

Multichannel Marketing and Hidden Markov Models

Chun-Wei Chang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Gary Erickson, Chair

Douglas MacLachlan

ZaoZao Zhang

Program Authorized to Offer Degree:

Michael G. Foster School of Business

University of Washington

Abstract

Multichannel Marketing and Hidden Markov Models

Chun-Wei Chang

Chair of the Supervisory Committee:

Professor Gary Erickson

Department of Marketing and International Business

Understanding how customers' channel preferences evolve is crucial to firms in managing multiple channels effectively. This dissertation examines the underlying influences which cause customers to migrate over time into different unobservable experience states with higher propensity to purchase in a specific channel. I apply a multiple-segment Hidden Markov Model (HMM) to discover the dynamic behavior of customers facing multiple channels. The study in Chapter 3 takes an alternative way to incorporate heterogeneity using structural segments, which allows heterogeneity in both state transition and channel choice, and offers substantive and interesting insights regarding multichannel shopping patterns. In the empirical application, I identify two segments and two states in the multiple-segment HMM and examine different learning patterns and rate of experience development for each segment. My results show that over time, customers do not tend to move away from bricks-and-mortar stores as some experts

expect as they gain more experience. Some customers perform multichannel-oriented behavior and show various evolving patterns. Customers also reveal different reactions to marketing communications for different combinations of channel tendency. Also, the proposed model suggests an effective way for a firm to dynamically segment and manage channel usage with its customer base. Based on empirically-derived insight regarding customer channel preference evolution with experience, marketers can allocate a firm's limited resources effectively and further refine marketing strategies. Furthermore, customer retention and churn has received increasing attention in the field of customer relationship management (CRM) in recent years. Chapter 4 provides a framework to estimate a relationship dynamics in a non-contractual setting whose customers' dropout time is not clearly stated and easily observed. I incorporate the effects of channel experiences and marketing communications on relationship dynamics, and use a nested structure to detect purchase preference and channel evolution simultaneously until "death" of a relationship, and identify a more (in)active-oriented channel. The proposed nested multinomial HMM addresses changes in preference of purchase incidence and channel choice across time with respect to various relationship states, and deals with the impact of marketing communications and channel experiences on customer retention as governed by transitions between relationship states.

Table of Contents

Chapter 1	1
Introduction.....	1
Chapter 2.....	4
Literature Review.....	4
2.1. Customer Experience and Multichannel Choice	5
2.2. Channel Migration Behavior	7
2.3. Multichannel Marketing and Customer Relationship Management	9
2.4. Non-contractual Customer Retention.....	11
2.5. HMM Applications and Customer Relationship Management	14
Chapter 3	18
Essay 1: The Effect of Unobservable Learning States on Customers' Channel Preference Evolution.....	18
3.1. Introduction	18
3.2. Model Development	24
3.2.1. Markov Chain Transition Matrix.....	27
3.2.2. State-dependent Choice Distribution.....	29
3.2.3. Initial State Distribution	30
3.3. Empirical Application	30
3.3.1. Data.....	31
3.3.2. Variables	34
3.3.3. Estimation Procedure and Model Selection.....	37
3.3.4. Estimation Results	41
3.4. Conclusions and Directions for Future Research	48
Chapter 4	55
Essay 2: The Effects of Varying Channel Experiences and Direct Marketing on Customer Retention	55
4.1. Introduction	55
4.2. Model Development	60
4.2.1. State-dependent Choice Distribution.....	63
4.2.2. Markov Chain Transition Matrix.....	66

4.2.3. Initial State Distribution	67
4.3. Empirical Application	68
4.3.1. Data.....	69
4.3.2. Variables	70
4.3.3. Estimation Procedure and Model Selection.....	73
4.3.4. Estimation Results	74
4.4. Counterfactual Analysis	80
4.5. Conclusions and Directions for Future Research.....	83
Chapter 5	89
Conclusions and Future Directions	89
Appendix A.....	93
Multiple-segment HMM Likelihood Function	93
Appendix B	96
Recovering the State Distribution of HMM.....	96
Bibliography	97

Chapter 1

Introduction

Multichannel marketing has become a prominent topic in recent years, with the advent of the Internet's World Wide Web (the Web) and an increasing amount of e-commerce. Firms and customers can reach each other through multiple channels for different purposes. For example, firms can send promotion codes through emails, send postcards and catalogs with new product introductions through regular mail, and conduct post-purchase surveys by phone. They can also try to attract new customers and retain existing customers for real and virtual stores. Likewise, customers can search for product information through the Internet, purchase products at a retail store, return products by mail, and provide post-purchase feedback and complaints by phone. Consequently, firms are faced with a number of multichannel marketing issues, such as how to manage multichannel shoppers and allocate a firm's limited resources, how to accurately segment customers according to their channel tendencies, how to build longer and lasting relationships with customers in a multichannel environment, how to decide which channel should be focused on, how to track tendency migration, and how channel experience and direct marketing affect customer relationship dynamics.

Most of the previous research which attempts to understand the determinants of channel choice has shown that the most important factor in channel selection is prior experience, which will affect subsequent purchase behavior. Additionally, previous behavioral research asserted

that category experience affects cognitive change rather than ultimate choice, but of this research most empirical studies examined only the correlation between category experience and ultimate choice: the gap between cognitive change and ultimate choice has been ignored.

I believe that customer experience affects “alternative evaluation” stage during the purchase decision process, which can be represented by a latent channel tendency viewed as a consideration set, and has a further influence on actual choices. However, there is no prior research on experience and learning effects on channel preference evolution exploring the gap and linking ultimate choice with unobserved channel tendency. The extent to which customers change their channel choices over time, and how customers’ channel preferences evolve as they learn and gain more experience in recognizing and incorporating the latency of channel tendencies, remain important research issues. Additionally, in a multichannel environment, firms are looking for a variety of ways to increase customer retention and avoid customer churn because the cost of customer acquisition is extremely high. Therefore, multichannel issues in the context of customer relationship management (CRM) about how to identify the most valuable customers, how to maintain long-term customer relationships through a variety of channel and marketing strategies, how to increase loyalty, and how to estimate customer retention across channels should be emphasized. Well-known approaches such as the Pareto/NBD and BG/NBD models, which explore issues of customer retention/churn and predict future purchase patterns to calculate customer lifetime value (CLV) in non-contractual settings, have some shortcomings. Research based on those approaches has examined individual customer retention on the basis of

an “aggregate” relationship with a firm, no matter what the customer purchased or through which channels the customer shopped. Little is known about whether channel experiences have various impacts on customer-firm relationships and which patterns of evolution are more likely to reinforce a customer’s relationship with a firm.

This dissertation examines these important issues and is structured as follows. Chapter 2 reviews previous literature related to customer experience and multichannel choice, channel migration behavior, customer retention, and Hidden Markov Models (HMM) as applied in CRM. Chapters 3 and 4 investigate the effects of experience on channel preference evolution over time and then explore several issues of customer retention in multichannel environments. In Chapter 3, I develop a non-homogeneous multiple-segment HMM to capture both dynamic variation and cross-sectional heterogeneity in customers’ channel choice, which allows me to identify a customer’s true state of latent channel tendency over time and examine the dynamic learning effects on the evolution of customer channel preference across time. Chapter 3 connects the indirect relationship between category experience and ultimate channel choice, and links unobserved alternative evaluation which forms a consideration set with the observed purchase stage in the purchase decision process. Chapter 4 investigates customer-firm relationships, examines which channels can build longer and lasting relationships with customers, highlights the impact of channel experiences and direct marketing in customer retention, and recovers the customer’s hidden relationship state which governs the purchase dynamics. Chapter 5 present conclusions, study limitations, and directions for future research.

Chapter 2

Literature Review

The literature related to Chapters 3 and 4 addresses the following important topics: customer experience and multichannel choice, channel migration behavior (Ansari et al. 2008; Dholakia et al. 2005), multichannel and customer relationship management (Hansotia and Rukstales 2002; Kumar and Venkatesan 2005; Rangaswamy and Van Bruggen 2005; Shankar et al. 2003; Thomas 2001; Verhoef and Donkers 2005), customer retention(Boehm 2008; Fader and Hardie 2010; Fader et al. 2005; Gupta and Lehmann 2003; Pfeifer and Farris 2004; Reinartz and Kumar 2000; Schmittlein et al. 1987; Schweidel et al. 2008), and HMMs (Brangule-Vlagsma et al. 2002; Du and Kamakura 2006; MacDonald and Zucchini 1997; Montgomery et al. 2004; Moon et al. 2007; Netzer et al. 2008). I first review the literature on the role of customer experience and its impact on channel choice and preference evolution, and issues arising from the linkage between experience and ultimate choice outcomes. I then review the literature on the role of a multichannel environment in CRM and its influence on customer-firm relationships. I next review studies of channel migration over time and of non-contractual customer retention, and the issues stemming from current retention models. Lastly, I review the literature on applying HMM in the field of marketing.

2.1. Customer Experience and Multichannel Choice

The majority of previous research has examined how product category experience affects brand preference, how channel experience affects multichannel choices, and how experience affects repeat purchase behavior. Verhoef and Donkers (2005) found that customer experience with the firm seems to be more important as the relationship persists. Favorable multichannel experiences may reinforce customer relationships, and then induce further purchase (Rangaswamy and Van Bruggen 2005). When customers make repeat purchases, prior channel choices affect subsequent channel choices (Dholakia et al. 2005). For example, customers tend to purchase within a particular channel when they make repeat purchases because of switching costs, risk aversion, and learning effects. Thomas and Sullivan (2005) used prior purchase channel choice as a factor to explain future channel choices. Ansari et al. (2008) created a set of variables associated with experience effects which include number of previous purchases (frequency) and duration between purchases (recency), and then explored how experience affects purchase incidence, purchase volume, and channel selection in the multichannel environment. As a result, greater knowledge or experience with a particular channel may lead to a greater sense of comfort with that channel and then affect subsequent choice.

Heilman et al. (2000) examined how category experience and learning affect dynamic choice processes indirectly through cognitive drivers like perceived risk and information search, and investigated the evolution of brand preference among customers new to a market. They believe category experience can be captured by a variable defined as the number of cumulative

product-related purchases and conclude that preference and price sensitivity change with increasing category experience by adding interaction terms between cumulative purchase and explanatory variables. Only a few prior studies have explored the effects of category experience on channel choices (Gupta et al. 2004; Schoenbachler and Gordon 2002). Gupta et al. (2004) argue that a quality risk associated with product performance is critical to channel choice if customers cannot touch and see substantial products prior to purchase, and experience with products therefore has an impact on consumers' channel choices. For example, customers with no product experience are more likely to purchase at a brick-and-mortar store because the customer can touch the products and examine their quality before making a purchase. This example illustrates the process of how category experience affects a customer's evaluative criteria in the "alternative evaluation" stage of the purchase decision process (Kotler 2000), and then affects purchase decisions. Schoenbachler and Gordon (2002) believe that positive past experience drives customers to shop through multiple channels. The more familiar customers are with a company, the more likely they are to shop using multiple channels. In short, customers are more likely to shop through multiple channels when they have a greater number of experiences with a company and its products.

Prior research asserts that the number of cumulative product-related purchases affects subsequent behavior associated with channel choice and tendency to multichannel purchase. On the basis, I believe that customer experience affects channel tendency in the alternative evaluation stage. Channel tendency, which depends on the level of experience for each customer

and has a further influence on actual choice at the purchase stage, is difficult to observe and not clearly state, especially from secondary data such as a firm's transaction records. It is problematic if marketers cannot identify correctly the state of channel tendency. The proposed model in Chapter 3 defines a set of latent states to overcome this problem and probabilistically identify a customer's state of channel tendency at any time. In Chapter 4, I examine the immediate impacts of channel experiences on both purchase incidences and channel preferences, and the enduring experience effects on the evolution of customer-firm relationships.

2.2. Channel Migration Behavior

The purchase decision process defined by Kotler (2000) consists of problem recognition, information search, evaluation of product options, purchase, and post-purchase support. Previous research exploring channel switching has focused on either the determinants of migration during various stages of the decision process (i.e., the use of different channels for different stages) or a particular stage of the decision process over time (e.g., migration over time toward online in the purchase stage). Verhoef et al. (2007) studied multiple channels and multiple stages, and considered channel-choice decisions in the search and purchase stages. They focused on the research-shopper phenomenon, defined as the tendency of customers to use one channel for search and another for purchase. Their results support the evidence that Internet search and store purchase is the most popular combination of research shopping, which is consistent with the report from DoubleClick (2004). This type of research-shopping is due to the characteristics of

poor lock-in of the Internet, significant cross-channel synergy between Internet search and store purchase, and relative search advantage and purchase disadvantage for the Internet. Gupta et al. (2004) summarized five factors (channel-risk perception; price-search intentions; search effort; evaluation effort; and delivery time) that affect consumer intentions to shop online or offline during the purchase-decision process.

Fewer research studies have considered customer migration behavior over time. Venkatesan et al. (2007) explored various interaction characteristics that influence a customer's channel adoption duration. They summarized three main customer-firm interactions which impact the amount of time until a customer adopts an additional channel. For example, higher interaction frequency leads to relationship development and shorter channel adoption duration. Their study investigated what factors affect the duration of adopting a second and a third channel. Dholakia et al. (2005) studied customer channel-switching behavior in terms of switching costs and perceived risks. Their results from descriptive statistics showed that a customer tends to be more likely to switch between similar channels than between dissimilar channels, and to remain loyal to particular channels for repeat purchases because of switching costs, risk aversion, and learning effects. They concluded that marketing efforts can influence a customer's channel-switching behavior. For example, more catalog promotions can lead to Internet customers switching to purchasing from a catalog rather than a brick-and-mortar store. However, their conclusions were based only on descriptive statistics.

Ansari et al. (2008) investigated what factors influence migration toward the Internet and how different combinations of marketing efforts affect this migration. They modeled purchase volume and channel selection simultaneously in the face of dynamics and customer heterogeneity in experience via individual-specific random effects. They found that experience effects were not associated with channel migration. However, they focused only on catalog and Internet channels, and did not include a physical channel such as a brick-and-mortar store. Their objective was to explain what determines migrating to the Internet and to examine the differences between migration and no migration groups. They did not consider the situation of reverse migration. Also, the way they examined the experience effect was to compare change in experience utility for both groups, while in Chapter 3 I explore how alternative evaluation evolves with experience, allowing for state transitions at any time, and incorporating the evolution of latent states of channel tendency that are not directly observed, and observed state-dependent choice. My objective is to see directly how experience affects the evolutionary path of channel tendency, and to recover a customer's latent channel tendency at the end of an observation period.

2.3. Multichannel Marketing and Customer Relationship Management

Customer relationship management (CRM) in multichannel marketing has become an increasingly important issue in recent years, considered as a means to enhance customer relationships and motivate customers to shop more frequently through a variety of channels (Hansotia and Rukstales 2002). Rangaswamy and Van Bruggen (2005) argued that multichannel

marketing can enable firms to build lasting customer relationships by simultaneously offering customers information, products, services and support through two or more channels, and that firms can improve their understanding of customers' decision making and develop strategies to enhance short and long term customer relationships. For example, firms could target high-value customers by providing more contact opportunities. Some research has linked multichannel environments to CRM issues and tried to support the hypothesis that encouraging customers to shop in multiple channels leads to higher customer profits. Venkatesan et al. (2007) argued that multichannel customers are exposed to a firm's services more frequently and therefore will be more satisfied and develop deeper relationships with the firm. Kumar and Venkatesan (2005) showed that multichannel customers provide higher revenues and higher share of wallet to firms, and are more likely to be active than single-channel customers.

Most of the research of CRM on multichannel environments focuses specifically on issues of customer satisfaction, customer lifetime value (CLV), and the impact of acquisition channels on relationship issues (Shankar *et al.* 2003; Thomas 2001; Verhoef and Donkers 2005). Verhoef and Donkers (2005) explored a variety of acquisition channels and found that acquisition channels differed with respect to customer retention and cross-buying in the early stage of customer relationship, but did not differ significantly for longer time periods. Previous research has also explored the relationship between the use of online banking and customer retention (Boehm 2008; Campbell and Frei 2010; Hitt and Frei 2002; Verhoef and Donkers 2005). All these studies found that Internet use has a positive effect on customer retention. There are some studies that have treated retention/churn as a binary independent variable defined by the occurrence of closing an account during the observation period, and treated Internet use as one of the predictors (Campbell and Frei 2010; Verhoef and Donkers 2005). However, the

retention/churn incidence is hard to observe in non-contractual settings, and the impact of various magnitudes of channel experiences on retention is still unknown in these settings. Although prior research has developed a good understanding of some CRM issues in multichannel marketing, it has not explored the dynamic impacts of various channel usage on retention probabilities, and the issue of customer retention in multiple channel environments in a non-contractual setting remains unexplored. Chapter 4 attempts to explore the effects on customer retention rate of multichannel experiences, by studying the transitions between a set of discrete customer-firm relationship states. The following subsection reviews literature related to retention models in non-contractual settings.

