Georgia State University
ScholarWorks @ Georgia State University

Marketing Dissertations

Department of Marketing

Fall 12-18-2015

A Double Loop Learning Model For Integrated and Proactive Customer Relationship Management

Jia Fan

A DOUBLE LOOP LEARNING MODEL

FOR INTEGRATED AND PROACTIVE CUSTOMER RELATIONSHIP MANAGEMENT

BY

JIA FAN

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Doctor of Philosophy

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY

ROBINSON COLLEGE OF BUSINESS

2015

ACCEPTANCE

This dissertation was prepared under the direction of the *JIA FAN* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

V Kumar

Naveen Donthu

Man Luo

Denish Shah

Praveen Kopalle

ABSTRACT

A Double Loop Learning Model
For Integrated and Proactive Customer Relationship Management

BY

JIA FAN

December, 2015

Committee Chair: *Dr. V. Kumar*

Major Academic Unit: Department of Marketing

The rapid development of information technology has changed how firms interact with their customers. On one hand, firms are better capable of collecting customer data, and equip themselves with more powerful analytical tools. On the other hand, customers are becoming more sophisticated in their purchase decision making and other non-purchase interactions, which create higher demand uncertainty for the firm. To survive in this complex and dynamic environment, firms need to manage their customer relationships with an integrated and proactive approach. Recent studies in adaptive learning helped the firm to answer the question of *How to learn* about customers so they can be proactive in their CRM practice. In this study, we introduce the concept of *Double Loop Learning*, where we added a strategic learning loop to the adaptive learning loop. With this double loop structure, we also answer the questions of *Why to learn* and

What to learn and *Who should be learn* simultaneously in an integrated framework. We use a Partially Observable Markov Decision Process (POMDP) approach to 1). Generate optimal marketing contact policy which balances exploration (learning how various modes of marketing contacts affect the transition of customer relationship state) and exploitation (maximizing short-term profit), and 2). Assess the Value of Learning (VOL) at individual customer level to give a feedback to the strategic learning loop where we can answer the questions of *Why, What* to learn at individual customer level. Theoretically, we introduced the concept of *Double Loop Learning* to marketing literature which is fundamental in that it achieves both effectiveness and efficiency in the marketing strategy development. Methodologically, we adopted a POMDP approach which enables us to access the value of information for connecting two loops in an integrated framework. In the first essay, we did extensive review on the CRM and Adaptive Learning literature, based on which we developed the conceptual framework for Double Loop Learning model. We also developed an analytical model to demonstrate the relationship between the VOL and Dynamic Customer Value (DCV) of the customers. In the second essay, we apply the proposed framework to an IT B2B firm. We show that the firm can achieve value gains by managing VOL and DCV simultaneously.

TABLE OF CONTENTS

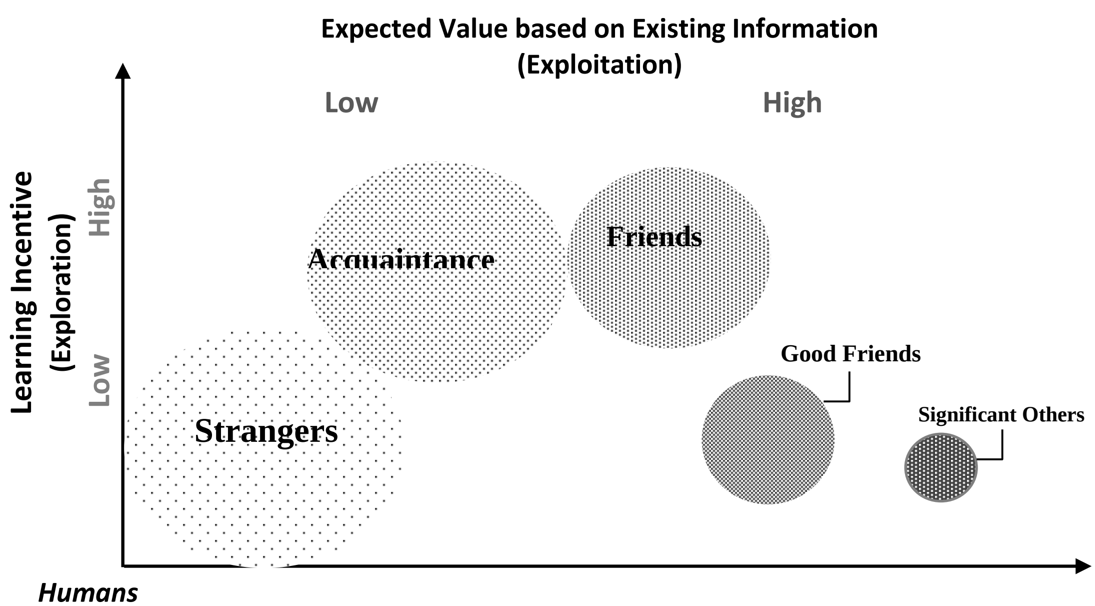
LIST OF TABLES

LIST OF FIGURES

MOTIVATION

In the past decade, Customer Relationship Management (CRM) has become a mainstream area among both marketing academics and practitioners. Continuous progress has been made by the marketing researchers towards better understanding and managing customer relationships. CRM models have evolved from aggregate level static models to the recent development of Hidden Markov Models (HMM). The role of marketing actions has been shifted from increasing short-term profitability to cultivating long-term relationships with the customers. We posit the key to successfully managing relationships is to managing customer relationship momentum. To managing momentum successfully, the role of marketing as a vehicle to learn the customers becomes critical.

Figure 1. Stages of Relationships



Managing relationships is a key aspect of our everyday life as a human being. The popular idea of Dunbar's number(Dunbar 1992) states that human beings can manage 150 stable relationships due to our cognitive limit. How do we do it? Cognitively, imaging we can put all our relationships on a map like Figure 1. We explore the relationship when we see opportunity of growing it into the next stage, and we also enjoy the relationship along the process. In other words, exploring and learning the relationship is the key to managing relationships dynamics.

The marketing department of a company often has to manage tens of millions of customers. We can hardly say its managing relationship as the marketing strategies are not incorporating learning aspect of relationship management. Despite the fact that companies are collecting overwhelming sized customer information, they barely know the dynamics of each individual customer. There is absence a strategic guideline on how to manage relationship through learning both effectively and efficiently.

Inspired by how we manage everyday relationships, we develop an integrated and proactive framework for the firm to manage customer relationships while tailoring learning at individual customer level. The proposed framework helps us to answer the questions of

*"**Why** do we learn the customers?"*

*"**How** do we learn the customers?"*

*"**What** do we use to learn the customers?"*

*"**Who** should we learn?"*

This Dissertation is composed of two essays. Essay 1 starts with a literature review on CRM and learning models, then propose then idea of double loop learning model, based on which we introduced the conceptual framework for integrated and proactive CRM. This essay also provide an analytical derivation for the size of learning effect and illustrated by some numeric examples. In essay 2, we demonstrated the application of this framework in a B2B setting where the firm has to decide how to allocate its marketing resources to maximize its dynamic customer values.

ESSAY1

A CONCEPTUAL FRAMEWORK FOR INTEGRATED AND PROACTIVE CRM

1.1 Introduction

While machine learning is a very popular concept in the business practice, it is rarely considered as a focal point of CRM mostly because it only answers the question of "*How do we learn the customers?"* The famous *Harvard Business Review* article by Argyris (1997) stated that learning by double loop method is the key to the success of an organization. We posit that a good framework for managing relationships through learning should also answer the questions of

"Why do we adaptively learn the customers?"

"How do we learn the customers?"

"What do we use to learn the customers?"

"Who should we learn?"

Based on the Double Loop Learning Model (DLLM), this essay develops a theoretical framework which incorporates the concept of value of learning the customers as a metric for measuring customer's potential as a high value customer. Through analytical analysis and numeric examples, we also provide a guideline on how adaptive learning should be incorporated into the practice of managing customer relationships.

This essay is organized as follows. Section 1.2 describes the relevant literature in CRM and adaptive learning models. Section 1.3 proposes an integrated and proactive framework for managing customer relationships. After laying out the conceptual framework, we explore the potential determinant of the size of learning effects. Section 1.4 explains dynamic customer value as a function of firm's knowledge about the customer. An analytical model was derived to explain the drivers of the value of learning customers, based on which we use some simulated numeric examples to show the nonlinear relationship between value of learning and customer value. Section 1.5 is the discussion on this essay.

1. 2 Literature Review

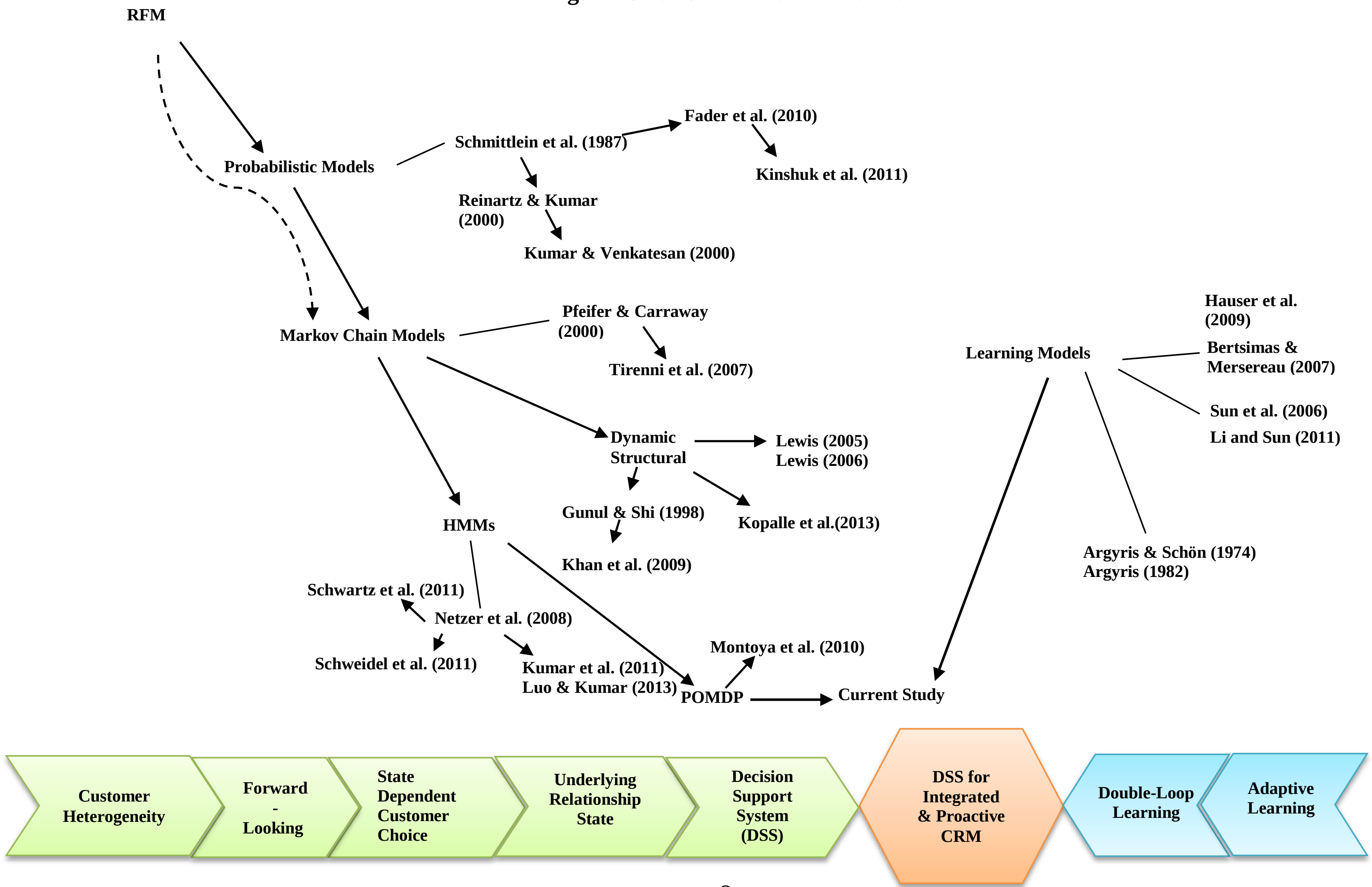
1.2.1 Evolution of CRM Models

As firms switch from product-centric paradigm to customer-centric paradigm, CRM started to become one of the most important fields in marketing. Information revolution made it possible for the firms to obtain customer level data which allowed marketing researchers to create disaggregated CRM metrics to identify the profitable customers and to allocate marketing resources more efficiently (Reinartz and Kumar 2003; Reinartz and Kumar 2000; Venkatesan et al. 2007). There are a few very good papers summarizing these CRM models (Berger and Nasr 1998; Gupta et al. 2006; Jain and Singh 2002). Here we are going to briefly describe the evolution of CRM models and explain the position of the proposed framework in the literature.

Figure 1 shows the evolution of CRM models. The first leap for CRM models is transitioning from backward-looking models like RFM to probabilistic models(Reinartz and Kumar 2000; Schmittlein et al. 1987) where the goal of customer management is to maximize net present value of future revenue stream. When the zero-order purchasing process assumption is relaxed by assuming it follows a first-order Markov process, the purchase process can be formulated as a Markov chain (Fader and Hardie 2009). Pfeifer and Carraway (2000) first proposed to model customer relationship as a Markov chain where the purchase probability is constantly updated by observations of the customers' behaviors. Markov chain CLV models treat the customer's purchase as a stochastic process and allows the customer to be inactive in some periods while still remains as a customer. Pfeifer and Carraway (2000) used *recency* as the state variable for the customers' future purchase probabilities.

The recent introduction of Hidden Markov Models (HMMs) shifted the paradigm from managing relationship through transactions to managing customer relationship itself. Studies (Table 1) have used HMM to understand the evolution of the underlying relationship state that governs customers' actions. Further, HMM was also used to evaluate the long term and short-term effects of marketing actions(Kumar et al. 2011; Montoya et al. 2010). Schweidel et al. (2011) studied customer service portfolio dynamics with a dynamic HMM to identify customers' underlying state with two sources of dynamics: portfolio inertia and service stickiness. Although it was a dynamic model, it did not provide a marketing intervention strategy due to lack of customer-firm interaction data. Kumar et al(2011b) uses a HMM to evaluate the short-term and long-term effects of marketing investment. While this study provided a strategy for optimizing marketing dollars, it did not assess the role of different marketing actions (for example, mail vs. telephone call) as tools of learning the underlying relationship states. Montoya et al.(2010) used POMDP to dynamically allocate detailing and sampling activities. They tracked physician's prescription behavior states and identify detailing as the most effective acquisition tool and sampling as the most effective retention tool. While all these HMMs estimates the drivers of underlying state of the customers and some of these studies did dynamic optimization. None of these studies had evaluated the effects of adaptive learning the customers and conditional planning on improving dynamic customer values.

