01$). However, when the direct and interaction effects of product categories on customer channel choice are introduced into the model, many of the direct effects of the other variables are diluted,

Table 2.8
Results of Channel Choice Model

		Multinomial Probit with MCMC #					
		Model I		Model II		Model III	
Intercept	Offline	-1.719*	(-1.907, -1.537)	-1.850*	(-2.099, -1.627)	-1.783*	(-2.504, -1.081)
	Online	-1.742*	(-1.899, -1.592)	-1.912*	(-2.138, -1.711)	-3.018*	(-3.866, -2.266)
Customer Demographics							
Age	Offline	-.047*	(-.053, -.041)	-.049*	(-.058, -.041)	-.031*	(-.054, -.007)
	Online	.028*	(.024, .032)	.026*	(.021, .031)	.038*	(.012, .064)
Age Square	Offline	.001*	(.001, .001)	.001*	(.001, .001)	.001*	(.000, .001)
	Online	-.001*	(-.001, -.000)	-.000*	(-.001, -.000)	-.001*	(-.001, -.000)
Income	Offline	-1.7e-6*	(-2.1e-6, 1.4e-6)	-1.6e-6*	(-2.0e-6, -1.2e-6)	-1.1e-6	(-2.3e-6, 9.6e-8)
	Online	-1.7e-6*	(-1.9e-6, 1.4e-6)	-1.5e-6*	(-1.8e-6, -1.3e-6)	-4.2e-7	(-1.6e-6, 7.6e-7)
Family Size	Offline	-.012*	(-.021, -.003)	-.008	(-.018, .002)	-.011	(-.062, .040)
	Online	-.016*	(-.022, -.009)	-.015*	(-.021, -.010)	-.009	(-.052, .032)
Education	Offline	-.007*	(-.008, -.007)	-.008*	(-.009, -.007)	-.009*	(-.006. -.013)
	Online	.006*	(.005, .006)	.005*	(.004, .006)	.007*	(.003, .010)
Shopping Traits							
No. of Categories	Offline	-.061*	(-.067, -.053)	-.108*	(-.122, -.096)	-.119*	(-.134, -.105)
	Online	-.022*	(-.026, -.019)	-.015*	(-.020, -.010)	-.022*	(-.028, -.016)
Experience	Offline	-.005*	(-.005, -.004)	-.006*	(-.006, -.005)	-.006*	(-.007, -.005)
	Online	-.001*	(-.001, -.000)	-.001*	(-.001, -.001)	-.004*	(-.003, -.005)
No. of Items	Offline	-.000*	(-.000, -.000)	-.000*	(-.001, -.000)	-.001	(-.003, .000)
	Online	-.000*	(-.000, -.000)	-.000*	(-.001, -.000)	-.000	(-.001, .001)
Seasonality of Purchases	Offline	.035*	(.001, .073)	.028	(-.019, .069)	-.139	(-.274, .288)
	Online	.034*	(.005, -.063)	.029	(-6.3e-5, .062)	.129*	(.013, .245)
Marketing Efforts							
No. of Mailers	Offline	-.061*	(-.064, -.059)	-.081*	(-.086, -.071)	-.001*	(-.001, -.000)
	Online	-.000	(-.001, 2.2e-5)	-.000*	(-.001, -3.8e-5)	-.109*	(-.114, -.104)
Inter Mailer Time	Offline	.002	(-.001, .004)	.002*	(.001, .004)	.001	(-.001, .003)
	Online	.002	(-.001, .002)	.001*	(.000, .002)	.001	(-.001, .002)
Product Categories							
Technology	Offline			-.364*	(-.401, -.330)	-.117	(-.772, .507)
	Online			.134*	(.113, .154)	.314	(-.078, .732)
Leisure	Offline			-.032	(-.066, .007)	-.492	(-1.065, .069)
	Online			-.032*	(-.050, -.015)	-.273	(-.613, .066)
Family and Personal	Offline			.158*	(.112, .207)	.609	(-.126, 1.385)
	Online			.240*	(.201, .280)	.949*	(.199, 1.711)
Hobby	Offline			.011	(-.024, .047)	-.027	(-.519, .489)
	Online			.027*	(.009, .046)	.361*	(.020, .716)
Miscellaneous	Offline			.056*	(.017, .092)	-.309	(-1.064, .342)
	Online			-.033*	(-.050, -.015)	-.081	(-.423, .278)

Table 2.8 (Continued)

		Multinomial Probit with MCMC #			
		Model I	Model II	Model III	
Moderating Effect of Product Categories					
Income X Technology	Offline			-8.3e-7	(-1.9e-6, 2.7e-7)
	Online			8.5e-7*	(2.2e-7, 1.4e-6)
Income X Leisure	Offline			7.1e-7	(-2.9e-7, 1.8e-6)
	Online			1.6e-6*	(1.1e-6, 2.1e-6)
Family Size X Family and Personal	Offline			-.057*	(-.089, -.025)
	Online			.009*	(.003, .015)
Education X Technology	Offline			-.004*	(-.007, -.002)
	Online			.003*	(.001, .006)
Education X Family and Personal	Offline			.004*	(.001, .008)
	Online			-.002	(-.005, .002)
Experience X Technology	Offline			.001	(-.000, .001)
	Online			-.001*	(-.002, .-001)
Experience X Leisure	Offline			1.6e-5	(-.001, .001)
	Online			-.001*	(-.002, .-001)
Experience X Family and Personal	Offline			.001	(.000, .001)
	Online			-.002*	(-.003, .001)
Items X Technology	Offline			-.002*	(-.004, -.000)
	Online			-.001*	(-.001, -.000)
Items X Hobby	Offline			-.001	(-.002, .008)
	Online			-.001*	(-.001, -.001)
Items X Family and Personal	Offline			.002*	(.001, .004)
	Online			.001*	(.000, .002)
Seasonality X Leisure	Offline			.145*	(.035, .264)
	Offline			.164*	(.089, .248)
Seasonality X Miscellaneous	Offline			.298*	(.185, .416)
	Offline			.078*	(.002, .155)
Sample Size		137,693	137,693	137,693	

Notes: # Posterior Means (95% Probability Interval)
Multichannel is the base channel.
*: Zero is not in the 95% probability interval of posterior mean.
The main and moderating effects of gender, the moderating effects of product categories on the relationships between age and channel choice and between marketing efforts and channel choice are not significant and are not shown to preserve space.

suggesting strong moderating effects of the product categories. I discuss the results of Model III, the best fitting model in predictive accuracy.

Results of Channel Choice Model

Customer Demographics

The directions of the estimates of age and age squared provide interesting insights into role of age in channel selection. Offline only users show a positive and significant ($p < .01$) relationship with age relative to multichannel users, consistent with my expectations. Thus, the likelihood of a customer being an offline only shopper relative to a multichannel customer is smaller for younger people, but greater for older people. However, the relationship between the likelihood of a customer being an online only customer relative to being a multichannel customer and age is also significant ($p < .001$), but follows an inverted U-shape. The, the likelihood of younger customers being online only customers than being multichannel customers, increases with age, but that of older customers diminishes with age. The effects of income and family size on customer channel choice, however, are not significant ($p > .05$). With regard to education, the greater the number of years of education of a customer, the more (less) likely she is a multichannel customer relative to an offline only (online only) customer ($p < .01$).

Customer Shopping Traits

Customers' shopping traits exhibit a strong correlation with their channel choices. Five of the eight parameters of the four shopping trait variables are statistically significant ($p < .01$) and have moderate to strong effects on the channel choices of customers. The relationship between the number of product categories bought and the choice of channel is interesting. The estimates for number of categories bought suggest that multichannel customers buy a greater number of product categories than do online only customers ($p < .01$), and than do offline only customers ($p < .01$). The likelihoods of a customer who buys an additional category being a multichannel customer are 11% and 2% more than that for offline and online only customers, respectively. The longer the shopping experience of a customer with retailers, the greater his/her likelihood of shopping across the different channels. The effect of the number of items bought by a customer on her channel choice decision is not significant ($p > .05$). The greater a customer's proportion of purchases during the holiday shopping season, the more likely she is an online only customer than a multichannel customer ($p < .01$). However, offline only customers are not significantly different from multichannel customers in the number of purchases made during the holiday season ($p > .05$). Thus, the likelihood of a customer making significant holiday purchases being an online only customer is 13% more than that of being a multichannel customer.

Marketing Efforts

Customers who receive more marketing mailers are more likely to be multichannel customers than they are single channel customers ($p < .01$). With an additional marketing mailer, the likelihood that the customer is a multichannel customer is 11% greater than being an online customer, holding other variables constant. However, inter-mailer time does not have a significant effect on channel choice ($p > .05$). Thus, the likelihood of a customer being a single channel or multichannel shopper is the same regardless of the average gap between two successive marketing mailers.

Product Category

Like the cases of other dummy variables, the effect of product category on channel choice should be interpreted as a likelihood ratio (log odds ratio) between the two channel choices in the event of that product category being bought by a customer versus the event in which the category is not bought. The interpretation of these estimates should not be confused with the summary figures in Table 2.5. The numbers in Table 2.5 show the percentages of each product category bought by the type of customers. For example, 71.6%, 3.9%, and 24.5% of the 111,213 customers who bought automotive products were offline only, online only and multichannel users, respectively. However, the estimates of each of these product categories in Table 2.7 compare the effect of the product category when bought versus not bought on the choice of channel.

I show the main and significant moderating effects of product categories in Table 2.7. Except in the case of family and personal and hobby items, I do not find any significant direct effects of product categories on customers' channel choice. Many of the moderating effects need careful interpretation. High income customers buying technology products are more likely to be online only users relative to those who did not purchase these product categories ($p < .01$). Similarly, high income customers who bought leisure products are more likely to be online only shoppers than those who did not purchase leisure products ($p < .01$). Customers with large families, who bought family and personal products, are likely to make purchases online only than offline only or in both the channels than offline only ($p < .01$). Highly educated buyers of technology products are most likely to be online only shoppers, followed by multichannel shoppers, and by offline only shoppers ($p < .01$). However, highly educated buyers of family and personal products prefer the offline only channel to using both the channels or only the offline channel ($p < .01$). Customers with longer shopping experience and purchasing technology or leisure or family and personal products are more likely to use both the channels than those using only the online channel ($p < .001$). Customers buying more number of items and purchasing technology products will likely use both the channels, followed by the online only channel and by the offline only channel. Similarly, customer buying more number of items and purchasing hobby related products are likely to use both the channels over the online only channel ($p < .001$). However, customers buying more number of items and purchasing family and personal products will likely use the offline only channel or the online only channel over using both the channels ($p < .01$). Customers who spend a higher proportion of their budget during the holiday shopping season and purchase leisure and miscellaneous product categories will likely be online only and offline only customers than being multichannel customers ($p < .001$). The moderating effects of product categories on the relationships between age and channel choice are not significant and are not shown in the table.

Results of the RFM Model

I use the predicted channel choices from the channel choice model in estimating the RFM model. I estimate the RFM model using OLS, 2SLS, 3SLS, and GMM. The results of the RFM model are presented in Table 2.9. In the first part of Table 2.9, I provide the results of the first system of equations involving frequency and monetary value. The Lagrange multiplier test for equality of mean and variance of number of

Table 2.9
Results of RFM Model

	OLS§	2SLS§	3SLS§	GMM
Dependent Variable: Frequency				
Intercept	-2.361***	2.464***	1.799***	2.047***
	(.048)	(.060)	(.056)	(.007)
Offline only channel dummy	1.425***	-.116*	-.117*	-.036***
	(.034)	(.040)	(.040)	(.004)
Online only channel dummy	-0.387*	-2.169***	-2.427***	-.860***
	(.132)	(.154)	(.153)	(.016)
Dollar spending (Monetary Value)	.004***	.002***	.002***	6.0e-5***
	(1.5e-5)	(2.9e-5)	(2.9e-5)	(2.5e-6)
No. of categories bought	1.803***	1.864***	1.884***	.132***
	(.006)	(.007)	(.007)	(.001)
No. of items bought	.040***	.070***	.070***	.001***
	(.000)	(.000)	(.000)	(4.7e-5)
No. of mailers	.042***	.265***	.271***	.006***
	(.001)	(.002)	(.001)	(.000)
Inter-mailer time	-.016***	-.087***	-.039***	-.020***
	(.002)	(.002)	(.001)	(.000)
Dependent Variable: Monetary Value				
Intercept	-165.010***	-23.543***	-27.187***	72.172***
	(3.291)	(5.474)	(5.467)	(8.698)
Offline only channel dummy	-99.397***	-160.388***	-156.526***	-148.043***
	(2.380)	(4.422)	(4.413)	(3.978)
Online only channel dummy	-163.805***	-29.824	-32.177	-212.582***
	(10.395)	(17.114)	(17.113)	(7.934)
Number of orders (Frequency)	52.089***	143.903***	142.060**	55.261***
	(.109)	(.562)	(.545)	(3.237)
No. of categories bought	14.589***	22.666***	22.354***	64.756***
	(.512)	(1.491)	(1.472)	(5.534)
No. of items bought	3.238***	3.631***	3.525***	1.626***
	(.019)	(.050)	(.050)	(.316)
Number of high-end purchases	96.995***	14.863***	27.905***	15.291***
	(.737)	(1.303)	(.843)	(2.729)
Life time highest purchase	2.076***	2.163***	2.151***	1.872***
	(.003)	(.005)	(.005)	(.026)
No. of mailers	21.764***	19.350***	19.339***	19.320***
	(.068)	(.327)	(.327)	(.862)
Model Fit (Adjusted R^2)	59.9%, 78.8%	44.1%, 79.2%	44.1%, 79.2%	44.8%, 79.3%
Dependent Variable: Recency	NBD OLS	NBD 2SLS	NBD 3SLS	NBD GMM
Intercept	3.354***	3.302***	3.300***	3.310***
	(.002)	(.002)	(.002)	(.002)
Offline only channel dummy	.020***	.024***	.024***	.024***
	(.002)	(.002)	(.002)	(.002)
Online only channel dummy	.096***	.143***	.144***	0.140***
	(.006)	(.006)	(.006)	(.006)
Number of orders (Frequency)	-.008***	-.006***	-.005***	-.006***
	(.000)	(.000)	(.000)	(.000)
Dollar spending (Monetary Value)	-.000***	-.000***	-.000***	-.000***
	(.000)	(.000)	(.000)	(.000)
Last order amount	.000***	.000***	.000***	.000***
	(.000)	(.000)	(.000)	(.000)
Mailers since last order	.012***	.009***	.009***	.010***
	(.000)	(.000)	(.000)	(.000)
Dispersion parameter	.173	.173	.174	.171
Sample Size (n)	137,693	137,693	137,693	137,693

Notes: § Frequency model is specified as linear model with assumption that errors are normally distributed.
Multichannel is the base channel. Significance levels: * (5%), ** (1%), *** (.1%).

orders suggested that the Poisson model is appropriate for frequency ($p < .01$), so I estimated the frequency and monetary value system of Poisson and linear regression models.

In the second part of the table, I present results of the recency model estimated by OLS using predicted values of frequency and monetary value from the respective models in the first part. For the recency equation, the equality of mean and variance of time since last order was rejected ($p < .05$), so I use the NBD model. The results of the OLS estimation are dramatically different ($p < .01$) from those of the limited information estimation and the full information estimation procedures, providing significant evidence for the endogeneity of frequency and monetary value. The results of the limited information estimation and the full information estimation methods are not very different. However, because I have a system of simultaneous equations in which frequency is measured with count data and monetary value with continuous data, I discuss the results of the GMM estimation.

The statistical significance and the direction of the two channel dummy variables ($p < .001$) provide support for my hypothesis that multichannel customers buy more frequently and spend more than do single channel customers. In the frequency equation, offline only customers have .036 ($p < .001$) and online only customers have .860 ($p < .001$) fewer orders than those of multichannel customers. The effect of inter-mailer time ($p < .001$) is in the expected direction. Similarly, in the monetary value equation, multichannel customers spend about \$148 ($p < .001$) and \$212 ($p < .001$) more than the offline only and online only customers, respectively. After controlling for the effect of channel, the values of an incremental order and of an additional category, are about \$55 and \$64, respectively ($p < .001$). These are the largest effects on the monetary value, after adjusting for channel differences. The other estimates are in the expected directions.

The estimates of the two channel dummy variables in the recency equation based on the predicted values of the covariates from the previous system of equations suggest that offline only and online only customers buy .024 ($p < .001$) weeks and .140 ($p < .001$) weeks, respectively, later than do multichannel customers. The estimates of other variables are in the expected directions.