2.4. Non-contractual Customer Retention

Retention in some CRM literature refers to a single and constant ratio used to represent the portion of retained customers and to calculate lifetime value, which means that the estimated retention probabilities do not vary over a customer's lifetime period (Blattberg and Deighton 1996; Gupta and Lehmann 2003). Gupta and Lehmann (2003) mentioned that it is difficult to estimate retention empirically and that, therefore, in many applications a constant retention over time is assumed or assigned. Since that study, there has been some research that provides a good explanation and application for the phenomenon of inconstant and increasing retention rates with time. Pfeifer and Farris (2004), for example, used sensitivity analysis to illustrate the importance of improved customer retention, which was assumed to increase to its expected lifetime value, and derived an equation for the retention elasticity of customer value. They assumed the

retention rate varied with tenure, but the true retention rate was still unknown. The above research studies all asserted that small increases in retention drive large increases in profits (S. Gupta and Lehmann 2003; Pfeifer and Farris 2004; Reichheld and Sasser Jr. 1990). None, however, provided a way to estimate retention probabilities in a customer's lifetime stream. Also, Fader and Hardie (2010) argued that an apparent increasing retention rate can result merely from ignoring underlying cross-sectional heterogeneity in relationship probabilities.

Inaccurate estimates of retention rates lead to biased estimates of the value of a customer base, and especially for a non-contractual service because a customer's tendency to retain or dropout is little observed. Most previous research in contractual settings used the family of hazard models to deal with retention duration, predict customer's lifetime, and examine the impact of predictors on relationship length (Boehm 2008; Schweidel et al. 2008). Fader and Hardie (2010) used shifted-beta-geometric (sBG) to model the cohort-level retention rates instead of aggregate retention rate because they found evidence that an aggregate retention rate was a biased estimate. These approaches are appropriate in contractual settings, and clearly stated dropout times are necessary to their application.

In non-contractual settings, issues around how to estimate customer lifetime value based on accurate retention, and how to count retained customers, remained until the Pareto/NBD (Reinartz and Kumar 2000; Schmittlein et al. 1987) and BG/NBD (Fader et al. 2005) models were proposed. In the context of CRM, the Pareto/NBD and BG/NBD models explore such issues as predicting future demand, customer churn, and retention rate by assuming customers

may transition from “active” state to “inactive” state at different rates. The two models attempt to estimate customer retention and dropout rate with slightly different assumptions, and provide good answers to questions about how many customers will be active or “alive” in the future given their past behavior. The Pareto/NBD and BG/NBD models both assume that customers evolve along two states (“active” and “inactive”) and cannot account for a switch back to active state once customers have dropped out. The models imply that there is no space between active and inactive, and that customers are inactive permanently once they are identified as inactive. This is an extreme restriction on estimating retention, and neglects important factors that influence and drive retention.

Managers are eager to understand the factors underlying retention rate rather than rely on aggregate information, e.g., in what circumstances a customer will come back after being inactive, and in what circumstances customer retention probability increases. Also, customers who are used to purchasing through one channel may have different retention probabilities than those who are used to purchasing through an alternative channel. Each channel has various impacts on a customer’s retention probability, and thus the effect of channel experiences on retention should be considered in modeling a multichannel environment. In this research, I use an HMM framework that allows for customers evolving flexibly¹ amongst more than two states instead of evolving through simply active or inactive. It also allows the exploration of varying

¹ Customers are free to transit among any states, and can switch back from an inactive state to an active state.

impacts of channel experience and marketing on retention by incorporating those factors into customer state transitions.

2.5. HMM Applications and Customer Relationship Management

HMM is a widely used methodology that has been applied to problems in a variety of fields, such as speech processing (Juang and Rabiner 1991; Rabiner 1989), biology (Leroux and Puterman 1992), medicine (Albert et al. 1994), and genetics (Churchill 1989). MacDonald and Zucchini (1997) illustrate many applications of HMM in several fields. HMM has been applied to marketing problems in recent years. Brangule-Vlagsma et al. (2002) attempt to explore how individual value systems change dynamically across time. Due to the latency of value segments, they used HMM to identify value segments and assumed that the observed value measurements depended on some latent value segments that followed a Markov process. They compared HMM with the classical latent class model (K-M model), which assumes fixed segments over time, and with the extension of the K-M model, which allows that value segments over time are independent, and found that HMM outperformed the other models. This implies that customers switch among segments in a structured way.

Montgomery et al. (2004) explored online browsing behavior by categorizing the sequence of pages or the path viewed by users. The path or the sequence of Web viewing can reflect a user's goals and is informative in predicting the user's future paths. Past movements and memory affect future movements, so models that account for memory fit and predict better than memory-less models. The latent states in HMM capture memory effects and the transitions in HMM capture longer-term dynamics and abrupt changes in browsing styles. Du and Kamakura (2006)

used HMM to identify unobserved and sequential household life stages from observed demographic profiles and depicted life paths that represented the sequences through which households move throughout their life stages. Their model is homogeneous and the transitions between the latent states are stationary. Their multivariate methodology enabled them to deal with discrete and continuous variables simultaneously, which combined multinomial and normal distributions. Due to the difficulty of accessing data on a competitor's promotion behavior, the competitor's efforts would be ignored in a promotion response model and the estimates would be biased. Moon et al. (2007) estimated own- and cross-promotion effects by treating unobserved competitive promotions as missing data to be imputed by a hidden Markov process. The random coefficient HMM they provided effectively estimated the impact of own- and cross-promotion efforts when competitive promotion data was not available, and led to a less biased estimate of promotion response.

In the context of CRM that focuses on retention and churn probability, unlike the Pareto/NBD and BG/NBD models, HMM does not impose a priori constraints on the number of states and transition paths; the total number of states is instead inferred by a model selection criterion. Schweidel et al. (2011) used HMM in a multi-service contractual setting to examine dynamics in acquisition and retention of service portfolios by incorporating multivariate choice (co-purchasing) and timing models (duration dependency). In spite of a few “active” states, they added an “end” state in which the service contract has been terminated to fully capture the dynamics in evolution due to the characteristics of service industries, and then identified customers who were more likely to terminate the relationship when they changed portfolios. Netzer et al. (2008) applied HMM in a non-contractual setting to gift-giving behavior in a

university alumni customer relationship dataset. Their model is non-homogenous and allows for time-varying covariates in the transitions. They found that the predictive ability of HMM outperformed non-dynamic models commonly used in CRM analysis, such as the latent class model and the recency-frequency model, which follows a binary logit formulation. Knox (2006) developed an HMM to analyze how direct marketing affects latent learning states over time and then revises customer channel preferences and buying behavior. However, he restricted their transitions to be absorbed in all states except state 1, which implies that not only will customers always move away from the first state, but also that they cannot switch flexibly to the other states. He concluded that customers begin purchasing offline and then migrate online, due to the rigorous constraint in transitions.

In Chapter 3, I propose a multiple-segment HMM applied to a multichannel environment. The proposed model is non-homogeneous with time-varying covariates, and accounts for dynamic variation and cross-sectional heterogeneity. Netzer et al. (2008) incorporate heterogeneity by allowing random effect parameters in the transition matrix and estimate them using a hierarchical Bayes estimation procedure. I take an alternative approach to incorporate heterogeneity, by using structural segments. My approach allows heterogeneity in both state transition and channel choice, and thereby allows a richer interpretation of customer state transition and state-dependent choice. It allows me to capture customer heterogeneity in experience effects as well as cross-sectional heterogeneity in channel choice. In Chapter 4, I incorporate the effects of channel experiences and marketing communications on relationship

dynamics, and use a nested structure to detect purchase preference and channel evolution simultaneously until “death” of a relationship, and identify a more (in)active-oriented channel.

Chapter 3

Essay 1: The Effect of Unobservable Learning States on Customers' Channel Preference Evolution

3.1. Introduction

With the recent advent of the Internet's World Wide Web and e-commerce, online sales grew 21% in 2007 and are expected to increase by about 50% over the next five years (Forrest Research and shop.org annual reports), and a multichannel environment is becoming increasingly prevalent. Firms may reach customers through channels including brick-and-mortar stores, catalogs, emails, telephones, and kiosks, to try to attract new and existing customers to their real and virtual stores. Likewise, customers can contact firms and make purchases through multiple channels. Multichannel marketing means that customers and firms can reach each other by different channels at different times for different purposes. For example, customers can search information at a Web site, purchase at a physical store, obtain technical support and other services via telephone, and return products by mail. As customers interact with firms through more and more channels, firms face the challenge of managing multiple channels effectively and understanding the behavior of multichannel customers. Multichannel customer management becomes a greater issue as firms move toward multichannel integration.

Researchers have devoted effort to understanding the value of multichannel customers and the drivers of customers' channel choice. Some of the research has explored what makes a customer switch between various channels during the stages of the purchase decision process, which may include problem recognition, information search, product evaluation, purchase, and post-purchase support, such as the "research-shopper" phenomenon (Verhoef et al. 2007), which covers the stage from information search to purchase. Prior research has shown that one of the important determinants of channel selection is previous experience, and indicates that prior experiences affect current choice and subsequent behavior. For example, a customer may first purchase a makeup product at a physical store when she doesn't know if the color of the product fits her or if she is allergic to product ingredients. After she learns and gains experience from transaction and usage, she may prefer to make purchases online because of a concern for convenience. The behavioral literature asserts that category experience affects cognitive changes rather than ultimate choice outcomes (Brucks 1985; King and Balasubramanian 1994). Heilman et al. (2000) linked choice with cognitive changes and explained how category experience affects ultimate choice as a result of cognitive changes. I believe experience and customer learning do not affect the ultimate channel selection directly, but instead have a direct influence on the alternative evaluation stage, which is one stage before the purchase stage and helps a customer establish the consideration set. Channel tendency can be viewed as a consideration set in the multichannel environment, and the ultimate channel choice depends on a customer's state of alternative evaluation, channel tendency. No matter whether the product is an experience good or a search good, customers make their own evaluations of the product and transaction process,

which may affect their channel tendency. However, little is known about how experience and learning affect changes in the formation of consideration sets (channel tendencies) over time and channel preference evolution in the long term. Customers may differ in their abilities to learn and gain experience from their usage, and what customers learn from each purchase experience and how this affects their channel choices are unobservable processes. Therefore, the extent to which customers change their channel choices over time and how customers' channel preferences evolve as they learn and gain more experience, recognizing and incorporating the latency of channel tendencies, remains an important research issue.

In this research, I integrate the unobservable channel tendency affected by learning process with heterogeneous customers' channel choice using a Hidden Markov Model (HMM) framework (MacDonald and Zucchini 1997) to examine how customers' channel preferences and reactions to marketing communications vary as their tendency evolves as a result of experience in a multichannel setting. This approach provides a good way to identify whether a customer is a multichannel shopper or a customer with any other tendencies, which is not an easy task but produces information essential to marketers. Using aggregate information can result in misleading classifications. For example, customer A makes five purchases at a retail store in the first five periods, and then makes five purchases online; customer B alternately makes purchases at a retail store and through an online channel. Based on aggregate information, customers A and B may be classified as multichannel shoppers or customers with retail/online tendencies because both have an equal preference for purchasing at a retail store and through an online channel. The

actual story for customer A, however, may be a switch from a retail-loyal tendency to an online-loyal tendency. Moreover, it is hard to say at the end of the observation period whether customer A is an online-loyal customer or a multichannel shopper.

The correct classification is essential for a firm because it affects the prediction of future choices and marketing optimization strategies. Theoretically, channel tendency is a construct or a state that cannot be directly observed or easily measured, particularly from a firm's transaction database. The HMM modeling approach can link the underlying and unobserved sequence of tendency states to observed outcomes of channel choice. The state of tendency is not explicitly defined a priori; rather, HMM provides a flexible structure that acts as an automatic classifier of learning patterns into groups in the model, which translates to particular patterns of customer channel preference. In HMM, this state-dependent choice behavior is defined such that the observed channel choice at any time is determined by the current state of channel tendency which follows a Markov chain. That is, a customer's channel choice at time t depends on the tendency state, i.e., the consideration set established at time t , which develops from what customers have learned from their previous experiences. A Markov transition matrix links the unobserved states to the Markov property, which means the tendency state at time t depends on the state at time $t-1$; and then the current tendency state affects current channel choice behavior. The HMM allows me to capture the latent dynamics of channel tendency and investigate customers' channel preference evolution.

The purpose of this research is to develop a non-homogeneous multiple-segment HMM with finite Markovian states in order to capture both cross-sectional heterogeneity and dynamic variation in customers' channel choice. I extend current HMM modeling by incorporating customer heterogeneity in the rate of learning progress, and in intrinsic channel preference and the effects of marketing communications on channel choice, through a discrete segment structure. Using the multiple-segment HMM, I demonstrate how a firm can classify its customers into segments based on their evolution path of channel tendency and predict their channel choices over time for each segment. The transition probabilities between the latent states in my multiple-segment HMM are determined by cumulative purchase of products across channels (Heilman et al. 2000), which captures the learning process from past experience of purchase and channel usage. Thus, the proposed model provides a way to control for unobserved heterogeneity in learning and experience, allows me to capture dynamic behavior and to identify a customer's true state of channel tendency over time, and accounts for cross-sectional heterogeneity in the data.

I apply the model to a dataset from a large multichannel retailer. In the empirical application, two segments and two states are identified. Choice behavior is assumed to differ for customers in the two segments and two tendency states. There are different combinations of tendency states for each segment, which shows that customers have various evolving patterns. I find customers in segment 1 may have chances to migrate from a multichannel state toward a retail state when they are not familiar with the products, whereas customers in a retail state have strong channel

loyalty. Customers in segment 2 switch between a multichannel and an online state, and then reach steady states as they learn about the channels and the product category. I do not find evidence that customers move away from brick-and-mortar stores and migrate toward the online channel. In addition, the effects of learning, intrinsic channel preference, and reactions to marketing communications vary among the different segments.

In sum, this study makes several contributions. I investigate the impact of category experience on the evolution of latent processes of alternative evaluation. Also, I am the first to model the dynamics of consumers' channel choice behavior by linking ultimate channel choice with latent consideration set. I illustrate how HMM can be applied to extract temporal patterns in customers' channel choices. The HMM I develop suggests an effective way for a firm to dynamically manage channel usage with its customer base, based on the channel tendencies of their customers that vary with experience levels. Third, my model allows segmentation of consumers according to their channel preference, reaction to marketing communications, experience development, and evolution of tendency states. The results show that the multiple-segment HMM indeed outperforms other benchmark models when there are both dynamic variation and cross-sectional heterogeneity. Finally, the results provide interesting insights into multi-channel strategies, which can help firms with their managerial plans to manage multichannel customers over time. Since the approach can more accurately classify customers according to their evolution pattern at any stage of customer experience, this study can help firms

better manage their evolving customer base, enhance their financial projections from different channels, and more effectively manage inventory through channel integration.

The rest of the essay is organized as follows. In Section 3.2., I introduce the structure of HMM, and illustrate the setting of my proposed model. The derivation of the likelihood of the multiple-segment HMM is included in an Appendix. In Section 3.3, I describe the empirical application of my proposed model in the context of learning, using panel data from a large multichannel retailer. Section 3.3 also includes the data collection process, variables used to calibrate the model, the model selection that compares my proposed model with benchmark models, and the empirical results. In Section 3.4, I discuss theoretical and practical contributions, and conclude with limitations and directions for future research.

3.2. Model Development

Segmentation is an important issue in customer relationship management (CRM). A widely applied segmentation method in marketing is the latent class model (LCM). The LCM assumes that an individual can be classified into one of several latent classes, and can be viewed as a means of modeling heterogeneity across individuals. It captures cross-sectional heterogeneity but does not allow a customer to shift between classes. The LCM can be viewed as a special case of HMM, i.e., HMM can account for evolution patterns through flexible transitions, but LCM cannot do this with its identity transition matrix. By incorporating experience as a covariate into the transition matrix, HMM can capture temporal learning effects and dynamic variation in

customers' experience levels. In a multichannel environment, it is crucial to identify customers as "multichannel" shoppers or "single-channel" shoppers, which definitions actually refer to latent channel tendencies. Therefore, it is not an easy task to identify tendencies by traditional choice modeling, or by simply counting the proportion of purchases among channels. For example, if a customer first makes a purchase at a retail store and then online, his tendency may be to migrate from the retail state to the online state, or to remain in a multichannel state, or to revert to a retail state tendency.