Figure 2. Evolution of CRM Models



1.2.2 Adaptive Learning in CRM

While Adaptive learning is a very hot topic in both computer science and operation research, very few studies in marketing has been focused on the effects of learning on CRM (Bertsimas and Mersereau, 2007). From the evolution of CRM models, all the models from aggregate level models to state dependent models are all passive learning models. Under demand uncertainties, studies have shown that adaptive learning model is better than passive learning model. Sun et al.(2006a) listed the major characteristics of adaptive learning models compare to passive learning models. In adaptive learning models, firms are CRM decision makers who learn about customers in real time fashion and update their beliefs on customer preferences continuously. They gain knowledge from customers' development path. The benefits of adaptive learning have been documented by a few studies in the marketing field. Table 1 shows selected work on adaptive learning in marketing literature. Sun et al. (2006a) proposed a two-step conceptual framework for the firm to adaptively learn about the customers for their CRM decisions. While conceptually very inspiring, it did not provide a detailed solution on how to act upon information and it did not quantify the value of learning. Cao and Sun (2007) analytically accessed the value of adaptively learns about customer's service preference in allocating two types of service offerings to customers. Sun and Li (2011) also demonstrated the value of learning and acting on customer information with simulation. In their analytical paper, Bertsimas and Mersereau (2007) formulated a problem on balancing exploitation and exploration in the context of allocation of marketing messages types. They posited that current CLV models ignored how future information gains will influence current marketing decision. They demonstrated how the action of learning (sending out poorly understood messages for the purpose of learning) provided extra insight into customers' preferences for multiple homogeneous customer segments. Gooley and Lattin (2000) also pointed out that sometimes it's better to "sacrifice potential early payoff for the prospect of gaining information about customers that will allow for more informed decision later." Hauser et al. (2009) used a POMDP model to learn about customers' cognitive styles from clickstream data. Their real time solution balanced learning cognitive styles and maximizing short-term profit simultaneously.

Table 1. Selected Literature in HMM and Learning (revise)

	Study	Context	Hidden State	Double-Loop Learning	Dynamic Customer Value Optimization	Adaptive Learning (How)	Value of Information
HMMs	**(Netzer et al. 2008)**	Alumni Relationship	Relationship				
HMMs	**(Montoya et al. 2010)**	Physician Prescription	Prescription behavior state		✓	✓	
HMMs	**(Schweidel et al. 2011)**	Service Portfolio	Service subscription propensity		✓		
HMMs	**(Kumar et al. 2011)**	It B2B product	Relationship		✓		
HMMs	**(Luo and Kumar 2013)**	It B2B product	Relationship		✓		
	This Study	**It B2B Integrated Product and Service**	**Relationship**	✓	✓	✓	✓
Adaptive Learning	**(Sun et al. 2006)**	DSL service call center; banking	Latent demand state			✓	
Adaptive Learning	**(Cao and Sun 2007)**	Analytical	Customer service preference		✓	✓	✓
Adaptive Learning	**(Sun and Li 2011)**	DSL Service call center	Customer service preference		✓	✓	

(Bertsimas and Mersereau 2007)	Analytical Interactive marketing	Marketing message effectiveness		✓	✓
(Hauser et al. 2009)	Web design	Cognitive style	✓	✓	✓

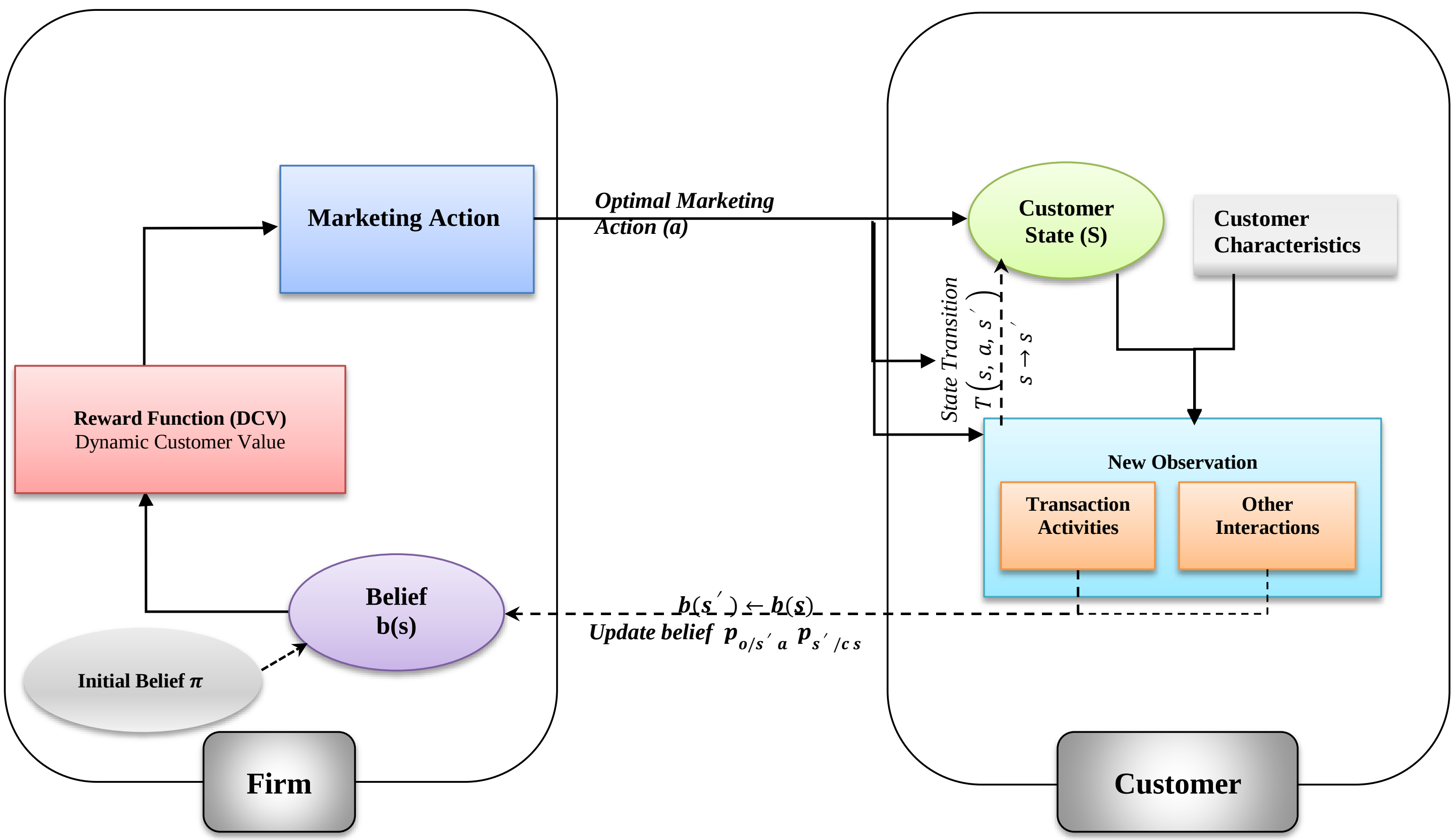
While these sophisticated learning models are very efficient on learning individual customer's references, there are a few questions that have not been addressed by the learning models in CRM literature. First, all these learning models have addressed the importance of learning customer's underlying preferences (service preferences, cognitive style, etc). Unlike HMMs in the marketing literature, none of these learning models has focused on customer's underlying relationships. Second, all these HMMs only focused on transaction aspect of customer relationships (Kumar et al. 2011; Luo and Kumar 2013; Netzer et al. 2008). However, other customer-firm interactions (for example, customers' service requests, product returns, customer initiated contacts, etc) could also help us to understand the dynamics of relationships as well. In other words, it is possible to use the information from various sources to help us better understand the underlying relationships. Third, all these learning models are trying to answer the question of "*How to learn customers?*" They help the firm to make their CRM process more efficient, but none of these models have addressed the issue of effectiveness. Adaptive learning might be an optimal strategy at aggregate level, but it could be suboptimal for individual customers, especially when learning involves costs. It assumes that a company should either dynamically learn the underlying relationship states for all customers, or not learning at all. None of studies above have answered the questions of *“Why do we adaptively learn the customers?” “What do we use to learn the customers?”* and *“Which customers should we learn?”*

1.2.3 CRM with Partially Observable Markov Decision Process (POMDP)

POMDP models are becoming popular in artificial intelligence(Kaelbling et al. 1998) for robot navigation. It also had applications in some other fields by operation research(Cassandra 1998). We apply POMDP to our CRM model by using the model shown in Figure 3. The general idea of learning in CRM is shown in Figure 3. There are five components in this POMDP framework: Core States, State Dependent Choice by a Business Customer, Core State Transition Matrix, Initial Distribution of Core States and Belief Updates after observing the State Dependent Choice.

The left hand side of the figure is the focal firm as a decision maker and the right hand side is the focal firm’s view about how a business customer makes its purchase decisions. Starting with an initial belief about the customer, the firm allocation its marketing resources to optimize its reward function, which is Dynamic Customer Value (DCV). When the optimal marketing action reaches the customer side, it has the short term impact on sales and long term impact on shifting relationship level. Depending on some customer characteristics, the customer will respond to the marketing actions. With the new observations, the firm update its belief regarding the customer. The process of CRM becomes a process of constantly learn the customers and allocate marketing resources based on the new knowledge about the customers.

Figure 3. CRM with POMDP



This framework solves the problem of *How to Learning* when we manage customer relationships. To answer the other three questions, we need to add another strategic learning loop to the CRM practice.

1.3 Double-Loop Learning for CRM: Conceptual Framework

1.3.1 Theoretical Background

To address the questions of efficiency and effectiveness of learning in CRM simultaneously, we introduce a Double Loop Learning (DLL) to marketing literature. DLL was first developed by Argyris (1976) which proposed an additional loop to the "adaptive learning" loop in single loop learning model so the firms can evaluate the effectiveness of the models we use in the inner loop.

Figure 4. Argyris's Double Loop Learning Model

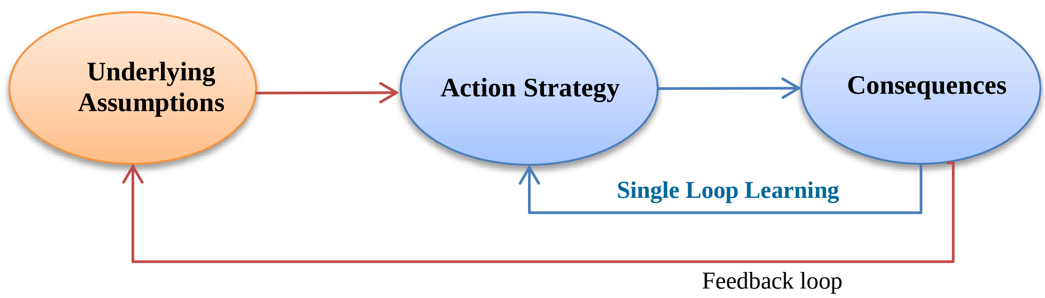


Figure 4 shows the original DLL where inner loop shows the adaptive learning process, and the outer loop shows the process of using the outcomes to re-examine the underlying assumptions for the adaptive learning model. This feedback loop is especially important if we want to develop a model that enables firms to become real decision makers. In practice, firms adjust their strategic plans when the existing strategy does not work, the existing marketing models are not flexible enough to allow these adjustment. For example, a firm normally makes marketing activity plans for 3 years. However, the manager decides to shorten the planning horizon to 1 year due the recent turbulent economy. The existing marketing models would not be able to adjust the change without re-estimation. Although DLL model is conceptually simple and elegant, it was rarely applied in empirical studies due to the challenge of generating the right feedback from the inner loop to evaluate the effectiveness of the adaptive learning model and test the underlying assumptions(Argyris 1976). Next, we are going to propose a Conceptual CRM framework with a DLL structure where we can access the value of information from the inner loop and use it as a feedback to the outer loop.

1.3.2 Conceptual Framework for CRM

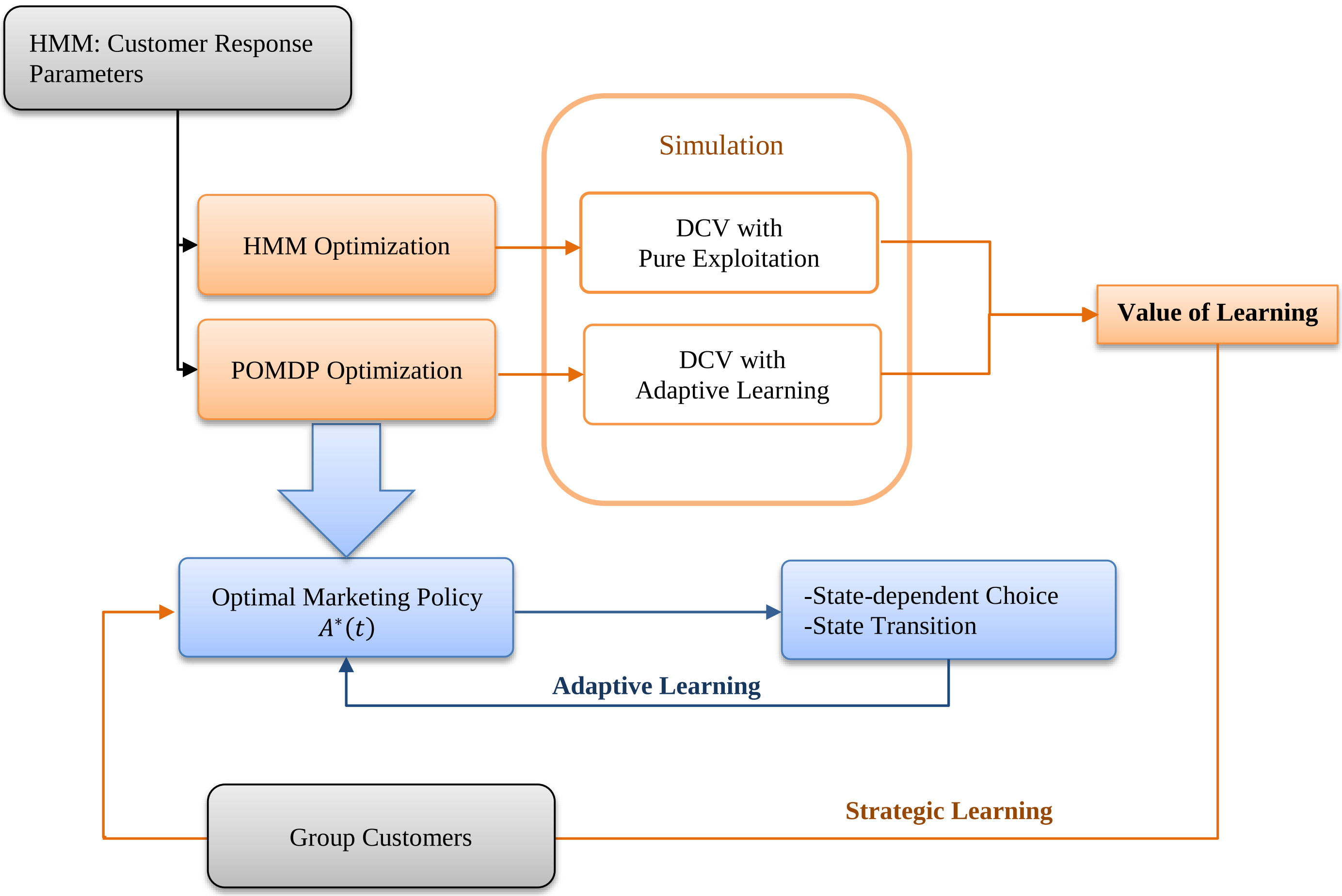
Figure 5 shows and integrated framework for managing customer relationships. Similar to the original DLL, it is composed of an inner loop of adaptive learning and an outer loop which we call strategic learning loop[1]. The focal firm has information regarding its product purchases and other customer-firm interactions for each individual customer. We first use an HMM to estimate the parameters for the state dependent choice probabilities and transition matrix for the

---

[1] Some studies call double learning strategic learning, here we use strategic learning only for the outer loop to differentiate it from the inner adaptive learning loop.

relationship state. Then we feed these parameters into a POMDP model and HMM to generate DCV for each model respectively. By comparing two models, we obtain the VOL for each individual customers.