Robustness Checks

Channel Choice Model

I check the robustness of channel choice model by comparing the results of Model III from Table 2.7 with other multinomial probit models in which I change the way I define multichannel shoppers. Furthermore, I estimate multinomial probit choice model using the standard MLE method. The results of this model are presented in Table 2.10 as Model IV. In Model IV, multichannel customers are defined in a way similar to those in Model I, II and III. However, in Model V and Model VI, I define multichannel customers as those who have purchased in each channel on at least two and three occasions, respectively. This exercise reduces the sample sizes for Model V and VI to 133,041 (96%) and 113,075 (82%), respectively compared to 137,693 (100%) for Models I to III. The estimates of Model III, IV, V and VI are in the same directions with similar statistical significance and moderately varying effect sizes. These results indicate

Table 2.10
Robustness Check: Results of Alternate Channel Choice Models

		Model IV		Model V		Model VI	
		MLE#		MCMC ##			
		One Purchase in Each Channel		Two Purchases in Each Channel		Three Purchases in Each Channel	
Intercept	Offline	-4.622***	(.349)	-1.979+	(-3.090, -.896)	-1.441+	(-3.276, -.392)
	Online	-.773	(.715)	-4.123+	(-5.660, -2.656)	-2.518+	(-4.342, -.879)
Customer Demographics							
Age	Offline	-.025*	(.012)	-.033+	(-.054, -.007)	-.060+	(-.122, -.002)
	Online	.029	(.025)	.048+	(.014, .083)	.034+	(.023, .093)
Age Square	Offline	.001***	(.000)	.001+	(.000, .001)	.001+	(.000, .001)
	Online	-.000	(.000)	-.001+	(-.001, -.000)	-.001+	(-.001, -1.6e-5)
Income	Offline	-1.0e-6	(5.8e-7)	-1.1e-6	(-2.7e-6, 5.3e-7)	-4.7e-7	(-3.2e-6, 2.7e-6)
	Online	-7.3e-7	(1.1e-6)	-6.7e-7	(-3.2e-6, 1.7e-6)	-3.9e-7	(-2.7e-6, 2.4e-6)
Family Size	Offline	-.044*	(.018)	-.027	(-.091, .038)	-.012	(-.121, .085)
	Online	-.021	(.028)	-.009	(-.082, .064)	-.044	(-.123, .031)
Education	Offline	-.012***	(.001)	-.010+	(-.014, -.006)	-.004+	(-.006, -.003)
	Online	.003	(.003)	.011+	(.005, .017)	.005+	(.002, .008)
Shopping Traits							
No. of Categories	Offline	-.041***	(.002)	-.124+	(-.141, -.106)	-.115+	(-.137, -.097)
	Online	-.345***	(.018)	-.043+	(-.055, -.032)	-.039+	(-.052, .029)
Experience	Offline	-.011***	(.000)	-.005+	(-.006, -.004)	-.009+	(-.011,-.006)
	Online	-.021***	(.001)	-.003+	(-.004, -.001)	-.002+	(-.001, -.003)
No. of Items	Offline	-.001**	(.000)	-.000	(-.002, .001)	-.000	(-.002, .002)
	Online	-.005*	(.002)	-.000	(-.002, .002)	-.000	(-.001, .002)
Seasonality of Purchases	Offline	-.041	(.064)	-.009	(-.084, .261)	-.047	(-.441, .376)
	Online	.266*	(.109)	.204+	(.058, .350)	.099+	(.025, .193)
Marketing Efforts							
No. of Mailers	Offline	-.002***	(.000)	-.112+	(-.119, -.106)	-.104+	(-.109, -.099)
	Online	-.377***	(.017)	-.002+	(-.003, -.001)	-.002+	(-.003, -.001)
Inter Mailer Time	Offline	.000	(.000)	.001	(-.001, .003)	.001	(-.001, .003)
	Online	.007	(.040)	.001	(-.001, .003)	.002	(-.001, .004)
Product Categories							
Technology	Offline	-.406*	(.160)	-.238	(-.965, .529)	-.059	(-1.068, .994)
	Online	.133	(.639)	.495	(-.206, 1.254)	.273	(-.258, .827)
Leisure	Offline	-.979***	(.141)	-.190	(-.923, .428)	-.514	(-1.559, .514)
	Online	-.026	(.671)	-.617+	(-1.208, -.002)	-.570	(-1.123, -.114)
Family and Personal	Offline	.134	(.344)	.604	(-.447, 1.703)	.114	(-.676, .977)
	Online	.740	(.688)	.680+	(.194, 1.169)	.362+	(.172, .746)
Hobby	Offline	-.827***	(.147)	-.188	(-.882, .434)	-.482	(-1.300, 2.264)
	Online	1.120*	(.560)	.409+	(.241, .578)	.286+	(.016, .557)
Miscellaneous	Offline	-.648***	(.146)	-.464+	(-1.242, -.306)	-1.055+	(-1.985, -.094)
	Online	-.497	(.762)	-.204	(-.821, .415)	-.583+	(-1.106, -.084)

Table 2.10 (Continued)
Robustness Check: Results of Alternate Channel Choice Models

		Model IV		Model V		Model VI	
		MLE #		MCMC ##			
		One Purchase in Each Channel		Two Purchases in Each Channel		Three Purchases in Each Channel	
Moderating Role of Product Categories							
Income X Technology	Offline	-1.9e-06***	(2.5e-7)	-1.4e-6+	(-2.6e-6, -2.1e-7)	-1.7e-6	(-3.3e-6, 8.5e-8)
	Online	2.6e-06	(2.5e-7)	1.6e-6+	(4.8e-7, 2.7e-6)	1.1e-6+	(2.5e-7, 2.0e-6)
Income X Leisure	Offline	5.5e-07	(1.0e-6)	1.2e-6	(-3.5e-7, 2.4e-6)	9.2e-7	(-2.9e-7, 2.2e-6)
	Online	2.6e-06***	(1.7e-7)	2.9e-6+	(2.1e-6, 3.9e-6)	1.8e-6+	(1.0e-6, 2.6e-6)
Family Size X Family and Personal	Offline	-.057 *	(.027)	-.054+	(-.092, -.014)	-.034	(-.103, .033)
	Online	.014	(.014)	.007+	(.006, .020)	.025+	(.004, .048)
Education X Technology	Offline	-.001	(.001)	-.003+	(-.006, -1.9e-5)	-.005+	(-.010, -.002)
	Online	.010***	(.003)	.006+	(.002, .011)	.002+	(.001, .004)
Education X Family and Personal	Offline	.004**	(.001)	.003+	(.001, .007)	.001+	(.000, .002)
	Online	-.001	(.003)	-.002	(-.008, .005)	-.000	(-.007, .007)
Experience X Technology	Offline	.002***	(.000)	.000	(-.000, .001)	.001	(-.001, .002)
	Online	-.004***	(.001)	-.001+	(-.002, .-001)	-.000	(-.001, .001)
Experience X Leisure	Offline	.002***	(.000)	.001+	(-.001, .001)	3.4e-5	(-.001, .001)
	Online	-.003***	(.001)	-.001+	(-.002, .-000)	-.001+	(-.002, -.000)
Experience X Family and Personal	Offline	.005***	(.000)	5.6e-5	(-.002, .002)	.000	(-.001, .001)
	Online	-.000	(.001)	-.001+	(-.002, -.000)	-.001+	(-.002, -.001)
Items X Technology	Offline	-.001***	(.000)	-.002+	(-.004, -.001)	-.001	(-.003, -.000)
	Online	-.009**	(.003)	-.001+	(-.004, -.000)	-.001+	(-.001, -.000)
Items X Hobby	Offline	-.001***	(.000)	-.001+	(-.002, .001)	-.001+	(-.002, -.000)
	Online	-.000	(.000)	-.001+	(-.002, -.001)	-.001+	(-.002, -.001)
Items X Family and Personal	Offline	.002***	(.000)	.002+	(.000, .003)	.001+	(.000, .002)
	Online	.005	(.002)	.002+	(.000, .003)	.001+	(.001, .001)
Seasonality X Leisure	Offline	.204***	(.033)	.122	(-.011, .242)	.040	(-.170, .243)
	Offline	.274**	(.091)	.286+	(.141, .441)	.192	(.055, .332)
Seasonality X Miscellaneous	Offline	.160***	(.035)	.231+	(.095, .363)	.225+	(.014, .445)
	Offline	.467**	(.117)	.102	(-.047, .257)	.094	(-.021, .212)
Sample Size		137,693		133,041		113,075	
Log Likelihood		-82,284.42					

Notes: # Estimates (Standard Error), ## Posterior Means (95% Probability Interval)
Multichannel is the base channel. Significance levels: * (5%), ** (1%), *** (.1%).
Zero is not in the 95% probability interval of posterior mean: +

the stability of multinomial probit model results across different measures of dependent variable. The results also indicate the stability of multinomial probit model across the augmented Bayesian and maximum likelihood estimation methods. Thus, the results of channel choice model remain stable across the estimation techniques, convergence algorithms, and measures of the dependent variable.

I also rule out a possible concern that operationalization of single channel and multichannel customer are ex-post, that is, based on the data and hence model may not truthfully classify the customers. First, the data used are based on shopping behavior over a long time window of four years, during which Internet had established itself and there were many multichannel retailers. It is reasonable to conclude that customers who shopped from only one channel during this long time period are most likely to be stable as single channel customers unless there are dramatic changes in marketing efforts and other covariates. Second, I compared the predicted channel choices from a comparable holdout sample of customers with their actual choices. The “hit rate” of predictions was above 80%, suggesting that my model and results are fairly robust.

I further check the robustness of channel choice model for multichannel customers by estimating the proportion of their online versus offline shopping. I use the econometric method associated with fractional response variable for estimating the degree of online shopping as a proportion of total number of orders placed (Papke and Wooldridge 1996). The method requires a logit transformation of the dependent variable followed by regressing the transformed dependent variable on the drivers of channel choice as included in Model III. In Table 2.11, I present the results of this estimation method. The directionality and significance of the results suggest that drivers of the proportion of online orders are similar to those reported in Model III.

RFM Model

I performed several robustness checks for the RFM model. Potential weak instruments pose significant challenges in simultaneous equation system (Bound, Jaeger, and Baker 1995). The Staiger and Stock (1997) test of F statistics for the first stage regression in data does not indicate presence of poor instruments. The F-statistics of the frequency and monetary value equations are 120,567 and 170,201, respectively. The adjusted R^2 for these regression equations are also healthy at 59% and 78%. Thus, any weakness of instruments induces less 10e-04% bias. The Hahn and Hausman (2002) test showed that forward and reverse estimates of the right hand side endogenous variable when estimated through 2SLS are not significantly different. I also compare the estimates from the 2SLS estimation with those from the LIML estimation, which performs better with weak instruments. These estimates were very similar, so I again rule out the presence of weak instruments. The results of the RFM model also remain stable when I use predicted choices obtained from Model V and VI. Thus, the results of RFM model based on a robustness check of the channel choice model remain similar to the ones reported in Table 2.8.

Other Robustness Checks

I checked to see if the Poisson model captured the frequency equation well. I estimated the frequency equation, using an alternative model for the count data, the negative binomial distribution (NBD) model. The substantive results remained the same.

Table 2.11
Robustness Check: Results of Channel Choice Model for Multichannel Customers

	Proportion of Online Spending (Standard Error)	
Customer Demographics		
Intercept	.441	(.357)
Income	-4.6e-7	(6.0e-7)
Family Size	-.056***	(.017)
Age	.018	(.012)
Age Square	-.000	(.000)
Education	.001	(.001)
Shopping Traits		
No. of Items	-.002***	(.000)
Shopping Experience	-.004***	(.000)
No. of Categories	-.064***	(.002)
Seasonality of Purchases	.009	(.071)
Marketing Efforts		
Mailers	-.007***	(.000)
Inter Mailer Time	.010***	(.000)
Product Categories		
Technology	.266	(.142)
Leisure	-.216	(.128)
Family and Personal	.134	(.344)
Hobby	-.422	(.350)
Miscellaneous	-.685***	(.131)
Moderating Role of Product Categories		
Income X Technology	4.3e-7*	(2.2e-7)
Income X Leisure	6.4e-7**	(1.9e-7)
Family Size X Family Personal	.010	(.015)
Education X Technology	.002***	(.001)
Education X Family and Personal	7.7e-6	(.002)
Items X Technology	-.001***	(.000)
Items X Hobby	-.001***	(.000)
Items X Family and Personal	.001**	(.000)
Experience X Technology	-.001**	(.000)
Experience X Leisure	-.000*	(.000)
Experience X Family and Personal	-.001*	(.000)
Seasonality X Leisure	.012	(.031)
Seasonality X Miscellaneous	.074*	(.032)
Sample Size	31,531	
Adjusted R Squared	15.11%	

Note: * (5%), ** (1%), *** (.1%).

MANAGERIAL AND RESEARCH IMPLICATIONS

Channel Equity/Value

From my results, managers can compare the contributions and the values of the customers in each channel. A comparison of the contributions and values of the channels appears in Table 2.12 and Figure 2.2. The volume share is the percentage of customers who were classified as offline only, online only and multichannel customers. Multichannel customers, who comprise about 23% of the customers in the data, contribute about 36% of total monetary value generated. In contrast, offline only and online only customers contribute lower shares of monetary value, that is, 63% and 1%, respectively compared to their volume shares of 72% and 5%, respectively. Similarly, my models predict that the average expenditures by an offline only, an online only, and a multichannel customer during this time window are $561, $237, and $1028, respectively. The marginal value of a customer in a channel type is the monetary value of a customer in that channel buying one marginal item in one product category through one order. The marginal values of an offline only, an online only, and a multichannel customer in the data are $82, $18, and $230, respectively. Table 2.12 also shows the standard errors of the average and marginal values by channel type. The standard errors, particularly those for the average channel value per customer, are reasonably small, reiterating that the equity per multichannel customer is, indeed, much higher than that per single channel customer.

These results show the significantly higher value or equity of a multichannel customer over a single channel customer. Thus, controlling for endogeneity, I find that the equity of an average multichannel customer is nearly twice as that of an average offline only customer, and about five and a half times as much as that of an average online only customer.

Figure 2.2
Comparison of Channel Contribution and Value

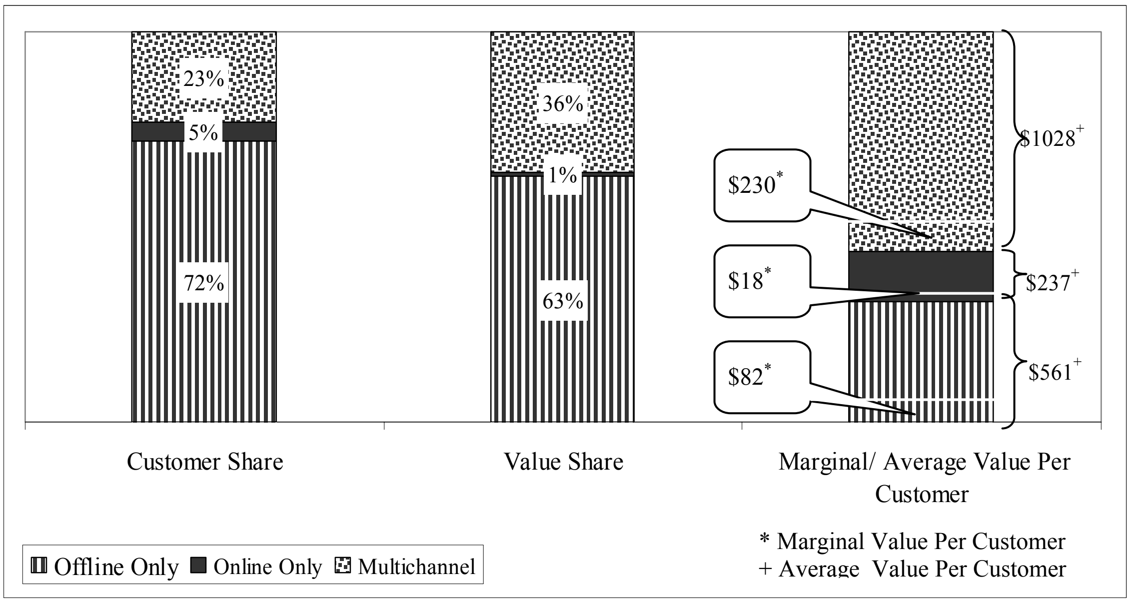


Table 2.12
Comparison of Contributions and Values of Channels

	Customer Share	Value Share	Average Value Per Customer	Marginal Value Per Customer	SE of Average Value Per Customer	SE of Marginal Value Per Customer
Offline Only	72%	63%	$561.14	$82.26	$2.52	$11.80
Online Only	5%	1%	$237.11	$17.72	$5.10	$13.68
Multichannel	23%	36%	$1,028.02	$230.30	$5.15	$11.15

Other Managerial Implications

In addition to providing insights on channel value, the results offer other valuable managerial implications. First, managers can get a better understanding of multichannel and single channel shoppers. Relative to single channel shoppers, multichannel shoppers are more affluent, have a longer shopping experience, buy more product categories, and have a lower proportion of annual purchases from holiday buying. They buy more (less) items of technology (family and personal) products than do single channel shoppers. They also receive more marketing mailers than do single channel customers. If a retailer wants to target multichannel shoppers, she can do so by aiming at households with high income and large basket sizes and offer or feature an array of technology products through a high dose of marketing mailers.

Second, managers can develop a better understanding of the relative values of their customers by their channel of purchase. Based on these values, they can suitably allocate their financial resources to each channel and customer cohort.

Third, the results show that the monetary values of an incremental order and an additional category are the largest, after adjusting for channel differences. Retailers should, therefore, focus on getting a customer to place an additional order in one category, regardless of their channel choice because that these variables are associated with the largest lifts in monetary value.

Implications Relative to Prior Research

The results extend findings from prior research in a number of ways. By estimating the value of customers by channel type and by identifying their characteristics through an analysis of a large scale database covering several product categories, retailers, and years, I offer generalizable results. First, I extend the findings of Kumar and Venkatesan (2005) and Montoya-Weiss, Voss, and Grewal (2003). Kumar and Venkatesan (2005) found that multichannel shoppers provide greater revenue and share of wallet than do single channel shoppers. I show that the value of multichannel customers is nearly two (five and one half) times that of offline (online) only customers. The result that multichannel customers provide greater value is consistent with Montoya-Weiss, Voss, and Grewal (2003), who predict that multiple channels offer greater scope for higher customer satisfaction.