Also, it is not appropriate to use a continuous dynamic structure such as a time-varying parameter model to capture dynamics in channel preference that is established in a discrete way (Netzer et al. 2008). In addition, customers with different channel tendencies show different choice patterns. This study uses latent states to infer customers' underlying channel tendencies through actual choice. Because of the state-dependency property of channel selection, and the characteristic that marketers can only observe a customer's actual purchase behavior, and not her state of tendency, HMM is an appropriate approach to modeling the structure of latent states and observed choice behavior. Also, HMM links underlying and unobserved states of tendency with observed outcomes of channel choice, and so is able to relax the limitation of static choice and thereby capture experience dynamics. A latent state is a measure of a customer's channel tendency and the transition between states is determined by the customer experience resulting from the cumulative purchase of products across channels. Moreover, HMM estimates the number of latent states based on the dynamics in the data.

The goal of this study is not to discover and test the factors that drive channel choices, which has been explored in the multichannel literature. One of the purposes of this study is to utilize HMM to discover the impacts of underlying learning or other unknown factors through customers' purchase-related experience which governs the evolutionary pattern in channel tendencies. My HMM has several advantages and extends current choice modeling in many aspects. First, my HMM models customer channel tendency as latent states which represent cognitive changes, and models the development of the experience process through transition probabilities. Second, it connects the indirect relationship between category experience and ultimate channel choices, and links unobserved alternative evaluation which forms a consideration set with the observed purchase stage in the purchase decision process. Third, my HMM allows for changes in consumer choice patterns through the development of experience. Variation in channel tendencies results from experiences updated and gained from each purchase. For example, customers may form a new consideration set as they learn with each transaction. Also, my model accounts for heterogeneity in that it allows customers to have different rates of experience development through a discrete structure. Finally, my model relaxes the assumption that the Markov chain is homogeneous, and assumes that the transition probabilities depend on time-varying covariates. It is important to incorporate those covariates into the Markov chain because the resulting Markov chain may then have a useful substantive interpretation (MacDonald and Zucchini 1997).

3.2.1. Markov Chain Transition Matrix

Unlike the Pareto/NBD and BG/NBD models that have only two states, I do not have limits on the number of states. I also estimate a full and flexible transition matrix that allows customers to stay in their current state or move to any other state, and do not impose an a priori restriction that customers will attain a steady state after the first transition. Given m states, I assume the transition matrix $Q(i_{t-1}, i_t)$ is defined as follows,

$$Q(i_{t-1}, i_t) = \begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1m-1} & q_{1m} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2m-1} & q_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ q_{m1} & q_{m2} & q_{m3} & \cdots & q_{mm-1} & q_{mm} \end{bmatrix} \quad (1)$$

where $q_{jk} = P(i_t = k | i_{t-1} = j)$ denotes the transition probability from state j at $t-1$ to state k at t ,

and $\sum_{k=1}^m q_{jk} = 1$, $0 \leq q_{jk} \leq 1$ for all $j, k = 1, \dots, m$.

The study attempts to explore the evolution of customers' channel tendency. Suppose there are three channels: retail store; online; and catalog. There are seven possible channel tendencies. The possible states can be described as “retail tendency”, “online tendency”, “catalog tendency”, “multichannel tendency”, and combinations of any two channels. I do not presume the number of states and their descriptions, and allow the data to inform us. I model the transition matrix as a multinomial logit model. I model the probabilities of a customer staying where he/she is and moving to any other state as a multinomial distribution. The propensity of transition is affected by the activities that influence customer channel tendency, the experience of the interactions

between customers and the firm, such as purchase transactions. I set the reference level customers choosing staying at the current state, so the parameters in my transition matrix can be represented as the ability to move away from the current state. Therefore, the elements in the transition matrix $Q(i_{t-1}, i_t)$ can be defined as follows.

$$q_{j,k} = \frac{\exp(\mu_k + \theta_k A_{nt})}{1 + \sum_{k \neq j} \exp(\mu_k + \theta_k A_{nt})}, \text{ for all } k \neq j \quad (2)$$

$$q_{j,j} = \frac{1}{1 + \sum_{k \neq j} \exp(\mu_k + \theta_k A_{nt})} \quad (3)$$

where θ_j is a vector of parameters for the impact of cumulative purchase on the probability of transition from state j , A_{nt} is the vector of time-varying covariates for customer n between purchase $t-1$ and purchase t , and μ_k are the constant terms for state j . There is an intrinsic propensity to migrate, which is captured by constant term. The transition probabilities are updated each period based on the purchase experience accumulated, which can help me determine experience effects on transitions. Note that the parameters in the transition matrix are state-specific.

In my multiple-segment HMM, the transition matrix can be seen as a block diagonal matrix as follows.

$$\begin{bmatrix} Q_1 & 0 & \dots & 0 \\ 0 & Q_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_L \end{bmatrix} \quad (4)$$

where Q_l ($l = 1, \dots, L$) is a square matrix of size $m \times m$, which represents the transition matrix in segment l . Therefore, it not only captures cross-sectional heterogeneity among segments, but allows for dynamic evolution within each segment.

3.2.2. State-dependent Choice Distribution

As I mentioned before, the channel choice probabilities are state dependent and consumers' sensitivity parameters are state-specific. The multichannel environment data allow me to model the state-dependent channel choice probabilities with a multinomial logit model.

$$P(Y_{nt} = v | S_{nt} = i) = \frac{\exp(\alpha_{iv} + X_{nt}\beta_{iv})}{\sum_{y=1}^Y \exp(\alpha_{iy} + X_{nt}\beta_{iy})}, i = 1, \dots, m \quad (5)$$

where α_{iv} is the state-specific intrinsic utility of channel v in state i , $\forall v \in \{1, 2, \dots, Y\}$, X_{nt} is a vector of explanatory variables for customer n at t , and β_i is a vector of the state-specific coefficient of variables X_{nt} for state i . $P(Y_{nt} = v | S_{nt} = i_{nt})$ represents the probability that customer n chooses channel v given state i at t . The time-varying covariates X_{nt} for customer n at t should consist of variables that have immediate impact on the customer's channel choice. The covariates are explained in the Empirical Application section.

3.2.3. Initial State Distribution

The initial state distribution can be defined as the stationary distribution of the transition matrix for a homogenous HMM. The transition matrix in my proposed model is a function of time-varying covariates, so I estimate the initial probability using the approach of Netzer et al. (2008). Define π as a vector of initial probabilities ($\pi = (\pi_1, \pi_2, \dots, \pi_m)'$) and the initial state distribution is calculated by solving the following equation.

$$\pi = \pi Q, \quad \sum_{i=1}^m \pi_i = 1, \quad (6)$$

where Q is the transition matrix with all covariates set to their mean values across customers and time periods.

Finally, the multiple-segment HMM model specifications in (1)-(6) are used to develop a likelihood function, which is outlined in the Appendix.

3.3. Empirical Application

I apply the proposed multiple-segment HMM to data on channel choices for clothing purchases by customers of a multichannel retailer. The model is well suited for the data, which include repeated purchases by individual customers that are well recorded. Also, customer experience with the retailer's products appears to be an important determinant of channel choice for a clothing category. Further, channel choice is observed in the data, but channel tendency is not.

3.3.1. Data

The data for this study are provided by a large multichannel retailer. Empirical estimation and evaluation of the model are done with customer membership data from the company. In addition to brick-and-mortar stores, the company has online and catalog sales channels. Customer transaction information from multiple channels is captured and integrated. Thus, the company's customer relationship system can query a complete purchase history for a particular member customer. The data integration process that acquires information from customer interactions through every channel is a critical aspect of the company's CRM success.

The dataset spans seven years, from December 1998 to July 2005, and consists of 157,156 customers with 728,362 observations in the clothing category. I focus on a single category, clothing, to fully control the effect of category experience on tendency state and channel preference, and also remove records involving transactions other than purchase (i.e., return). Because data with historical records of marketing communications start in November 2002, I truncate the dataset for this empirical application to the period November 2002 to July 2005. I randomly sample 10% of the customers who made purchases at least once from November 2002 to December 2004 and split the dataset for calibration and validation: I use the observations from November 2002 to December 2004 to calibrate the model, and the observations from January 2005 to July 2005 test validity. This provides 1518 customers with a complete purchase history in each period, with 14,263 observations in the calibration period and

3686 observations in the validation period. The mean number of purchases per customer in the calibration sample is 9.4. I also calculate other descriptive statistics for both samples (Table 1).

Of the 1518 customers in the calibration sample, 817 (53.82%) customers shopped only in a single channel, and of the 701 (46.18%) customers who made purchases through multiple channels, 609 shopped between two channels, and 92 shopped in all three channels. Of the customers who stayed within one channel, 802 shopped only in brick-and-mortar stores, 14 shopped only online, and one shopped only by mail order (Table 2); the data also showed how many customers shifted between brick-and-mortar stores and online, brick-and-mortar stores and catalog, and online and catalog. I also extracted the average number of purchases for different groups of customers (Table 2). Generally speaking, customers who shopped through multiple channels tended to have a higher purchase frequency than those who shopped through single channel.

Table 1 Descriptive Statistics

Descriptive Statistics	
Calibration Sample (11/2002~12/2004)	
Number of Observations	14263
Number of Customers	1518
Mean Observations Per Customer	9.4
Holdout Sample (01/2005~07/2005)	
Number of Observations	3686
Number of Customers	1518
Mean Observations Per Customer	2.43

Table 2

	number of customers	number of purchases (mean)	percentage
Shop Within 1 Channel	817	8.52	53.82%
within Retail	802	8.51	
Online	14	8.93	
Catalog	1	6	
Shop Between 2 Channels	609	10.25	40.12%
Retail-Online	544	10.41	
Retail-Catalog	54	8.65	
Online-Catalog	11	10.18	
Shop Between 3 Channels	92	11.59	6.06%

*time frame: November 2002 ~ December 2004

The data displayed in Table 2 do not reveal if a customer who had shopped through two or three channels would become a single channel loyal shopper or a multichannel shopper. Shopping through two channels does not mean that a customer is a multichannel shopper. For example, he may try the catalog channel in the beginning of the relationship, and subsequently shop only through brick-and-mortar stores. In this case, “single channel” may describe this customer better than “multichannel”.

I developed an observed switching matrix among the data’s three channels (Table 3). It provides aggregate information about channel switching, but does not reveal the dynamics of evolution. It shows that, in general, customers who shopped at a retail store at t are more likely to remain in the retail store channel at $t+1$, whereas there was a greater chance that customers who shopped through online and catalog channels at t switched to alternative channels at $t+1$. In my study, I use an alternative way to examine evolution of channel preference over time as a result

of category experience, and to discover customers' true state of channel tendency at the end of an observation period.

Table 3 Observed Switching Matrix Between the Channels

t→ t+1	Retail Store	Online	Catalog
Retail Store	90.39%	8.53%	1.09%
Online	50.92%	45.53%	3.55%
Catalog	55.97%	18.11%	25.93%

3.2.2. Variables

In this section, I distinguish the variables that constitute A_{nt} for the transition relationships in (2) and (3) from those that form X_{nt} for the channel-choice probabilities in (5). The vector A_{nt} is the set of variables which are assumed to have an enduring impact on customer's choice, whereas the vector X_{nt} influences the state dependent channel-choice probabilities and is assumed to affect only short-term choice behavior. In the empirical application, I view category experience as a variable having an enduring impact on channel choice, and marketing communications as a variable having immediate effects.

3.3.2.1. Choice Behavior

The observable channel choice is the dependent variable. My empirical application studies three channel choices: brick-and-mortar retail store; online; and catalog. The variable ($Y_{nt} = v$) shown

in equation (5) with $v = \{1,2,3\}$ represents an individual n making a purchase in time period t through brick-and-mortar retail store, online, and catalog channels, respectively.

3.3.2.2. Variables Affecting the Transition Matrix

Previous research has asserted that the number of cumulative product-related purchases affects subsequent behavior associated with channel choice and tendency to multichannel purchase. It has also been asserted that customers are more likely to shop through multiple channels when they have a greater number of experiences with a company and its product. Previous literature also indicates that customer experience does not affect ultimate choice directly, but indirectly through cognitive changes. Therefore, one of the purposes of this study is to utilize HMM to discover the impacts of underlying learning or other unknown factors on the customer purchase-related experience that governs the dynamics in channel tendency.

The HMM requires that latent states be discrete. When experience is accumulated to a certain level, a customer's state of channel tendency is likely to show different switching patterns. Further, the progressing rate may also vary by states and segments. The multiple-segment HMM allows for parameters being state-specific and segment-specific. Also, the variables should occur and be observed prior to customer channel choice. In particular, I specify a customer's state of tendency as a discrete state variable that depends on last period cumulative purchases, and define A_{nt} in (2) and (3) as:

Category experience_{nt} = cumulative purchases made by customer n by time $t - 1$

3.3.2.3. Variables Affecting Immediate Choice

The variables that consist of X_{nt} in (5) are assumed to have immediate and short-term effects on state-dependent channel choices. The firm in my study conducts marketing communications through direct mailings to their selected customers by sending new product information, flyers, sales and event notices, and so on. I want to know if marketing communications received in the past impact a customer's channel selection, and capture the marketing effect by the variable *marketing communications_{nt}*, which represents the number of marketing communications a customer n received within 14 days before her channel selection occasion t .

3.3.2.4. Variables for the Latent Class Model

The latent class models (LCMs) are calibrated with and without category experience. The non-dynamic LCM does not account for customers' experience effects, and therefore the vector of covariates for non-dynamic LCM consists only of the covariates that impact state-dependent choices used in the HMM (i.e., the variables included in X_{nt}). For the LCM with experience effects, the covariate vectors consist of the covariates impacting state-dependent choice (i.e., marketing communications), and category experiences that affect channel tendency in the transitions of my HMM. Category experience is considered in the LCM in order to examine channel preference over time as in Heilman et al. (2000), who treated experience level as

continuous but not discrete, and used a simple specification of cumulative purchases over time to represent learning and experience accumulation. Therefore, channel choice given segment l can be represented as follows.

$$P(Y_{nt} = v|l) = \frac{\exp(a_{vl} + Z_{nt}\gamma_{vl})}{\sum_{y=1}^3 \exp(a_{yl} + Z_{nt}\gamma_{yl})},$$

where γ_{vl} is a vector of segment-specific parameters, and Z_{nt} is a vector of covariates consisting of marketing communications and category experiences.

3.3.3. Estimation Procedure and Model Selection

The parameters for the multiple-segment HMM are estimated by maximum likelihood estimation, which is accomplished through numerical optimization in the GAUSS program. The Bayesian Information Criterion (BIC), which I use in my proposed model to select the number of segments and states, can be represented as follows,

$$BIC = 2 * LL - Npara * \log(Nobs)$$

where LL is the log-likelihood, $Npara$ is the number of parameters, and $Nobs$ is the number of observations.

I estimated sixteen models: basic multinomial logit with(out) experience without heterogeneity and evolving customer preference; two-state to four-state HMM without heterogeneity; two-segment to four-segment LCM with experience; non-dynamic LCM; and

multiple-segment HMMs that contain the combinations of two/three/four segments and two/three states. Table 4 shows the number of parameters Pd , and BIC for the models. I compared model performance using in-sample log-likelihood and holdout sample log-likelihood in addition to BIC. For LCMs, 3-segment was the best for model specification with and without experience (non-dynamic LCM). That result implies that there indeed existed cross-sectional heterogeneity among customers, because LCM outperformed the two basic multinomial logit models without heterogeneity and evolving customer preference. It also implies that there existed dynamic learning behavior for customer channel choice, because HMM without heterogeneity outperformed basic multinomial logit models. At this point, I know that the simple HMM overlooks cross-sectional heterogeneity but LCM neglects dynamic changes over time, and the multiple-segment HMM is the model that can capture the two sources of variation. Therefore, it is not surprisingly that my proposed multiple-segment HMM outperforms both the HMM without heterogeneity and the LCMs. Based on the measures of BIC and holdout log-likelihood, two-segment two-state HMM is the best-fitting model for the multiple-segment HMMs. I also compare the predictive ability for each model by examining the holdout hit rate. Not only does the proposed two-segment two-state HMM have the best fit compared to other benchmark models, but it also has the best prediction ability for the validation sample.