Figure 5. A Double-Loop Learning Model for Integrated and Proactive CRM



For the group of customers who have high VOL, we move them into the inner loop of adaptive learning where we closely monitor the customer-firm interactions. In each step, we apply the optimal policy we obtain from the POMDP optimization step after updating information regarding the customer's state. For the group where adaptive learning will not provide too much of extra information, it is safe to apply the dynamic forward-looking model without learning where we don't need to closely manage the customer relationship dynamics through conditional planning. The optimal marketing action for each step in the planning horizon is determined at the beginning of the planning horizon. Next section, we will demonstrate how this framework works with a simple numeric example.

1.4 An Analytical Model for DCV and VOL

CLV is the concept based on the idea of heterogeneity in how much a customer worth to the firm. Similar to CLV, we believe that VOL is also different across customers. However, just like any relationships, the value from learning the other party is not necessarily linearly correlated with "value" of the relationship between the two parties. The idea of the proposed framework is to accounting for the heterogeneity in the VOL across customers. To get some idea of what are the drivers of VOL, we derive an analytical model to look for the characteristics of customers with high VOL.

1.4.1 Analytical Analysis on VOL

The goal of this analytical model is to demonstrate the relationship between DCV and VOL as well as some potential characteristics of the customers with the potential for learning. Therefore, we use very simple functional forms for the transition function and emission function that are consistent with the common practice of HMM for this analytical analysis.

Problem setup:

1. Assuming there a customer has two relationship state: low and high. The state variable$b = Prob(s = 1)$. i.e,. the probability of being in high state . The initial state distribution is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, in other words, $b_0 = 0$.
2. There are three emission observations, the probabilities of the observations are state-dependent. Assuming the marketing effect stays the same for the emission observations, The underlying utilities of the three observation can be written as

Low: $U_0 = \alpha_0 + \beta_0 g_0(M) + \varepsilon_0$

High: $U_1 = \alpha_1 + \beta_1 g_1(M) + \varepsilon_1$

where, $\alpha_1 > \alpha_0$ $\varepsilon \sim f(0{,}1)$

The underlying

$$P_i^O = \begin{cases} P_i^{No} = 1 - F\big(\alpha_i + \beta g_i(M)\big) & if\ O = no\ purchase \\ P_i^{Low} = F\big(\alpha_i + \beta g_i(M)\big) - F\big(\alpha_i - c + \beta g_i(M)\big) & if\ O = low\ purchase \\ P_i^{High} = F\big(\alpha_i - c + \beta g_i(M)\big) & if\ O = high\ purchase \end{cases}$$

Where $i = 0{,}1$

For this optimization problem, there are two dynamics strategies to choose from, the first strategy is *Pure Exploitation Strategy*[2] where dynamics of the belief purely relies on the current

---

[2] We use *Pure Exploitation* and *HMM*, *Adaptive Learning* and *POMDP* strategy interchangeably throughout the paper.

knowledge about the system dynamics. Under this strategy, the decision maker develops the optimal strategies at the beginning of the planning horizon either without belief monitoring or continuously monitoring the customers' interactions and updating beliefs. The bellman equations for these two dynamic models are shown below,

$$V^{HMM}(b) = Max_M\left\{R(b, M) + \gamma EV\left(\varphi\left(b^{'}\ |M, b\right)\right)\right\}$$

$$V^{POMDP}(b) = Max_M\left\{R(b, M) + \gamma EV\left(\varphi\left(b^{'}\ |O, M, b\right)\right)\right\}$$

$R(b, M)$ is the expected instant reward based on current belief b after marketing action M.

$$\begin{aligned} R(b, M) &= P(s = 1) * \left[P_1^{Low}R^{Low} + P_1^{High}R^{High}\right] + P(s = 0) * \left[P_0^{Low}R^{Low} + P_0^{High}R^{High}\right] \\ &= b * \left[P_1^{Low}R^{Low} + P_1^{High}R^{High}\right] + (1 - b) * \left[P_0^{Low}R^{Low} + P_0^{High}R^{High}\right] \end{aligned}$$

The difference between these strategies comes from whether or not to continuously learn about customers' state through customer's future interactions.

For HMM, the belief update function is:

$$\varphi\left(b^{'}\ |M, b\right) = b\pi_{11} + (1 - b)\pi_{01}$$

which only depends on the transition matrix estimated from the HMM.

For POMDP model, the belief update is

$$\begin{aligned} \varphi\left(b^{'}\ |O, M, b\right) &= \frac{[b\pi_{11} + (1 - b)\pi_{01}] * P_1^O}{[b\pi_{11} + (1 - b)\pi_{01}] * P_1^O + [b\pi_{10} + (1 - b)\pi_{00}] * P_0^O} \\ &= \frac{\varphi\left(b^{'}\ |M, b\right) * P_1^O}{\varphi\left(b^{'}\ |M, b\right) * P_1^O + [b\pi_{10} + (1 - b)\pi_{00}] * P_0^O} \end{aligned}$$

The size of learning effects is determined by

$$\left|\varphi(b' \mid O, M, b) - \varphi(b' \mid M, b)\right| = \left|\frac{P_1^O}{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O} - \varphi(b' \mid M, b)\right|$$

$$= \left|\frac{\varphi(b' \mid M,b)*P_1^O - \varphi(b' \mid M,b)*\{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O\}}{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O}\right|$$

$$= \left|\frac{\varphi(b' \mid M,b)[1-\varphi(b' \mid M,b)]*P_1^O - \varphi(b' \mid M,b)*[1-\varphi(b' \mid M,b)]*P_0^O}{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O}\right|$$

$$= \left|\frac{\varphi(b' \mid M,b)[1-\varphi(b' \mid M,b)]*(P_1^O - P_0^O)}{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O}\right|$$

$$= \left\{\frac{1}{\frac{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O}{\varphi(b' \mid M,b)[1-\varphi(b' \mid M,b)]}}\right\} * \left|P_1^O - P_0^O\right|$$

The first term of the equation is the inverse of total volatility of the system weighted by the emission observation probabilities. The second term is the size of the state-dependence for observation O. From can see:

When $\varphi(b' \mid M, b) \to 0 \; or \; 1$, $\left|\varphi(b' \mid O, M, b) - \varphi(b' \mid M, b)\right| \to 0$

There's small value of learning if there is very high tendency of moving to one state. In other words, the value of learning is high for a system with moderate level of volatility.

Given P_1^O, P_0^O

$Max_{\varphi} \left|\varphi(b' \mid O, M, b) - \varphi(b' \mid M, b)\right| = Min_{\varphi} \frac{\varphi(b' \mid M,b)*P_1^O + [1-\varphi(b' \mid M,b)]*P_0^O}{\varphi(b' \mid M,b)[1-\varphi(b' \mid M,b)]} =$

$Max\left[\frac{P_1^O}{[1-\varphi(b' \mid M,b)]} + \frac{P_0^O}{\varphi(b' \mid M,b)}\right]$

FOC: $-[1-\varphi(b' \mid M, b)]^{-2} P_1^O + \varphi(b' \mid M, b)^{-2} P_0^O = 0$

$$\rightarrow \frac{\varphi\left(b^{'} \;|M,b\right)}{1-\varphi\left(b^{'} \;|M,b\right)} = \sqrt{\frac{P_1^O}{P_0^O}}$$

When the ratio of the tendency of going to each state is proportional to the square root of the ratio of state dependent emissions, the belief difference in two strategies is maximized.

$$\frac{\sqrt{P_1^O}}{\varphi\left(b^{'} \;|M,b\right)} = \frac{\sqrt{P_0^O}}{1-\varphi\left(b^{'} \;|M,b\right)}$$

The equation above could also be expressed as: when the square root of the emission probabilities of each state averaged by the total force of going to each state is about the same, the value of learning is the highest. It could also be viewed as "equal state elasticity of emission probabilities" condition.

When we take a closer look at the equations,

$\varphi\left(b^{'} \;|M,b\right) = b\pi_{11} + (1-b)\pi_{01} = T(b,M), \quad P_i^O = F(\alpha_i, M)$

Both of which depends on the belief on the level of customer relationship and the role of marketing action. Therefore, the value of learning also depends on the effectiveness of marketing actions in the transition force and the emission functions.

1.4.2 Numeric Demonstration

We have shown analytically that the value of learning is determined by the state dependence of the emission probabilities and the transition force of moving towards each of the states. In this section, we will demonstrate the levels of value of learning by numeric examples with different levels of transition and emission.

Assuming a customer's purchase behavior follows are two state, five action and three observation HMM model. We started with a simple model with ordered logit model emission function and multinomial logit model transition function. Assuming the cost of marketing is 0, however, the effects of marketing is concave which means "over marketing" will hurt the sales. The parameters we used for this simulation is shown below.

Emission Model: $Y_{Low} = .45 + .3 * MKT - .2MKT^2$

$$Y_{High} = .2 + .5 * MKT - .15MKT^2$$

Transition Model: $TR_{Low-High} = 2 + 1.5 * MKT - .235MKT^2$

$$TR_{High-High} = 4 + 2 * MKT - .235MKT^2$$

The initial state is set to be 0, which means all customers start with low state. The reward for the three levels of purchases are: 0, 1, 2 respectively. The error terms for both models follow standard extreme value distribution. Using these parameters, we did value iteration to get a contraction mapping for both HMM and POMDP models with different belief update functions shown in equation (). Due to the continuous nature of the state variable, i.e., $b \in [0,1]$, we create 100 grid points in the vector $[0,1]$, and generate optimal marketing action on each of these points. With the piecewise linear and convex property of value function, we found the optimal marketing action for each belief segments. For this particular customer, policies for the HMM and POMDP model are the same:

$$Policy^{HMM\ \&POMDP} = \begin{cases} 2 & b \in [0, 0.378) \\ 3 & b \in [0.378, 1] \end{cases}$$

When the probability of being in high state is lower than 0.378, the optimal marketing level is 2, and then the probability of being in high state is higher than 0.378, the optimal marketing level is 3. Then we went on to do a policy simulation for these two models. We simulated 10 sequences of purchases, state and optimal marketing actions for 100 periods. Figure 6 shows 2 of the sequences for HMM pure exploitation optimization. The marketing action is always spending level 3 of marketing as the state variable reached the steady state at 0.66. As the firm assumes that it has all the information regarding the state, it does not update its belief on the state level with new observation on purchases. Figure 7 shows three of the sequences we simulated with POMDP adaptive learning model. With all these three sequences, the “baseline” state stays at a high level when the customer is purchasing at level 1, however, when there is a dip in purchase, the belief state is updated accordingly, and the optimal marketing actions are adjusted based on the belief state. The average state level for pure exploitation model for the 10 simulated process is.66, and average DCV is 83.5. The average state level for adaptive learning model is 0.86, and average DCV is 96.1 In other words, the VOL for this customer is roughly 12.6.

Figure 6. HMM: Pure Exploitation with 2 Sequences

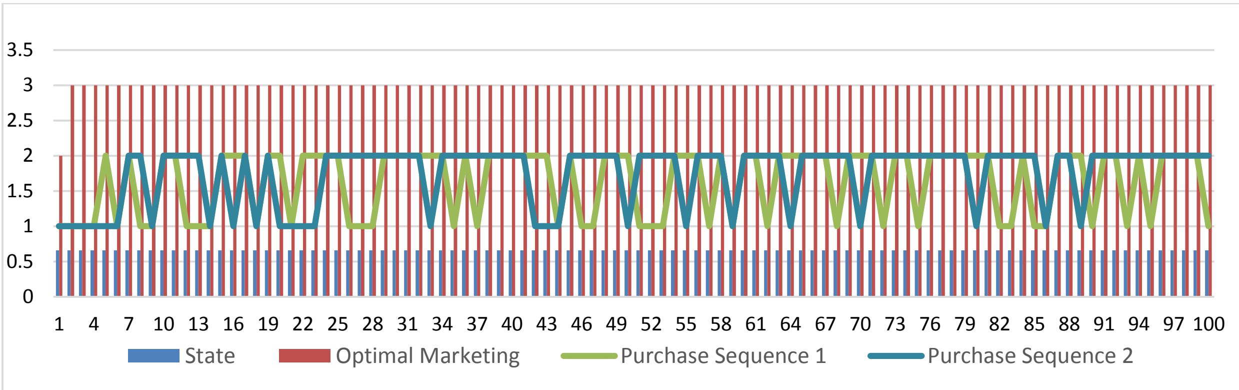
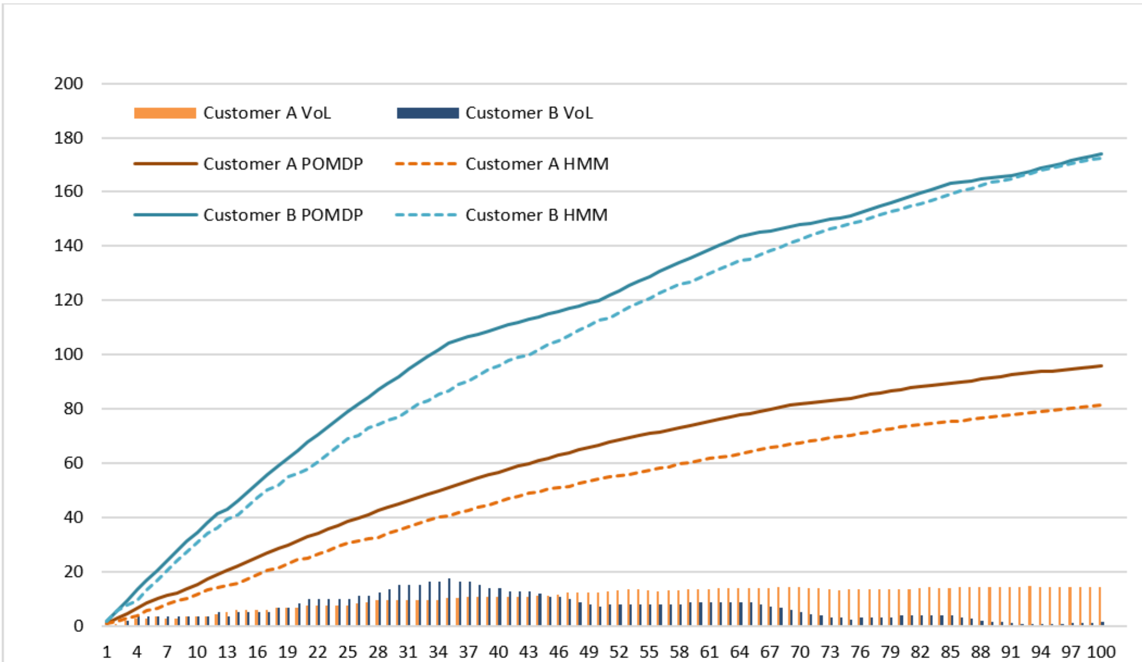