Second, with regard to channel choice, my findings add to Ansari, Mela, and Neslin (2007), Knox (2006), and Venkatesan et al. (2007). While Ansari, Mela, and Neslin (2007) found that marketing efforts and experience drive channel selection, I show that in addition to these variables, demographic variables such as age and education and shopping traits such as number of categories and seasonality of purchases also determine channel choice. Moreover, consistent with Knox (2006), I find that response to marketing efforts varies across customer channel segments. Furthermore, I build on the finding of Venkatesan, Kumar, and Ravishanker. (2007) that purchase frequency and marketing communications have the strongest influence on the adoption of the next channel adoption by showing that in addition to the direct effects of marketing efforts and customer shopping traits on channel choice, product categories significantly moderate the effects of customer demographics and of shopping traits on channel choice.

Third, my analysis of the RFM of customer purchases offers new results. The results show that after adjusting for the effects of channel, an incremental order of a product category has the greatest influence on the monetary value of a customer's purchases.

LIMITATIONS, FUTURE RESEARCH, AND CONCLUSIONS

This research has limitations that could be addressed by future research. First, I have studied only observed purchase behavior. I do not have data on how customers use the channels in making their final purchase decisions. Analyzing such data together with transaction data could shed additional light into our understanding of single channel vs. multichannel shopping and extend Verhoef, Neslin, and Vroomen (2007). Second, the data used in the study are from firms which offer products through a catalog or a Web site. Although my sample is very large and provides high external validity, it does not include purchases from brick-and mortar retailers such as Wal-Mart, Target, and K-Mart. It would be useful to extend the analysis to purchases from firms which offer their products through physical stores in addition to the catalog and the Web channels, consistent with Thomas and Sullivan (2005). Third, the data used in the study are cross-sectional. If large scale longitudinal data across product categories and retailers are available, the dynamics of channel choice can be modeled. Fourth, due to the nature of this data, the relationships I uncovered are associative, not causal. As suggested by Neslin et al. (2006), new types of data and analyses are need to offer definitive conclusions on what causes single channel or multichannel shopping. Finally, I did not examine marketing resource allocation across the channels. Future research could address this issue as well.

In conclusion, I sought answers to important questions relating to the growing multichannel shopping phenomenon: What factors drive channel choice? What is the role of product category on the effects of these factors? How does channel choice affect the recency, frequency, and monetary value of customers' purchases? How valuable are single channel and multichannel shoppers to the firm? I addressed them by developing a model of channel choice that explains its drivers and consequences such as recency, frequency, and monetary value. The empirical analysis on a unique large database that cuts across several product categories and retailers over multiple years allows me to make some empirical generalizations. The results show significant direct effects of customer demographics and shopping traits, and marketing mailers, and moderating effects of product category on channel choice. Surprisingly, inter-mailer time has no effect on channel choice. I also find that multichannel customers are much more valuable than single channel customers as they provide greater monetary value, spend more frequently, and purchase more recently than single channel customers. The equity or value generated by an average multichannel customer is nearly two (five and one half) times that generated by an offline (online) only customer. These findings offer important guidelines to managers in formulating their multichannel marketing strategies and could serve as an impetus for further research on this growing phenomenon.

CHAPTER III

BRICKS, CLICKS, AND FLICKS: OPTIMAL ALLOCATION OF MARKETING EFFORTS BY CUSTOMER-CHANNEL SEGMENTS

As firms increasingly offer their products through multiple channels such as store (bricks), the Web (clicks) and catalog (flicks) and more consumers buy them through different channels, the allocation of marketing efforts targeted at customers across channels is becoming a critical issue for many marketers. I propose a model for optimal allocation of marketing efforts across multiple channels. I first develop a set of models for consumer response to specific marketing efforts for each customer-channel segment. This set comprises four models, the first for purchase frequency, the second for purchase quantity, the third for product return propensity, and the fourth for contribution margin of an item. For the purchase frequency model, I use an extended Beta Geometric/Negative Binomial Distribution model, which includes the effect of marketing and other relevant covariates on purchase frequency. I model purchase quantity and product return propensity using Conditional Negative Binomial Distribution models. I model the contribution margin of an item using a Gamma-Gamma model. Based on the parameter estimates from the models, I then derive the optimal marketing effort allocation to each customer-channel segment. I estimate the models using customer level purchase, cost, and promotional data from a large marketer of shoes and apparel accessories across multiple channels, namely, the direct mail, the store, and the web. I solve the optimization model using simulations. The results show that consumer response to firm marketing efforts varies significantly across the customer-channel segments. The results also suggest that marketing efforts influence purchase frequency, purchase quantity and monetary value in different ways. The results show that firm profits can be substantially improved (by as much as 32%) by more optimally reallocating marketing efforts across the different customer-channel segments.

INTRODUCTION

The allocation of marketing resources across various marketing instruments and customer segments is a topic of immense interest to marketing academics and practitioners alike. Studies in marketing have examined managerial allocation of marketing resources across markets (Mantrala, Sinha, and Zoltners 1992), media (Naik, Mantrala, and Sawyer 1998), mailing campaigns (Elsner, Krafft, and Huchzermeier 2004), acquisition and retention efforts (Blattberg and Deighton 1996; Reinartz, Thomas, and Kumar 2005), and customers (Venkatesan and Kumar 2004).

Organizations deploy marketing resources through multiple channels such as physical store (bricks), the Web (clicks), and catalog (flicks). Customers may choose one or more of these channels for their purchases and can be segmented based on their channel choice, leading to customer-channel segments (Kumar and Venkatesan 2005; Thomas and Sullivan 2005). Multichannel shoppers constitute a valuable customer segment for marketers (e.g., *Doubleclick* 2004; Kumar and Venkatesan 2005). By knowing how different customer-channel segments respond to marketing efforts, managers can better allocate their resources across these segments.

However, little is known about the allocation of marketing resources across these customer-channel segments (Berger et al. 2002; Neslin et al. 2006).[6] The primary advantage of allocating marketing resources at the customer-channel segment level is that it provides firms with the ability to use channel as a segmentation tool and use marketing instruments in different channels with differing intensities.

In this paper, I address three important research questions relating to the allocation of marketing efforts across customer-channel segments. First, how much marketing efforts should a firm expend for each customer-channel segment? Second, can a firm improve its profitability by incorporating multichannel shopping behavior in its resource allocation decisions? Third, can we decompose the responsiveness of profits to marketing efforts in each customer-channel segment into the responsiveness of purchase frequency, purchase quantity and contribution margin to marketing efforts?

I develop models to estimate the responsiveness of different customer-channel segments to marketing efforts. Based on these response parameters, I develop a resource allocation model to optimize the allocation of marketing resources across customer-channel segments and maximize firm profits. I derive firm profits by aggregating the profit contributions from each customer through customer-channel segments. I decompose customer profits into profits due to purchase frequency (number of orders), purchase quantity per order, product return propensity, and contribution margin per item by developing models for each of these components. For purchase frequency, I use an extended Beta Geometric/Negative Binomial Distribution model that includes the effect

[6] I do not address allocation across channels, which may involve the fixed costs of setting up and running channels.

of marketing covariates on purchase frequency. I model purchase quantity and product return propensity, using Conditional Negative Binomial Distribution models. I model contribution margin per item using a Gamma-Gamma model.

This modeling approach extends related existing research in many ways. First, existing customer level approaches (e.g., Fader, Hardie, and Lee 2005a, b; Kumar and Venkatesan 2005) decompose model contribution margin at an order level by splitting it into two elements: purchase frequency and contribution margin per order. This approach extends these approaches by modeling contribution margin at an item level by decomposing it into three components: purchase frequency, purchase quantity per order, and contribution margin per item. Such an approach enables me to differentiate between order size effect and up-selling effect. These effects may be different and need to be captured separately to derive an optimal customer-channel marketing effort allocation model. These effects, however, have not been accounted for by prior studies. Second, prior stochastic models of customer purchase behavior in an interactive marketing context (e.g., Fader, Hardie, and Lee 2005a, b) do not include marketing covariates and hence do not offer optimal marketing allocation results. The set of models used here includes the effects of marketing efforts and offers optimization of marketing resources. Third, previous resource allocation models tend to ignore customers' product return propensities, leading to inflated firm profits. Anecdotal evidence suggests that product returns can range from 4% to 25% of the orders, depending on the product category (Fenvessy 1992; Hess and Mayhew 1997). By incorporating product return propensity into my model, I use a more appropriate metric for computing firm profits.

My results show that the responsiveness of different customer-channel segments to marketing efforts varies substantially across the different components of profits and across different customer-channel segments. They also show that a firm can improve its total profits by optimally allocating marketing resources to these customer-channel segments based on the heterogeneous response behavior of these segments to the firm's marketing efforts.

The remainder of this chapter is organized as follows. In the next section, I briefly review related research and develop the conceptual framework. In the third section, I develop a disaggregated model for firm profits, which I decompose into four models. In the fourth section, I briefly discuss the customer level transaction data I use for estimating the models. In the fifth section, I discuss model estimation and the results. I also evaluate the predictive validity of each model for each customer-channel segment. I conclude by summarizing the work and discussing the implications of the findings.

LITERATURE REVIEW AND CONCEPTUAL DEVELOPMENT

There is growing research on multichannel shopping and marketing (e.g., Ansari, Mela, and Neslin 2007; Balasubramanian, Raghunathan, and Mahajan 2005; Gensler, Dekimpe, and Skiera 2004; Knox 2005; Inman, Shankar, and Ferraro 2004; Kumar and Venkatesan 2005; Kushwaha and Shankar (2007); Montoya-Weiss, Voss, and Grewal 2003; Thomas and Sullivan 2005; Venkatesan, Kumar, and Ravishanker 2007; Verhoef and Donkers 2005; Verhoef, Neslin, and Vroomen 2005; Villanueva, Yoo, and Hanssens 2003). Another stream of relevant research focuses on marketing efforts in a direct marketing context (Bult and Wansbeek 1995; Bitran and Mondschein 1996; Elsner, Krafft, and Huchzermeier 2005; Gönǖl and Shi 1998; Gönǖl and ter Hofstede 2006). The results from these studies suggest that customers can be segmented based on their channel usage and response behavior.

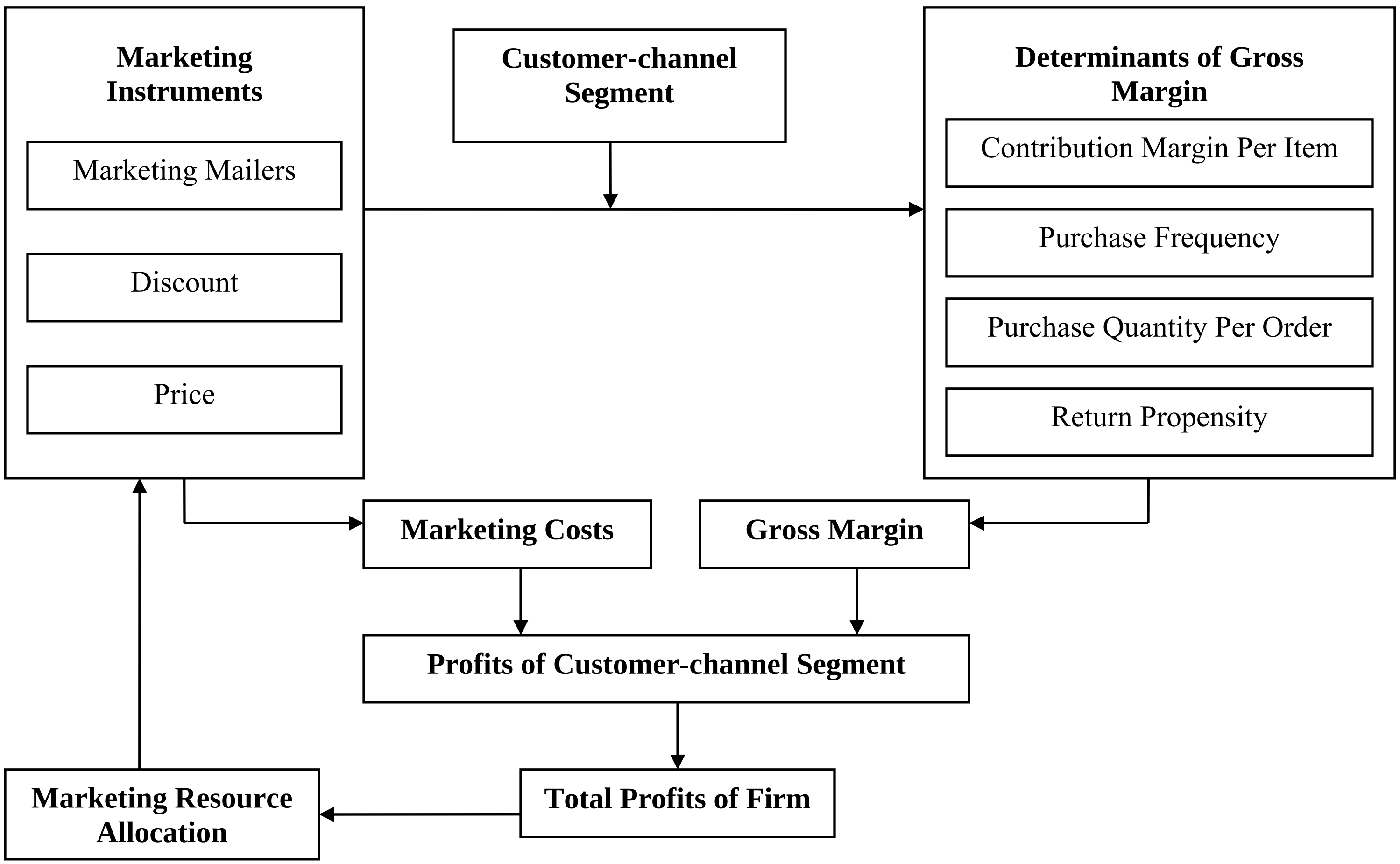
The literature that is most relevant to this essay is the research stream that focuses on resource allocation decisions. Venkatesan and Kumar (2004) identify low and high value customer segments and determine optimal allocation of marketing resources to these segments. Some research has also examined allocation between customer acquisition and retention (Reinartz, Thomas, and Kumar 2005), and between retention and reacquisition of lost customers (Thomas, Blattberg, and Fox 2004). However, the problem of resource allocation has not been addressed at the customer-channel segment level. The need to address resource allocation decisions at customer-channel segment level has been raised by practitioners and academia alike. Studies by IBM and McKinsey & Company call for developing resource allocation metrics across channels (Achabal et al. 2005; Myers, Pickergill, and van Metre 2004). Berger et al. (2002) and Neslin et al. (2006) emphasize the need to develop models for the allocation of marketing resources across channels. Libai, Narayandas, and Humby (2002) show that a channel-segment based approach to resource allocation can bring significant improvements in firm profitability.

I address the resource allocation problem across customer segments derived from channel choice behavior. I base the resource allocation decisions on the predicted future profits of the firm. The primary advantage of developing such a forward looking resource allocation metric at the customer-channel segment level is the ability to use channel as a segmentation tool.

I present the conceptual framework in Figure 3.1. In this framework, a firm's marketing instruments (marketing mailer, promotional discount, and price) influence purchase behavior (purchase frequency, purchase quantity per order, and contribution margin per item). I anticipate that the effect of marketing instruments on customer's responsiveness will be different across the different components of purchase behavior or profits and across the customer-channel segments. The primary thesis of this work is that (1) segmenting by channel choice is an efficient way of segmenting customers; (2) different customer-channel segments respond differently to a firm's marketing efforts; and (3) firms can improve their profits by optimally allocating marketing resources to these customer-channel segments based on the responsiveness of these segments to the firm's marketing efforts.

There are important theoretical reasons to expect differential responsiveness of different customer-channel segments to marketing efforts. First, different channels offer different opportunities for customers to interact with the firm and these interactions could lead to different responses to marketing efforts (Berger et al. 2002). For example, a customer who shops only at a store may receive a marketing mailer and buy a wide assortment and large quantities of the items of selected products to amortize the cost of a trip to the store (Bell and Lattin 1998; Bhatnagar and Ratchford 2004; Messinger and Narasimhan 1997). In contrast, a customer who buys only on the Web, may not buy an item in large quantities in the first order before verifying the correctness of choice of the

Figure 3.1
Conceptual Framework for Multichannel Resource Allocation Decisions



item after delivery (Bart et al. 2005). Second, different channels provide different levels of involvement for customers, leading to greater or smaller allocation of attentional resources for information processing and thereby differential levels of responsiveness to marketing activities (Assael 1998; Greenwald and Leavitt 1984). Third, each channel calls for a unique level of cognitive effort on the part of customer to be able to react to marketing messages (Balasubramanian, Raghunathan, and Mahajan 2005). Fourth, the relationships among communication intensity, commitment and trust are different for different channels (Morgan and Hunt 1994).