Table 4: Selecting the Number of Segments and States & Model Comparison

Model	Number of Segments	Number of States	Pd	In-Sample Log-Likelihood	Holdout log-Likelihood	BIC	Holdout Hit Rate
Logit no experience	1	1	4	-6931.79	-1864.82	13901.85	75.25%
Logit with experience	1	1	6	-6926.03	-1856.89	13909.46	77.10%
HMM	1	2	12	-5772.49	-1669.20	11659.76	81.39%
	1	3	24	-5668.74	-1663.27	11567.04	80.35%
	1	4	40	-5635.18	-1663.20	11652.97	80.35%
Latent Class with experience	2	1	13	-5802.33	-1669.47	11729.02	77.20%
	3	1	20	-5650.87	-1665.34	11493.04	77.20%
	4	1	27	-5634.09	-1664.36	11526.45	77.20%
Non-Dynamic Latent Class	2	1	9	-5812.35	-1700.08	11710.78	77.20%
	3	1	14	-5657.03	-1666.13	11447.98	77.20%
	4	1	19	-5637.26	-1665.20	11456.27	77.20%
Multiple-segment HMM	2	2	25	-5545.83	-1652.21	11330.80	82.85%
	2	3	49	-5539.43	-1653.54	11547.56	82.82%
	3	2	38	-5534.44	-1660.19	11432.36	78.64%
	3	3	74	-5528.51	-1655.98	11764.87	78.39%
	4	2	51	-5526.18	-1652.26	11540.20	82.85%

3.3.4. Estimation Results

I calculate the estimated parameters and corresponding standard errors for the two-segment two state HMM (Table 5). Plugging these estimates back into equations (2), (3), and (6) to calculate segment-specific and state-dependent probabilities in choice and transitions, I get a better understanding of channel tendency with learning from experiences.

3.3.4.1. Channel Tendencies

I describe the channel tendencies of the states in each of the two segments by examining the state-and-segment-specific intrinsic channel utility estimates for the channel-choice relationships (Table 5). I also calculate the conditional probability of channel choice at the mean of the marketing covariate. For segment 1, the conditional probability of choosing retail store is 58.3%, online is 8.3%, and catalog is 33.4% given state 1, whereas the conditional probabilities given state 2 are 95.8%, 3.7%, and 0.5% for retail, online and catalog, respectively. The probabilities of choosing retail store and online given state 1 become identical when the marketing covariate is at value 10.2. For segment 2, the conditional probability of choosing retail store is 7.5%, online is 86.5%, and catalog is 6.0% given state1, while the conditional probabilities given state 2 are 66.3%, 33.0%, and 0.7% for retail, online, and catalog, respectively. Therefore, for segment 1, state 1 can be labeled as a multichannel state, while state 2 is definitely a retail state. In segment 2, state 1 tends primarily toward online, while state 2 is primarily a retail state, and

secondarily online. The states reflect different degrees of preference for retail, online and multichannel shopping.

The marketing parameter estimates (Table 5) indicate differences in the reaction to the number of marketing communications received in the past 14 days across the two segments, and the states within each segment. In segment 1, marketing communications have significantly positive effect on channel choices given state 1, but have no effect on customers in state 2. These results show that receiving more marketing communications relatively increase the probability of shopping through an online channel given state 1. In the retail state (state 2) in segment 1, customers have a strong preference for retail stores, and do not respond to marketing communications significantly. In Segment 2, customers in the online state (state 1) do not respond to marketing communications significantly, either. For the multichannel state (state 2) in segment 2, the number of marketing communications appears to have had negative impact on retail and online channels, but less so for retail than for online.

3.3.4.2. Learning State Transitions

The intercept term μ in the transition relationships represents the intrinsic propensity of transitioning to the other state, and the sign of μ indicates how (un)sticky the state is. The larger the value of the intercept term, the more likely is a jump to the other state. A state with a negative intercept with greater absolute value is stickier than one with a positive intercept. A positive μ means that customers are more likely to move to the other state than remain in the

same state, indicating a propensity to switch between states. I calculate the intrinsic transition propensities for the two segments without considering the effect of category experience (Table 6); the covariate of category experience in (2) and (3) is set to zero. For segment 1, a customer in the multichannel state (state 1) has a slightly higher intrinsic probability of migrating toward the retail state (state 2) than that for staying in the same state, whereas a customer in the retail state (state 2) is “sticky” to his current state. Segment 2 customers have a slightly higher intrinsic probability of changing from the online state (state 1) to the multichannel state (state 2), and a higher probability of remaining in the multichannel state (state 2) than moving to the online state.

Table 6 Transition Matrix (Intrinsic Propensity to Transition)

Segment 1			Segment 2		
t→ t+1	state1	state2	t→ t+1	state1	state2
state1	45.02%	54.98%	state1	48.47%	51.53%
state2	2.78%	97.22%	state2	12.29%	87.71%

The coefficients for lagged cumulative purchases represent the impact of learning experience on the transition probabilities. Positive coefficients imply that the probability of switching to the other state increases with experience, whereas negative coefficients imply a higher tendency to remain in the same state. The experience coefficient estimates are all significantly negative for both states in segments 1 and 2, indicating that purchase experience reinforces a customer’s current tendency and significantly increases the tendency of remaining in the same state.

I address the implications of the estimated purchase experience effects for the two segments (Tables 7 and 8) by examining the transitions for customers in each segment after they have made 5 and 30 purchases, respectively, to represent low and high customer experience. For segment 1, experience decreases the tendency to switch, and customers in the multichannel state (state 1) are more likely to remain so with purchase experience, while customers in the retail state (state 2) show a higher preference for staying in retail than for switching, regardless of category experience. These results reveal that customers approach a steady state with purchase experience. For segment 2, the online and multichannel state are essentially absorbing states; with purchase experience, customers in the multichannel state (state 2) are highly likely to remain so, and those in the online state (state 1) become less likely to switch to multichannel. The results show that all states for both segments are absorbing, and imply that purchase experience reinforce customers' channel tendency in various rates of development.

Table 7 Transition Matrix (5 cumulative purchases)

Segment 1			Segment 2		
t→ t+1	state1	state2	t→ t+1	state1	state2
state1	96.78%	3.22%	state1	89.04%	10.96%
state2	0.16%	99.84%	state2	2.46%	97.54%

Table 8 Transition Matrix (30 cumulative purchases)

Segment 1			Segment 2		
t→ t+1	state1	state2	t→ t+1	state1	state2
state1	99.94%	0.06%	state1	98.90%	1.10%
state2	0.01%	99.99%	state2	0.37%	99.63%

3.3.4.3. Customer Heterogeneity

“There is definitely some shift in spending from offline to online.” U.S.News and World Report (2008).

The statement above does not appear to reflect the reality of the empirical setting I studied. There is a tendency to engage in state switching from multichannel to retail in segment 1, and state switching between online and multichannel state in segment 2 when customers are new to the company and the product category. For both of the segments in my study, customers' current channel tendencies are enhanced with greater purchase experience with the company. In segment 1, the one with 77.13% of the customers, customers in retail state have a strong tendency to stay in retail only without switching to the alternative state, whereas customers with a multichannel tendency have roughly equal chances to remain the same or to migrate toward the retail state with little experience, and then show a tendency, enhanced by experience, to retain a preference for multichannel shopping. Customers in segment 2, which has 22.87% of the customers, switch between preferring to shop in brick-and-mortar stores and shopping online, and this switching propensity decreases with experience. Drawing upon these findings, it becomes an easy task for a firm to identify a customer's latent channel tendency at various levels of experience.

There are also differences between customers in the two segments in their reaction to marketing communications. It is important to understand the effect of marketing communications on channel choices, especially for customers in the multichannel state, in order to accurately

predict in which channel a customer is going to shop. For example, my results show that more marketing communications increase the propensity to shop online relatively more than retail for the multichannel customers in segment 1. For the multichannel state in segment 2, marketing communications offset part of online preference, and make the retail preference relatively higher.

3.3.4.4. Posterior Analysis

The model used in my study can help me identify which segment a customer is in at the end of an observation period, by calculating the posterior probability. To fully understand the differences across segments, I identify customer characteristics using aggregate purchase behavior measures for each segment through posterior analysis. The posterior probabilities of membership in segment l can be obtained by the following formula,

$$p_l = \frac{ss_l L_{nTl}}{\sum_{l=1}^2 ss_l L_{nTl}},$$

where p_l is the posterior probability of customer n being in segment l , ss_l is the relative size of segment l , and L_{nTl} is the likelihood of customer n 's purchasing history given membership in segment l . I then run a logistic regression for the segment, using as explanatory variables relationship length, RFM measures, frequency, recency, and monetary value, as well as a distance dummy variable called "intrade", which has a value of one if a customer lives within a 10-mile trading area of a retail store and zero if not. There are 84.91% of the customers who live within the 10-mile trading area. The dependent variable, segment, has a value of one if a

customer is in the first segment and zero if in the second segment. Richer interpretation and better identification of customer characteristics are possible if the posterior logistic analysis includes customer demographic information such as age, household income, gender, education, and so on, but this information was, unfortunately, missing in my data.

Table 9 Posterior Analysis

	Estimate	S.E.	
Intercept	0.3616	(0.3222)	
Intrade	1.1886	(0.1571)	***
Relationship length	0.0003	(0.0001)	**
Frequency	-0.0628	(0.0162)	***
Recency	0.0003	(0.0014)	
Monetary Value	0.0008	(0.0018)	

Posterior logistic analysis results (Table 9) shows that customers who live in the trading area are more likely to be in segment 1, and less likely to be in segment 2. The parameter estimates of the variable relationship length, represented by the duration since a customer's first purchase, imply that a customer with a longer relationship or who has joined the loyalty program earlier is more likely to be a segment 1 customer. The other variables I include in the analysis are recency, average monetary value, and frequency of purchases. These variables correspond to the RFM variables that are typically used to assess a customer's lifetime value. Segment 1 is the biggest segment, with 77.13% of the market. Customers in segment 1 tend to live closer to a brick-and-mortar store, purchase with less frequency, and to have longer relationships with the firm. The recency variable, which is captured by average duration between purchases, does not

make a significant difference for either segment, nor does monetary value. Customers in segment 2, with 22.87% of the total market, live far away from brick-and-mortar stores, join the program later, and purchase with higher frequency than customers in segment 1.

3.4. Conclusions and Directions for Future Research

In this study, I propose a modeling framework to address changes in customers' channel preference over time as a result of changes in tendency as their levels of experience increase. Theoretically, this study extends the marketing literature by connecting the indirect relationship between category experience and ultimate channel choices, and linking unobserved alternative evaluation, which forms a consideration set, with the observed purchase stage in the purchase decision process. This study utilizes the characteristics of HMM to discover the impacts of underlying learning or other unknown factors through the customers' purchase-related experience that governs the dynamics in channel tendencies. In addition, the proposed multiple-segment HMM extends the family of HMMs by providing an alternative way to deal with dynamic accumulation of experience by capturing cross-sectional heterogeneity in both state transition and channel choice behavior, through a discrete segment structure.

In the empirical application, I estimate a non-homogeneous multiple-segment HMM which incorporates time-varying covariates into the transition between the latent states, and uses observed channel choice and purchase information to make inferences about customers' underlying states of channel tendency. Compared with alternative models, the proposed approach

provides a better fit and predictions, which illustrates that it is important to consider over-time variation and cross-sectional heterogeneity when studying channel choice behavior. Also, this approach can provide better managerial insights by its ability to recover latent channel tendency and segment membership, which can have further influences on marketing strategies.

In addition to better prediction of future channel choices, this approach is able to depict the evolutionary pattern among possible discrete tendency states, and utilize full information to recover latent tendencies. The traditional generalized linear model (GLM) may do well for the former, but is lacking at the latter. Also, it is not appropriate to use a continuous dynamic structure such as a time-varying parameter model to capture dynamics in channel preference which is established in a discrete way.

The dynamic model of customer learning in a multichannel environment is estimated with data consisting of longitudinal transaction records from a large nationwide retailer in the U.S. In the empirical application, I identify two segments and two states and examine different patterns of experience development for each segment. The state represents customers' underlying channel tendency which cannot be observed directly. The empirical study offers substantive and interesting insights regarding multichannel shopping patterns. Customers in both segments show a tendency for customer with not much experience switching, and later migrating toward a steady state when they are more experienced customers. Also, there are clear differences in the patterns of customer channel preference evolution between the two segments. Segment 1 customers with multichannel tendency may switch to retail-loyal tendency, whereas segment 2

customers switch between an online and a multichannel state when they are not familiar with the product category. As experience increases, customers' channel tendency may reach steady states because the two states in each segment are absorbing. This is reasonable because customers may try various channels to get familiar with the transaction process, purchase atmosphere, and product category, when they are new to the company. Category experience will reinforce customers' channel choice behavior and help them move toward a particular tendency, i.e., single-channel or multichannel. In addition, customers in the two segments learn from experience at different rates, and show different reactions to marketing communications. Based on the empirical results and with information about a customer's level of experience and receipt of marketing communications, I forecast future channel choices with improved predictive ability.

From a practical perspective, marketers can measure consumers' channel preference and reactions to marketing communications more accurately by using my model to overcome the difficulties of observing customers' channel tendency and experience effects. In addition to predicting channel choices, this approach uncovers underlying channel tendencies with distinct properties that provide the basis of segmentation. Firms can use my proposed model to classify customers into various states dynamically and into particular segments probabilistically, and assess the dynamic learning effect on customer channel preference evolution across time. Based on empirically derived insights regarding customer channel preference with experience, marketers can allocate a firm's limited resources more effectively and further refine marketing strategies. For example, the channel tendencies in different segments can have different

representations, e.g., one segment with retail and multichannel tendency, and the other segment with online and retail/online tendency. Also, the channel tendencies at different levels of experiences are not consistent: a customer can be a retail/online shopper (switching between 2 channels) with low experience level, and become an online only shopper with increasing experience.

Accurately classifying customers into particular segments with a particular channel tendency (i.e., single-channel shopper, retail store or online, and multichannel shopper) for the stage of experience development is important for a firm to optimize current marketing efforts and develop further strategies, such as differentiation in pricing and promotions. The results of this study can be applied to allow a firm to compare costs and profits for different segments and determine which segments can provide the most profitability and therefore should be targeted. Also, marketers can access information about which channels customers in different segments (low or high value) prefer, and then provide those customers with fewer (or more) communications or services. Further, marketing strategies for single channel shoppers and multichannel shoppers can be distinguished. If a customer in the segment with strong online tendency has a small chance of migrating toward alternative channels, a firm can send more marketing communications regarding products and promotions online, and save costs by not sending communications regarding retail events, unless empirical results show that the customer responds positively to communications regarding retail events as well. In addition, the marketer can incorporate price into the model and see how customers in different segments and states

respond to the firm's pricing strategies, although I do not model this in this study because the price information among the three channels is not sufficiently detailed.

I believe that my model has research implications for customer relationship management in general, as well as in channel preference evolution in particular. Possible areas of future application may include the estimation of churn rate of specific services by bank or Internet phone card companies, and the examination of preference evolution of competing or complementary brands. All of these problems involve the evolution of latent states.

There are several limitations in my research. The current model only considers a single category, and does not account for learning spillover from purchase experience in other categories. The research question is simplified to examine the learning pattern in a single product category. It would be interesting to model the impact of multicategory learning on multichannel choice and compare the rates of experience development among different product categories. Another future extension to my study would be to apply the modeling approach to other industries and other product categories. Similar channel preference patterns might not result. Although I expect that some customers will tend to channel switch and others to prefer a particular channel with different levels of experience, various product categories may have different impacts of experience on channel tendency. Additionally, the current study only considers learning spillover from purchase experience in a single company, and does not account for experiences gaining from purchasing similar product categories in other companies and

retailers. Future extension to this study may investigate the impact of outside learning on the evolution of customer channel tendency.

In addition, the marketing communications sent by the firm in this study are all through direct mailings. Further study could test the impact of various communications media on the evolution process and channel preference. For example, marketing communications through emails may have different impacts than those through direct mailings. A further limitation of this study is that the information about the firm's communication is not clear and complete, so I do not know what kinds of communication were sent: flyers for new product introduction; notices of special events or annual sales; promotion coupons; catalog; membership reward information; and so on. Different marketing communication activities may have different impacts. Further research can address the impact of various types of marketing communications on the evolution of channel choice through customer learning using full information with the consideration of the costs associated with particular promotion activities. An ability to measure the impact sizes of specific promotion activities in each channel can help firms develop a segment specific communication strategy.

Future research can also try to use purchase patterns from other categories and individual level demographics to explain differences in channel choice pattern for each segment. Finally, the latent tendency states in my study depend only on cumulative purchase. Future research should investigate other aspects of the learning process to fully understand the impact of

dynamic learning on multichannel choice. Thus, while my study has revealed many insights regarding multichannel choice patterns in a particular setting, there is more to be learned.