Figure 7. POMDP: Conditional Planning with 3 Sequences

Figure 7 shows how the focal firm interacts with the customer, and updates its information on the customers and allocate marketing spending accordingly. This is the simplified scenario to show how pure exploitation and adaptive learning work differently in practice. From the sequences, we can see that when the customer is in an inactive state,

To show the relationship between Customer Value and Value of Information, assuming there is a Customer B with the same transition parameters as the previous customer, but high state dependence. i.e., higher emission probabilities and higher value associated with the choice. We obtained policies for HMM model and POMDP model, then simulated policies for each models for 100 periods. Figure 8 shows the customer values by HMM and POMDP, and the VOL for customer A and B. The orange lines are values for Customer A, and the dark green lines are values for Customer B. We can see Customer B value is roughly twice as Customer A value. However, for a 100 period policy simulation, the total net present value from HMM and POMDP are about the same for Customer B. In other words, there is very small VOL for Customer B. From Figure 8, VOL is also dependent on the planning horizon. For the planning horizon of 36 months, the firm should focus on adaptively learn about Customer B who has high VOL for this period of time. However, the firm should focus on A if the planning horizon is 48 months or higher.

Figure 8 Value of Learning (VOL) vs. Dynamic Customer Value (DCV)



1.5 Discussion

As learning is a key aspect of managing human relationships, we posit that learning is also a key aspect of managing customer relationships. In this essay, we developed a conceptual framework to integrate learning customers into CRM. We used a POMDP modeling framework to address the question of "*How to learning customers?*" dynamically. We also adopted the classical DLLM to guide us on answering the questions of question of *"Why do we adaptively learn the customers?" "What do we use to learn the customers?"* and *"Which customers should we learn?"* in an integrated framework.

To prove learning customers is key element of CRM practice, we used analytic analysis to show the factors that create high opportunities for learning. By using this analysis, we also showed that value of learning is another dimension of customer valuation in addition to customer values. To demonstrate the learning effects, we also did a numeric example with simulation to show how adaptive learning strategy works better than pure exploitation strategy. We also demonstrated the nonlinear relationship between the value of learning and DCV. In the next essay, we are going to show how we adopt this framework in the practice of managing business customer relationships.

ESSAY 2

INTEGRATED AND PROACTIVE CRM: AN APPLICATION TO IT B2B INDUSTRY

In the previous essay, we proposed an integrated and proactive framework for managing customer relationships. We proposed to use POMDP to solve the question of *How* to learn customers, and created measurements VOL and DCV to answer the questions of *Why, What* and *Who* questions. In this Essay, we applied this framework to a IT B2B industry to illustrate how to use this DLLM framework to capture various demand uncertainties in a complex B2B purchase scenario.

2.1 Industry Background and Data Description

2.1.1 Industry Background

In a B2B setting, the customer purchase is a very complex decision process. The demand for products from a particular provider is driven by their internal needs for the products, but it's only fulfilled by the provider company if they are in good relationship. In other words, B2B customer purchase is governed by underlying relationship state between the focal firm and its business customers(Kumar et al. 2011; Netzer et al. 2008). The lumpy purchases by the business customers also creates high demand uncertainties from both the uncertain relationship state as well as infrequent demand. The evolution of this underlying relationship state depends on business customers' past purchase experience with the focal firm as well as all the other non-transaction encounters between them.

Most marketing research on customer purchase behaviors has been focused on either goods or services(Rust and Chung 2006). However, a customer's relationship with a company is more likely to be formed by its total experience with the firm including transaction, marketing contacts and other service interactions. Since so many factors are involved in the process of "building relationship" and "improving sales", it's very challenging to dynamically allocate marketing spending at each individual customer level accounting for the evolution of customer relationship. In B2B marketing practice, the allocation of marketing spending is either very general or very subjective. It is general when the B2B firms try to send direct mails to customers, they have the uniform strategy on the timing of the contact. It is subjective because for the case of direct phone calls, sales representatives use their own judgement on who and what time they should call, normally without the knowledge about the customers' other interactions with the company except for sales. All these characteristics are reflected in the data we have from a B2B IT firm.

2.1.2 Data Description

The data for this analysis comes from an IT B2B firm. We randomly selected 160 customers who had made their first purchase in 1998. These customers were also from middle sized firms whose employment size is from 50 to 500. From the focal firm, the marketing spending pattern is shown in the figure below. Figure 9 shows the marketing spending on Cohort 1998 over the 66 months observation periods. The purple area shows the number of customers who had made their first purchase in each month in 1998. We found that although the "new customers" generate pretty high revenues, the marketing spending on these customers are surprisingly low. The marketing spending on this cohort only started to pick up after they reach two year tenure with the company. We found similar pattern for another cohort of customers who started with the company in 1999. It could be caused by the customer representative's tendency to contact with old customers more often. In regards to endogeneity of marketing spending, the overall trend shows some spike in marketing spending after a dip in revenue. However, the correlation between marketing spending and last period sales at individual customer level is fairly weak at 0.17. Figure 10 shows the marketing spending by customers. We can see that the marketing spending is not always proportional to revenues. Both Figures show some level of subjectivity on marketing resource allocation.

Figure 9. Marketing Spending on Cohort 1998

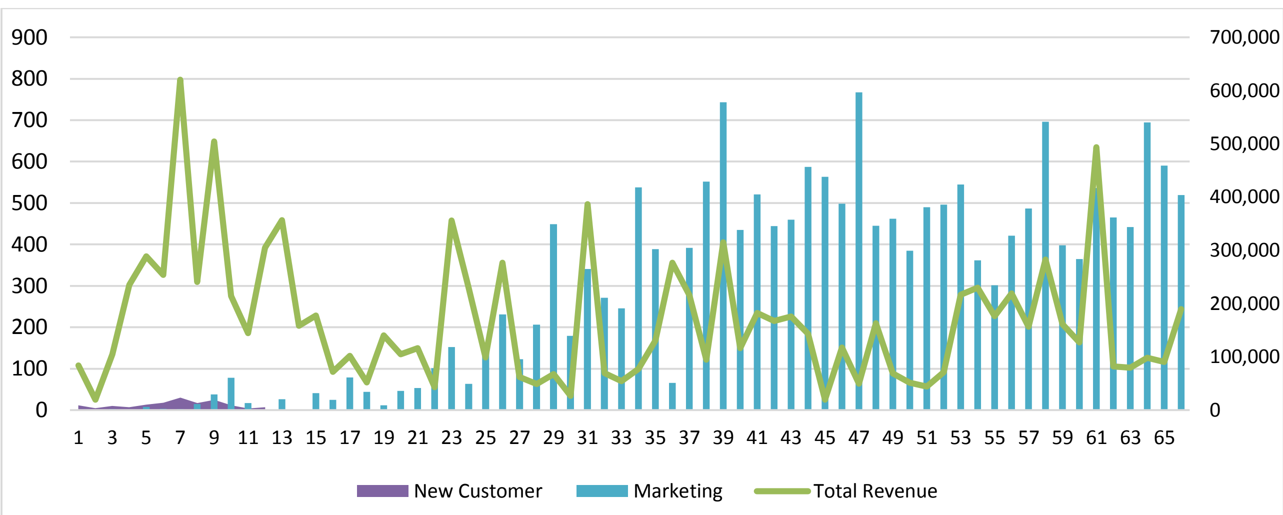
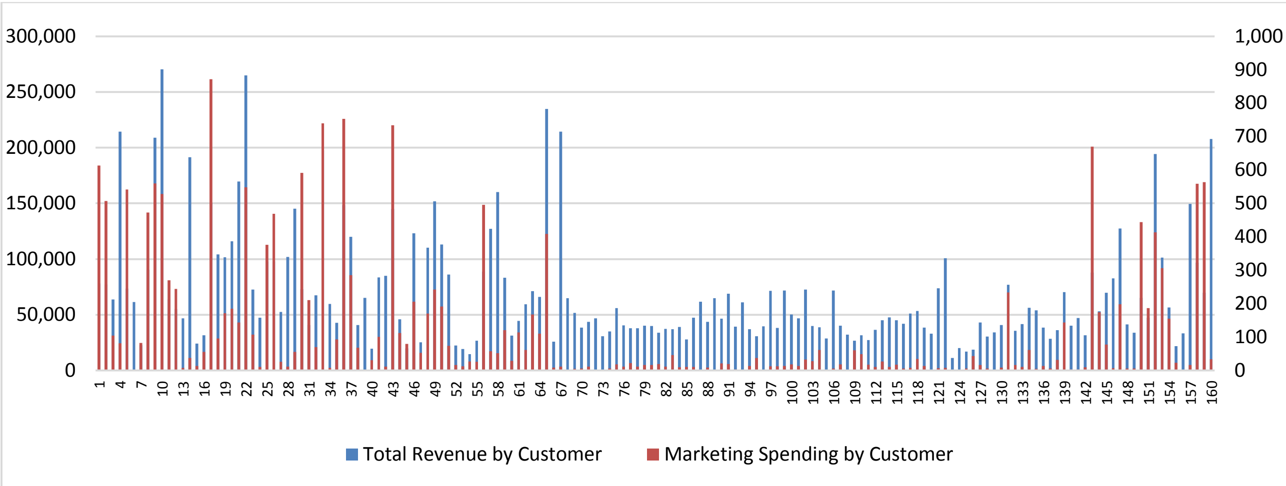


Figure 10. Marketing Spending by Customer



While from the supply side, the marketing spending allocation is pretty general and subjective. The demand is heterogeneous due to various factors.

Business Customer Industry and User Size

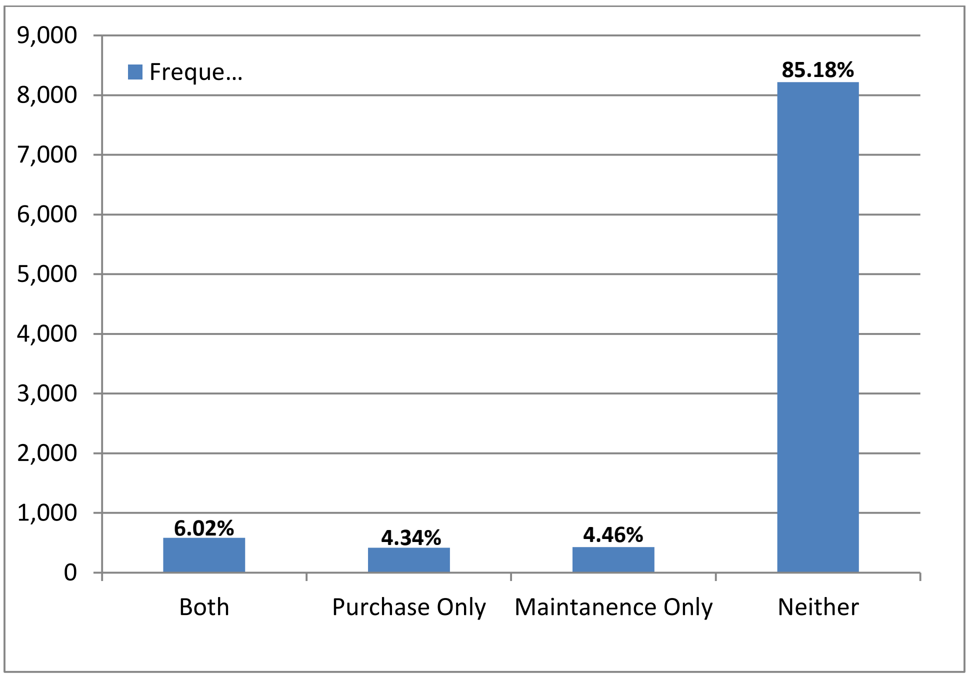
The business customer's industry influences what types of machines they purchase, therefore, influence the size of their purchases. There are 9 industries where the customers were coming from. They are: *Retail, Manufacture, Consumer Package Goods, Wholesale, Health, Travel, Media-Entertainment, Auto and Government.* Among these industries, customers in *Health* industry have the highest average purchase quantity at 29,953, while retail has the lowest purchase quantity at 10,573.

While we selected middle sized firm with employment size between 50 and 500 for the analysis, the actual demand for IT products are better reflected by the actual number of users of the products. The average number of users from an establishment is 4, and the maximum number of users is 15.

Maintenance Service Interactions

Besides purchases, customers also interact with the focal firm with maintenance service interactions. Figure 11 shows the number of months by the types of customer-firm interactions. There are 581 (6.02%) months with both interactions, 461(4.34%) months with purchase only, 430(4.46%) with maintenance only, and 8,218 (85.18%) with no interactions between customer and the firm.

Figure 11. Number of Months by Purchase and Maintenance Interactions



Promotion Offer Interaction

In the B2C scenario, promotion is one of the major drivers for sales. However, In B2B scenario, promotion normally happens as a result of the negotiation between sales representative and the business buyer during the purchase. In other words, *Promotion* generally does not lead the customer to make the purchase decision, however, it could help the focal firm to build relationships with the business customers. Around 60% of the customers have received promotion for at least once. The depth of promotion is generally less than 10%. Among all the purchases, about 20.4% of the purchases received promotion. We don’t see high association between the promotion and level of purchases. It’s because that unlike B2C scenario, the decision makers are generally not the one who pay for the purchases, therefore, they are less price sensitive.