MODEL DEVELOPMENT

Resource Allocation Model

The firm's objective function, total profits, is the sum of profits generated by each customer-channel segment of the firm. Let Π be the total profits of a firm over a given time horizon and Π_k be the profit of the k^th customer-channel segment over the same horizon. Because the time horizon for each segment is the same, for ease of exposition, I drop the time subscript from the equations. The total profits are given by:

$$\Pi = \sum_{k=1}^{K+1} \Pi_k (m_{ik}) \tag{3.1}$$

where m_{ik} is the number of marketing mailers sent to customer i from customer-channel segment k and is the resource allocation variable.[7]

[7] I do not need to include a resource constraint on the marketing variable because any solution that violates non-negativity or ceiling constraint on the marketing variable will be rejected on grounds of face validity.

A firm that markets through k channels will have k single-channel segments and one customer-channel segment of 'multichannel' users. The profits derived from a customer-channel segments are the sum of the gross margins contributed by each customer belonging to the customer-channel segment k less the sum of the costs of marketing to all the customers in that customer-channel segment.

$$\Pi_k = \sum_{i=1}^{n_k} [TM_{ik}(m_{ik}) - c_m m_{ik}] \tag{3.2}$$

where TM_{ik} is the total gross margin contributed by customer i of customer-channel segment k and it is a function of m_{ik}, c_m is the unit cost of a marketing mailer which does not vary across customers, and n_k is size of customer-channel segment k.

The total margin earned from customer i of customer-channel segment k is given by

$$TM_{ik}(m_{ik}) = NIB_{ik}(m_{ik}) \times \overline{CM}_{ijk}(m_{ik}) \tag{3.3}$$

where NIB_{ik} is the net number of items bought by customer i belonging to customer-channel segment k and $\overline{CM}_{ijk}$ is the average gross contribution margin per item of customer i belonging to customer-channel segment k over the J items bought by that customer. $\overline{CM}_{ijk}$ is given by:

$$\overline{CM}_{ijk}(m_{ik}) = \frac{1}{J}\sum_{j=1}^{J} CM_{ijk}(m_{ik}) \tag{3.4}$$

CM_{ijk} is the gross contribution margin of customer i belonging to customer-channel segment k for the j^{th} item bought and is given by:

$$CM_{ijk} = P_j - C_j - D_{ijk} \tag{3.5}$$

where P_j is the price of item j, C_j is the cost of item j, and D_{ijk} is the discount offered to customer i belonging to customer-channel segment k for item j. Note that while price and cost vary across items, the promotional discount can vary across both items and customers. In other words, some customers may buy an item during a promotion period, while others may buy during a non-promotion period, leading to different contribution margins for the same item across different customers.

The net items bought by customer i belonging to customer-channel segment k is given by:

$$NIB_{ik}(m_{ik}) = [IPO_{ik}(m_{ik}) - IRPO_{ik}] \times NO_{ik}(m_{ik}) \tag{3.6}$$

where IPO_{ik} is the number of items bought per order by customer i belonging to customer-channel segment k, $IRPO_{ik}$ is the number of items returned per order by customer i belonging to customer-channel segment k, and NO_{ik} is the number of orders by customer i belonging to customer-channel segment k.

The firm's profit from the customer-channel segment k of size n_k customers is given by:

$$\Pi_k = \sum_{i=1}^{n_k}[(IPO_{ik}(m_{ik}) - IRPO_{ik}) \times NO_{ik}(m_{ik}) \times \overline{CM}_{ijk}(m_{ik}) - c_m m_{ik}] \tag{3.7}$$

Using this disaggregated approach, I decompose the total margin derived from a customer i into four different sub models: purchase frequency (NO_{ik}), purchase quantity per order (IPO_{ik}), product returns per order ($IRPO_{ik}$), and gross contribution margin ($\overline{CM}_{ijk}$). This approach enables me to understand the effect of marketing efforts on order size and on up-selling.

In the above model specification, I assume that the purchase frequency of a customer is independent of her contribution margin per item.[8] However, I do not assume that the customer's purchase quantity per order and returns per order are independent of her purchase frequency. It is reasonable to assume that customers with a greater purchase quantity per order may buy less frequently than those with a smaller purchase quantity per order. Similarly, the probability of returning an item is high when purchase frequency is high. Thus, I condition the predicted purchase quantity per order and predicted returns per order model on the predicted purchase frequency.

The objective function has a total margin component and a cost component, both of which are a function of the firm's marketing efforts. I want to identify the optimal values of marketing efforts that would maximize Π_k for each customer-channel segment and Π for the firm across the customer-channel segments. The optimization equation is given by:

$$\text{Allocation Rule}(m_k) =$$

$$\left[\max_{m_k} \sum_{k=1}^{K+1}\sum_{i=1}^{n_k}[(IPO_{ik}(m_{ik}) - IRPO_{ik}) \times NO_{ik}(m_{ik}) \times \overline{CM}_{ik}(m_{ik}) - c_m m_{ik}\right] \tag{3.8}$$

Marketing cost is a linear function of the amount of marketing efforts undertaken by a firm, consistent with Mantrala, Sinha, and Zoltners (1992) and Venkatesan and Kumar (2004). However, the total margin contributed may or may not have a similar relationship with marketing effort. Thus, I wish to identify the response parameters

---

[8] It could be argued that a low income customer might order low margin items more frequently. However, it is also possible that a high income customer might order low margin items less frequently. Therefore, in the overall population, the relationship between purchase frequency and contribution margin per item may not be highly correlated.

associated with marketing efforts. Given the response parameters, marketing costs, and the total margin of a customer-channel segment, I can determine the level of marketing efforts that would maximize the profits from that segment.

In the model, I include three different types of marketing instruments, namely, marketing mailer, promotional discount, and price. These instruments may have different elasticities. From a resource allocation perspective, however, I optimize only the number of mailers sent to a customer-channel segment. Such an assumption is consistent with industry practice. For example, JC Penney, which spends about $357 million on marketing mailers, makes important decisions on the customers to whom mailers should be sent based on the channels through which they shop (*J.C. Penney 10K Statement* 2006).

I also assume the prices of the products and the discounts offered to customers as exogenous to the modeling system. Note that the aim of the study is to optimize resource allocation decision at the customer-channel segment level, so I do not pursue resource allocation at an individual customer level within each customer-channel segment.

I approach the resource allocation model in four steps. First, I estimate each of the four models (purchase frequency, purchase quantity, product return propensity, and contribution margin) for every customer-channel segment to get the response and shape parameters. I anticipate that the response parameters for a given marketing instrument will be different across different customer-channel segments. Second, using these response and shape parameters, I predict the different components of purchase behavior in the prediction window. Third, I evaluate the predictive ability of the model for each customer-channel segment. Finally, through simulation, using these estimated response and shape parameters and the cost of each type of marketing instrument for each customer-channel segment, I get the optimal values of $\sum_{i=1}^{n_k} m_{ik}$, the levels for marketing efforts that should be expended toward the customer-channel segment k. I compare the predicted profits derived from my model with those actually generated by the firm in a holdout sample. The difference reflects the profit improvement generated by this modeling approach.

Purchase Frequency Model

The two most popular methods to estimate purchase frequency are the Pareto/NBD (Schmittlein, Morrison and Colombo 1987) and the generalized gamma model (Allenby, Leone and Jen 1999). While the Pareto/NBD model assumes a Poisson distribution of customer transaction rate, the generalized-gamma model assumes a gamma distribution of customer transaction rate. The Pareto/NBD model is theoretically appealing, but its implementation requires tedious evaluation of multiple Gaussian hyper geometric functions. The availability of faster computing has partially resolved this problem. Some studies in marketing have successfully implemented the model (e.g. Schmittlein and Peterson 1994; Reinartz and Kumar 2000; Fader, Hardie, and Lee 2005b). Fader, Hardie and Lee (2005a) developed a derivative of the Pareto/NBD model, namely, the Beta Geometric/NBD (BG/NBD) model, which is significantly easier to implement, yet performs similar to the Pareto/NBD model. None of the studies that use the Pareto/ NBD and the BG/ NBD model in marketing, however include the effects of covariates on purchase frequency. The other derivatives of the Poisson class of models are the Conditional NBD and the HB version of NBD (see Jen, Chou, and Allenby 2003). While the Pareto/NBD and the BG/NBD are four-parameter models, the Conditional NBD and the HB NBD are two-parameter models, so hence their model fits are not comparable with those of the four-parameter models.

I use the BG/NBD model to estimate and predict customer purchase frequency, consistent with Fader, Hardie, and Lee (2005a). The BG/NBD model has two distinct parts. First, the probability of a customer remaining active is captured by the BG part of the model. Second, the transaction rate of a customer who is active is captured using the NBD part of the model. The probability of a customer remaining active and the transaction rate are assumed to be independent across customers. Additional underlying assumptions of the model are:

a) A customer can become inactive immediately after her purchase. This customer drop out is distributed across transactions according to a geometric distribution,
b) the dropout rate across customers is beta distributed with shape parameters 'a' and 'b',
c) while active, a customer's transaction rate follows Poisson distribution, and
d) the transaction rate across customers is gamma distributed with shape parameter 'r' and scale parameter α .

The model requires three pieces of information from each individual customer: the number of repeat transactions (x), the time since the first purchase *(T)*, and the time of last purchase (t_x). While the customer dropout process is inherently a stochastic process, the transaction rate of customers who are active can be modeled for the effect of covariates. I introduce covariates to the model as a function of the shape parameter of gamma heterogeneity in following form.

$$\alpha_{ik} = \frac{\exp(\beta z'_{ik})}{r} \tag{3.9}$$

where z'_{ik} is the matrix of covariates and β is a vector of response parameters. I introduce three marketing covariates in the model: number of marketing mailers, the discount offered per item, and the average price of each item. The model derivation and estimation details are shown in the Appendix 1. The modified log likelihood function for customer-channel segment k is given by:

$$LL_k(r,a,b,\beta \mid X = x_{ik}, t_{xik}, T_{ik}) =$$

$$\sum_{i=1}^{n_k} Ln\left[\frac{\left[\dfrac{\Gamma(r+x_{ik})\left(\dfrac{\exp(\beta z'_{ik})}{r}\right)^{r}}{\Gamma(r)}\right]}{\left[\dfrac{\Gamma(a+b)\Gamma(b+x_{ik})}{\Gamma(b)\Gamma(a+b+x_{ik})}\right]}\left[\left(\frac{1}{\dfrac{\exp(\beta z'_{ik})}{r}+T_{ik}}\right)^{r+x_{ik}} + \delta_{x>0}\left(\frac{a}{b+x_{ik}-1}\right)\left(\frac{1}{\dfrac{\exp(\beta z'_{ik})}{r}+t_{xik}}\right)^{r+x_{ik}}\right]\right] \tag{3.10}$$

where x_{ik}, t_{xik}, and T_{ik} are the purchase frequency, time of last purchase, and time since first purchase of customer i of customer-channel segment k. The shape parameters of the joint distribution are ‘r’, ‘a’, and ‘b’.

The conditional expectation of purchase frequency of a customer in a non-overlapping prediction time window ‘t’, given the shape parameters, the response parameters, and the observed purchase behavior of the same customer in the estimation window of length ‘T’ is given by: (see Fader, Hardie, and Lee 2005a for details)

$$E(Y(t) \mid r, a, b, \beta, X = x, t_x, T) =$$

$$\frac{\frac{a+b+x-1}{a-1}\left[1-\left(\frac{\frac{\exp(\beta z^{''})}{r}+T}{\frac{\exp(\beta z^{''})}{r}+T+t}\right)^{r+x} {}_2F_1\left(r+x, b+x; a+b+x-1; \frac{t}{\frac{\exp(\beta z^{''})}{r}+T+t}\right)\right]}{1+\delta_{x>0}\frac{a}{b+x-1}\left(\frac{\frac{\exp(\beta z^{''})}{r}+T}{\frac{\exp(\beta z^{''})}{r}+t_x}\right)^{r+x}} \tag{3.11}$$

where ‘Y(t)’ is the expected number of transactions for a given customer in a future time window ‘t’, given her purchase behavior in a non-overlapping previous time window ‘T’ and $z^{''}$ are value of covariates for the given individual in the prediction window. ${}_2F_1$ is a Gauss Hypergeometric function (sum of a hypergeometric series) with two parameters of Type 1 and one parameter of Type 2. The expected number of transactions as given in Equation (3.11) incorporates the probability of a customer remaining active for the time duration ‘t,’ multiplied by the expected transaction rate and the time duration of the window.

The ${}_2F_1$ is Gauss Hypergeometric function is given in Equation 3.12 (see Fader, Hardie, and Lee 2005c for details). Note that one hypergeometric function is evaluated for each customer in the prediction window.

$${}_2F_1(P1, P2; P3; P4) = \sum_{s=0}^{\infty} ST_s \tag{3.12}$$

where, P1 = r + x, P2 = b + x, P3 = a + b + x -1, and P4 = $\dfrac{t}{\dfrac{\exp(\beta z'')}{r} + T + t}$ are four parameters of the hypergeometric series as given in Equation (3.11). ST_s denotes the s^{th} term of the hypergeometric series which is given by:

$$ST_s = \frac{(P1)_s (P2)_s (P4)^s}{(P3)_s s!} \tag{3.13}$$

where $(P1)_s$ is the ascending factorial given by (P1)(P1+1)(P1+2)…..(P1+s-1). The other two terms, $(P2)_s$ and $(P3)_s$ are defined in a similar fashion.

Purchase Quantity Model

While many studies investigate purchase frequency and purchase quantity separately, no study models the effect of marketing efforts on customer order size. Lewis(2006) and Lewis, Singh, and Fay (2006) model consumer order size in an online context as a function of shipping fees and find that the shipping fee structure can act as a motivation for consumers to increase their order sizes. The literature on customer basket size in the grocery industry has investigated the drivers and consequences of variation in customer basket size. Research has shown that availability of a surprise coupon for a pre-planned purchase product category helps to increase the size of a customer's basket (Heilman, Nakamoto, and Rao 2002). The basket size of a customer impacts her store selection through fixed and variable costs of shopping (Bell, Ho, and Tang 1998). The customer's expected basket size per grocery shopping trip is related to the choice of store format (Bell and Lattin 1998).

To measure the average size of an order on a given purchase occasion, I apply the commonly used count data regression approach. Purchase quantity per order for consumers follows a Poisson distribution. However, the restrictive assumption of the mean and the variance being equal in the Poisson distribution cannot capture overdispersed data. The NBD distribution is an ideal substitute for the Poisson distribution when data exhibit overdispersion (Cameron and Trivedi 1998). I use the Conditional NBD model developed by Morrison and Schmittlein (1988) to model customer purchase quantity per order. The NBD model assumes that purchase quantity per order is Poisson distributed with gamma heterogeneity across customers with shape parameter 'c' and scale parameter ϕ. The covariates in the model are introduced as a function of the scale parameter.

$$\phi_{ik} = \frac{\exp(\eta z'_{ik})}{c} \tag{3.14}$$

where η is the response parameter of covariates z'_{ik}.

The likelihood function of the NBD model for customer-channel segment k is given by

$$LL_k(c, \eta \mid F = f_{ik}) = \sum_{i=1}^{n_k} Ln\left[\frac{\Gamma(f_{ik}+c)}{\Gamma(f_{ik}+1)\Gamma(c)} \frac{(\eta z'_{ik})^c (c)^{f_{ik}}}{(\exp(\eta z'_{ik})+c)(c+f_{ik})}\right] \tag{3.15}$$

where f_{ik} is the purchase quantity per order of customer i belonging to customer-channel segment k.

The purchase quantity per order of a customer is not independent of her purchase frequency. For example, consumers who order more frequently may have smaller order sizes and vice-versa. I correct for the dependence by conditioning the expected purchase quantity per order of a given customer on her purchase frequency (Morrison and Schmittlein 1988). The conditional expectation of purchase quantity per order for consumers with at least one repeat purchase is given by:

$$E(G(y) \mid Y = y > 0; c, f, \eta, x, y) = \frac{(c+f)y}{\frac{\exp(\eta z'')}{c} + x} \tag{3.16}$$

The conditional expectation of purchase quantity per order for consumers with no repeat purchase is

$$E(G(y) \mid Y = 0; c, \eta, x, y) = \frac{(1-\gamma)\left(\frac{cy}{\frac{\exp(\eta z'')}{c} + x}\right)\left(\frac{\frac{\exp(\eta z'')}{r}}{\frac{\exp(\eta z'')}{c} + x}\right)^{c}}{\gamma + (1-\gamma)\left(\frac{\frac{\exp(\eta z'')}{c}}{\frac{\exp(\eta z'')}{c} + x}\right)^{c}} \tag{3.17}$$

where G(y) is the conditional expectation of purchase quantity per order for a customer with purchase frequency ‘y’ in the prediction window, ‘x’ is the purchase frequency of the customer in the estimation window, ‘f’ is the average purchase quantity per order in estimation window, z'' are the values of covariates in the prediction window, and γ is the proportion of consumers in the dataset with no repeat purchase.

Product Return Propensity Model

By modeling customer product return per order, I correct for the bias created in total margin derived from a customer when product returns are ignored. Prior studies in marketing do not account for the product return propensities of consumers and hence tend to overestimate purchase quantity, contribution margin, and thereby, their customer valuation. Product return per order is conceptually very similar to purchase quantity per order. However, I treat product return per order as a purely stochastic process. I assume return per order to be Poisson distributed with gamma heterogeneity, having a shape parameter 'd' and a scale parameter μ. The likelihood function for the NBD model for customer-channel segment k without covariates is very similar to the likelihood function in Equation (3.15).

$$LL_k(d, \mu \mid H = h_{ik}) = \sum_{i=1}^{n_k} Ln\left[\frac{\Gamma(h_{ik} + d)}{\Gamma\left(\frac{h_{ik}+1}{d}\right)}\left(\frac{\mu}{\mu+1}\right)^{d}(\mu+1)^{h_{ik}}\right] \tag{3.18}$$

where h_{ik} is the number of returns per order for a given customer i belonging to customer-channel segment k.