Chapter 4

Essay 2: The Effects of Varying Channel Experiences and Direct Marketing on Customer Retention

4.1. Introduction

In recent years, a multichannel environment is becoming increasingly prevalent: more firms may reach customers through different channels, and customers have more alternative channels for selecting and buying than ever before. For example, firms and their customers can interact via brick-and-mortar stores, catalogs, telephones, emails and virtual stores. Multichannel marketing has become a critical means to motivate customers to shop more frequently through a variety of channels (Hansotia and Rukstales 2002) and to build lasting customer relationships (Hansotia and Rukstales 2002; Rangaswamy and Van Bruggen 2005). Additionally, in an increasingly integrated multichannel environment, firms are looking for a variety of ways to increase customer retention rates and avoid customer churn because the costs of customer acquisition are extremely high. Also, small increases in retention drive large increases in profits (Gupta and Lehmann 2003; Pfeifer and Farris 2004; Reichheld and Sasser 1990). Gupta and Lehmann (2003) assert that an increase in retention of 5% yields a dramatic 22% to 37% increase in lifetime value. Customers with greater loyalty have a higher share of wallet, produce more profits, and have longer retention durations. That is, the links between multichannel operations and customer retention remain important to marketers. Multichannel issues in the context of customer relationship management (CRM) – how to identify the most valuable customers, how to maintain

long-term customer relationships through a variety of channels, and how to increase loyalty and estimate customer retention across channels – should be emphasized and well-studied.

Customer retention and churn have received increasing research attention in recent years. Researchers want to make more accurate predictions of customer lifetime value (CLV), which requires accurate estimates of retention rate, and knowing how and when a customer terminates his relationship with a firm instead of assuming that he stays with the firm for life. Such research is harder for firms in non-contractual service settings because the termination of relationships is difficult to observe, and thus retention rates used to calculate CLV are not easily evaluated. In addition, firms have an interest in identifying which customers are inactive and not valuable, and to reduce their marketing costs by removing these customers from communication lists. Previous research related to multichannel marketing has explored the drivers of multichannel choices (Balasubramanian et al. 2005; Hansotia and Rukstales 2002; Neslin et al. 2006; Rangaswamy and Van Bruggen 2005; Schoenbachler and Gordon 2002) and whether multichannel customers provide higher revenues and higher share of wallet to a firm than single-channel customers (Kumar and Venkatesan 2005). This prior research has not estimated retention probabilities or the effects of multiple drivers on customer retention.

Most of the prior research on customer retention has focused on the importance of increase in retention to customer lifetime value (Gupta and Lehmann 2003; Pfeifer and Farris 2004; Reichheld and Sasser 1990) through sensitivity analysis, or by attempting to estimate the impacts of various factors on relationship length in a contractual setting (Boehm 2008; Schweidel et al. 2008). Boehm (2008) does not provide a way to estimate retention in a customer's lifetime stream, and Schweidel et al. (2008) require accurate information about termination of memberships. There are some studies on the impact of Internet use (specifically,

online banking) on customer retention (Boehm 2008; Campbell and Frei 2010; Hitt and Frei 2002; Verhoef and Donkers 2005). Those studies are based on contractual settings to estimate the effects of multiple factors including Internet use on membership or service termination. Overall, the various impacts of multichannel experiences and direct marketing on customer retention in non-contractual settings have not been well explored.

These issues are hard to resolve because of the difficulty in identifying a customer's true relationship with a firm. Customers may seem inactive when they have not actually terminated the relationship with a firm, and customers with longer purchase durations, need to be distinguished from customers who are truly inactive. Previous research that attempted to derive customer retention from an "aggregate" relationship with a non-contractual firm has not accounted for what or through what channel a customer purchases. For example, a ten-purchase customer who makes seven purchases in a retail store and three purchases via the Internet should be distinguished from a customer who makes five purchases via the Internet and five purchases in a retail store. Customer experience from various channels should have different influences on a customer's relationship state. The Pareto/NBD and BG/NBD models, which explore repeat purchase behavior while accounting for unobserved customer dropout rates in the field of CRM, ignore the possibility that customers who shop in a particular channel may have different retention/churn probabilities than those who shop in alternative channels, and do not allow for a customer coming back to an active relationship once identified as inactive. Few research studies have explored the relationship between choice preference and relationship dynamics in non-contractual multichannel environments. Little is known about whether different channel experiences have different impacts on customer-firm relationships, and which patterns of evolution are more likely to reinforce a customer's relationship with a firm.

The purpose of this research is to investigate which channels can build longer and more lasting relationships with customers, examine the impact of channel experiences and direct marketing on customer retention, and identify customers' relationship states that govern purchase dynamics. I examine both important decisions – purchase incidences and channel choices – simultaneously, because customers need to decide whether to make a purchase and where to buy in the multichannel environment, and then recover the underlying relationship state which represents a customer's tendency to stay active or drop out, but which cannot be directly observed or easily measured. What researchers can observe is a sequence of state-dependent decisions that rely on the hidden relationship states. The Hidden Markov Model (HMM) is an approach that can link the underlying and unobserved sequence of relationship states with observed outcomes of purchase incidences and channel choices. Relationship states are not explicitly defined a priori; rather, HMM provides a flexible structure that automatically classifies of evolution patterns into groups and translates to particular patterns in choice preference.

The state-dependent choice behavior is defined such that the observed choice behavior at time t is determined by the unobservable relationship state at time t , and the relationship state over time follows a Markov chain. A Markov transition matrix links the unobserved states with the Markov property, which means the relationship state at time t depends on the state at time $t-1$; the current state then affects current purchase incidence and channel choice behavior. I integrate those two choices with the unobservable relationship states by using a nested multinomial HMM and examine how nested choice preference evolves as a customer-firm relationship changes as a result of customer-firm interactions, i.e., channel-related experiences and marketing communications. The non-homogeneous HMM with finite Markovian states developed in this research captures the variation in customers' purchase incidence and channel preference while

recovering the relationship dynamics. The multiple-segment and nested multinomial HMM also incorporates heterogeneity through a discrete segment structure. It also allows me to first identify a customer's true state of relationship over time, second, determine the retention probability affected by channel experience, third, identify customers who are more likely to churn by the end of observation period, and finally, examine the impact of alternative channel experiences and direct marketing on moving customers toward various relationship states.

I apply the model to a dataset from a large multichannel retailer. The empirical application identifies one segment with two relationship states, corresponding to inactive and active states, respectively. Retail store-related experience seems to lessen the probability of being inactive, and encourage customers to stay in or migrate toward an active state. Generally speaking, retail and online experience increase customer retention, whereas catalog experience produces asymmetric patterns for the two states. Further, retail store is the most active-oriented channel when a customer is in an inactive state, whereas online channel has the greatest power to reinforce customer relationships and encourage customers to stay in an active state when a customer is currently in an active state. Mail order-related experience does not necessarily help slow down the probability of moving toward inactive state. In addition, the impacts of channel experiences and reactions to marketing communications on purchase incidence and channel utility vary among states.

This study makes several contributions. First, it provides a way to understand customer retention in a non-contractual and multichannel environment. Second, the proposed approach provides a framework to estimate the relationship dynamics in a non-contractual setting where a customer's dropout time is hard to observe and not clearly stated. It suggests an effective way for a firm to dynamically manage life-long relationships with its customer base. Third, I extend the

current research on customer retention by accounting for the diversity in the impact of channel experience and direct marketing on retention probabilities. Fourth, I address changes in preference of purchase incidence and channel choice across time with respect to various relationship states, and deal with the impact of channel experiences on customer retention as governed by transitions between relationship states. Fifth, I investigate temporal-based evolution with a few latent states that represent a set of customers' tendency to retain or terminate a relationship with a firm. Finally, the research will help firms more effectively allocate resources to those channels that lead to longer relationships with customers, develop differentiated marketing strategies to motivate customers to migrate toward a particular channel, and discourage customers from shopping through inactive state-oriented channels.

The rest of the essay is organized as follows. The structure of multiple-segment and nested multinomial HMM is introduced in Section 4.2. In Section 4.3, I describe the empirical application of my proposed model in the context of customer-firm relationships using panel data from a large multichannel retailer. In this section I present the data collection process, variables used to calibrate the model, the model selection procedure, and the empirical results. Section 4.4 is a counterfactual analysis that provides insights into customer-firm relationships by simulating marketing communications strategies. In Section 4.5, I discuss theoretical and practical contributions, and conclude with limitations of this study and directions for future research.

4.2. Model Development

Retention/churn modeling has been used in the field of CRM to address issues in contractual settings such as mobile services, television, financial services, and Internet subscriptions, where

the timing of relationship termination is clearly observed. The models used to estimate and predict time of dropout with limited information for a non-contractual service are the Pareto/NBD and BG/NBD models, which only allow for two states (“active” and “inactive”) and assume customers do not switch back to active relationship once they have become inactive. Two of the RFM measures (i.e., recency and frequency) are sufficient statistics to implement either approach for predicting future demand, regardless of what and where a customer buys. Both models ignore the varying impacts of important drivers, specifically channels and direct marketing, on customer retention in multichannel environments. For example, customers that are accustomed to shopping in a single channel may have different retention probabilities than customers who shop in multiple channels. As previous literature asserts, multichannel shoppers are more active (Kumar and Venkatesan 2005), and have lower churn rates and higher propensity to buy more (Stone et al. 2002). Internet use may have a positive effect on customer retention (Verhoef and Donkers 2005, Campbell and Frei 2010, Boehm 2008, Hitt and Frei 2002). Also, different purchasing patterns across channels should be accounted for, e.g., a customer who makes seven purchases in a retail store and three purchases via the Internet should be distinguished from a customer who makes six purchases via the Internet and four purchases in retail store, although both make a total of ten purchases. My proposed model deals with these issues by incorporating the interactions between a customer and a firm in a given relationship state, which governs the dynamics of purchase behavior.

Most importantly, the relationship pattern that governs purchase dynamics is unobserved, but can be defined as sets of hidden states which vary with respect to the magnitude of activity. The HMM can overcome the limitations I mentioned above, account for varying impacts of channel experience on customer retention, and link underlying relationship states with observed

outcomes of choices. Further, it can account for dynamics in purchase behavior through the evolution of a relationship, and let dynamic changes in a customer relationship follow a Markov process. Various choice preferences result from different relationship states, which are updated each period. Given the state-dependent property of choices, and the fact that marketers can observe a customer's purchase behavior but not the relationship pattern, HMM can model the structure of latent states and observed choice behavior, and so is able to relax the limitation of static choice and capture dynamics. A latent state is a measure of a customer-firm relationship status, and the transition between states is determined by interactions between a customer and a firm. In addition, HMM does not have a two-state restriction but can estimate the number of latent relationship states based on the dynamics in the data, and it allows flexibility in transitions. I incorporate time-varying covariates because the resulting Markov chain then has a useful substantive interpretation (MacDonald and Zucchini 1997). Therefore, the Markov chain in my model is non-homogeneous because it is updated by time-varying covariates, i.e., channel experience and marketing communication. Estimating the parameters for those covariates helps explain the impact of channel experience and direct marketing on transitions because the probabilistically determined transitions between states are affected by those covariates, which represent relationships. The proposed model accounts for heterogeneity in that it allows customers to have different retention rates and choice preferences through a discrete segment structure. In addition, the nested structure HMM in the state-dependent choice distributions proposed here consists of repeated binary choices and multiple channel choices. Not only can the model capture preference evolution from active to inactive states but it also captures channel migration while in the active state.

The state-dependent choice distribution, Markov chain transition matrix, and the initial state distribution are the three main components of HMM, which are described in detail in the following subsections.

4.2.1. State-dependent Choice Distribution

Customers are involved in two decisions each period: whether to make a purchase, and where to purchase (e.g., retail store, online, or catalog). The two choices of purchase incidence and channel selection depend on a customer's relationship state with the firm. I model the state-dependent choice probabilities with the nested multinomial logit structure, which allows me to account simultaneously for purchase incidences and channel choices. The two choices the n^{th} customer makes in time period t given state $S_{nt} = i$ are defined as follows.

$$Y_{nt}|(S_{nt} = i) = Y_{nt|i} = (B_{nt|i}, C_{nt|i}),$$

where $B_{nt|i}$ equals 1 if the n^{th} customer makes a purchase, and zero otherwise. $C_{nt|i}$ is the channel choice. The nested structure can be divided into two parts: purchase probability, $P(B_{nt|i})$, and channel choice probability conditional on purchase, $P(C_{nt|i}|B_{nt|i})$. Therefore, the joint probability of purchase incidence and channel selection can be represented as follows.

$$P(Y_{nt|i}) = P(Y_{nt}|S_{nt} = i) = P(C_{nt|i}|B_{nt|i})P(B_{nt|i}).$$

4.2.1.1 Conditional Channel Utility

Consider first the channel choice conditional on the purchase. The random utility of channel choice, which include deterministic and random component, is

$$U_{nt|i}^v = \alpha_{iv} + X_{nt}\beta_{iv} + \varepsilon_{iv}, \quad v \in \{1, \dots, C\}, \quad (7)$$

and it is specified as a multinomial logit model by utility maximization as follows.

$$P(C_{nt|i} = v | B_{nt|i}) = \frac{\exp(\alpha_{iv} + X_{nt}\beta_{iv})}{\sum_{c=1}^C \exp(\alpha_{ic} + X_{nt}\beta_{ic})}, \quad i = 1, \dots, m, \quad (8)$$

where α_{iv} is the state-specific intrinsic utility of channel v in state i , $\forall v \in \{1, \dots, C\}$, X_{nt} is a vector of explanatory variables that are common across channels for customer n at t , and β_{iv} is a vector of the state-and-channel-specific coefficient of variables X_{nt} for channel v in state i . The term $P(C_{nt|i} = v | B_{nt|i})$ represents the probability that customer n chooses channel v given state i at t while making a purchase. The time-varying covariates X_{nt} for customer n at t should consist of variables that have immediate impact on a customer's channel choice.

4.2.1.2. Purchase Probability

Next, consider the purchase probability of customer n at time t in state i , $P(B_{nt|i})$. Assume the utility of purchasing R for customer n at t in state i is as follows.

$$R_{nt|i} = \delta_i + Z_{nt}\gamma_i + e_{nt}, \quad (9)$$

where δ_i and γ_i are unknown parameters, and Z_{nt} is a vector of covariates which contribute to the purchase decision incidence, such as household characteristics. A customer will make a purchase if and only if the maximum channel utility is greater than his utility of not purchasing.

The inclusive value for purchasing which is the expected maximum utility of making a purchase at t in state i , is defined as follows.

$$CV_{nt|i} = \ln \sum_{c=1}^C \exp(U_{nt|i}^c)^2. \quad (10)$$

Therefore, the purchase probability is

$$P(B_{nt|i} = 1) = \frac{\exp(\delta_i + Z_{nt}\gamma_i + CV_{nt|i}\xi_i)}{1 + \exp(\delta_i + Z_{nt}\gamma_i + CV_{nt|i}\xi_i)}, \quad 0 \leq \xi \leq 1, \quad (11)$$

where the parameter ξ_i is restricted to be one to get the non-nested model³, and unrestricted to allow some degree of heteroscedasticity. Thus, I can denote the unconditional channel choice probability as follows.

$$P(C_{nt|i} = v) = P(C_{nt|i} = v | B_{nt|i} = 1) P(B_{nt|i} = 1) + 0 \times P(B_{nt|i} = 0). \quad (12)$$

Therefore, the state-dependent choice function for customer n at shopping occasion t in state i is

$$P(Y_{nt|i}) = [1 - P(B_{nt|i} = 1)]^{\omega_{nt}^0} \prod_{c=1}^C P(C_{nt|i} = c)^{\omega_{nt}^c}, \quad (13)$$

where $\omega_{nt}^0 = \begin{cases} 1, & \text{if a customer } n \text{ does not make a purchase at } t \\ 0, & \text{otherwise} \end{cases}$,

$\omega_{nt}^c = \begin{cases} 1, & \text{if a customer } n \text{ makes a purchase through channel } c \text{ at } t \\ 0, & \text{otherwise} \end{cases}$.

² Greene (2003).

³ With $\xi_i = 1$, (5) reverts to a basic multinomial logit model (Greene 2003).

For the purpose of state identification, I restrict the intrinsic purchase probability (δ_i) to be non-decreasing, thus $\delta_1 \leq \delta_2 \leq \dots \leq \delta_m$.

The covariates in the nested multinomial structure are discussed in detail in the Empirical Application section.