Time since Last Purchase

As stated in the last section, one of the biggest characteristics of business purchases is that they are very infrequent. It depends on a firm’s natural demand on the particular products and services. Since we selected middle sized business customers, their average inter-purchase time range from 6 months to 14 months. To disentangle the effects of natural demand for the products and relationship dynamics, we created time since last purchase as one of the independent variables in both transition and emission equation.

Purchase Level

Adaptive learning with continuous state and continuous emission observation is extremely challenging(Porta et al. 2006a).The reason is the explosion of state space when both state and observation are continuous. For adaptive learning purposes, the continuous emission created extra computation burden to discretize the new observation for conditional planning. Studies in Artificial Intelligence(AI) have developed various ways to tackle the problem including point-based value iteration and policy directed observation aggregation with discretization(Porta et al. 2006a). However, AI field generally works on problem with one fixed transition and emission functions. While in this study, we want to focus on the heterogeneity in DCV and VOL. In other words, we have individual customer level transition and emission functions. It would be computationally impossible to implement conditional planning with continuous observations for many customers at the same time. Therefore, we discretized the purchases into five levels by clustering. Table 2 shows the count of months by observed purchase levels and the average purchase amount by levels across all customers.

Table 2 Levels of Purchases

	Month Count	Percentage	Average Purchase
No purchase	8648	89.6%	0
Level 1 Low	445	4.6%	3,813
Level 2 Median	309	3.2%	7,421
Level 3 High	196	2%	20,123
Level 4 Super High	50	0.6%	63,331

We can see that purchases are very infrequent with only about 10% months with purchases. For the months with purchases, most cases are low to median level with less than 10,000 purchase amount.

2.2 Hidden Markov Model for B2B Purchases

To obtain the optimal marketing strategies, we first use HMM to estimate the customer demand response model. There are three major components of HMM models, they are: initial state distribution, state dependent choices and transition dynamics. We are going to elaborate each component in the following sections.

2.2.1 Initial State Distribution (π_{is})

The core state is used to summarize the firm's view about each individual customer. Assuming the evolution of the customer relationship intensity is an underlying Markov process that governs each individual customer's purchase behavior. Assuming there are NS discrete relationship states for all the customers.

$$S \in \{0,1,\cdots NS\} = \{very\ weak, \cdots very\ strong\}$$

The actual number of levels NS will be determined by the data.

Initial state distribution is the focal firm's original belief on the customer's state membership at the beginning of the study period. The initial state distribution depends on the interactions between the focal firm and its business customer prior to the first actual transaction. As most firms only start to keep track of the customers after they make the first purchase, research using HMM either assumes there is an equal probability for a customer to be in any state (Schweidel et al. 2011) or it assumes all customers start with the lowest relationship state (Montoya et al. 2010; Poupart et al. 2006). In our case, about 10% of customers were touched by the focal firm a few times before they made their first purchase. We use this pre-purchase information to empirically estimate the initial state distribution using a standard logit model (Kumar et al. 2011; Netzer et al. 2008; Scott 2002).

Let π_{is} denotes the probability that customer i is in state s at time 0, where $\pi_{is} \geq 0$ and $\sum_{s=1}^{NS} \pi_{is} = 1$. The initial state probabilities can be estimated as,

$\pi_{is} = \Pr(S_{i0} = s) = \frac{\exp(\rho_\pi PreMkt_\pi)}{1+\sum_s^{NS} \exp(\rho_\pi PreMkt_\pi)}$, $s \neq NS$ The probability of a customer being in the lowest state is, $\pi_{i0} = \Pr(S_{i0} = 0) = \frac{1}{1+\sum_s^{NS} \exp(\rho_\pi PreMkt_\pi)}$ for identification purpose.

2.2.2 Sate Dependent Purchase Probability (p_{ijlt})

Conditional on the relationship state level in month t, a customer makes purchase decisions based on customer's intrinsic demand for the products as well as the marketing interactions between the focal firm and the customer. Due to the lumpy nature of business purchases, we discretize the purchase quantity into multiple levels. We adopt the general logit model with varying cut points to capture the heterogeneity in the size of purchases across industries.

The level of hardware purchases by customer i at time t under relationship state j can be generally expressed as,

$$Y_{ijt}^* = \beta_j + \beta_{Cj} * g(CUSTCHAR_i) + \beta_{Mj} * f(MARKETING_{it}) + \beta_{Oj} * h(OTHER_{it}) + \varepsilon_{jt}$$

$$\varepsilon_{jt} \sim GEV(0, \sigma_j)$$

$$Y_{ijt} = 0 \; if \;\; Y^* < 0; Y_{ijt} = l \; if \;\; \mu_{il-1} < Y^* < \mu_{il} \; l = 1,2,3 \ldots; \qquad Y_{ijt} = L \; if \;\; Y^* > \mu_{iL}$$

Despite the fact that business purchases are lumpy and infrequent, business customers are almost constantly contacted by their business suppliers. In other words, the interactions between the focal firm and its customers are happening every month. Therefore, we assume the business customers consider about purchasing every month. Specifically, the underlying utility of hardware purchase is determined by the intrinsic utility of purchase under relationship state j, customer firm *i's* characteristics, including the total number of employment of the firm as well as the actual number of users of the products. The utility also depends on the other customer-firm interactions at the particular time point, including promotional offers as an outcome of the negotiation between the firm and the client representative including discount, free hardware, etc. These interactions are generally controlled by sales representative at the point of purchase, which are not part of the pre marketing planning.

The cut points of the underlying random utilities of purchase for customer i is determined by the industry variables,

$$CUT_i = \tau_1 * AUTO + \tau_2 * CON\_PKG + \tau_3 * GOV + \tau_4 * HEALTH + \tau_5 * MANUFACTURE + \tau_6 * MEDIA + \tau_7 * WHOLESALE + \tau_8 * TRAVEL$$

Since we only observe the purchase behavior if the underlying utility is greater than 0. We restrict

$$\mu_{i0} = 0$$

To ensure the condition $\mu_{l<}\mu_m \; if \; l < m$, i.e., the cut points for lower level purchase is lower than the cut points of higher purchases, we restrict the parameter for the cut points to be the form,

$$\mu_{i1} = \exp(\nu_1) * exp[CUT_i]$$

$$\mu_{il} = [\exp(\nu_1) + \cdots \exp(\nu_l)] * exp[CUT_i]\ for\ l > 1$$

Assuming the deterministic part of HW^*_{ijt} is $\beta_i X_{ijt}$, from the ordered logit choice model, the probability of customer i, purchasing level l at time t under relationship level j can be expressed as,

$$P_{ijlt} = \frac{exp\left(\frac{\beta X_{it} - \mu_{il-1}}{\sigma_j}\right)}{1 + exp\left(\frac{\beta X_{it} - \mu_{il-1}}{\sigma_j}\right)} - \frac{exp\left(\frac{\beta X_{it} - \mu_{il}}{\sigma_j}\right)}{1 + exp\left(\frac{\beta X_{it} - \mu_{il}}{\sigma_j}\right)}$$

Where σ_j is the standard deviation of the emission random utility function in state *j*. This general form also consist of the case of $l = 0$ and $l = L$.

For$l = 0$, $\mu_{i0} = 1$,setting $\mu_{i-1} = -\infty$, the equation is reduced to $P_{ij0t} = \frac{1}{1+exp\left(\frac{\beta X_{it}}{\sigma_j}\right)}$

For$l = L$, setting $\mu_{iL} = \infty$, the equation is reduced to $P_{ijLt} = \frac{exp\left(\frac{\beta X_{it}-\mu_{il-1}}{\sigma_j}\right)}{1+exp\left(\frac{\beta X_{it}-\mu_{il-1}}{\sigma_j}\right)}$

2.2.3. Relationship State Dynamics ($q_{itjj'}$)

As stated in 2.2.1. we assume there is a discrete underlying relationship state between the focal firm and the business customer from *very weak* to *very strong.* The transition of the relationship state from one period to the next depends on the business customer-firm interactions, which includes the past purchase experience, service interactions as well as the marketing contacts by the focal firm.

Following (Luo and Kumar 2013), we define $TR_{itjj'}$ as business customer *i*'s propensity for transition from state *j* to *j'* at time *t*.

$$TR_{itjj'} = \gamma_{jj'} + \gamma_{Ijj'}\, g(INDUSTRY\ CHAR) + \gamma_{Mjj'}\, f(MARKTING) + \gamma_{Ijj'}\, h(INTERACTIONS) + \zeta_{itjj'}$$

$$\zeta_{itjj'} \sim GEV(0, \delta_{jj'})$$

Assuming the deterministic part of transition utility for transitioning from *j* to *j'* is $\gamma_{jj'} Y_{jj'}$, where $Y_{jj'}$ is the set of variables in transition utility function and $\gamma_{jj'}$ is a vector of their parameters. The transition probability can be expressed as,

$$q_{itjj'} = \frac{exp[\gamma_{jj'}\ Y_{jj'}]}{1 + \sum_{j \neq j'} exp[\gamma_{jj'}\ Y_{jj'}]}$$

For identification purpose, the probability of staying in state *j* is

$$q_{itjj} = \frac{1}{1 + \sum_{j \neq j'} exp[\gamma_{jj'}\ Y_{jj'}]}$$

In sum, there are four sets of parameters to be estimated: ρ_π from the initial distribution model; $\{\tau, \beta, \sigma\}$ from the state dependent choice model (emission function); $\{\gamma, \delta\}$ from the transition model. The specific independent variables and their function forms will be determined by the estimation.

2.2.4 The Likelihood of an Observed Sequence of Choices

As the observed customer choices are governed by the underlying relationship state, which evolves as a Markov chain. A customer's current choice depends on the full history of choices it made in the past. In other words, the likelihood of choices have to sum over possible path that the customer could take in the course of state transition(Netzer et al. 2008). The total likelihood of an observed sequence of choices up to time T by customer i is,

$$L_i(Y_{i1}, \dots Y_{iT_i}) = \sum_{s_0=1}^{NS} \sum_{s_1=1}^{NS} \cdots \sum_{s_T=1}^{NS} \left[\pi_{is_0} \prod_{t=1}^{T} q_{itjj'} \prod_{t=1}^{T} P_{ijlt}{}^{I(Y_{ijlt}=1)} \right]$$

where, π_{is_0} is the initial state distribution from Section 2.2.1, $q_{itjj'}$ is the state transition matrix from section 2.2.2, and P_{ijlt} is the emission choice probabilities from section 2.2.3. The total likelihood is the sum of all the possible path of the state evolution and observed sequences. The parameters are estimated by maximum likelihood estimation.

2.3 Estimation Results and Implications

After trying various specifications for emission and transition functions, we adopted a two state model where a customer is has higher utility from purchasing, and more responsive to marketing actions. We choose the two-state model for two reasons: 1). the two-state HMM performs almost good as the three-state HMM model based on the BIC. 2). The state variable will have two dimensions, so the policy is a direct mapping from state to actions without having to simulate marketing policies. Additionally, the curse of dimensionality is especially prominent for POMDP models where the planning complexity grows exponentially with the dimension of the state.

The final emission functions for the two states are:

$$Y_{i0t}^{*} = \beta_{01} + \beta_{02} CUST\_NUM_i + \beta_{03} COVERAGE + \beta_{04} LNMARKET\_EX_{it} + \varepsilon_{0t}$$

$$Y_{i1t}^{*} = Y_{i0t}^{*} + \exp(\beta_{11}) + exp(\beta_{12}) LNMARKET\_EX_{it} + \varepsilon_{1t}$$

The final emission functions for transitioning to high state and staying in high state are:

$$TR_{it01} = \gamma_{01} + \gamma_{101} IND_{INTPUR_i} + +\gamma_{201} COVERAGE_i + \gamma_{301} * MARKEX_{it} + \gamma_{401} * MARKEX_{it}^2 + \gamma_{501} PROMO_{it} + \gamma_{601} LNTSLP_{it} + \zeta_{it01}$$

$$TR_{it11} = \gamma_{11} + \gamma_{111} * MARKEX_{it} + \gamma_{211} * MARKEX_{it}^2 + \gamma_{311} MAINTEIN_{it} + \gamma_{411} LNTSLP_{it} + \zeta_{it11}$$

We estimated the model with maximum likelihood estimation. The estimation results are shown in Table 3. The results show that the customers have much higher intrinsic propensity to make a when they are in high state. Some customer characteristics variables like actual customer count in the business organization, whether or not the customer is under the focal firm's service coverage have positive effects on purchase level. At the low state, marketing expenditure has weak but significant effect on sales. However, marketing expenditure is very effective on boosting sales when the customer is in high state.

For state transition, a customer has more force move to high state when it's a new customer (tenure<2 years), and a customer has higher propensity to go to high state if it's under the service coverage. The longer the industry level inter-purchase time, the lower propensity for the customer to move to high state. The marketing effects are concave for both transition functions. Promotion offers helps the firm to move customer from low to high state and maintenance interactions helps the firm to stay in high state. Time since las purchase has positive effects on transitioning to high state.

The industry indicator variables are all significant in the cutoff function. In other words, the purchase level varies across business customers' industries.