Return per order of a customer is related to her purchase behavior and hence I do not assume independence of these processes. Customers who order more often are more likely to return a product because of several stochastic unobserved processes as wrong items, wrong fits, defective units, and unsatisfactory product evaluation. To correct for this dependence bias, I condition the expected return per order in the prediction window on past purchase frequency and expected order frequency. The conditional expectation of the number of returns per order for consumers with at least one repeat purchase is given by:

$$E(I(y) \mid Y = y > 0; d, h, \mu, x, y) = \frac{(d+h)y}{\mu + x} \tag{3.19}$$

The conditional expectation of the number of returns per order for consumers with no repeat purchase is given by:

$$E(I(y) \mid Y = 0; d, \mu, x, y) = \frac{(1-\gamma)\left(\frac{dy}{\mu + x}\right)\left(\frac{\mu}{\mu + x}\right)^{d}}{\gamma + (1-\gamma)\left(\frac{\mu}{\mu + x}\right)^{d}} \tag{3.20}$$

where ‘I(y)’ is the conditional expectation of return per order for a customer with purchase frequency ‘y’ in the prediction window, ‘x’ is the purchase frequency of the customer in the estimation window, ‘h’ is the average return per order in estimation window, and γ is the proportion of consumers in the dataset with no repeat purchase.

Contribution Margin Model

Contribution margin per item can be modeled at a customer level, order level and an item level. Studies in marketing have focused on the first two levels. Customer level approaches have used hierarchical models to capture the effect of firm-specific marketing intervention and customer-specific shopping traits. Venkatesan and Kumar (2004) use a panel data regression with lagged contribution margin to correct for model misspecification to capture the contribution margin generated by a customer at the order level. Fader, Hardie and Lee (2005b) use a “regression to the mean” approach to capture the monetary value at the order level. Their modeling approach is superior to previous models because it incorporates heterogeneity across multiple orders for a customer and heterogeneity in average monetary value across customers. However, they model contribution margin at the order level and do not incorporate covariates in their model. I extend their model to incorporate the effect of marketing efforts and model the contribution margin per item.

I extend the Gamma-Gamma model used by Fader, Hardie and Lee (2005b) to capture the effect of marketing efforts on average contribution margin per item for a given customer. If a customer orders ‘x’ number of times in a given window, where the size of each order is ‘f’, the total number of items ordered by a customer are x*f. If w_{11}, w_{12}, w_{21}……., w_{lj} are the contribution margins of the f^{th} item bought by a customer on the x^{th} purchase occasion, the average contribution margin per item for that customer is given by:

$$w_{xf} = \frac{\sum_{l=1, j=1}^{x, f} w_{lj}}{x \times f} \tag{3.21}$$

My goal is to first test the effect of marketing efforts on w_{xf} and then to predict the expected value of the contribution margin per item for a given customer in the holdout window. If I know exactly which item a customer is going to buy in a prediction window, the prediction of average contribution margin per item for that customer is not required. However, for firms marketing multiple product categories with hundreds of SKUs in each category, it is almost impossible to predict customer’s exact item choice in a prediction window. Alternatively, I predict the average contribution margin per item for a given customer with the assumption that as x*f tends to infinity, w_{xf} tends to the expected value of the contribution margin per item for a given customer. Because this approach would require observation of customer behavior for an extremely long time window, I assume that there is heterogeneity in contribution margin across the items ordered by a customer and also across customers. The following are the assumptions of the Gamma-Gamma model.

a) If w_{11}, w_{12}, w_{21}......., w_{lj} are gamma distributed with shape parameter 'p' and scale parameter ν, the average value of the contribution margin per item for a given customer (w_{xf}) will be gamma distributed with shape parameter p*x*f and scale parameter ν*x*f.

b) To account for heterogeneity in the value of w_{xf} across customers, the model assumes that ν is distributed across customers according to a gamma distribution with shape parameter 'q' and scale parameter θ.

c) If 'p' is assumed to be constant across customers, the joint marginal distribution of w_{xf} will be distributed with shape parameters 'p' and 'q', and scale parameter θ.

I introduce covariates in this "regression to the mean" approach model as a linear form of the scale parameter θ of the joint marginal distribution in the following form.

$$\theta_{ik} = \omega z'_{ik} \tag{3.22}$$

where ω are the response parameters of the marketing efforts z'_{ik}. I introduce two marketing covariates into the model: the number of marketing mailers and discount offered per item. Although I include price in the purchase frequency and purchase quantity per order model as a covariate, I do not introduce it as a covariate in the contribution margin per item model. If a firm does not engage in dynamic pricing, the contribution margin and the price of a product are linearly related. Hence introduction of price as a covariate in the model would lead to a nearly perfect prediction of the contribution margin by price. The likelihood function of the marginal distribution with the effect of covariates for customer-channel segment k is given by:

$$LL_k(p,q,\omega \mid w_{ikxf}, x_{ik}, f_{ik}) = \sum_{i=1}^{n_k} Ln\left[\frac{\Gamma(px_{ik}f_{ik}+q)}{\Gamma(px_{ik}f_{ik})-\Gamma(q)} \frac{(\omega z_{ik}^{'})^q (w_{ikxf})^{(px_{ik}f_{ik}-1)} (x \times f_{ik})^{(px_{ik}f_{ik})}}{(\omega z_{ik}^{'} + w_{ikxf}x_{ik}f_{ik})^{(px_{ik}f_{ik}+q)}} \right] \quad (3.23)$$

where 'w_{ikxf}' is the average contribution margin per item for customer i belonging to customer-channel segment k with 'x_{ik}' orders and 'f_{ik}' purchase quantity per order. To optimize marketing efforts in a future time period 't', I need to predict the conditional expected contribution margin per item of a randomly chosen individual. The conditional expectation of contribution margin per item for a customer given the model parameters, response parameters, observed contribution margin per item, purchase frequency and purchase quantity per order for the customer in estimation window is given by (see Fader, Hardie, and Lee 2005b for details):

$$E(U \mid p,q,\omega,x,f,w_{xf}) = \left(\frac{q-1}{Pxf+q-1}\right)\left(\frac{\omega z^{''} p}{q-1}\right) + \left(\frac{pxf}{pxf+q-1}\right) w_{xf} \quad (3.24)$$

where 'U' is the expected value of contribution margin per item for a given customer. Note that the covariates only influence the population mean (variation across customer) and not the mean value of margin per item for a given customer (variation across items for a given customer). The value of 'U' in the above equation is the weighted average of population mean and the mean value of contribution margin per item for a given customer (w_{xf}). The terms $q-1$ and $p*x*f$ divided by their sums are the weights placed on the population mean and the mean of value of margin per item for a given customer, respectively.

In the prediction window of duration 't,' for a given customer with purchase frequency 'x', purchase quantity per order 'f', items returned per order 'h', and contribution margin per item 'w_{xf}', the total contribution margin derived from that customer will be (f – h) * x * w_{xf} . The response parameters of this customer to marketing effects in purchase frequency, purchase quantity per order, and contribution margin per item model are given by β , η and ω , respectively. For the same customer in the prediction window of duration 't' the purchase frequency, purchase quantity per order, items returned per order, and expected contribution margin are given by 'Y', 'G', 'I', and 'U' respectively. The total margin derived from this customer in the prediction window is: (G – I) * Y * U. These values in the prediction window are the customer's responses to the firm's marketing efforts in the prediction window and the shape parameters of each model as estimated in the estimation window.

DATA

I estimate the models using customer transaction data obtained from a large shoe and accessories manufacturing and marketing firm. The firm operates in the high-end market of the product categories and has been in this business for almost a century. The firm markets its products through physical stores (bricks), the Web (clicks), and catalogs (flicks). The transaction data used in this study are from customers to whom the firm sells through its own retail network. The customer response file begins on January 1, 2003 and ends on August 7, 2005. The firm operates two different types of direct marketing campaigns. In addition to mailing 10 catalogs a year, the firm also offers promotional discounts semi-annually. The promotional flyers for these promotional discount campaigns are mailed to customers and prospects before the beginning of the sales period.[9] The prices of products in the promotion period are marked down by a fixed percentage. This information on the firm's marketing efforts is available at the individual customer level from January 1, 2004 until August 7, 2005. The sales promotion campaigns are implemented by the firm uniformly across all three channels. The prices of products are consistent across the channels during both the promotion period and the non-promotion period. A customer is tracked in the database by a unique customer identification key assigned to her when she makes her first purchase with the firm. The database tracks customers' choices of the channel, the SKU purchased, the price paid, the date of transaction, and the date of return for every transaction. Each SKU bought by a customer constitutes one record in the customer response file. The product level file contains information on the product category, the SKU, the retail price, and the cost. The customer response file, marketing information file, and the product category file together constitute the dataset used in this study.

Of the 135 weeks of data, only 84 weeks beginning January 1, 2004 contain information about direct marketing efforts. I use the first 56 weeks of data as the estimation sample and the following 28 weeks of data as the holdout sample for testing the predictive validity of the models. I identify the cohort of first time buyers in the estimation window and use them for the estimation and prediction samples. There are 213,142 new customers with over half a million records in the estimation window. The discount offered to a customer for a given item is calculated as the retail price of the item

[9] Because marketing mailers are sent out during holidays or seasonal time frame, seasonality and marketing mailers are highly correlated. Because marketing mailers are a key part of my model, I do not include seasonality in the model.

minus the dollar amount paid for that item by the customer. Similarly, margin is calculated as the dollar amount paid for an item less the cost of the item. The price paid for a given SKU by customers will differ in the promotion and the non-promotion periods, thereby bringing variation in the discount variable in the data (see Equation 3.5).[10]

The summary statistics on some of the key variables of the data are shown in Table 3.1. A cursory look at the table suggests that multichannel customers outspend single channel customers by a factor of at least two. A similar pattern is evident for purchase frequency and total number of items bought. Store orders are likely to be larger orders than web and catalog orders. While customer buying in a store may have larger order sizes, customers shopping on the web and through the catalog purchase the items with a larger contribution margin. The 'web only' and 'catalog only' customers have smaller order sizes, but more than make up for this drawback by ordering more expensive items from the firm's portfolio. The number of items returned per order is high for customers purchasing through the direct channels. The marketing efforts expended by the firm seem to be very evenly distributed across the customer-channel segments. Multichannel and 'store only' customers' exhibit the highest tendency to shop across multiple categories. This trend in return per order is in line with that reported by other direct marketers.

[10] Data on shipping costs, which are relevant for the catalog and Web channels, and which vary by customer location, are not available. Because they are customer-transaction specific and are not decision variable under the firm's control, their omission from the empirical analysis is not a serious issue.

Table 3.1
Means and Standard Deviations of Key Variables in the Data

	Catalog Only	Store Only	Web Only	Multichannel	All Customers
Cohort Size	13628	179612	13500	6402	213142
Total Spending ($)	266.83	261.21	252.45	580.66	271.93
	(256.45)	(316.04)	(316.20)	(467.02)	(324.52)
Total Margin ($)	175.44	163.09	165.43	368.09	170.95
	(186.42)	(199.26)	(262.44)	(294.73)	(209.40)
Number of Orders	1.41	1.36	1.33	2.92	1.42
	(.78)	(1.03)	(.68)	(1.37)	(1.04)
Total Items	2.45	3.26	2.29	6.28	3.26
	(2.63)	(4.48)	(2.25)	(5.32)	(4.35)
Number of Categories Bought	1.24	1.54	1.22	1.86	1.51
	(.48)	(.63)	(.48)	(.73)	(.62)
Order Size	1.70	2.37	1.67	2.11	2.28
	(1.09)	(1.31)	(.91)	(.39)	(1.30)
Return Per Order	.22	.08	.16	.20	.10
	(.54)	(.36)	(.47)	(.46)	(.39)
Spending Per Item ($)	119.32	96.35	125.6	104.88	99.68
	(75.53)	(63.66)	(217.72)	(54.62)	(83.25)
Margin Per Item ($)	79.27	59.97	83.52	66.72	62.70
	(48.44)	(41.11)	(211.54)	(35.58)	(66.98)
Discount per Item ($)	2.70	3.76	4.58	3.24	3.73
	(10.50)	(9.65)	(17.70)	(6.73)	(10.33)
No. of Mailers	6.88	4.54	6.20	6.68	4.86
	(3.33)	(3.12)	(3.44)	(3.32)	(3.25)

Note: Standard errors in the parentheses.

MODEL ESTIMATION AND RESULTS

I estimate the models using the maximum likelihood estimation method. This flexible approach allows the specification of user defined likelihood function, places constraints on the lower bounds of the shape and scale parameters, and selects appropriate starting values. I use an improvement in the log likelihood function by 1.0e-03 for at least two consecutive steps as the convergence criteria for the models. I gave different starting values to the model parameters and checked if the algorithm converged to same value of parameters and same value for the log likelihood function. For obtaining the standard errors of the parameter estimates, I derived the Hessian matrix of each model. The Hessian matrix is the second derivative of the log likelihood function with respect to each parameter. I used the inverse of Hessian to calculate the information matrix. The square roots of the vector of the diagonal elements of this information matrix are the standard errors associated with the parameter estimates.

I also estimated the 2F1 Gauss hyper geometric function in purchase frequency model. Equation 3.13 was used to evaluate the terms of hypergeometric series while Equation 3.12 was used to compute the value of the function for each customer. The first 500 terms of the series were evaluated. Though the machine on which the function was evaluated had a machine epsilon of 1.0e-300, the terms converged to zero much before the first 500 terms, thus only the first 500 terms were used in the calculation of hyper geometric function.

Purchase Frequency Model Results

The results of purchase frequency model are presented in Table 3.2. The effect of marketing mailers on customer purchase frequency is positive and significant for each customer-channel segment. The results indicate that the 'multichannel' (.41, $p < .001$) and the 'catalog only' (.24, $p < .001$) customers are most responsive to marketing mailers. Thus a 'multichannel' ('catalog only') customer who receives a marketing mailer is 50% (30%) more likely to order than a customer who does not receive the marketing mailer. The effect of marketing mailers on purchase frequency of the 'store only' (.06, $p < .001$) and the 'web only' (.06, $p < .001$) customers is positive and significant. However, a ‘store only’ (‘web only’) customer receiving a mailer is only 6% (5%) more likely to order than a customer in that segment who does not receive the mailer.

Table 3.2
Results of Purchase Frequency Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
r	.214***	.205***	.219***	.148**
	(.020)	(.004)	(.020)	(.049)
a	3.92e-05	5.92e-08	1.01E-08	2.39e-06
	(4.77)	(.010)	(.160)	(3.600)
b	2.14	2.759	7.813	7.27e-05
	(8809.650)	(45.697)	(3295.650)	(1613.860)
Intercept	3.762***	3.207***	3.731***	9.950***
	(.060)	(.016)	(.060)	(.030)
Mailers	.235***	.061***	.057***	.409***
	(.028)	(.010)	(.010)	(.000)
Price	-.054	.198***	-.234***	-.140**
	(.040)	(.015)	(.030)	(.050)
Discount	.551	1.522***	.363	.212
	(.770)	(.208)	(.380)	(.430)
Log-Likelihood	-28248.03	-422189.35	-24564.50	-54708.09

Note: Standard errors in parentheses. * significant at 5%, ** significant at 1%, *** significant at .1%

The effect of the prices of SKUs bought by customers on their purchase frequency is generally in the expected direction. The anticipated spending on planned purchases by a consumer is endogenous to her disposable income. Thus, the relationship between price of SKUs bought (budget) and purchase occasions is likely to be negative. The ‘web only’ customers exhibit the highest degree of price sensitivity (-.24, $p < .001$). Previous studies have shown that customers who order through the web buy less expensive items to mitigate the risk of using a direct channel. Although a similar effect of price on purchase frequency of a ‘catalog only’ customer is negative, it is not significant (-.05, $p > .05$). One of the surprising findings is the positive relationship between the average price of an item bought by a customer and the purchase frequency of 'store only' customers (.20, $p < .001$). This firm markets to the high-end customer segment. Because of their higher disposable incomes, middle-aged customers are more likely to belong to this segment than are other customers. Previous research suggests that older customers prefer traditional physical stores over direct channels. Thus, the customer self-selection process could possibly explain this positive relationship. In other words, customers who come for repeat purchase are the ones who are inherently more likely to purchase than are other customers and thus end up making several repeat purchases. The effect of price on purchase frequency of 'multichannel' customers (-.14, $p < .01$) is in the expected direction.

The results also suggest that offering discounts to a 'store only' customer drives their purchase frequency higher (1.52, $p < .001$). The odds ratio of discount on purchase frequency for a 'store only' customer is extremely high (4.5). The results do not indicate a similar effect of discount on purchase frequency of customers belonging to other customer-channel segments.