4.2.2. Markov Chain Transition Matrix

Unlike the Pareto/NBD and BG/NBD models that consider only two states, the HMM model does not have an a priori restriction on the number of states. Also, it allows for customers staying in a current state or moving to any other state, by estimating a full and flexible transition matrix. In addition, it does not impose an a priori restriction that customers attain a steady state after the first transition. Given m states, I assume the transition matrix $Q(i_{t-1}, i_t)$ is defined as follows.

$$Q(i_{t-1}, i_t) = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1m-1} & q_{1m} \\ q_{21} & q_{22} & \cdots & q_{2m-1} & q_{2m} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ q_{m1} & q_{m2} & \cdots & q_{mm-1} & q_{mm} \end{bmatrix}, \quad (14)$$

where $q_{jk} = P(i_t = k | i_{t-1} = j)$ denotes the transition probability from state j at $t-1$ to state k at t , and $\sum_{k=1}^m q_{jk} = 1$, $0 \leq q_{jk} \leq 1$ for all $j, k = 1, \dots, m$.

This study attempts to explore the evolution of the customer-firm relationship. The relationship can be ordered by the propensity to move from inactive to active, and the transitions of the order-ranked states are affected by those activities that influence customers' relationships

with a firm, e.g., interactions between customers and the firm, such as purchase experiences and marketing communications. Therefore, I model the transition matrix as an ordered logit model (Greene 2003). The elements in the transition matrix $Q(i_{t-1}, i_t)$ can be defined as follows.

$$q_{j,1} = \frac{\exp(\mu_j^1 - \theta_j A_{nt})}{1 + \exp(\mu_j^1 - \theta_j A_{nt})}, \quad (15)$$

$$q_{j,2} = \frac{\exp(\mu_j^2 - \theta_j A_{nt})}{1 + \exp(\mu_j^2 - \theta_j A_{nt})} - \frac{\exp(\mu_j^1 - \theta_j A_{nt})}{1 + \exp(\mu_j^1 - \theta_j A_{nt})}, \quad (16)$$

$$q_{j,k} = \frac{\exp(\mu_j^k - \theta_j A_{nt})}{1 + \exp(\mu_j^k - \theta_j A_{nt})} - \frac{\exp(\mu_j^{k-1} - \theta_j A_{nt})}{1 + \exp(\mu_j^{k-1} - \theta_j A_{nt})}, \quad (17)$$

$$q_{j,m} = 1 - \frac{\exp(\mu_j^{m-1} - \theta_j A_{nt})}{1 + \exp(\mu_j^{m-1} - \theta_j A_{nt})}, \quad (18)$$

$$\text{for } j \in \{1, \dots, m\}, k \in \{2, \dots, m-1\}, \mu_j^1 < \mu_j^2 < \dots < \mu_j^{m-1} < \mu_j^m,$$

where θ_j is a vector of parameters for the impact of marketing and experience on transitions from state j , A_{nt} is the vector of time-varying covariates for customer n between time $t-1$ and time t , μ_j^k is the threshold value to a more active state ($k \geq j$) or a more inactive state ($k < j$) for a customer in state j . Note that the parameters in the transition matrix are state-specific.

4.2.3. Initial State Distribution

The initial state distribution can be defined as the stationary distribution of the transition matrix for homogenous HMM. In this study, the transition matrix in the proposed model is a function of

time-varying covariates, so I estimate the individual initial probability as do Netzer et al. (2008). Define π_n as a vector of initial probabilities for a customer n ($\pi_n = (\pi_{n1}, \pi_{n2}, \dots, \pi_{nm})'$) and the initial state distribution as calculated by solving the following equation.

$$\pi_n = \pi_n Q_n, \sum_{i=1}^m \pi_{ni} = 1, \quad (19)$$

where Q_n is the transition matrix with all covariates set to their mean values for each customer across time periods.

The model specifications of multiple-segment HMM in (7)-(19) are used to develop a likelihood function, which is outlined in the Appendix.

4.3. Empirical Application

The objective of this study is to show temporal-based preference evolution with a few latent relationship states, examine the impact of choice dynamics on customer retention, find out (in)active-oriented channel experiences, and predict future demand. I apply my nested multinomial HMM to data on observed purchase incidence and channel choice for clothing purchases by customers of a multichannel retailer. My model is well suited for the data in a non-contractual setting, which include well-recorded repeated purchases and individual customer exposure to marketing communications, and to exploring the customer-firm relationship, which is not observed in the data but which appears to be an important determinant of purchase behavior.

4.3.1. Data

The data for this study is provided by a large multi-channel retailer which has a strong customer service and relationship orientation. Empirical estimation and evaluation of the model are done with customer membership data from the company. In addition to brick-and-mortar stores, the company has online and catalog sales channels. Customer transaction information from multiple channels is captured and integrated. Thus, the company's customer relationship system can produce a complete purchase history for a particular member customer. The data integration process that acquires information from customer interactions through every channel is a critical aspect of the company's CRM success.

The dataset include complete purchase history over a seven-year period from December 1998 to July 2005, and marketing communication records from November 2002 to July 2005. To incorporate the effect of marketing communications on choice dynamics, the dataset for this empirical application is truncated to the period from December 2002 to July 2005. I keep only data that involved transactions within the clothing category in order to exclude multi-category effects on relationship evolution. I choose customers who made their first purchase before December 2002, and purchased at least once during the calibration period of December 2002 to December 2004. This provides a complete monthly history for 59,498 customers. The data contains monthly information about purchase incidence, channel choice, and number of marketing communications sent to a customer, from December 2002 to July 2005. I randomly choose 1% of the customers (595 individuals), and use the observations from December 2002 to December 2004 (25 months) to calibrate the model, and the observations from January 2005 to July 2005 (seven months) for validation. The data for the 595 customers have 14875 observations in the calibration period, and 4165 in the validation period. I calculate the mean of

the proportion of making a purchase per individual in the calibration period sample (34.42%), and in the holdout period sample (25.57%) and other descriptive statistics (Table 10).

Table 10 Descriptive Statistics

Descriptive Statistics	
Calibration Sample (12/2002~12/2004)	
Number of Observations	14875
Number of Customers	595
Mean of Purchase Probability	34.42%
Holdout Sample (01/2005~07/2005)	
Number of Observations	4165
Number of Customers	595
Mean of Purchase Probability	25.57%

4.3.2. Variables

The variables for state-dependent choice behavior in (8) and (11) include Z_{nt} and X_{nt} , which are distinguished from the variables constituting A_{nt} for the transition relationships in (15) to (18). The vector A_{nt} is the set of variables which impact the transition probabilities and which are assumed to have an enduring impact on customer retention and the customer-firm relationship. Vectors X_{nt} and Z_{nt} are assumed to affect only short-term choice behavior. Vector Z_{nt} influences the state-dependent purchase incidences, whereas vector X_{nt} influences the state-dependent channel choice probabilities. In my empirical application, I view marketing communications as a variable with both immediate and lasting impact, and channel experience updated each period as a variable with enduring impact on evolution and immediate impact on choice.

4.3.2.1. Choice Behavior

I consider two kinds of observable choice behavior – purchase incidence and channel choice – as dependent variables. I study the nested structure consisting of incidence of purchase (0 or 1) and three channels (brick-and-mortar retail store, online, and catalog). The variable ($C_{nt|il} = v$) shown in equation (2) with $v = \{1,2,3\}$ represents an individual n making a purchase in time period t given state i and segment l through brick-and-mortar retail store, online, and catalog, respectively.

4.3.2.2. Variables Affecting the Transition Matrix

Prior research has asserted that multichannel experiences impact customer relationships with firms (Kumar and Venkatesan 2005; Stone et al. 2002). Previous research has also found that customer experience seems to become more important as a relationship persists (Verhoef and Donkers 2005). Thus, channel-related experiences may impact customer-firm relationships, and further, customers who usually shop in a particular channel may have different drop out and retention rates than those who usually shop in another channel. In this study, I assume that the primary driver of the relationship state is a customer's channel-related experience. Also, marketing communications may have both immediate and long-term effects, and can be considered as one of the explanatory variables in the transition matrix. My proposed model is a discrete HMM that requires the latent relationship states to be discrete. When the impact of channel-related experience and marketing communications has accumulated to a certain level, customers are likely to change their current relationship state to a more active or inactive state. I specify the relationship state as a function of marketing communications and cumulative purchases associated with each channel, and define A_{nt} from (15) to (18) as follows.

$Retail_exp_{nt}$ = cumulative retail store purchases made by customer n by time $t-1$

$Online_exp_{nt}$ = cumulative online purchases made by customer n by time $t-1$

$Catalog_exp_{nt}$ = cumulative catalog purchases made by customer n by time $t-1$

$marketing_{nt-1}$ = the number of marketing communications sent to customer n at time $t-1$

4.3.2.3. Variables Affecting Purchase Incidence

The variables in state-dependent choice behavior include Z_{nt} and X_{nt} , which are shown in equations (8) and (11). The vector Z_{nt} influences the state-dependent purchase probabilities, whereas the vector X_{nt} influences the state-dependent channel choice probabilities. Marketing communications and channel experiences which may have immediate impacts on a customer's decision whether to buy are included in Z_{nt} to allow examination of the immediate influences on purchase incidences. I define Z_{nt} as follows.

$Retail_exp_{nt}$ = cumulative retail store purchases made by customer n by time $t-1$

$Online_exp_{nt}$ = cumulative online purchases made by customer n by time $t-1$

$Catalog_exp_{nt}$ = cumulative catalog purchases made by customer n by time $t-1$

$marketing_{nt}$ = the number of marketing communications sent to customer n at time t

4.3.2.4. Variables Affecting Channel Choice

The firm in my study routinely conducts marketing communications through direct mailings to every member, by sending new product information, flyers, and promotions and events notices. I

investigate whether the number of communications sent to a customer per month impacts channel selection. I also consider channel experiences as one of the explanatory variables that have an immediate influence on channel selection. The vector X_{nt} , which affects channel choices, is represented as follows.

$marketing_{nt}$ = the number of marketing communications sent to customer n at time t

$Retail_exp_{nt}$ = cumulative retail store purchases made by customer n by time $t-1$

$Online_exp_{nt}$ = cumulative online purchases made by customer n by time $t-1$

$Catalog_exp_{nt}$ = cumulative catalog purchases made by customer n by time $t-1$

4.3.3. Estimation Procedure and Model Selection

The parameters for the nested multinomial HMM are estimated by maximum likelihood estimation (MLE), which is accomplished through numerical optimization in GAUSS. The Bayesian Information Criterion (BIC), which is used to select the number of segments and states for my proposed model, is represented as follows.

$$BIC = 2 * LL - Npara * \log(Nobs)$$

where LL is the log-likelihood, $Npara$ is the number of parameters, and $Nobs$ is the number of observations.

I compare the proposed model to a basic multinomial logit without heterogeneity and evolving customer preference, and to HMM with no segment specification. I compare performance using the holdout sample log-likelihood in addition to BIC (Table 11). Based on the measures of BIC and holdout log-likelihood, the one-segment two-state HMM is the best-fitting

model among the multiple-segment nested multinomial HMMs, and it also outperforms other benchmark models.

Table 11: Selecting the Number of Segments and States & Model Comparison

Model	Number of Segments	Number of States	BIC	Holdout log- Likelihood
Logit	1	1	22875.14	-2953.22
	1	2	20735.59	-2553.55
HMM	1	3	20853.59	-2556.63
	1	4	21150.69	-2554.87
Multiple-segment	2	2	20961.31	-2558.52
HMM	2	3	21405.11	-2557.57

4.3.4. Estimation Results

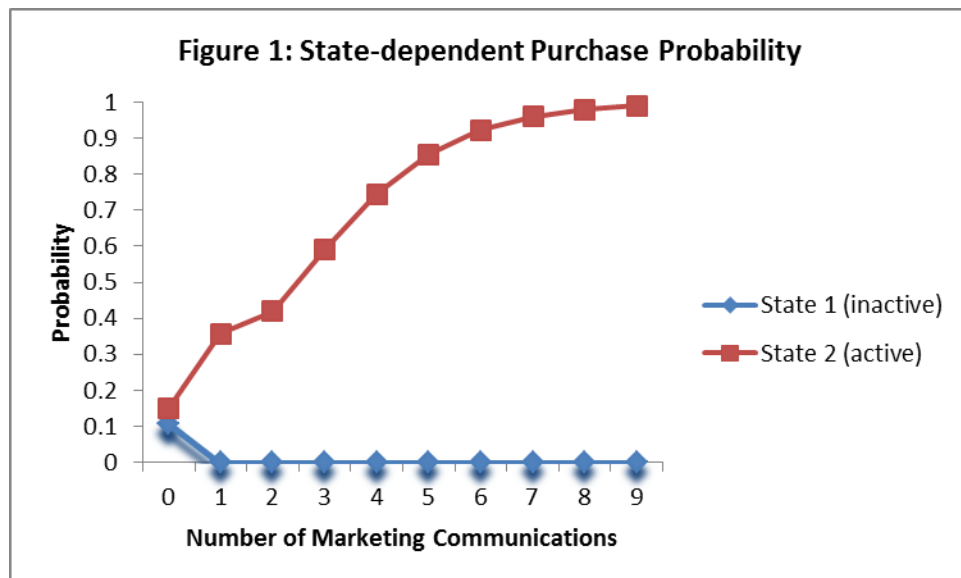
Table 12 shows the estimated parameters and corresponding standard errors for the one-segment two-state nested multinomial HMM. State-specific probabilities in purchase and (un)conditional channel choice are calculated by plugging these estimates back into equations (8), (11), and (12). I can get a better understanding of customer-firm relationship state with channel experiences and marketing communications based on state transitions by plugging these estimates back into equations (15)-(18).

Table 12 Estimated Parameters

	State1		State2	
	Parameter Estimates		Parameter Estimates	
Purchase Probability				
δ	-2.4690	(0.092)	-2.0108	(0.099)
CV	0.4989	(0.005)	0.3901	(0.009)
Retail_exp _{nt}	1.7498	(0.104)	0.4012	(0.049)
Online_exp _{nt}	0.2029	(0.046)	1.2396	(0.079)
Catalog_exp _{nt}	0.4724	(0.027)	1.0640	(0.120)
marketing _{nt}	-9.6490	(0.336)	0.7006	(0.092)
Channel Utility				
Retail	0.8982	(0.043)	2.3539	(0.120)
Online	-0.2316	(0.030)	0.3319	(0.164)
marketing _{nt} _retail	-2.6267	(0.107)	0.0254	(0.035)
marketing _{nt} _online	-0.5419	(0.020)	0.0898	(0.041)
Retail_exp _{nt} _retail	1.1639	(0.059)	0.8565	(0.049)
Retail_exp _{nt} _online	0.5620	(0.040)	0.2867	(0.044)
Online_exp _{nt} _retail	-0.2857	(0.025)	0.0290	(0.075)
Online_exp _{nt} _online	0.9198	(0.037)	1.4254	(0.076)
Catalog_exp _{nt} _retail	-1.2646	(0.106)	-1.6572	(0.078)
Catalog_exp _{nt} _online	-0.9457	(0.077)	-1.3646	(0.086)
Transitions				
μ	1.4705	(0.082)	0.9655	(0.089)
Retail_exp _{nt}	0.2397	(0.026)	0.1884	(0.032)
Online_exp _{nt}	0.0628	(0.032)	0.2376	(0.039)
Catalog_exp _{nt}	0.0695	(0.045)	-0.1012	(0.042)
marketing _{nt-1}	-0.0628	(0.043)	-0.1057	(0.079)
	state 1		state 2	
mean of initial probabilities	0.7087		0.2913	

4.3.4.1. Heterogeneity in Relationship States

The relationship states which reflect different degrees of activity can be interpreted by examining the state-specific intrinsic propensity to purchase δ (Table 12). With covariates set to zero, the purchase probability is 10.8% in state 1 and 15.7% in state 2. I then examine the intrinsic propensity to purchase calculated at the mean of the covariates: the probability in state 1 is close to zero, and the probability in state 2 is 97%. Figure 1 shows different patterns between the two states in purchase probability with an increasing number of marketing communications. State 2 shows high purchase probability when the number of marketing communications is increasing, whereas state 1 shows extremely low purchase probability no matter how many marketing communications. Therefore, state 1 can be labeled as an inactive state, while state 2 is definitely an active state.



The estimates of experience parameter (Table 12) indicate differences in the reaction to channel experiences across the two states. Although all three channel experiences help increase the propensity to purchase for both relationship states, retail experience is the most influential in state 1. In state 2, online experience helps increase the propensity to purchase the most, followed by catalog experience.