Table 3. Estimation Results for 2-State Hidden Markov Model

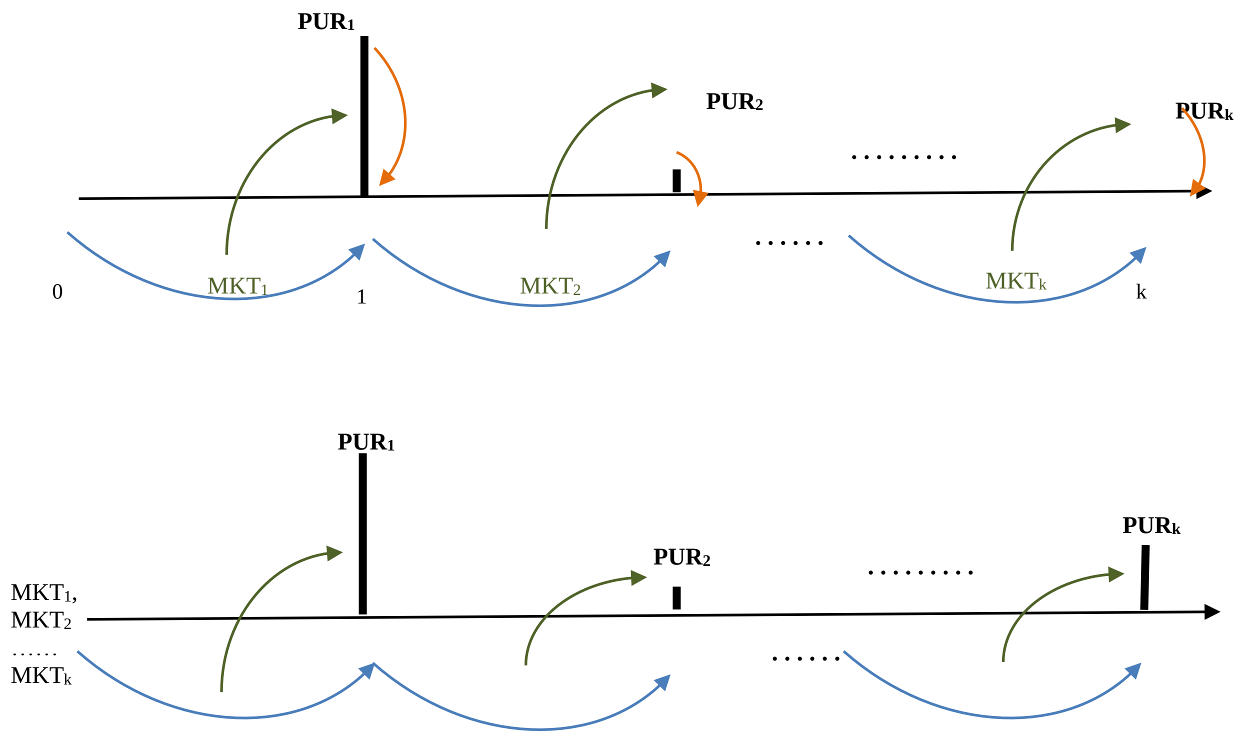
	Parameter	Estimates	Std. error	T-Value
Initial State	Ln(pre-purchase marketing)	45.080	33.147	1.360
Transition Function	$S_{0\text{-}1}$ Intercept	-2.631	0.284	-9.275
	$S_{0\text{-}1}$ Coverage Type	0.347	0.141	2.463
	$S_{0\text{-}1}$ Industry Inter-purchase Time	-0.123	0.049	-2.531
	$S_{0\text{-}1}$ New Customer	0.197	0.129	1.525
	$S_{0\text{-}1}$ Marketing Expenditure/10	0.062	0.034	1.825
	$S_{0\text{-}1}$ (Marketing Expenditure/10)^2	-0.016	0.014	-1.181
	$S_{0\text{-}1}$ Promotion Offer	2.565	0.575	4.460
	$S_{0\text{-}1}$ ln(Time since last purchase)	1.018	0.527	1.932
	$S_{1\text{-}1}$ Intercept	-2.951	0.246	-11.981
	$S_{1\text{-}1}$ Marketing Expenditure/10	0.295	0.098	3.022
	$S_{1\text{-}1}$ (Marketing Expenditure/10)^2	-0.032	0.021	-1.524
	$S_{1\text{-}1}$ Maintenance Interaction	2.563	0.575	4.461
	$S_{1\text{-}1}$ ln(Time since last purchase)	0.797	0.351	2.269
Emission Function	S_0 Intercept	-3.710	9.738	-0.381
	S_0 Customer Count	0.133	0.096	1.387
	S_0 Coverage Type	1.572	0.538	2.920
	S_0 ln(Marketing Expenditure)	0.225	0.126	1.780
	Addl. S_1 Intercept (exp)	4.378	2.696	1.624
	Addl. S_1 ln(Marketing Expenditure) (exp)	5.591	2.347	2.382
	S_0 std. dev. (exp)	-0.021	1.909	-0.011
	S_1 std. dev. (exp)	2.656	2.632	1.009
Cut-off Function	Travel	-0.153	0.084	-1.811
	Auto	-0.180	0.106	-1.694
	Consumer Package Goods	-0.347	0.146	-2.375
	Government	-0.271	0.086	-3.142
	Health	-0.159	0.083	-1.925
	Manufacturer	-0.142	0.059	-2.399
	Media	-0.282	0.130	-2.161

	Wholesale	-0.144	0.049	-2.922
	Cut1 (exp)	1.537	0.204	7.530
	Cut2 (exp)	0.680	0.060	11.367
	Cut3 (exp)	0.881	0.075	11.710
Log Likelihood	-5409.87			

2.4 Optimizing Marketing Expenditure with HMM and POMDP

From the estimation of HMM, we have learned the how customer's relationship evolve over time, as well as the short term effects of marketing actions on sales and long term effects of marketing actions on cultivating customer relationships. With all these information, the goal of the firm is to dynamically allocate marketing resources for optimizing dynamic customer value. The popular strategy is to fully rely on the information from HMM and allocate marketing resources accordingly. This strategy is called *Dynamic HMM* or *Pure Exploitation* strategy which assumes we know everything about the customer's response to focal firm's actions and purely rely on this information for resource allocation. This strategy does not require the focal firm to keep track of the evolution of customer's relationship state. However, the business customer purchase is very lumpy and infrequent, using *Dynamic HMM* strategy to allocate all the marketing resources in the future period can lead to missing some important information during the future customer-firm interactions. Therefore, we simulate all the possible future outcomes to see how much benefit the focal firm can obtain by continuously learning all the customer-firm interactions and monitoring the relationship state evolution.

Figure 12. Timing of Allocation: Adaptive Learning vs. Pure Exploitation



2.4.1 Transition Probabilities and Emission Probabilities

From the parameter estimates in section 2.3, we calculate the average initial belief state distribution, state transition and emission probabilities. From the Table below, we can see that the customers normally start with the low initial state as there were only about 9% customers who had been contacted before they made their first purchase. For those who had been contacted before purchase, their initial probability of being in high state is about .545. For the rest of the customers, they all start from low state with initial probability of being high state 0.

Table 4 Average Initial State Distribution, Transition and Emission Probabilities

		Pi
Initial State Level	Low	0.921
	High	0.079

		State Level (t+1)	
		Low	High
State Level(t)	Low	0.9516	0.0484
	High	0.8869	0.1131

		Emission Purchase Level				
		0	1	2	3	4
State Level	Low	0.9531	0.0346	0.0084	0.0033	0.0005
	High	0.0043	0.2568	0.4004	0.2688	0.0697

From the transition probability table, there is only 4.8% probability a customer transit from low to high state, whereas the probability of staying in high state is 11%. These results are consistent with the common belief that it's easier to maintain a happy customer than to transfer a new customer to a loyal customer. The low total probability of being in high state also reflects the nature of low purchase incidence in this data. The emission probabilities shows the customer decisions are very state-dependent for this data set. When a customer is in high state, it almost always makes are purchase, whereas it only has only about 5% change of making a purchase when it's in low state.

2.4.2 Optimization with Pure Exploitation (Dynamic HMM)

If the focal firm adopts a *Pure Exploitation* strategy, it means it only acts on the current belief assuming it knows all the future state dynamics. Specifically, we know that the state dynamics is by equation below.

$$b_{i,t+1}\left(s' \mid b_{it}, MKT_{it}\right) = b_{it} * q_{it11} + (1 - b_{it}) * q_{it01}$$

The dynamic optimization problem with Pure Exploitation can be expressed by Bellman Equation,

$$DCV_i^{HMM} = V_i^*\left(b_{i,t}\right)$$
$$= Max_{MKT_{it}} \left\{ \sum_{s=1}^{NS} b_{it}(s) * \left[E_{b_{it}}\left[Y_{ijlt} * \overline{REV_{\iota l}}\ |b_{it}\right] - MKT_{it}\right] \right.$$
$$\left. + \rho \sum_{s=1}^{2} b_{it}(s) \left[V_i^*\left(b_{i,t+1}\left(s^{'}\ \ |b_{it}, MKT_{it}\right)\right)\right] \right\}$$

The first term of the equation is just the instant expected revenue from the control variable, marketing spending (MKT_{it}). The second term is the expected future value associated with transitioning from state b_t to b_{t+1} with marketing actions. ρ is the discount factor. The first and the second term of the Bellman equation reflects the idea of balancing short-term profit and building the long-term relationship with the customers.

2.4.3 Optimization with Adaptive Learning (POMDP)

While Dynamic HMM strategy is forward looking and dynamic, it assumes the focal firm knows evolution of relationship dynamics. In other words, this strategy does not adaptively obtain more information regarding the customer relationships in the future. Alternatively, the focal firm can adaptively learn about the customer state while allocating marketing expenditure.

At the beginning of each period, the focal firm has a belief on the distribution of the customer's relationship state membership. Specifically, in our case, $b_{i,t} = Prob(S_{it} = 1)$. After observing customer's state dependent choice Y_{ijlt} obtained from 2.2.2, combining with the state transition matrix obtained from 2.2.3, the focal firm incorporates this new information into its knowledge about its customer's relationship state. Each period, the focal firm's belief on each customer's relationship state membership is updated by Bayes' rule,

$$b_{i,t+1}\left(s^{'} = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) = \frac{\sum_{s=1}^{J} b_{it}(s) q_{itjj^{'}} \ P_{ijlt}}{\sum_{s=1}^{J} \sum_{l=1}^{J} b_{it}(s) q_{itjj^{'}} \ P_{ijlt}}$$

Specifically, for our two-state model, assuming the total force of transitioning to high state is $b_{i,t+1}\left(s^{'} | b_{it}, MKT_{it}\right) = \varphi_{it}$, the belief update function after having spending *MKTit* and observing level *l* of purchase can be expressed as,

$$b_{i,t+1}\left(s^{'} = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) = \frac{\varphi_{it} * P_{i1lt}(MKT_{it})}{\varphi_{it} * P_{i1lt}(MKT_{it}) + (1 - \varphi_{it}) * P_{i0lt}(MKT_{it})}$$

DCV with POMDP can be expressed as,

$$\begin{aligned} DCV_i^{POMDP} &= V_i^{*}\left(b_{i,t}\right) \\ &= Max_{MKT_{it}} \left\{ \sum_{s=1}^{2} b_{it}(s) * \left[E_{b_{it}}\left[Y_{ijlt} * \overline{REV_{il}} \, | b_{it}\right] - MKT_{it}\right] \right. \\ &\quad \left. + \rho \sum_{j=1}^{2} \sum_{l=0}^{4} b_{it}(s = j) P\left(Y_{ijlt} | b_{it}(s), MKT_{it}\right) \left[V_i^{*}\left(b_{i,t+1}\left(s^{'} = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right)\right)\right] \right\} \end{aligned}$$

The first term is the expected instant revenue, the second term is the discounted expected value from transitioning to the next belief level after new observations in the future period. It is an expected value on all the potential observations in the future. In other words, the marketing expenditure in the current period also account for the fact that the firm will adaptively learn about the customer in the future.

2.4.4 Optimization Algorithm

There are two commonly used algorithms for the dynamic optimization problems. They are, *policy iteration* and *value iteration.* Some very good review and summary papers in both operation research and computer science fields are available for the solution method for POMDP (Bertsekas 2000; Kaelbling et al. 1998). The major challenge of solving this type of problem is the continuous state space (Montoya et al. 2010; Sun and Li 2011). We adopted similar approach as Montoya et al. (2010) and Sun and Li (2011) for value function interpolation. To solve our dynamic allocation problem, we used value iteration with approximation for an infinite horizon to find the unique fixed point of the bellman equations in 2.4.2 and 2.4.3(Bertsekas 2000).

With the parameters we obtained from the estimation, we calculated the initial state distribution, emission function and transition function for each individual customer. There are two types of independent variables in the transition and emission force functions: customer level variables including *Number of Customer*, *Industry Indicators*, *Coverage, New Customer Indicator* and time-varying variables including *Marketing Expenditure*, *Maintenance, Promotion*. All the parameter estimates of the customer level variables become intercept in the emission function as they don't change over time with the iteration process. *Marketing Expenditure* is the control variable for this optimization problem. *Maintenance*, *Promotion* are the dummy variables indicating if there were promotion and maintenance in the past customer-firm interactions. These two variables does not influence purchase directly, however, they influence the transition force function of the state variable. These two variables are not explicitly modeled in this study, calculated the percentage of months that a customer requested maintenance service and the percentage of transactions that a customer received promotion offers. The expected percentage multiplied by the parameter estimate of these two variables also become part of the intercept in the transition force function. After the calculations both emission and transition functions are functions of the control variable, *Marketing Expenditure*. We calculated the two-period optimization problem as the starting point of the value iteration, then approximate the future value through value interpolation(Keane and Wolpin 1994a). The algorithm iterates through the steps of value maximization and belief update (pure exploitation with HMM and adaptive learning with POMDP), then converges to a unique fix point. The details of the algorithms for solving the POMDP and HMM problems are in Appendix A.

2.4.5 Optimal Policies for HMM and POMDP

After value iteration described in 2.4.4, we obtained optimal policies for each individual customers for both HMM and POMDP. As we wanted to simulate multiple purchase sequences for each individual customers for DCV and VOL calculations. We selected 4 customers from the data, and simulated 100 random purchase sequences for each customer.

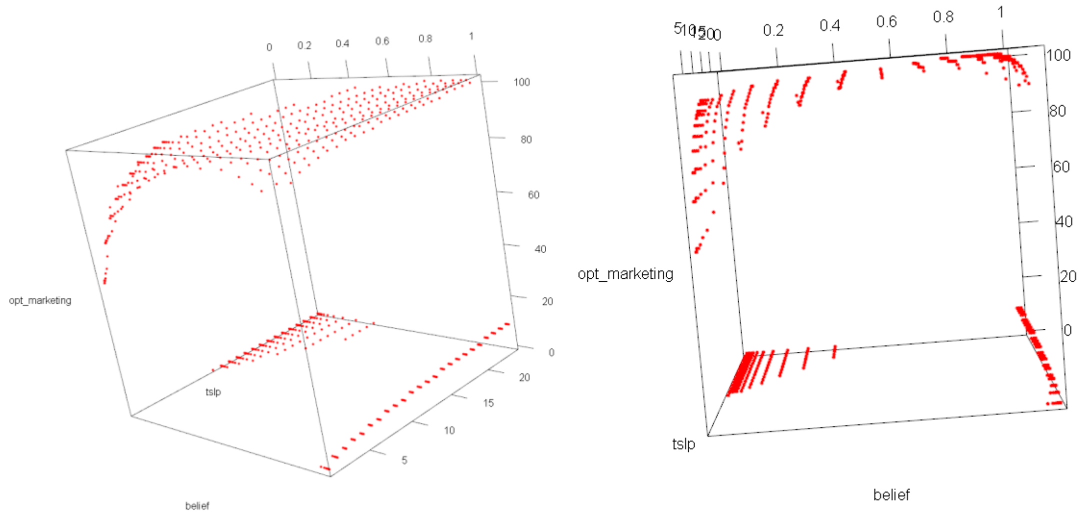
We simulated optimal marketing actions for 36 and 120 months for each customer to show the short-term and long-term effects of adaptive learning. For simulation, we first obtain policies for each customer from the value iteration. Figure 13 shows an example of optimal policy for a customer. The optimal policy is a mapping from two state variables: belief state and *TSLP* to the optimal marketing expenditure. Since the belief state is divided into 40 grids and

TSLP is from 1 to 24, there are 960 combinations of the two state variables. The policy is the optimal marketing expenditure for each of these 960 combinations.