Purchase Quantity Model Results

The results of the purchase quantity model indicate that marketing mailers positively influence the purchase quantity per order of customers belonging to the 'catalog only', the 'web only' and the 'multichannel' segments. Multichannel customers receiving marketing mailers are 93% more likely to have larger orders than those customers who do not receive the mailers (.66, $p < .001$). I find a similar influence of marketing mailers on the purchase frequency of multichannel customers. This finding indicates that in response to a firm's marketing efforts, multichannel customers not only purchase more frequently, but also order more items per purchase occasion than do single channel customers. The effect sizes of marketing mailers on purchase quantity per order of 'catalog only' customers is also large and significant (.41, $p < .001$). The results of the model are reported in Table 3.3.

Table 3.3
Results of Purchase Quantity Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
c	2.079***	6.027***	3.298***	312.026***
	(.051)	(.005)	(.048)	(.794)
Intercept	.872***	3.051***	1.943***	10.399***
	(.059)	(.020)	(.067)	(.026)
Mailers	.406***	.074	.055***	.659***
	(.052)	(.042)	(.019)	(.026)
Price	-.017***	-.034***	-.021**	-.024***
	(.008)	(.004)	(.009)	(.004)
Discount	.799	1.747***	.465	.708
	(.699)	(.256)	(.377)	(.437)
Log-Likelihood	-11861.55	-166188.84	-10758.61	-10444.30

Note: Standard errors in parentheses. * significant at 5%, ** significant at 1%, *** significant at .1%

Price has a negative relationship with purchase quantity per order for 'store only', 'catalog only' and 'web only' customers. This result indicates that when customers order more expensive items, they purchase fewer items on that purchase occasion. This relationship is strongest for the 'store only' customer segment (-.03, $p < .001$). Evaluating this result in conjunction with the results of the purchase frequency model suggests that 'store only' customers purchase more expensive items more frequently, but buy fewer items on each purchase occasion than do other customers. However, other customer-channel segments exhibit higher price sensitivities. When purchasing more expensive items, ‘catalog only’, ‘web only’, and ‘multichannel’ customers not only purchase less frequently, but also have smaller order sizes on each purchase occasion than do other customers.

The effect of discount on purchase quantity per order is similar to its effect on purchase frequency. The ‘store only’ customer-channel segment is highly responsive to discounts. Other customer-channel segments, except ‘catalog only’ customers, also have positive response coefficients to discounts, but these coefficients are not statistically.

Product Return Propensity Model Results

The shape and scale parameters of the Conditional NBD model for product return propensity per order are presented in Table 3.4. All the parameter estimates are significant and the model fits are good. Therefore, the predicted values should help accurately estimate the net purchases in the prediction window.

Table 3.4
Results of Product Return Propensity Model

Parameters	Catalog Only	Store Only	Web Only	Multichannel
d	1.381***	.686***	.731***	290.483***
	(.135)	(.021)	(.052)	(13.849)
μ	9.567***	8.221***	4.234***	4164.948***
	(.546)	(.202)	(.308)	(6.200)
Log-Likelihood	-8463.85	-61442.02	-6608.67	-3657.39

Note: Standard errors in parentheses. * significant at 5%, ** significant at 1%, *** significant at .1%

Contribution Margin per Item Model Results

The results of the contribution margin model suggests a positive relationship between number of marketing mailers received and the contribution margin of an average item bought by ‘store only’ (3.25, $p < .001$) and ‘catalog only’ (.08, $p < .01$) customers. Thus, ‘store only’ and ‘catalog only’ customers receiving marketing mailers are more likely to purchase items with higher contribution margins than those who do not receive the mailers. The extremely large effect of marketing mailer on the contribution margin of items bought by ‘store only’ customers could be partly attributed to store-specific unobserved effects. Store-specific unobservables such as personal selling efforts and customer interactions with the sales staff could be providing an extra impetus for the ‘store only’ customer to make purchases of more expensive items. The results of the model are presented in Table 3.5.

Table 3.5
Results of Contribution Margin Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
p	.979***	.991***	1.875***	1.989***
	(.037)	(.006)	(.045)	(.081)
q	5.713***	6.588***	5.083***	5.189***
	(.094)	(.032)	(.083)	(.092)
Intercept	163.513***	232.447***	110.995***	63.684***
	(2.908)	(1.302)	(2.010)	(1.261)
Mailers	.082**	3.255***	.479	-.147
	(.038)	(.243)	(.290)	(.169)
Discount	-31.438	21.457***	-11.635	-6.359
	(22.417)	(1.958)	(8.361)	(18.875)
Log-Likelihood	-71819.25	-910465.90	-70512.53	-31075.02

Note: Standard errors in parentheses. * significant at 5%, ** significant at 1%, *** significant at .1%

The results also suggest a positive relationship between the amount of discount offered and the contribution margin per item for the ‘store only’ customer-channel segment (21.46, $p < .001$). These weakly significant results for the contribution margin per item model can be explained by the fact that the categories marketed by the firm are lifestyle products, so customers may have strong preferences and loyalty toward specific products.

Optimization Model Results

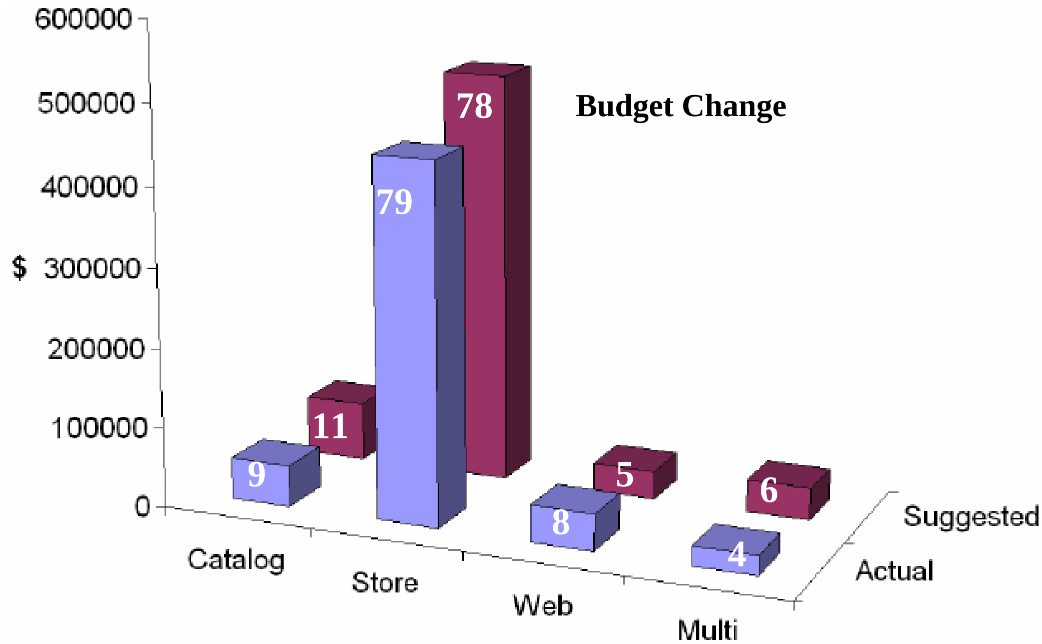
The results of the optimization model are presented in Table 3.6 and Figure 3.2. The results suggest that if the allocated budget for this cohort of customers is increased by 16%, the profits from the entire cohort of customers can be improved by as much as 32%. The results suggest that more marketing resources should be allocated to ‘catalog only’ and ‘multichannel’ customer segments. Currently, 9% and 4% of the budget are allocated to the ‘catalog only’ and ‘multichannel’ segments, respectively. The optimization results suggest that these figures should be increased to 11% and 6% of the budget, respectively. This reallocation represents an increase of 42% and 70% in dollar values of budgets for the ‘catalog only’ and ‘multichannel’ segments, respectively. The results of the optimization model are representative of the responsiveness of these segments to marketing mailers of the firm, especially with respect to their purchase frequency and purchase quantity. The results in Tables 3.2 and 3.3 suggest that of all the customer-channel segments, these two customer-channel segments are most responsive to a firm’s marketing mailers.

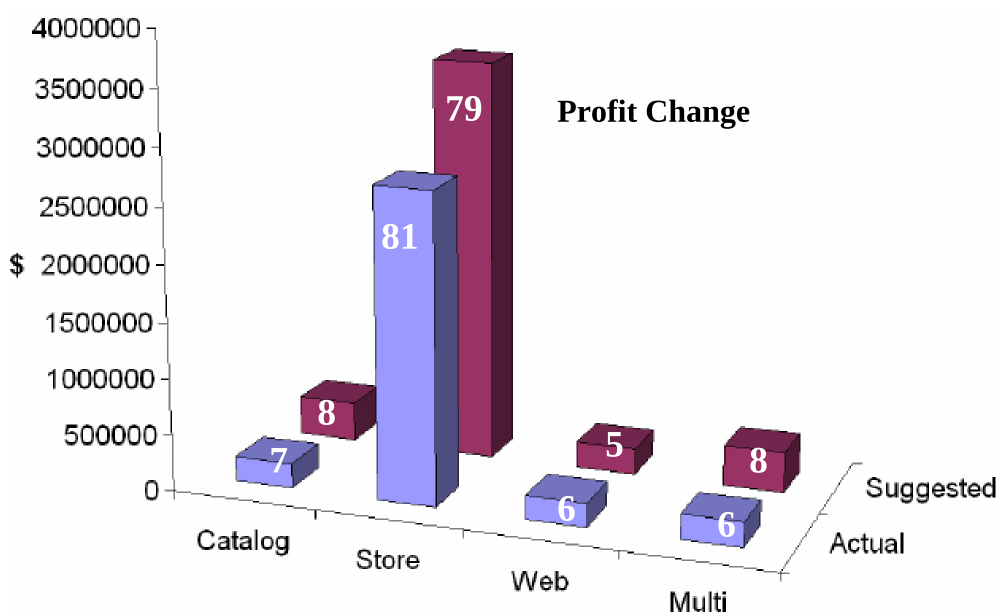
Table 3.6
Optimization Results

Segment	Budget			Profits		
	$ Value	% of Total	% Change Over Actual	$ Value	% of Total	% Change Over Actual
Actual						
Catalog	51,538	9.05		223,828	6.58	
Store	448,486	78.73		2,743,031	80.69	
Web	46,035	8.08		210,695	6.20	
Multichannel	23,584	4.14		222,116	6.53	
Total	569,643	100.00		3,399,672	100.00	
Suggested						
Catalog	73,455	11.09	42.4	346,680	7.71	54.9
Store	513,690	77.58	14.5	3,554,341	79.07	29.6
Web	34,898	5.27	-24.2	233,381	5.19	10.8
Multichannel	40,141	6.06	70.1	360,546	8.02	62.3
Total	662,183	100.00	16.2	4,494,949	100.00	32.2

The optimization results also suggest that budget allocation to the ‘web only’ segment should be reduced from 8% to 5% of the total budget. This decline is also reflected in the decline in absolute dollar value (24%) allocated to the customer-channel segment. For the ‘store only’ segment, the budget allocated should be reduced from 79% to 78% of the total marketing mailer budget. However, in absolute number, the dollar amount allocated to this segment increases by 15%. The budget increase of 16% in dollar value outweighs the decline in proportional allocation to this segment. The suggested allocation for these segments are reflective of their responsiveness to the firm’s marketing mailers as reflected in results presented in Tables 3.2, 3.3, and 3.5.

Figure 3.2
Optimization Results





Note: Numbers on bar reflect percentage of total across all customer-channel segments

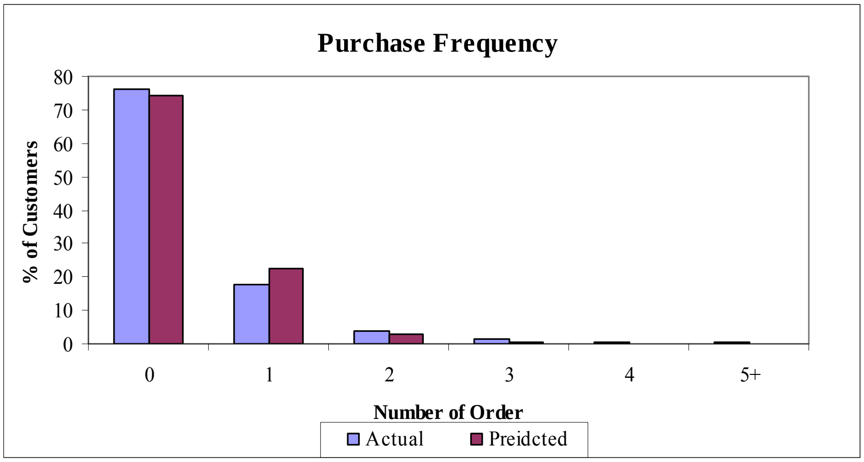
The optimization results presented in Figure 3.2 suggest that in response to the recommended increase in marketing budget, the profits from this cohort of customers over the 28-week prediction window would increase from $3.4 million to $4.5 million. This increase in profits reflects the extraction of an additional $5 from each customer from this cohort. This increase, when extrapolated over a 52-week period for the entire customer base, translates to profits of over $7.75 million. The share of profits contributed by ‘catalog only’ and ‘multichannel’ customers increases from 7% by each segment to 8% by each segment. This reflects a 55% and 62% improvement in the profits from the ‘catalog only’ and ‘multichannel’ segments, respectively. Although the profits from the ‘store only’ and ‘web only’ segments increases by 30% and 11%, respectively, the contributions from these segments to the firm’s profits declines to 79% and 5%, respectively. The marginal increase in the profits from the ‘web only’ segment is primarily due to a decline in the budget allocated to the segment. The results also indicate that for every one dollar increase in marketing budget, the additional profits contributed by the ‘store only’, ‘multichannel’ and ‘catalog only’ segments are $12, $8, and $6, respectively.

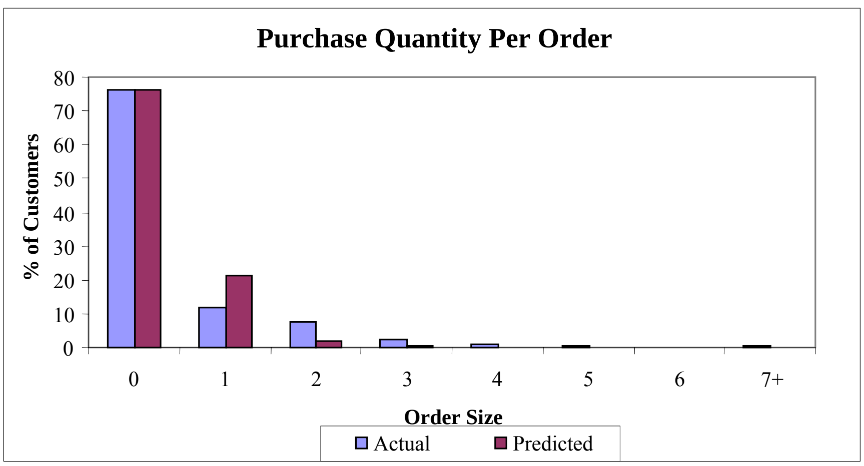
MODEL VALIDATION

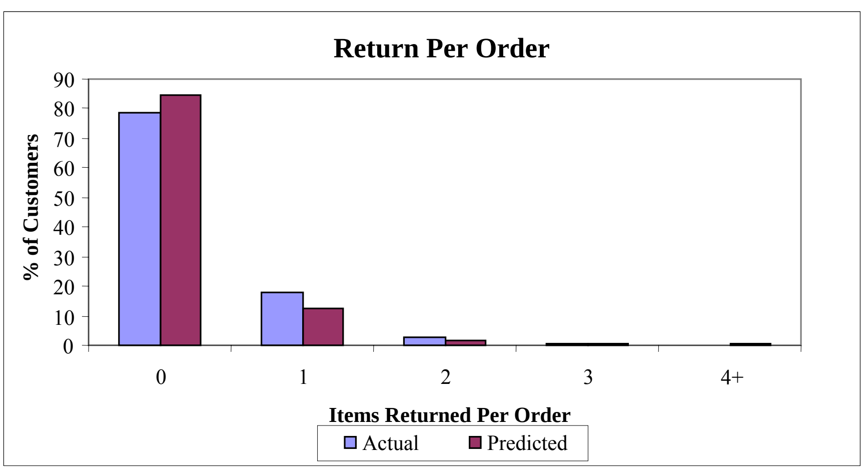
Predictive Validation ‘Catalog Only’ Segment

For ‘catalog only’ customers, charts of the predicted and actual values of components of behavior are presented in Figure 3.3. For purchase the frequency model,