In the nested structure, channel choice is on the second layer of decisions if a purchase indeed occurs. I calculate the conditional probability of channel choice with all covariates set to zero in order to examine the intrinsic channel utility conditional on purchase across states. For an inactive state, the conditional probability of choosing a retail store is 57.79%, online is 18.67%, and catalog is 23.53% given purchase. For an active state, the conditional probabilities are 81.47%, 10.78%, and 7.74% for retail, online, and catalog, respectively. With all covariates at their mean values, the conditional probability of choosing a retail store is 1.26%, online is 50.25%, and catalog is 48.49% in the inactive state, while the conditional probabilities given an active state are 90.73%, 7.90%, and 1.30% for retail, online, and catalog, respectively. A customer in an inactive state (state 1) is more likely to make a purchase through multiple channels instead of a specific channel when a purchase occurred, whereas a customer in an active state (state 2) is more likely to make a purchase at a retail store. Thus, customers in state 1 prefer multichannel shopping, and customers in state 2 prefer shopping at a retail store.

Furthermore, the immediate effects of marketing communications and channel experiences on channel choice vary between states. I examine the number of communications sent to a customer per month and channel experiences to see their impacts on the selection of channels in which a customer decides to buy. Marketing communications make the catalog channel more favorable to customers in state 1 than retail store and online channels, while

making the online channel more favorable to customers in state 2. In state 1 (inactive state), retail store experience increases the propensity to purchase through retail store and online channels, online experience increases the probability of online shopping but decreases the probability of retail store shopping, and catalog experience does not increase the propensity to purchase through either retail store or online channels. In state 2 (active state), retail store experience increases the propensity to purchase through retail store and online channels, online experience increases the probability of online shopping but does not have a significant impact on retail store shopping, and catalog experience helps increase only the probability of catalog shopping.

4.3.4.2. State Transitions

The parameter μ_j (Table 12) represents the threshold between inactive state (state 1) and active state (state 2). The sign and absolute value of the threshold parameter imply how easily a customer moves from state 1 to state 2. The larger the value of the threshold, the less likely is a jump from inactive to active, and the more likely is a jump from active to inactive. A negative threshold value implies that it is easy to pass the threshold from inactive to active, and thus that a customer is more likely to remain active when already in the active state, or switch toward active when in the inactive state. A positive threshold value implies that a customer is more likely to switch toward inactive or remain inactive. A customer's intrinsic propensity to transition can be calculated by determining the threshold parameter μ_j with marketing and experience covariates set to zero (Table 13). In my application, the thresholds for each state are all positive, which represents the same information as intrinsic propensity to transition. The likelihood of staying in the inactive state is 81.31%, and the likelihood of moving from active to inactive state is 72.42% when impacts of channel experiences and marketing communications are not considered. It

means that the probability of moving toward active for an inactive state (state 1) and the retention rate for an active state (state 2) are 18.69% and 27.58%, respectively. When experience and marketing impacts are not taken into account, a customer in the inactive state has a higher intrinsic probability of remaining inactive than of migrating toward an active state, whereas a customer in the active state has a higher intrinsic probability of moving toward an inactive state than of staying active.

Table 13 Transition Matrix (Intrinsic Propensity to Transition)

$t \rightarrow t+1$	inactive	active
inactive	81.31%	18.69%
active	72.42%	27.58%

The coefficients for channel experiences and marketing communications represent the impact on transitions of, respectively, retail store, online, and catalog experiences, and marketing communications received in the previous month. The sign of the coefficients implies whether channel experiences and marketing communications help a customer remain or move toward an active state, and the value of the coefficients implies the magnitude of the influence. A negative coefficient means that the channel experiences or marketing communications accelerate the probability of being inactive; but positive coefficient means that the parameter helps a customer remain in or move toward the active state. In the inactive state (state 1), retail experience significantly increase the probability of an inactive customer moving toward an active state (state 2). In the active state (state 2), retail and online experiences increase the probability of an active customer remaining active, while catalog experience decreases the probability of retaining an active relationship with the firm. In general, retail and online experience increase the probability

of being active for both active and inactive customers. Catalog experience does not have a significant impact on relationship transitions for inactive customers, and further accelerates the probability of being inactive for an active customer. In addition, marketing communications do not have significant enduring effects on transitions toward the active state.

I explore the implications of the estimated channel experience and marketing effects (Table 14). The transition probabilities are calculated at the mean of the covariates. The propensity for entering active state (27.08%) becomes higher than the intrinsic propensity to transition (18.69% – see Table 13) when a customer’s current state is inactive, while the propensity for staying active increases to 33.58% from 27.58% when a customer is currently in an active state. Examining the parameter estimates, I find that retail experience indeed accelerates the rate of being active in both relationship states, and online experience accelerates the active rate more for customers already in an active state than for those in an inactive state. Reverse patterns result for catalog experience for both relationship states.

Table 14 Transition Matrix (Mean covariates)

$t \rightarrow t+1$	inactive	active
inactive	72.92%	27.08%
active	66.42%	33.58%

4.4. Counterfactual Analysis

A customer’s relationship state at any given period of time can be recovered probabilistically by “smoothing” or “filtering”. The goal is to recover the relationship state at ending period T , and to use full information available up to time T to recover the relationship state. It is helpful to know the state of a customer’s relationship at the end of observation period T , because marketers can

then build differentiated strategies for the future period based on that state. In this section, I describe how I recover customer relationship states at the end of the observation period (December 2004), and simulate a 7-period counterfactual plan for marketing communications.

The recovery of the relationship state at the end of an observation period is calculated by using equation (vii) in Appendix B. There are 37.8% of the customers in the active state, and 62.2% in the inactive state. Supposing the firm conducts marketing communications at the average rate for past periods for each individual, I generate data of customer choices for time $T+1$ based on the relationship state at time T . I then update the customer's relationship state at time $T+1$, and generate observations of customer choices at time $T+2$, and so on through time period $T+7$. From period $T+1$ to $T+7$, the estimated number of purchase incidence is 1575, and the estimated number of direct mail marketing communications sent to 595 customers is 11,164. Assuming that each purchase produces revenue of \$30, the estimated total revenue for the seven periods is \$47,250. Assuming the cost of each direct mailing (design, printing and mailing) is \$0.90, the total costs for 11,164 marketing communications would be \$10,047.60, and therefore the profit would be \$37,202.40 (Table 15 - Policy (A)). Policy (A) for marketing communications does not differentiate between customer-firm relationship states. In this simulation, 6,922 out of the 11,164 marketing communications are sent to inactive (state 1) customers who earlier have been found not to have influential reactions to direct mails. The 6922 marketing communications do not lead inactive customers to make purchases. This result implies that decreasing the amount of marketing communications with inactive customers does not decrease a firm's sales revenue. The Policy (B) simulation shows that when the firm reduces by half the amount of marketing communications sent to customers in the inactive state, profit increases from \$37,202.40 to \$40,318.20 (Table 15). The Policy (C) simulation illustrates the

scenario of the firm not conducting any marketing communication; the estimated number of purchase incidences is reduced to 1065, producing \$31,950 in sales revenue. Policy (C) produces 32.38% fewer sales and 14.1% less profit compared to policy (A).

This analysis illustrates the usefulness of differentiating between active state and inactive state customer-firm relationships, and illustrates how the simulation of differentiated marketing strategies can be used to increase a company's profit. In addition, I examine the optimal marketing communications by simulating the amount of marketing communications sent to customers in the active state from zero to four for period $T+1$, because customers in the inactive state do not respond positively to marketing communications. Profit will reach the highest (\$6,345.00) when there are two marketing communications sent to each customer at period $T+1$ (Table 16).

Table 15: Policy Simulation

Policy	Revenue (a)	Number of marketing communications (b)	Cost per mail (c)	Total Costs (d) = (b) x (c)	Profit = (a) -(d)
	\$47,250	11164	\$0.90	\$10,047.60	\$37,202.40
(A)	Inactive state	6922	\$0.90	\$6,229.80	
	Active state	4242	\$0.90	\$3,817.80	
	\$47,250			\$10,047.60	\$37,202.40
(B)	Inactive state	3460	\$0.90	\$3,114.00	
	Active state	4242	\$0.90	\$3,817.80	
	\$47,250			\$6,931.80	\$40,318.20
(C)	\$31,950	0	\$0.90	\$0.00	\$31,950

Table 16: Optimal Marketing Communications

Number of marketing communications sent to each customer at period $T+1$	Profit
0	\$4,590.00
1	\$5,820.00
2	\$6,345.00
3	\$6,142.25
4	\$5,945.00

4.5. Conclusions and Directions for Future Research

I have proposed a framework to estimate relationship dynamics in a non-contractual setting where a customer's dropout time is difficult to observe and not clearly defined. The proposed nested multinomial HMM addresses changes in preference of purchase incidence and channel choice across time with respect to various relationship states, and deals with the impact of marketing communications and channel experiences on customer retention as governed by transitions between relationship states. It captures dynamics in relationship transitions and purchase behavior through a discrete structure. The model extends prior studies of customer retention in non-contractual service settings by estimating retention probabilities, which are driven by various channel experiences and direct marketing, and incorporating channel evolution for active state customers, while allowing customers to move back to an active state after entering an inactive state.

Because a customer with a longer purchase cycle should be distinguished from a customer who has terminated the relationship with the firm forever, I develop a non-homogeneous multiple-segment and nested multinomial HMM that incorporates time-varying covariates into the transition between the latent relationship states, which represent the tendency of staying active or moving toward inactive state. The nested structure captures two layers of

choice decision simultaneously, purchase incidence and channel choice, and thus allows capture of the channel evolution when a customer is active. I use observed choice behavior and purchase information to make inferences about customers' underlying relationship states. My dynamic model of customer-firm relationship in a multichannel environment is calibrated and verified with data consisting of longitudinal transaction records from a large nationwide retailer in the US.

In the empirical application, I identify one segment with two relationship states, and examine different impacts (lasting and immediate) of channel experiences and marketing communications on relationship dynamics and state-dependent choice. Compared to alternative models, the proposed two-state HMM is the best-fitting model. I also use a hold-out sample to show that the proposed model outperforms the alternative models. The states revealed by the model represent customers' underlying relationship states which cannot be observed directly. State 1 is an inactive state, while state 2 is definitely an active state under the assumption of at least one marketing communication with the customer. Customers may move between these two relationship states based on their previous channel experiences and the number of marketing communications sent.

The empirical study provides substantive and interesting insights about multichannel shopping patterns. Multichannel experiences and marketing communications have both direct and indirect impacts on purchase incidence and channel selections. All three channel experiences (retail store, online, and catalog) help increase the propensity to purchase for both relationship states. Marketing communications have a positive influence on purchase incidence for active (state 2) customers. The model's nested structure revealed information about channel preference of customer that made a purchase. Customers in an inactive state are more likely to make a purchase through multiple channels, and marketing communications do not make the retail store

and online channels more favorable to those customers. Customer in an active state prefer the retail store channel more than other channels, but marketing communications to these customers increase their propensity to purchase through the online channel. For customers in both inactive and active states, the probability of shopping at a retail store increases with retail channel experience, whereas the probability of online shopping increases with both retail and online experience. Catalog experience does not have a positive influence on retail store or online shopping for customers in either relationship state.

Additionally, the result shows that customers who are already in the inactive state have a higher propensity to stay in the inactive state. Increased retail store experience retards the probability of moving toward an inactive state, and helps customers stay in the active state. More online experiences also help reduce a customer's rate of churn and help move a customer back to the active state. Catalog experience does not significantly increase an inactive customer's retention probability, but does decrease an active customer's retention probability. In general, retail and online experience increase customer retention, whereas catalog experience results in asymmetry patterns for the two states. More specifically, the retail store is the most active-oriented channel for inactive state customers, whereas the online channel is the most active-oriented channel for active state customers. Marketing communications do not have a significant effect on transitions toward active state. Based on the transitions of relationship state, one can clearly understand the changes in customer retention that are a consequence of channel experiences and direct marketing.

Using the model described in this dissertation will allow marketers to measure customer relationships and choice dynamics more accurately by overcoming the difficulties of observing customer relationship states. Marketers can predict better customer lifetime value by using more

accurate retention probabilities instead of assuming that customers stay with a firm for life. Multichannel managers are interested in the impacts of a variety of channels on customer retention and lifetime in addition to the value of retention probabilities. They can use my model to identify which channels can help build longer and lasting customer-firm relationships by examining the effects of direct marketing and channel experience on relationship evolution, and refine marketing strategies to encourage customers to move toward active-oriented channels. Further, firms can use the model to identify customers who are more likely to cancel their subscriptions and those with a greater possibility of terminating their relationship a firm; firms can then decide whether to retain these nearly inactive customers based on cost-benefit evaluation. Marketing strategies for active customers can be distinguished from strategies for nearly inactive customers. This approach described in this dissertation provides insights about the drivers of purchase incidence and channel choice. It may help firms retain active customers, and draw inactive customers back to active status. This approach helps firms classify customers into various relationship states dynamically and into particular segments probabilistically, as well as assess the dynamic effects of relationship evolution over time.

Accurately classifying customers into particular segments with a particular relationship state is important for a firm to optimize marketing efforts and further refine strategies, such as differentiated pricing, promotions and communications strategies. By comparing costs and profits for different segments, firms can target the most profitable segments. Marketers can incorporate price into the model and see how customers in different segments and relationship states respond to pricing strategies, although I do not model this because my dataset has insufficiently detailed price information among three channels. Possible areas of future applications for this research include any service providers that would benefit from an estimation

of retention or churn rate, such as banks, phone carriers, non-profits, and retailers. The model can also be applied to service providers with multi-brand, multi-category, or multi-channel services, who wish to understand their customers' preferences, client-firm relationship evolution, and the multiple impacts on retention.

There are several limitations in this research. First, because the data in my application only shows one active state and one inactive state, channel evolution while a customer is active cannot be illustrated. In addition, the current study only considers channel experience gaining from a single company, and does not account for channel experiences gaining from purchasing similar product categories in other companies and retailers. Future extension to this study may investigate the impact of competitors' activities and channel experiences gaining from other companies on relationship dynamics. Another limitation is that my dataset captures only direct mailing marketing communications, so online communications are not integrated into the study. Further study can test the impact of various communication media on relationship evolution and choice preference. Email marketing communications, for example, may have different impacts on customer retention than direct mailings. A further limitation is that the information about the firm's direct mail marketing communications is not clear and complete, so I do not know whether they were notices of annual sales or holiday sales, flyers for new product introduction, coupons for membership reward, catalogs, and so on. Neither does my data reveal the firm's goal for communications that encourage customers to shop in a certain channel. Different goals and different kinds of marketing communications may have different impacts. Further research can address the impact of various types of marketing communication on the evolution of relationships, and consider the costs associated with particular promotion activities. This can help a firm develop segment-specific and state-specific communication strategies. Additional

research is required to investigate more fully the evolution of inactive to active relationships, to fully understand the impact of customer-firm relationships on choice behavior in multichannel environments.

Chapter 5

Conclusions and Future Directions

In the first study of this dissertation, I present a model to display customers' channel preference as experience level increases. In the second study, I propose a framework to estimate relationship dynamics in a non-contractual setting in which a customer's dropout time is not observable and is not clearly stated. This dissertation contributes to the contemporary understanding of preference dynamics and customer-firm relationships in a multichannel environment.

The proposed models in Chapters 3 and 4 overcome the difficulties of observing evolving channel tendency and a customer's relationship state, respectively. Instead of guessing about the evolution of a customer's channel tendencies among seven possible channel combinations, the first study accurately identifies the evolutionary path of channel tendencies for each segment. Customers in segment 1 evolve between multichannel tendency and retail-loyal tendency, whereas customers in segment 2 evolve between an online and a multichannel state. Although the progressing rates of experience development are different for customers in both segments, they still lead to a pattern that shows a tendency for a customer with not much experience to switch, and later to migrate toward a steady state when more experienced. This implies that category experience reinforces customers' channel choice behavior and helps them move toward a particular tendency. In addition to developing an understanding of transitions in customer channel tendency, the study also incorporates the immediate effects of marketing communications on channel choices.

In addition, the second study extends current research on customer retention in a non-contractual service setting by estimating retention probabilities that are driven by various channel experiences and direct marketing, and allowing a customer to move back to an active state after entering an inactive state. The study also takes account of both immediate and enduring effects of channel experiences and marketing communications, i.e., immediate effects on purchase incidence and channel choice, and enduring effects on relationship dynamics. My data shows that all three channel experiences help increase the propensity to purchase, for both inactive and active states, whereas marketing communications have asymmetric impacts on both relationship states. Further, the model's nested structure reveals information about channel preference, and the immediate impacts of channel experiences and marketing communications if a customer indeed makes a purchase. In general, retail and online experience increase customer retention, whereas catalog experience results in asymmetric patterns for the two states. Specifically, the retail store is the most active-oriented channel for the inactive state, whereas the online channel is the most active-oriented channel for an active state. Moreover, marketing communications do not have a significant effect on transitions toward an active state. The results imply that the impact of marketing communications is immediate rather than enduring.