Figure 13 is the optimal marketing policy for POMDP and HMM models for a customer. Figure a shows the optimal policy for POMDP model. We can see when the customer is in very low belief state and with long *TSLP*, the firms should stop spending on this customer. When it's in low belief but short *TSLP*, i.e., it had a fairly recent purchase, the firm should spend medium level marketing. From that point, the optimal marketing first goes up as the belief state goes up, then goes down as the belief becomes very high. As the belief approaches 1, the optimal marketing level drops to around 9.

Figure 13 Optimal Policy for POMDP and HMM Model for Customer A

a. POMDP Policy



b. HMM Policy

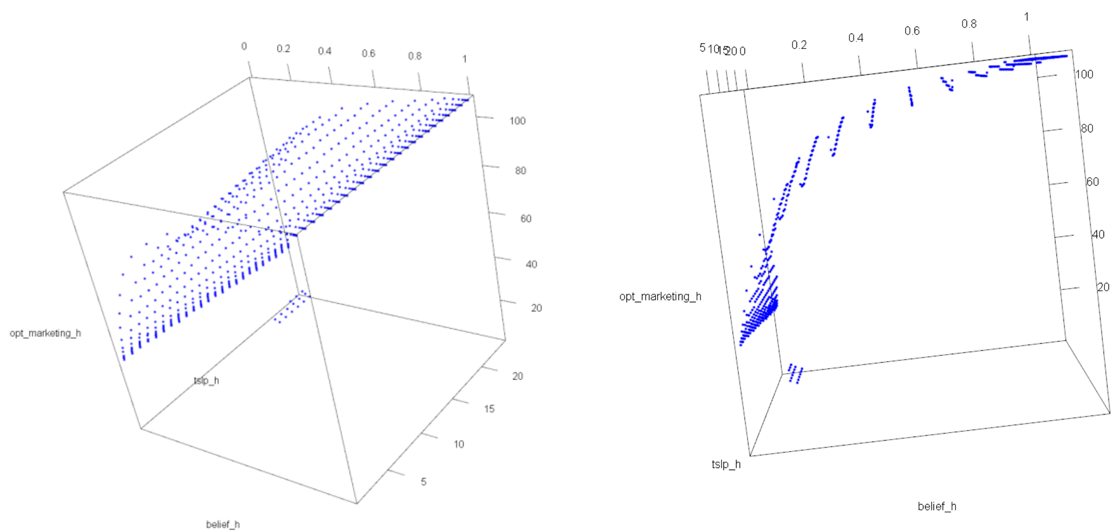


Figure 13b is the optimal policy from *Pure Exploitation* HMM strategy. From the figure, these two policies look similar. However, the actual policies and their implications are very different:

1). The optimal marketing for low belief state ranges from 27 to 34, which is lower than that of POMDP policy (52 to 93 for *TSLP<11*, 0 otherwise) regardless of *TSLP*. In other words, POMDP strategy first tries to improve customer relationship by higher marketing, then stop investing in marketing on this customer no purchase for 11 periods. Whereas HMM strategy recommends to keep a low marketing strategy for as long as possible.

2). The optimal marketing is monotonically increasing with the belief state, and reaches the highest of 107 at the highest relationship state for HMM strategy. The POMDP policy is slightly concave where it reaches the highest level of 101 when the belief is medium to high level (0.68-0.97) and decreases a little after the belief approaching to the highest level (0.97+). The decrease in state is also more prominent right after the customer made a purchase. It is consistent with the infrequent nature of business purchase where the customer is unlikely to make purchase right after another.

3). For HMM strategy, optimal marketing is almost independent of *TSLP*. Whereas for POMDP strategy, the optimal marketing is concave in *TSLP* for a given level of state. In other words, you don't need invest in marketing right away when you know for sure a customer is in high state, and had just made a purchase. When a customer has been consistently in low state and hadn't make a purchase for a long time, you should stop invest marketing in this customer.

2.5 Value of Learning (VOL) with Simulation

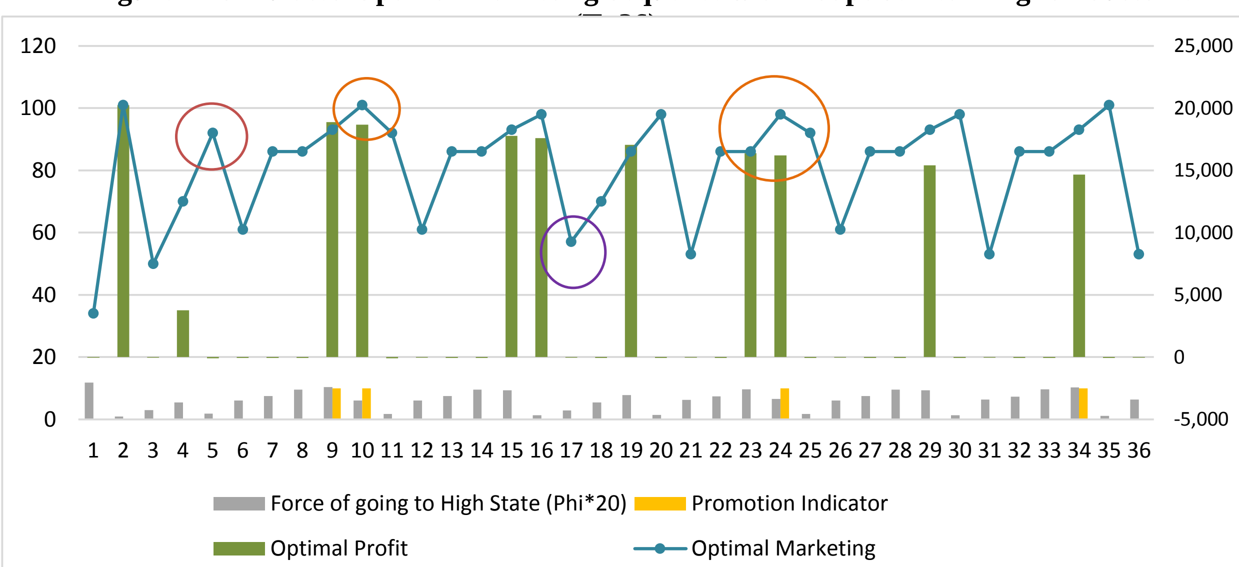
In the previous section, we generated the optimal marketing policies for POMDP and HMM strategies. The POMDP strategy incorporates more information while making decisions, it generally leads to a higher customer value. However, it is also harder to operationalize logistically. The double-loop learning framework we developed in the previous sections can help us to tailor the learning strategies by answering the questions of *Why do we learn? How do we learn? What do we use to learn?* and *Who should we learn?* By using simulation, we proactively answer all these questions before we allocate our marketing resources.

The general idea of optimal interaction sequence simulation is, we use the optimal policy to obtain the best possible response to the customer-firm interactions. In each period, started with a belief state, obtain optimal marketing action based on the policy, then generated state dependent choice with the optimal marketing action, and eventually update the state with the new observed choice and state transition function with the optimal marketing action. The next period will start with the new updated belief from last period. Besides the uncertainty on the customer state, there are three sources of demand shocks that the firm has to response to. The GEV distributed error term from the random utility when we simulated purchase level, the occurrence of promotion that comes with the purchase and occurrence of maintenance service interaction will be influence the transition force of the relationship state. The details of the Simulation procedure is in Appendix B.

2.5.1 Adaptive Learning: *How to learn?*

Using the optimal marketing policies in the previous section, we first did optimal marketing sequence simulation for customer A. Figure 14 is an example of the simulated sequence. The gray bars are the total force of going to high state (φ_{it}). Compare to actual state at t, this is a better indicator of the state dynamics at a given t. The yellow bars are the randomly simulated indicator of promotion offers for a given purchase. The promotion offers helps to increase φ_{it}. As we mentioned in the description of optimal policy for POMDP, the firm generally reduces to low state after making a purchase, therefore, the marketing is normally low following each purchases. Especially at t=17, marketing dropped immediately after two consecutive purchases (purple circle) because it is very unlikely for this business customer to make purchase three months in a row. However, sometimes when the transition force to high state is high enough, especially when the purchase received a promotion offer, there might be some chance to cross-sale another category with the boosted belief state. The orange circles in the figure shows this type of “strike while the iron is hot” tactic. At t=4, there’s a low level purchase with medium level φ_{it}, there is high marketing expenditure the following month to either upsell or to build relationships.

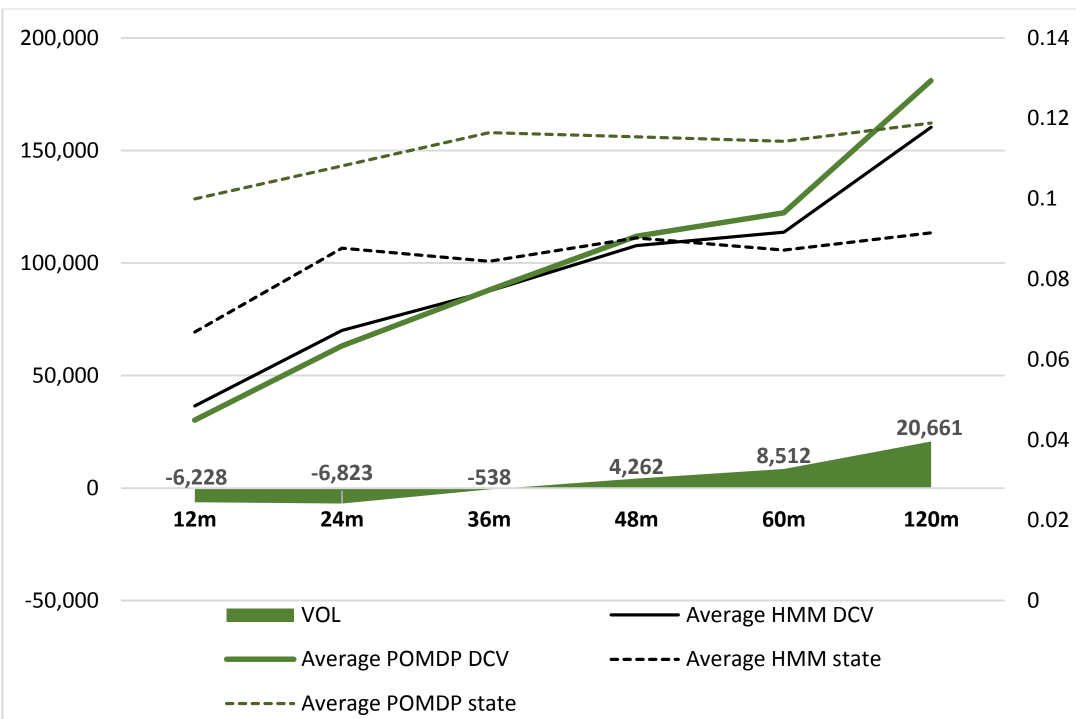
Figure 14 Simulated Optimal Marketing Sequence with Adaptive Learning for Customer A (T=36)



2.5.2 VOL: *Why do we learn?*

The adaptive learning strategy requires the focal firm to closely monitor the customer-firm activities and respond almost instantly. To justify this strategy, we need to quantify the value gain from this practice. Based on the optimization and simulation procedures we described in the previous two section, we obtained DCV^{HMM}, DCV^{POMDP} and VoL^{s} for the selected customers. W simulated learning effects for Short-term (12, 24 months), inter-mediate term (36, 48 months), long term (60 months) and very long term (120 months). Figure 15 shows DCV^{HMM}, DCV^{POMDP} and VoL for Customer A. For each planning horizon, we simulated 100 random sequences and Figure 15 shows the average values of these samples.

Figure 15. Simulation for Customer A (Sample Size=100)

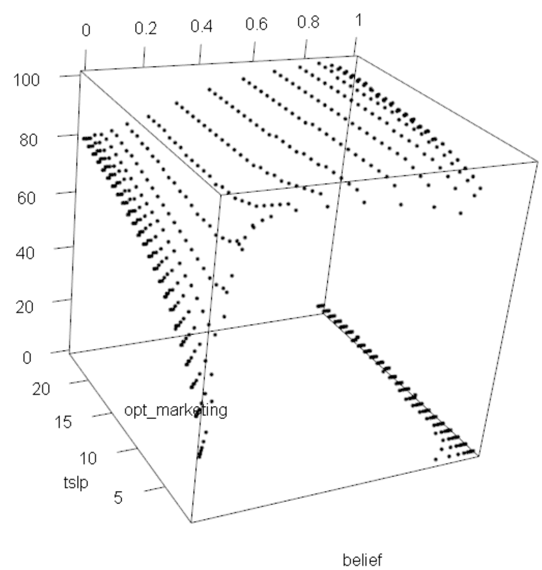


As from the optimal policies, POMDP strategy invests more marketing than HMM strategy, the adaptive learning strategy does not generate as much DCV as *pure exploitation* strategy. Especially for the infrequent B2B purchases, the investment in marketing on building the relationships and learning the customers may not generate direct results in the short run. Therefore, the VOL is negative until the relationship reaches 3-4 year. The management with myopic orientation may not want to adopt the adaptive learning strategy. However, the management with long-term focus can easily recover the loss by the 5^{th} year of the practice.

2.5.3 VOL from other Customer and firm Interactions: *What do we use to learn?*

We have shown in Figure 15 that the promotion offers can help to build business customer relationships. Unlike the B2C scenario where promotion directly increase sales, B2B promotion offers are highly dependent on the individual relationship between the customer and the sales representative and idiosyncratic situation at the time of the purchase. Since it is carrying some information regarding the customer-firm relationship dynamics, we want to quantify the benefit of integrating this costumer-firm interaction into our marketing planning practice. We first generated the optimal marketing without the promotion effects, then simulated the DCV without promotion effects. From the data, Customer A has relatively high probabilities of getting a promotion. It gets promotion offers 24% of its purchases. The optimal marketing policy for Customer A is shown in Figure 16. Compare to Figure 13a. The optimal policy with promotion offers, the optimal marketing is at a lower level for the optimization without promotion information. The reason is that, without accounting for promotion effects on improving relationships, this customer is considered at a lower belief state level which is corresponding to a lower marketing level.