Figure 3.3
Predictive Validity for Catalog Only Customer-Channel Segment





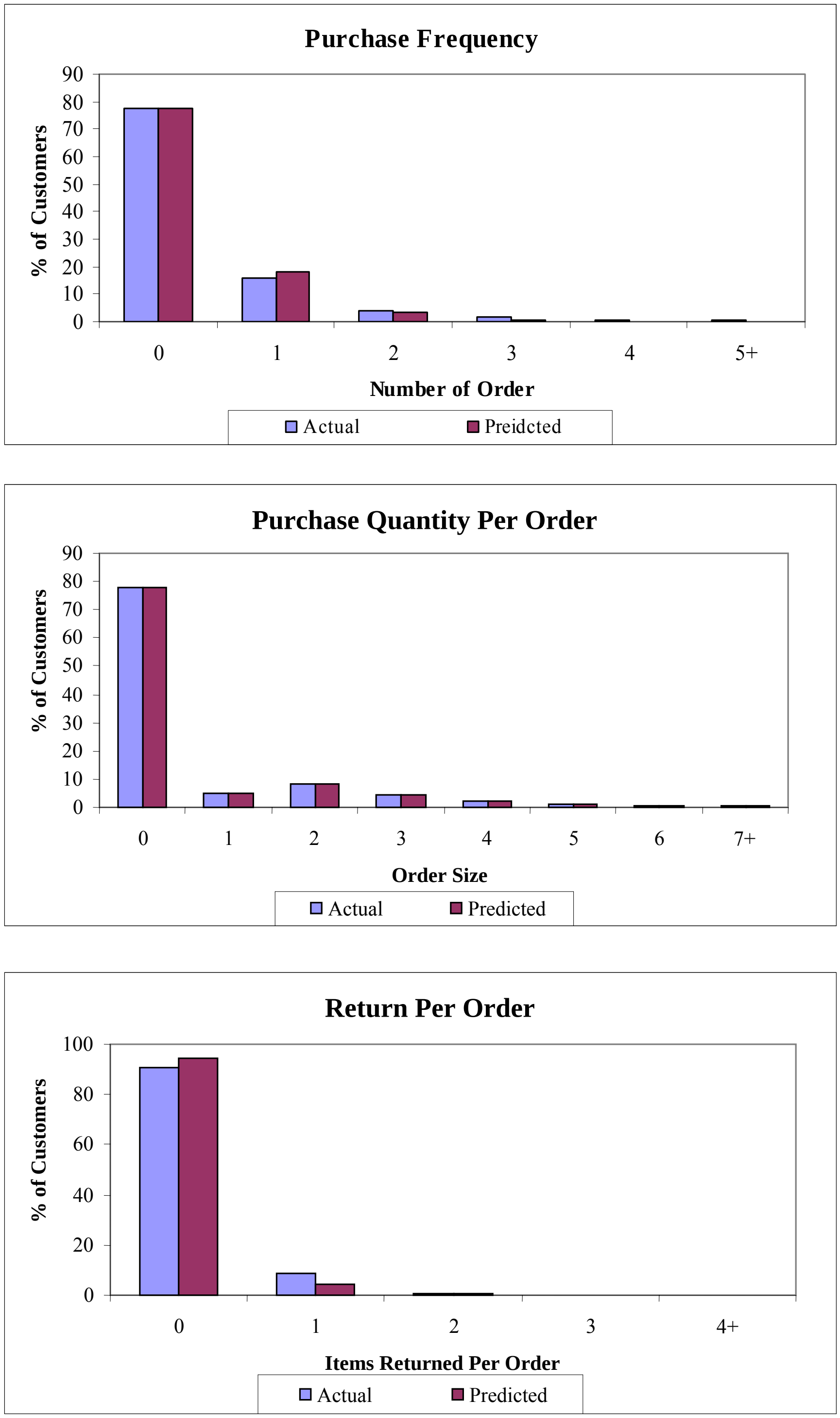


the beta geometric part of the model which estimates customer drop out rates, performs very well. The NBD part of the model fits well for customers making two or more purchases. The predicted value of customers making only one repeat purchase overestimates the actual figure by 4%. This modest underestimation is also carried forward in the purchase quantity per order model because the predicted purchase quantity per order is conditional on the predicted purchase frequency. This error combined with the stochastic error of the purchase quantity per order model swells the total error to 10%. In the model capturing product returns per order, customers who are not likely to return are optimistically represented, thereby under-representing the customers who are more likely to return an item for every order they place. Toward the right tail of the distribution, the predicted values are very close to the actual values. The mean (standard deviation) of margin per item as predicted by the contribution margin per item model is $77.91 ($22.72) and compares well to the actual value of $79.49 ($49.09). The predicted and the actual values are within 2% of each other.

Predictive Validation of ‘Store Only’ Segment

The model fits of the ‘store only’ customer-channel segment are extremely good as shown in Figure 3.4. The purchase frequency model performs very well in predicting the customer dropout rate. The NBD part of the model also captures the data very well and predicts the purchase frequency of customers within 2% of the actual values. The model fit of the purchase quantity per order model is also excellent with the prediction results within half a percent of the actual values across the distribution. The prediction results of items returned per order model are also on the optimistic side. It over

Figure 3.4
Predictive Validity for Store Only Customer-channel Segment

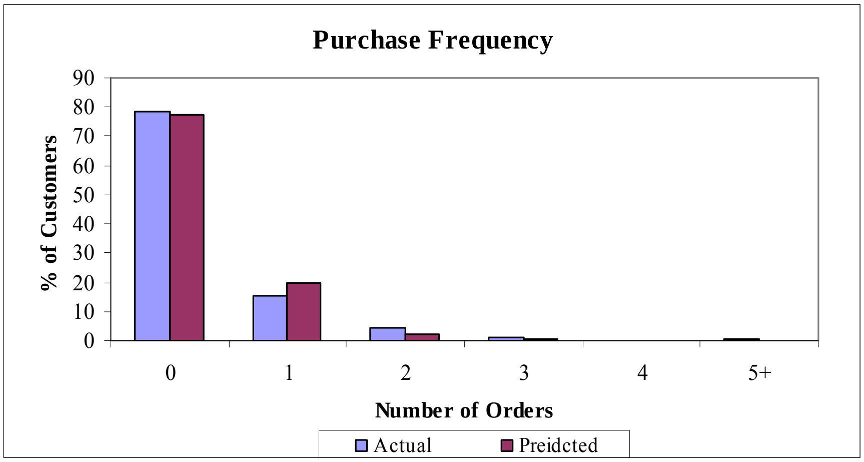


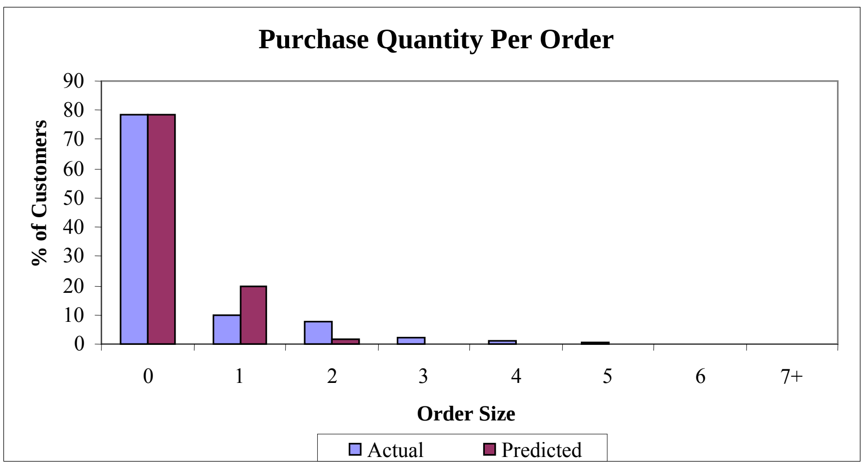
represents the percentage of customers who never return an item for every order they place by 3%. The mean (standard deviation) of margin per item predicted by contribution margin per item model is \$55.27 (\$13.62) compared to the actual value of \$60.84 (\$40.95). This under representation is likely because of large variance in the value of margin per item for ‘store only’ customers. This finding is evident from the fact that the shape parameter (q) of the gamma distribution, which captures heterogeneity across customers for this segment, is the largest among all customer-channel segments.

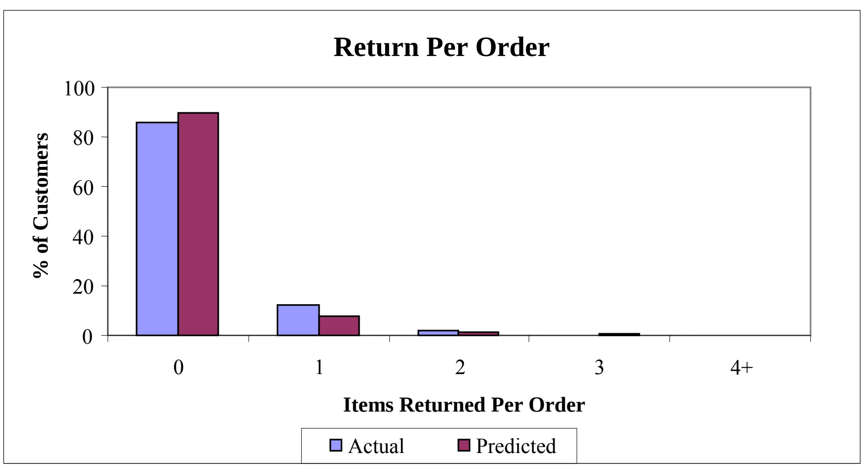
Predictive Validation of ‘Web Only’ Segment

The model fits of the ‘web only’ customer-channel segment, as shown in Figure 3.5, are very similar to those for the ‘catalog only’ customer-channel segment. The inherent similarity in behavior across customers who use only direct channels is evident. A maximum deviation of 4% is observed between the actual and predicted values of purchase frequency for customers making only one repeat purchase. This number rises to 9% for the purchase quantity per order model. The items returned per order model predicts fewer people who are likely to return an item for every order they place. The ‘web only’ customer-channel segment purchases items with the highest contribution margin per item. The mean (standard deviation) of the contribution margin per item predicted by the contribution margin per item model is \$77.96 (\$55.27) compared to the actual value of \$81.32 (\$130.74). The predicted and actual values are close to each other.

Figure 3.5
Predictive Validity for Web Only Customer-channel Segment



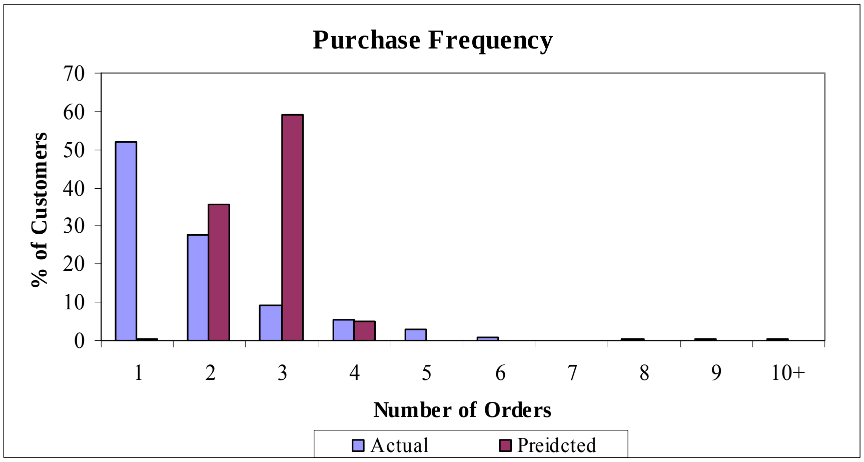


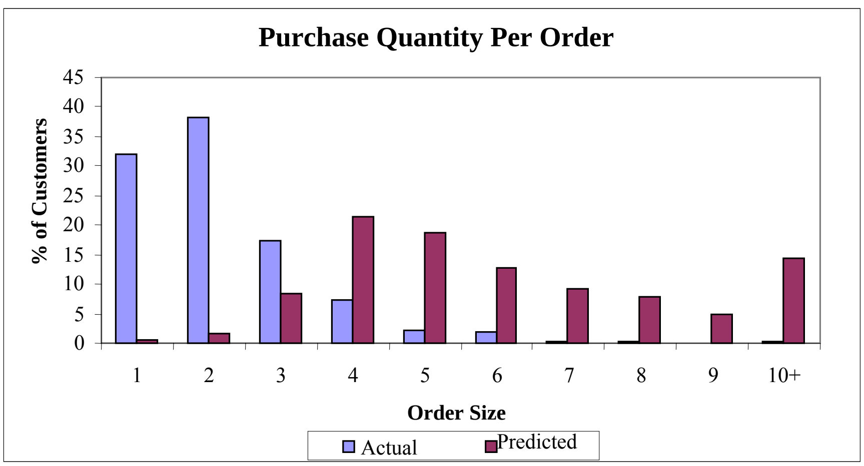


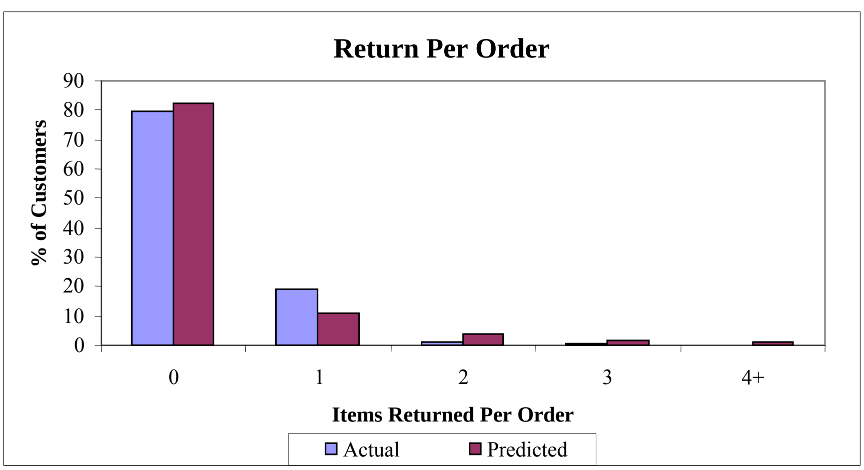
Predictive Validation of Multichannel Segment

The results for the multichannel customer-channel segment appear in Figure 3.6. The model fits of the purchase frequency model and the purchase quantity per order model are moderate. The BG/NBD model used for estimating purchase frequency assumes that customer drop out will occur at some point. However, multichannel customers are those customers who have ordered at least twice using two different channels. Because every customer in the multichannel customer segment ordered more than once, the BG/NBD model does not capture the phenomenon as well. Similarly, the purchase quantity per order model which measures items purchased per order only in the repeat orders, does not capture the phenomenon adequately. The fit of the product return per order model is very good and the prediction accuracy is within 3% of the actual value. The contribution margin per item model also fits the data very well. The mean (standard deviation) value of the margin per item predicted by the contribution margin per item model is $64.85 ($25.86) compared to the actual values of $65.82 ($32.84). These values are extremely close to each other and the fit appears to be best among all customer-channel segments evaluated. The very low variance in the actual values of the margin per item helps in achieving a better fit for this dataset.

Figure 3.6
Predictive Validity for Multichannel Customer-channel Segment







IMPLICATIONS, LIMITATIONS, FUTURE RESEARCH AND CONCLUSION

Managerial Implications

The findings of the study have important managerial implications. An 'a priori' segmentation of customers based on their channel choices is theoretically and managerially relevant. The model can serve as a decision support tool for resource allocation decisions. The findings can help managers identify how much marketing efforts should be expended in each channel. The model is generalizable and can be implemented in variety of contexts because it is based on predicted future profits. By decomposing profits from the customer into multiple components, the model enables managers to identify the influence of marketing efforts on the customer's purchase frequency, purchase quantity, and contribution margin.

I performed a post-hoc analysis to decompose profits. Of the three average quantities predicted, namely, purchase frequency, purchase quantity, and contribution margin, I held two at their actual values, while I set the third at the predicted value for calculating the improvement in profits. If AcPF, AcPQ, AcCM, and PrPF, PrPQ, PrCM are the actual and predicted purchase frequencies, purchase quantities, and contribution margins, respectively, the actual firm profit is given by: AcPF*AcPQ*AcCM. The difference in the profits realized by replacing one quantity by its predicted counterpart is the dollar value of contributions made by the replaced quantity. Thus, the additional dollar worth of profit generated by the increase in purchase frequency as a response to increased mailing is given by (PrPF*AcPQ*AcCM - AcPF*AcPQ*AcCM). For the entire customer cohort, the results reveal that 43%, 33%, and 25% increases in profits are contributed by the improvements in contribution margin, purchase quantity, and purchase frequency, respectively, which translate into gains of $2.20, $1.68, and $1.27, respectively. Thus, through my modeling approach, I am able to differentiate between improvements in purchase quantity and contribution margin. Previous modeling approaches could only suggest that improvement in dollar value of an order would contribute an aggregate figure (78%) toward improvement in profits. The average profits generated by a customer of each segment and the decomposition of those profits are shown in Table 3.7.

Table 3.7
Profitability Decomposition

	Actual Profits ($)	Optimal Profits ($)	Change ($)	Contribution Margin (%)	Purchase Quantity (%)	Purchase Frequency (%)
Catalog	16.42	25.44	9.01	17.83	49.41	32.76
Store	15.27	19.79	4.52	46.72	31.40	21.89
Web	15.61	17.29	1.68	27.94	22.54	49.52
Multichannel	34.69	56.32	21.62	11.83	51.55	36.62
All customers	15.95	21.09	5.14	42.63	32.59	24.77

A similar decomposition of average profits per customer reveals some interesting insights. The increase in average profits is highest for ‘multichannel’ customers. Little over half (52%) and one third (37%) of this increase are because they purchase more items on each purchase occasion and purchase more frequently. The decomposition for the ‘catalog only’ customer also reveals similar results. However, for ‘store only’ customers, I find that an increase in profits is primarily contributed by an increase in the contribution margin. Thus, 'store only' customers respond differently to marketing mailers than do 'catalog only' and 'multichannel' customers. The 'web only' customers' responsiveness to marketing mailers is fairly low. The improvement in profits from this segment is moderate ($1.68). Almost half of this improvement is primarily from a higher purchase frequency.