The findings in this dissertation can allow marketers to assess the dynamic learning effect on customer channel preference evolution over time, and thereby both allocate a firm's limited resources more effectively and further refine marketing strategies. Accurately classifying customers according to channel tendency is essential for a firm to optimize current marketing

efforts and develop overall strategies, such as differentiation in pricing and promotions. This research can allow a firm to compare costs and profits for different segments and discover which segments can provide the most profitability and thus should be targeted. Also, marketers can obtain information about which channels customers in different segments (low or high value) prefer, and provide those customers with fewer (or more) communications and services. In addition, by examining the effects of direct marketing and channel experience on relationship evolution, marketers can identify which channels can help build more enduring customer-firm relationships, and make further marketing strategies to encourage customers to move toward those active-oriented channels. Further, firms can use the model to identify inactive customers, who are more likely to cancel their subscription or have a greater possibility of terminating their relationship with firms. Firms can then decide whether to retain these nearly inactive customers based on cost-benefit evaluation. By comparing costs and profits for different segments, firms can target those segments that will provide the most profitability.

This model will translate well to other service-based industries, including any provider that would benefit from an estimation of retention or churn rate, such as banks, phone carriers, non-profits, and retailers. The model can also be applied to providers with multi-brand, multi-category or multi-channel services, who wish to understand their clients' preferences, client-firm relationship evolution, and the multiple forces affecting client retention. This research also opens other possible directions for future research. First, it would be interesting and beneficial to model the impact of multi-category learning on the evolution of channel tendency and multichannel choice, and compare the rates of experience development among different product categories.

Second, this study could be extended by applying the modeling approach to other industries and product categories. Third, further study could test the impact of various communications media on the evolution of channel preference, relationship dynamics, and choice preference. Also, the current study only considers experience and marketing activities in a single company, and does not account for marketing activities and experiences gaining from other companies and retailers. Future extension to this study may investigate the impact of competitors' activities and channel experiences from other companies on channel preference evolution and relationship dynamics. Additionally, future research could investigate other aspects of the learning process to fully understand the impact of dynamic learning on multichannel choice. Future research could also explore other aspects of inactive to active relationship evolution to understand the impact of customer-firm relationships on choice behavior in multichannel environments.

Appendix A

Multiple-segment HMM Likelihood Function

The HMM is a model composed of an unobservable Markov chain S_t and another observable stochastic process that produces a set of observations (Netzer et al. 2008). Conditional on the unobservable process, the observable process is a sequence of independent random variables such that the distribution of Y_{nt} given S_{nt} only depends on S_{nt} , where Y_{nt} is the sequence of the observed channel choices in Chapter 3. In Chapter 4, Y_{nt} is the sequence of the observed combination of purchase incidences and channel choices (B_{nt}, C_{nt}) for customer n , and S_{nt} is the set of latent relationship states. The joint probability of an observed sequence of choices Y is given by summing over all possible states over time, as follows.

$$\begin{aligned}
 L_{nT} &= P(Y_{n1} = y_{n1}, Y_{n2} = y_{n2}, \dots, Y_{nT} = y_{nT}) \\
 &= \sum_{\{S\}} P(Y_{n1} = y_{n1}, Y_{n2} = y_{n2}, \dots, Y_{nT} = y_{nT} | S_{n1} = i_1, S_{n2} = i_2, \dots, S_{nT} = i_T) P(S_{n1} = i_1, S_{n2} = i_2, \dots, S_{nT} = i_T) \\
 &= \sum_{i_1=1}^m \dots \sum_{i_T=1}^m \left[\prod_{t=1}^T P(Y_{nt} = y_{nt} | S_{nt} = i_t) \prod_{t=2}^T P(S_{nt} = i_t | S_{nt-1} = i_{t-1}) \times \pi_n \right] \quad (i)
 \end{aligned}$$

A forward recursive algorithm can be applied by rearranging L_T in a more useful matrix form which follows MacDonald and Zucchini (1997), as follows.

$$L_{nT} = \pi_n \Lambda_n(i_1, y_1) Q_n(i_1, i_2) \Lambda_n(i_2, y_2) Q_n(i_2, i_3) \dots Q_n(i_{T-1}, i_T) \Lambda_n(i_T, y_T) \mathbf{1}' \quad (ii)$$

where n is the n^{th} customer, $\Lambda_n(i_t, y_t)$ is a $m \times m$ diagonal matrix with $(P(y_{nt} | i_{nt} = 1), \dots, P(y_{nt} | i_{nt} = m))$ on the diagonal, π_n is $1 \times m$ vector of initial probability for each state, and $\mathbf{1}'$ is $m \times 1$ vector of ones. Therefore, the log-likelihood function for the HMM

becomes the sum of individual log-likelihood which can be represented as $\sum_{n=1}^N \ln L_{nT}$ in one-segment specification.

My multiple-segment HMM is composed of an unobservable learning process of Markov chain S_t , a set of latent classes, and another observable process. Conditional on a customer being in segment l and the unobservable learning process S_t , the observable process Y_t given l and S_t is independent. That is, for a group of T observations of a specific customer, the joint probability of an observed sequence of choices is now as follows.

$$\begin{aligned}
& P(Y_1 = y_1, Y_2 = y_2, \mathbf{L}, Y_T = y_T) \\
&= \sum_{\{l\}} \sum_{\{S\}} P(Y_1 = y_1, Y_2 = y_2, \mathbf{L}, Y_T = y_T \mid S_1 = i_1, S_2 = i_2, \mathbf{L}, S_T = i_T, l) P(S_1 = i_1, S_2 = i_2, \mathbf{L}, S_T = i_T \mid l) P(l) \\
&= \sum_{l=1}^L P(l) \sum_{i_1=1}^m \mathbf{L} \sum_{i_T=1}^m \left[\prod_{t=1}^T P(Y_t = y_t \mid S_t = i_t, l) \prod_{t=2}^T P(S_t = i_t \mid S_{t-1} = i_{t-1}, l) \times P(S_1 = i_1 \mid l) \right] \\
&= \sum_{l=1}^L ss_l L_{Tl} \tag{iii}
\end{aligned}$$

where $P(l) = ss_l$ is the proportion of segment l .

Therefore, the log-likelihood function for the multiple-segment HMM becomes as follows.

$$\ln L_{msHMM} = \sum_{n=1}^N \ln \left[\sum_{l=1}^L ss_l f_l(Y_{n1}, \mathbf{L}, Y_{nT}) \right] = \sum_{n=1}^N \ln \left[\sum_{l=1}^L ss_l L_{nTl} \right] \tag{iv}$$

for $n = 1, \mathbf{L}, N$, and $l = 1, \mathbf{L}, L$

where

$$ss_l = \frac{\exp(\pi_l)}{1 + \sum_{l=1}^{L-1} \exp(\pi_l)}, \quad (\text{v})$$

and

$$L_{nTl} = \sum_{i_1=1}^m \mathbb{L} \sum_{i_T=1}^m \left[\prod_{t=1}^T P_n(Y_t = y_t \mid S_t = i_t, l) \prod_{t=2}^T P_n(S_t = i_t \mid S_{t-1} = i_{t-1}, l) \times P_n(S_1 = i_1 \mid l) \right] \quad (\text{vi})$$

where n is the n^{th} customer, l is the l^{th} segment, π_l is the size parameter to be estimated, and ss_l is the likelihood of a consumer in segment l , which is the relative size of the l^{th} segment ($0 < ss_l < 1, \sum ss_l = 1$). The parameters in L_{nTl} vary by segment.

Appendix B

Recovering the State Distribution of HMM

A customer's relationship state at any given period of time can be recovered probabilistically by “smoothing” or “filtering”. Given a customer's history of observed behavior from period 1 to T , the probability distribution of the relationship state for the n^{th} customer at time T is as follows.

$$\begin{aligned}
 &P(S_{nT} = i | Y_{n1}, Y_{n2}, \dots, Y_{nT}) \\
 &= \pi_n \Lambda_n(i_1, y_1) Q_n(i_1, i_2) \Lambda_n(i_2, y_2) Q_n(i_2, i_3) \cdots \tilde{Q}_n(i_{T-1}, i_T = i) \tilde{\Lambda}_n(i_T = i, y_T) / L_{nT}
 \end{aligned} \tag{vii}$$

where

$\Lambda_n(i_t, y_t)$ is a $m \times m$ diagonal matrix with $(P(y_{nt} | i_{nt} = 1), \dots, P(y_{nt} | i_{nt} = m))$ on the diagonal, π_n is $1 \times m$ vector of initial probability for each state,

$\tilde{Q}_n(i_{T-1}, i_T = i)$ is the i^{th} column of $Q_n(i_{T-1}, i_T)$,

$\tilde{\Lambda}_n(i_T = i, y_T)$ is $P(y_T | i_T = i)$,

T is the end of the observation period, and

L_{nT} is the joint probability of an observed sequence of choices Y per equation (ii) in Appendix A.

Bibliography

- Albert, PS, HF McFarland, ME Smith, and JA Frank (1994), "Time Series for Modelling Counts from a Relapsing-Remitting Disease: Application to Modelling Disease Activity in Multiple Sclerosis," *Statistics in Medicine*, 13(5-7), 453–66.
- Ansari, Asim, Carl F Mela, and Scott A Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research (JMR)*, 45(1), 60–76.
- Balasubramanian, Sridhar, Rajagopal Raghunathan, and Vijay Mahajan (2005), "Consumers in a multichannel environment: Product utility, process utility, and channel choice," *Journal of Interactive Marketing*, 19(2), 12–30.
- Blattberg, Robert C., and John Deighton (1996), "Manage Marketing by the Customer Equity Test," *Harvard Business Review*, 74(4), 136–44.
- Boehm, Martin (2008), "Determining the impact of internet channel use on a customer's lifetime," *Journal of Interactive Marketing*, 22(3), 2–22.
- Brangule-Vlagsma, Kristine, Rik G.M. Pieters, and Michel Wedel (2002), "The dynamics of value segments: modeling framework and empirical illustration," *International Journal of Research in Marketing*, 19(3), 267–85.
- Brucks, Merrie (1985), "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research*, 12(1), 1.
- Campbell, Dennis, and Frances Frei (2010), "Cost Structure, Customer Profitability, and Retention Implications of Self-Service Distribution Channels: Evidence from Customer Behavior in an Online Banking Channel," *Management Science*, 56(1), 4 –24.
- Churchill, GA (1989), "Stochastic-Models for Heterogeneous DNA-sequences," *Bulletin of Mathematical Biology*, 51(1), 79–94.

- Dholakia, Ruby Roy, Miao Zhao, and Nikhilesh Dholakia (2005), "Multichannel retailing: A case study of early experiences," *Journal of Interactive Marketing*, 19(2), 63–74.
- DoubleClick Inc. (2004), "Multi-channel Shopping Study -- Holiday 2003."
- Du, Rex Y, and Wagner A Kamakura (2006), "Household Life Cycles and Lifestyles in the United States," *Journal of Marketing Research*, 43(1), 121–32.
- Fader, Peter S., and Bruce G. S. Hardie (2010), "Customer-Base Valuation in a Contractual Setting: The Perils of Ignoring Heterogeneity," *Marketing Science*, 29(1), 85 –93.
- Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), "'Counting Your Customers' the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24(2), 275 – 284.
- Greene, William (2003), *Econometric analysis*, Upper Saddle River N.J.: Prentice Hall.
- Gupta, Alok, Bo-chiuan Su, and Zhiping Walter (2004), "An Empirical Study of Consumer Switching from Traditional to Electronic Channels: A Purchase-Decision Process Perspective," *International Journal of Electronic Commerce*, 8(3), 131–61.
- Gupta, Sunil, and Donald R Lehmann (2003), "Customers as assets," *Journal of Interactive Marketing*, 17(1), 9–24.
- Hansotia, Behram J., and Bradley Rukstales (2002), "Direct marketing for multichannel retailers: Issues, challenges and solutions," *Journal of Database Marketing*, 9(3), 259–66.
- Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright (2000), "The Evolution of Brand Preferences and Choice Behaviors of Consumers New to a Market," *Journal of Marketing Research*, 37(2), 139–55.
- Hitt, Lorin M., and Frances X. Frei (2002), "Do Better Customers Utilize Electronic Distribution Channels? The Case of PC Banking," *Management Science*, 48(6), 732 –748.

- Juang, BH, and LR Rabiner (1991), "Hidden Markov-models for Speech Recognition," *Technometrics*, 33(3), 251–72.
- King, Maryon F., and Siva K. Balasubramanian (1994), "The Effects of Expertise, End Goal, and Product Type on Adoption of Preference Formation Strategy," *Journal of the Academy of Marketing Science*, 22(2), 146–59.
- Knox, George Aldo Hadley (2006), "Modelling and managing customers in a multichannel setting," *University of Pennsylvania, Working Paper*.
- Kotler, Philip (2000), *Marketing management*, Upper Saddle River N.J.: Prentice Hall.
- Kumar, V., and Rajkumar Venkatesan (2005), "Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior," *Journal of Interactive Marketing*, 19(2), 44–62.
- Leroux, BG, and ML Puterman (1992), "Maximum-Penalized-Likelihood Estimation for Independent and Markov-Dependent Mixture Models," *Biometrics*, 48(2), 545–58.
- MacDonald, Iain, and W. Zucchini (1997), *Hidden Markov and other models for discrete-valued time series*, London ;;New York: Chapman & Hall.
- Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23(4), 579 –595.
- Moon, Sangkil, Wagner A Kamakura, and Johannes Ledolter (2007), "Estimating Promotion Response When Competitive Promotions Are Unobservable," *Journal of Marketing Research (JMR)*, 44(3), 503–15.
- Neslin, S. A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, and P. C. Verhoef (2006), "Challenges and Opportunities in Multichannel Customer Management," *Journal of Service Research*, 9(2), 95–112.

- Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27(2), 185–204.
- Pfeifer, Phillip E, and Paul W Farris (2004), "The elasticity of customer value to retention: The duration of a customer relationship," *Journal of Interactive Marketing*, 18(2), 20–31.
- Rabiner, LR (1989), "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the Ieee*, 77(2), 257–86.
- 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.
- Reichheld, Frederick F., and W. Earl Sasser Jr. (1990), "Zero defections: Quality comes to services," *Harvard Business Review*, 68(5), 105–11.
- Reinartz, Werner J., and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64(4), 17–35.
- Schmittlein, D. C., D. G. Morrison, and R. Colombo (1987), "Counting Your Customers: Who-Are They and What Will They Do Next?," *Management Science*, 33(1), 1–24.
- Schoenbachler, Denise D., and Geoffrey L. Gordon (2002), "Multi-channel shopping: understanding what drives channel choice," *Journal of Consumer Marketing*, 19(1), 42–53.
- Schweidel, David A., Eric T. Bradlow, and Peter S. Fader (2011), "Portfolio Dynamics for Customers of a Multiservice Provider," *Management Science*, 57(3), 471–486.
- Schweidel, David A., Peter S. Fader, and Eric T. Bradlow (2008), "A Bivariate Timing Model of Customer Acquisition and Retention," *Marketing Science*, 27(5), 829–843.

- Shankar, Venkatesh, Amy K. Smith, and Arvind Rangaswamy (2003), "Customer satisfaction and loyalty in online and offline environments," *International Journal of Research in Marketing*, 20(2), 153–75.
- Stone, Merlin, Matt Hobbs, and Mahnaz Khaleeli (2002), "Multichannel customer management: The benefits and challenges," *Journal of Database Marketing*, 10(1), 39.
- Thomas, Jacquelyn S. (2001), "A Methodology for Linking Customer Acquisition to Customer Retention," *Journal of Marketing Research*, 38(2), 262–68.
- Thomas, Jacquelyn S., and Ursula Y. Sullivan (2005), "Managing Marketing Communications with Multichannel Customers," *Journal of Marketing*, 69(4), 239–51.
- USNews.com (2008), "Online Retailing Offers a Bright Spot," *U.S. News and World Report*.
- Venkatesan, Rajkumar, V Kumar, and Nalini Ravishanker (2007), "Multichannel Shopping: Causes and Consequences.," *Journal of Marketing*, 71(2), 114–32.
- 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.
- Verhoef, Peter C., Scott A. Neslin, and Björn Vroomen (2007), "Multichannel customer management: Understanding the research-shopper phenomenon," *International Journal of Research in Marketing*, 24(2), 129–48.

Vita

Chun-Wei Chang was born in Taipei, Taiwan, on August 16th, 1978. She earned a Bachelor of Arts degree in Economics at the National Chengchi University and a M.B.A. with a concentration in Marketing from National Taiwan University. She then entered the University of Michigan Ann Arbor in the fall of 2003 and earned a Master of Science in Statistics in 2005. In 2012 she earned degrees of Master of Science and Doctor of Philosophy at the University of Washington in Business Administration.