Limitations and Future Research

The research has certain limitations that could be addressed by future research. First, although I test for the endogeneity of customer channel choice is not a serious problem in this data based on tests, channel choice could be a function of marketing effort in other contexts. Models developed for such contexts capture this possibility. Second, I treated segments of all combinations of channel choices to be one multichannel segment. Future research could examine if there are any differences in response to marketing efforts across different types of multichannel segments (e.g., bricks and clicks, clicks and flicks, and bricks and flicks). Third, I focused on marketing mailer as the decision variable because that was the most important decision variable to the company that provided the data. Future research could extend the decision variables to price and discounts.

Conclusion

In conclusion, I addressed three important research questions and proposed a model for the optimal allocation of marketing efforts across multiple customer-channel segments. I developed and estimated a set of marketing covariate-driven stochastic models for purchase frequency, purchase quantity, product return propensity, and contribution margin for each customer-channel segment. Based on the parameter estimates from the models, I then derived the optimal marketing effort allocation to each customer-channel segment for future holdout period using simulations. My results show that the responsiveness to marketing efforts of the firm varies substantially across the different components of customer behavior and across the different customer-channel segments. My model shows that firms can improve their profits by as much as 32% by optimally allocating their marketing resources to these customer-channel segments based on the heterogeneous responses of these segments to the firm's marketing efforts.

CHAPTER IV

SUMMARY

In the two essays of this dissertation, I seek answers to important research and managerial questions relating to the growing multichannel shopping phenomenon. Who are multichannel shoppers? What drives multichannel shopping behavior? How valuable are multichannel shoppers to the firm? How responsive are multichannel customers to a firm's marketing efforts? How should resources be allocated to different customer-channel segments? I address these questions by developing an integrated framework of multichannel marketing that explains the drivers and consequences of multichannel shopping behavior. Based on this framework, I develop an optimization model to allocate marketing resources across the different customer-channel segments to maximizes firm profitability.

I expect the findings from this dissertation to guide managers in making more informed decisions with respect to their multichannel marketing strategies. In particular, the findings will help managers identify and target the high value customers and allocate their marketing resources by channel. The results from the first essay show that multichannel customers are more valuable than single channel customers; demographic characteristics and shopping traits of these multichannel customers differ significantly from single channel only customers; and these effects may differ depending upon the product category under consideration. These findings can be used by mangers to identify multichannel customers and target appropriate product categories and promotional efforts to them. The findings from the second essay suggest that catalog only and multichannel customer segments are most responsive to marketing communications, while web only and store only segments respond more to price and discounts, respectively. These findings will help managers target specific segments for marketing communication, price reduction and discounts selectively to customer-channel segments. In addition, the resource allocation model will help them with the optimal levels of marketing efforts for each customer-channel segments.

REFERENCES

Achabal, D.D., M. Badgett, J. Chu, and K. Kalyanam (2005), "Cross-Channel Optimization: A Strategic Roadmap for Multichannel Retailers", *White Paper*, IBM Consulting Services

Allenby, Greg, Robert Leone, and Lichung Jen (1999), "A Dynamic Model of Purchase Timing with Application to Direct Marketing", *Journal of the American Statistical Association*, 94(446), 365-374.

Ansari, Asim, Carl Mela, and Scott Neslin (2007), "Customer Channel Migration," *Journal of Marketing Research*, forthcoming.

Assael, Henry (1998), *Consumer Behavior and Marketing Actions*, Sixth Edition, Cincinnati, OH: South-Western College Publication.

Balasubramanian, Sridhar, Rajagopal Raghunathan, and Vijay Mahajan (2005), "Consumers in a Multi-channel Environment: Product Utility, Process Utility, and Channel Choice," *Journal of Interactive Marketing*, 19 (2), 12-30.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, and Glen L. Urban (2005), "Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study," *Journal of Marketing*, 69 (4), 133-152.

Bell, David R., Teck-Hua Ho and Christopher Tang (1998), "Determining Where to Shop: Fixed and Variable Costs of Shopping", *Journal of Marketing Research*, 35 (3), 352-369.

______ and James M. Lattin (1998), "Shopping Behavior and Consumer Preferences for Retail Price Format: Why Large Basket Shoppers Prefer EDLP", *Marketing Science*, 17 (1), 66-88.

Berger, Paul D., Ruth N. Bolton, Douglas Bowman, Elten Briggs, V. Kumar, A. Parasuraman, and Creed Terry (2002), "Marketing Actions and Value of Customer Assets: A Framework for Customer Asset Management", *Journal of Services Research*, 5 (1), 39-54.

Bhatnagar, Amit, and Brian T. Ratchford (2004), "A Model of Retail Format Competition for Non-Durable Goods", *International Journal of Research in Marketing*, 21, 39-59.

Bitran Gabriel R. and Susana V. Mondschein (1996), "Mailing Decisions in Catalog Sales Industry", *Management Science*, 42 (9), 1364-1381.

Blattberg, Robert C. and John Deighton (1996), "Manage Marketing by the Customer Equity Test," *Harvard Business Review*, 74 (4), 136-144.

Bolton, Ruth N. and Crina O. Tarasi (2006), "Managing Customer Relationships," *Review of Marketing Research*, 3, 3-38.

Bound, John, David A. Jaeger, and Regina M. Baker (1995), "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association*, 90 (430), 443-450.

Bult, Jan Roelf, and Tom Wansbeek (1995), "Optimal Selection for Direct Mail," *Marketing Science*, 14 (4), 378-394.

Cameron, A. Colin and Pravin K. Trivedi (1998), *Regression Analysis of Count Data*, NewYork, NY: Cambridge University Press.

Currim, Imran S. (1982), "Predictive Testing of Consumer Choice Models Not Subject to Independence of Irrelevant Alternative," *Journal of Marketing Research*, 29 (2), 208-222. 0020.

Dholakia, Ruby Roy, Miao Zhao, and Nikhilesh Dholakia (2005), "Multi-channel Retailing: A Case Study of Early Experiences," *Journal of Interactive Marketing*, 19 (2), 63-74.

The DMA (2005), *2005 Multichannel Marketing Report*, The Direct Marketing Association, New York, NY.

Doubleclick (2004), *2004 Holiday Season Multi-Channel Retail Research*, Doubleclick Inc. NewYork, NY.

Du, Rex Yuxing, Wagner A. Kamakura, and Carl F. Mela (2007), "Size and Share of Customer Wallet," *Journal of Marketing*, forthcoming.

Elsner, Ralf, Manfred Kraft, and Arnd Huchzermeier (2004), **"**Optimizing Rhenania's Direct Marketing Business Through Dynamic Multilevel Modeling (DMLM) in a Multicatalog-Brand Environment", *Marketing Science*, 23 (2), 92-106.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005a), "Counting Your Customers the Easy Way: An Alternative to Pareto/ NBD Model", *Marketing Science*, 24 (2), 275-284.

______, ______, and ______ (2005b), “RFM and CLV Using Iso-Value Curves for Customer Base Analysis”, *Journal of Marketing Research*, 42 (4), 415-430.

______, ______, and ______ (2005c), “Implementing the BG/NBD Model for Customer Base Analysis in Excel”, http://www.brucehardie.com/notes/004/, Accessed on May 15, 2006

Fenvessy, Stanley (1992), “Fulfillment Planning: An Overview” in *The Direct Marketing Handbook*, Edward Nash, ed. New York: McGraw-Hill.

Gensler Sonja, Marnik G. Dekimpe, and Bernd Skiera (2004), “Channel Power in Multi-Channel Environments,” Working Paper, Erasmus Research Institute of Management, Erasmus University, Rotterdam, The Netherlands, ERIM Report Series ERS-2004-041-MKT.

Gönül, Füsun F. and Meng Ze Shi (1998), “Optimal Mailing of Catalog: A New Methodology Using Estimable Structural Dynamic Programming Models,” Management Science, 44 (9), 1249-1262.

______ and Frenkel ter Hofstede (2006)**, “**How to Compute Optimal Catalog Mailing Decisions,” *Marketing Science*, 25 (1), 65-74.

Greene, William H. (2003), *Econometric Analysis*, Upper Saddle River, NJ; Prentice Hall.

Greenwald, Anthony G. and Clark Leavitt (1984), “Audience Involvement in Advertising; Four Levels,” *Journal of Consumer Research*, 11 (1), 581-592.

Hahn, Jinyong and Jerry Hausman (2002), “A New Specification Test for the Validity of Instrumental Variables,” *Econometrica*, 70 (1), 163-189.

Heilman, Carrie M, Kent Nakamoto, and Ambar Rao (2002), “Pleasant Surprises: Consumer Response to Unexpected Instore Promotions”, *Journal of Marketing Research*, 39 (2), 242-251.

Hess, James D., and Glenn E. Mayhew (1997), “Modeling Merchandise Returns in Direct Marketing,” *Journal of Direct Marketing*, 11 (2), 20-35.

Hoch, Stephen J., Eric T. Bradlow, and Brian Wansink (1999), “The Variety of an Assortment,” *Marketing Science*, 18 (4), 527-546.

Imai, Kosuke and David van Dyk (2005), “A Bayesian Analysis of the Multinomial Probit Model Using Marginal Data Augmentation,” *Journal* of *Econometrics*, 124 (2), 311-334.

Inman, Jeffery J, Venkatesh Shankar, and Rosellina Ferraro (2004), “The Roles of Channel-Category Associations and Geodemographics in Channel Patronage Decision,” *Journal of Marketing*, 68 (2), 51-71.

J.C. Penney 10K Statement (2006), ‘J.C. Penney Company Inc: Form 10K”, http://thomson.mobular.net/thomson/7/2348/2579/, Accessed on May 26, 2007.

Jen, Lichung, Chien-Heng Chou and Greg M. Allenby (2003), "A Bayesian Approach to Modeling Purchase Frequency," *Marketing Letters*, 14 (1), 5-20.

Johnson, Carrie (2004), “Multichannel Retailing Best Practices: Which Retailers Offer the Best Cross Channel Customer Experience,” Research Report, Forrester Research.

Jones, Andrew (2000), “Health Econometrics” in *Handbook in Health Economics Volume 1A*, A.J. Culyer and J.P. Newhouse, eds. New York, NY: Elsevier.

Jupiter Research (2006), *Jupiter Research Internet Shopping Model*, New York, NY.

Knox, George (2006), “Modeling and Managing Customer in a Multichannel Setting,” *Dissertation*, University of Pennsylvania, Philadelphia, PA.

Kumar V. and Rajkumar Venkatesan (2005), “Who are Multichannel Shoppers and How Do They Perform?: Correlates of Multichannel Shopping Behavior,” *Journal of Interactive Marketing*, 19 (2), 44-61.

Levy, Michael and Barton A. Weitz (2006), *Retailing Management*, Fifth Edition, New York, NY: McGraw Hill.

Lewis, Michael (2006), “The Effect of Shipping Fees on Purchase Quantity, Customer Acquisition, and Store Traffic”, *Journal of Retailing*, 82 (1), 13-23.

______, Vishal Singh, and Scott Fay (2006), “An Empirical Study of Impact of Non Linear Shipping Fees on Purchase Incidence and Expenditure Decisions”, *Marketing Science*, 25 (1), 51-64.

Libai, B., D. Narayandas, and C. Humby, (2002), "Toward an Individual Customer Profitability Model: A Segment-Based Approach." *Journal of Service Research*, 5 (1), 69-76.

Maddala, G.S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, NewYork, NY: Cambridge University Press.

Mantrala, Murali K., Prabhakant Sinha and Andrew Zoltners (1992), “Impact of Resource Allocation Rules on Marketing Investment-Level Decisions and Profitability”, *Journal of Marketing Research*, 29 (2), 162-175.

Marketing Science Institute (2006), “Research Priorities 2006-2008: A Guide to MSI Research Programs and Procedures,” *Marketing Science Institute Research Priorities*, Report, Cambridge, MA.

Messinger, Paul,and Chakravarthi Narasimhan (1997), “A Model for Retail Format Based on Consumers’ Economizing of Shopping Time,” *Marketing Science*, 16 (1), 1-23.

McCulloch, R. and Peter Rossi (1994), “An Exact Likelihood Analysis of the Multinomial Probit Model,” *Journal of Econometrics*, 64, 207-240.

______, N. G. Polson, and Peter Rossi (2000), “A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters,” *Journal of Econometrics*, 99, 173-193.

Montoya-Weiss, Mitzi M., Glenn B. Voss, and Dhruv Grewal (2003), “Determinants of Online Channel Use and Overall Satisfaction with a Relational, Multichannel Service Provider,” *Journal of the Academy of Marketing Science*, 31 (4), 448-458.

Morgan, Robert M. and Shelby D. Hunt (1994), “The Commitment-Trust Theory of Relationship Marketing,” *Journal of Marketing*, 58 (3), 20-38.

Morrison, D.G. and Schmittlein, D.C. (1988), "Generalizing the NBD Model for Customer Purchases: What Are the Implications and Is It Worth the Effort?," *Journal of Business and Economic Statistics*, 6 (2), 145-166.

Myers, Joseph B., Andrew D. Pickersgill, and Evan S. VanMetre (2004), “Steering Customers to the Right Channels,” *McKinsey Quarterly*, 4, 36-48.

Naik, Prasad A, Murali Mantrala K. and Allen G. Sawyer (1998), “Planning Media Schedules in the Presence of Dynamic Advertising Quality”, *Marketing Science*, 17 (3), 214-235.

National Telecommunication and Information Administration (2004), “A Nation Online: Entering the Broadband Age,” US Department of Commerce, Economics and Statistics Administration, Washington, DC.

Neslin, Scott A., Dhruv Grewal, Robert Leghorn, Venkatesh Shankar, Marije L. Teerling, Jacquiline S. Thomas, Peter C. Verhoef (2006), "Challenges and Opportunities in Multichannel Management," *Journal of Service Research*, 9 (2), 95-112.

Nobile, A. (1998), "A Hybrid Markov Chain for the Bayesian Analysis of the Multinomial Probit Model," *Statistics and Computing*, 8, 229-242.

Papke, Leslie and Feffrey M. Wooldridge (1996), "Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates," *Journal of Applied Econometrics*, 11 (6), 619-632.

Popkowski Leszczyc, Peter T. L. and Harry J. P. Timmermans (1997), "Store-Switching Behavior," *Marketing Letters*, 8 (2), 193-204.

Rangaswamy, Arvind and Gerrit H. Van Bruggen (2005), "Opportunities and Challenges in Multichannel Marketing: An Introduction to the Special Issue," *Journal of Interactive Marketing*, 19 (2), 5-11.

Reinartz, Werner and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implication for Marketing," *Journal of Marketing*, 64 (4), 17-35.

______, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing,* January, 69(1), 63-79.

Schmittlein, David C., David G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who they Are and What Will They Do Next", *Management Science*, 33 (1), 1-24.

______, and Robert Peterson (1994), "Customer Base Analysis: An Industrial Purchase Process Applications", *Marketing Science*, 13 (1), 41-67.

Staiger, Douglas and James H. Stock (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65 (3), 557-586.

Stock, James H. and Mark W. Watson (2003), *Introduction to Econometrics*, NewYork, NY: Pearson Addison Wesley.

Thomas, Jacquelyn, Robert Blattberg, and Ed Fox (2004), "Recapturing Lost Customers," *Journal of Marketing Research,* 41 (1), 31-45.

______ and Ursula Sullivan (2005), "Managing Marketing Communication with Multichannel Customers," *Journal of Marketing*, 69 (4), 239-251.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (4), 106-126.

______, V. Kumar, and N. Ravishanker (2007), "Multichannel Shopping: Causes and Consequences," *Journal of Marketing*, forthcoming.

Verhoef, Peter C. and Bas Donkers (2005), "The Effect of Acquisition Channels on Customer Loyalty and Cross-Buying," *Journal of Interactive Marketing*, 19 (2), 31-43.

______, Scott A. Neslin, and Bjorn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing*, forthcoming.

Villanueva, Julian, Shujin Yoo, and Dominique Hanssens (2006), "The Impact of Marketing-Induced vs. Word-of-Mouth Customer Acquisition on Customer Equity," MSI Working Paper, 06-119.

Wall Street Journal (2004), "Shoppers Who Blend Store, Catalog and Web Spend More," 244 (46), A7.

Windmeijer, F.A.G. and J.M.C. Santos Silva (1997), "Endogeneity in Count Data Models: An Application to Demand for Health Care," *Journal of Applied Econometrics*, 12 (3), 281-294.

Zellner, Arnold and H. Theil (1962), "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica*, 30 (1), 54–78.

VITA

TARUN LALBAHADUR KUSHWAHA
Kenan-Flagler Business School
The University of North Carolina
McColl Building, Campus Box 3490
Chapel Hill, NC 27599-3490
tarun_kushwaha@unc.edu

EDUCATIONAL BACKGROUND

Mays Business School, Texas A&M University, College Station, TX
Doctor of Philosophy, Marketing, August 2007
Master of Science, Marketing, May 2002

Nirma Institute of Management, Nirma University, India
Post Graduate Diploma in Management (M.A.S.), Marketing, April 1998

St. Xavier's College, Gujarat University, India
Bachelor of Science, Physics, June 1996

CURRENT EMPLOYMENT

Assistant Professor of Marketing, Kenan-Flagler Business School, The University of North Carolina, Chapel Hill, NC, July 2007 onwards

INDUSTRY EXPERIENCE

Management Trainee and Assistant Manager, Advertising and Media Planning, Nirma Chemical Works Limited, India, May 1998 – May 2000.

.