Figure 16. Policy without Promotion Offers for Customer A



With this marketing policy, we simulated 100 samples of planning horizon $T=60$ as 5 year is a reasonable planning horizon for B2B industry. By comparing the $DCV^{No\ Promo}$ from the restricted model of "No adaptively learning through promotion activities" to the full model DCV^{POMDP}, we obtain the VOL from tracking promotion activities. Empirically, for our sample size of 100, it's calculated as,

$$VOL^{Promo} = \frac{1}{100}\sum_{j=1}^{100} DCV_j^{60\_POMDP} - \frac{1}{100}\sum_{j=1}^{100} DCV_j^{60\_No\ Promo}$$

Table 5. Simulation Results for Adaptive Learning without Promotion for Customer A (*T=60, Sample=100)*

Simulation for *T=60*	**Adaptive Learning Full Model**	**Adaptive Learning without Promotion Info**	**Difference (VOL)**
Average State	0.11425	0.09953	0.01472
Average Optimal Marketing	76	61	15
DCV	122,236	98,605	23,631

Table 5 shows the results of the simulation for optimal marketing without monitoring promotion offer information. The average level of belief state is lower by about 13%. Under this restricted learning strategy, the average optimal marketing spending is about 61, which is about 20% lower than the optimal spending of 76 for the full model. The value of monitoring the belief state through the promotion offers is the difference of the DCV from these two models, which is 23,631. The firm can achieve about 24% more value from monitoring the promotion activities for the planning horizon of 5 years.

2.5.4 VOL vs. DCV: *Who should we learn?*

After answering the questions of *Why, What, How* to manage customer relationship for a customer in the previous sections, We want to move on to answer the question of *Who* should we learn? Following the procedures we described in 2.5.1 to 2.5.3, we calculated VOL and DCV for four selected customers with different emission and relationship dynamics. Table 6 shows the DCVs and VOL for the four customers. Customer A has the highest VOL with medium level DCV. We can say this customer is about to become good friend to the company. Customer B has low DCV, medium VOL. It is going to become an acquaintance to the firm. Customer C has high

DCV, medium level VOL. It is the company's best friend. Customer D has low DCV and low VOL, who is like a stranger with low potential to the company. Both policies suggest to de-market this customer.

Table 6 VOL and DCV for Four Customers (*T=60, Sample=100*)

	DCV^{POMDP}	DCV^{HMM}	VOL	*Relationship Dynamics Type*
Customer A	122,236	113,724	8,512	Friend → Good Friend
Customer B	73,092	69,488	3,604	Stranger →Acquaintance
Customer C	227,626	224,834	2,792	Best Friend
Customer D	71,014	70,742	272	Stranger with low potential

Figure 13 in Section 2.4.5 has shown details about the difference in these two policies for Customer A. Figure 14 also shows the how to use marketing to build relationships with Customer A by continuously monitoring the customer relationship evolution. Figure 17 shows the optimal HMM and POMDP policies for the rest of the three customers. The red dots are POMDP policy and the blue dots are the optimal HMM policy. The pink lines are the difference in optimal marketing based on these two policies. Therefore, if the pink line is in the upper cube, it means POMDP policy suggests higher marketing expenditure than HMM policy; if the pink line is in the lower cube, HMM policy suggests higher marketing.

Customer B

The difference in treating Customer B under these two policies are mostly in two regions. The first region is when it's in low belief/high *TSLP* situation, HMM policy suggests to demark this customer with 0 optimal marketing spending while POMDP policy suggests to cultivate this customer with marketing investment. The second region is when the customer is in very high state, POMDP policy suggests to lower the marketing level as the customer is going to make a purchase while HMM policy suggests to keep on investing the highest level when the customer is in very high state. By shifting the marketing resource from “reinforcing high state” to “cultivating low state”, the firm gains 3,604 for adaptively learn about the customer state for 5 years.

Customer C

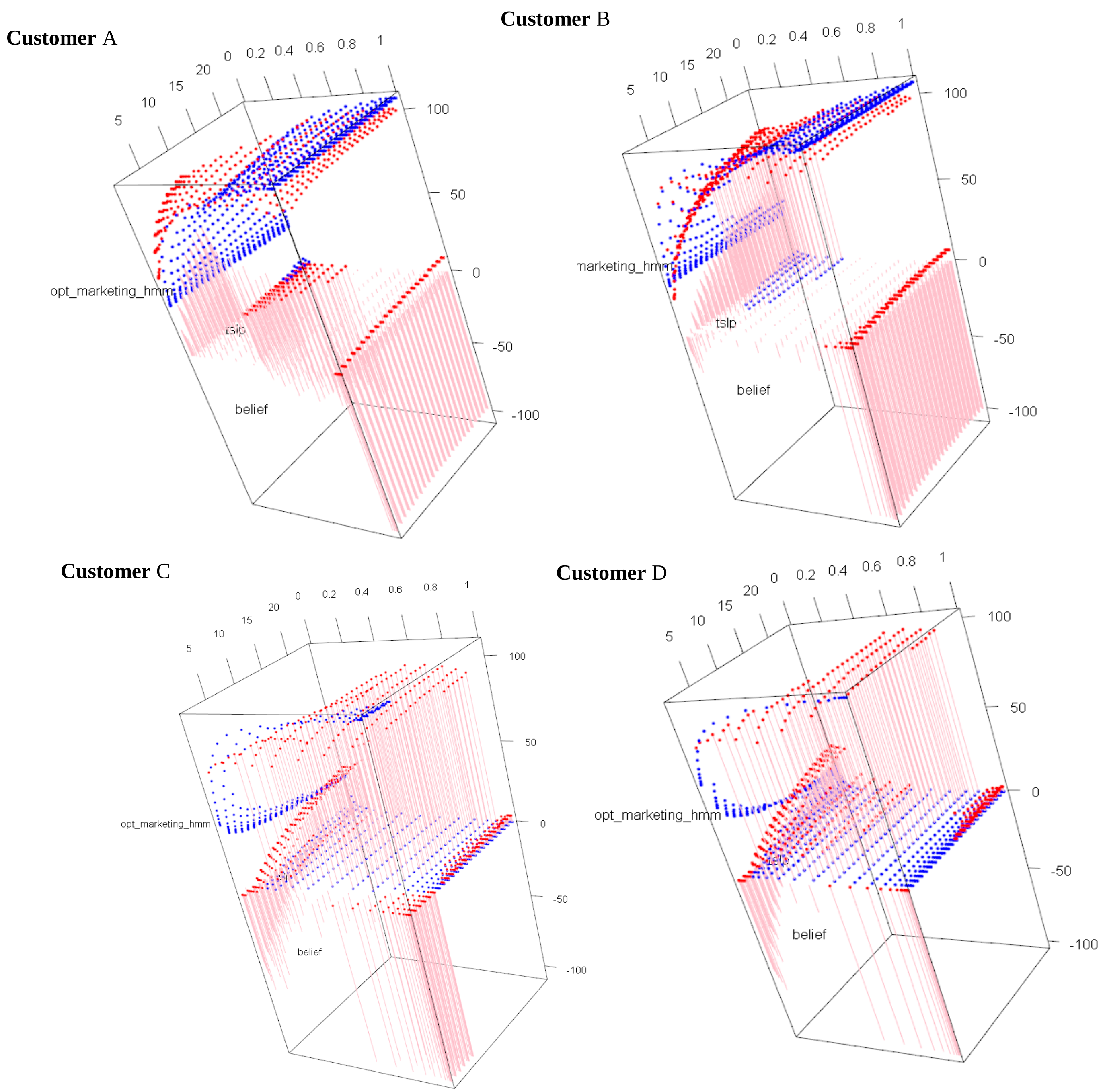
For customer C, POMDP invests more when it’s in medium level relationship state while HMM demark this customer unless it’s in low or high relationship state. When the customer is in low belief/low *TSLP* state, HMM policy suggests to invest more marketing, whereas POMDP policy suggests to wait until the customer “warm up” to the medium level belief state and *TSLP*, to push the sales.

Customer D

Customer D has low DCV and VOL. Both policies suggest that the firm should not spend on building relationship with marketing when this customer is in low relationship state or it has not been made a purchase for a long time.

By DCV standard, the firm should focus more on customer A and customer C. However, VOL tells us that we should cultivate the customers who have the most potentials to become more valuable customers when the firm manage the relationship with these customer in an integrated and proactive way.

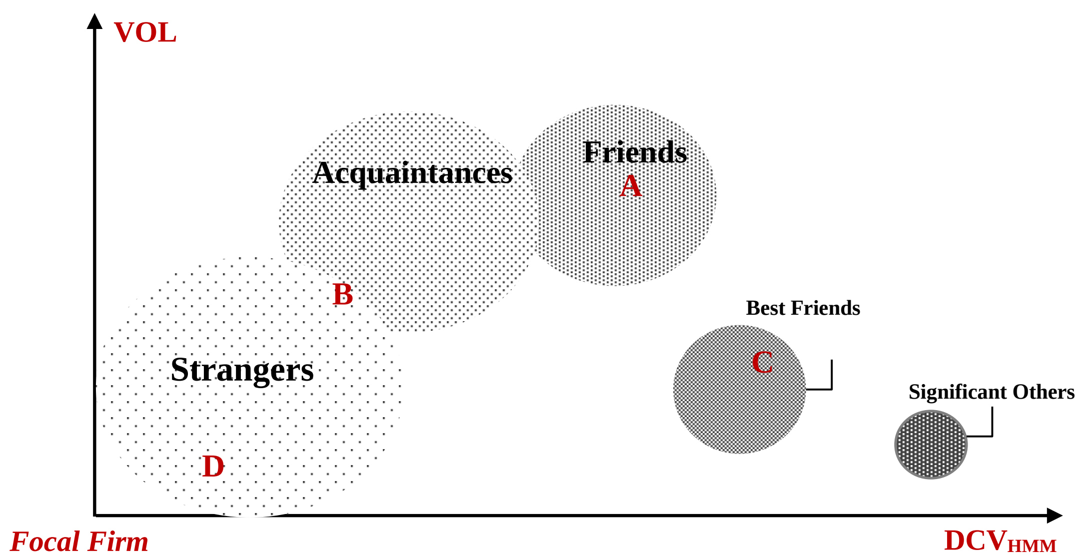
Figure 17. POMDP vs. HMM Policies for 4 Customers



2.5.5 Discussion

From Table 6 and Figure 18, we answered the *How*, *Why*, *What*, *Who* questions for the four selected customers from the data. If we take a second look at the relationship state map in the Motivation section, we can map these four customers into the *DCV/VOL* map to categorize their relationship stages. Applying the same approach, we can evaluate the entire customer base and pinpoint each customer on this map.

Figure 18 The Stages of Relationships for 4 Customers



2.6 Conclusions, Limitations, and Further Research

2.6.1 Contributions

In this dissertation, we propose a DLLM framework for managing the momentum of customer relationships through learning. This method is integrated as it helps us to incorporate transaction and non-transaction information to learn about customer relationships. It is also proactive as it provides a decision support system for optimal marketing actions accounting for the future gain from shifting relationship to a higher state as well as learning the customers.

The main contribution of this study is in suggesting a DLLM framework for managing customer relationships through learning. Adaptive learning models are starting to gain attention in the marketing literature. However, it has not been seen as key element of managing customer relationships. The proposed framework quantify the value gains from adaptively learn about the customers to justify its importance in the process of developing relationships with customers. The model goes beyond transaction aspect of customer-firm encounters and use the other non-transactional customer-firm interactions to learn about the customers. We also quantify the value of incorporating these information into our decision making. It also echoes the idea of interaction orientation which found that high interaction orientation of the firm leads to high performance (Kumar et al. 2004; Ramani and Kumar 2008). By using the data from a B2B IT firm, we demonstrated that the focal firm can achieve value gain from proactively learn about the customer’s relationship state through incorporating new information on the customer-firm interactions.

The proposed framework extends the CRM literature by incorporating VOL as another dimension on managing relationships. Instead using the conventional practice of categorizing the customers statically into four categories with CLV: True Friends, Barnacles, Butterflies, and Strangers (Reinartz and Kumar 2000; Venkatesan and Kumar 2004). We focus on managing customer relationship momentum. By mapping the customers into the VOL and DCV dimension, we have dynamic and forward-looking view of the relationship development. It also echoes the idea of obtaining competitive advantage through customer knowledge management(Garc et al. 2002).

It extends the machine learning and POMDP literature by providing a guideline from obtain customer response to operationalize the conditional planning in practice. Machine learning has become a very popular filed in business practice. Many algorithms were developed by computer scientists to help with various types of decision making. However these algorithms generally developed for given system dynamics. The proposed framework provides a road map to first learn the system dynamics of a key construct through various sources of observations. Then apply the outcomes to create optimal decision support system. Additionally, it also propose the idea of tailor learning activities through VOL.

2.6.2 Limitations and Future Research

This research also has limitations. While the focus of this study is to propose the framework to guide the learning practice in CRM. The HMM model we used in Essay 2 is relatively parsimonious. Due to small data size, we only accounted the customer heterogeneity in emission and transition as level effects through customer characteristics. Future research could use a larger dataset and incorporate heterogeneity in customer response parameters through hierarchy bays model. Due to the curse of dimensionality, we restricted the model to have two states (belief state and *TSLP)* and five levels of discrete observations. The problems in other scenarios could be much more complicated. One potential challenge is when the observation is continuous. In addition to the continuous belief state, the belief updating in the conditional planning step could be very complicated(Porta et al. 2006b).

The empirical application of the model in this study focused on dealing the uncertainties in managing relationships through learning. The virtue of the proposed model is that the firm constantly learning about customers' needs and the environment in which the customers make their purchase decisions. It could be extended to addressing other demand uncertainties. For example, it could be used to incorporate customers' potential strategic behaviors with which the model becomes a stochastic game between the focal firm and its customers. It could also be used to incorporate the uncertainty from the competition. In addition, we specify the objective function for the firm as a discounted total profit from the customer assuming that the manager is risk neutral. The model could be extended to add risk attitude of the manager in the objective function. Specifically, the managers' objective function could be a concave utility function with respect to the revenues over time.

To summarize, we propose an integrated and proactive framework for managing customer relationships. In incorporates learning as a key aspect of CRM. It answers the questions of *Why, What, Who* and *How* to adaptively learn customers in the process of managing customer relationships. It incorporates the concept of Double Loop Learning into the CRM and machine learning literature. It empirically developed the measure of VOL as an additional dimension on managing customer relationships. Hope both researchers and practitioners find this research useful.

REFERENCES

Argyris, Chris (1976), "Single-Loop and Double-Loop Models in Research on Decision Making," *Administrative Science Quarterly*, 21 (3), 363-75.

Berger, Paul D. and Nada I. Nasr (1998), "Customer lifetime value: Marketing models and applications," *Journal of Interactive Marketing*, 12 (1), 17-30.

Bertsekas, Dimitri P. (2000), Dynamic Programming and Optimal Control: Athena Scientific.

Bertsimas, Dimitris and Adam J. Mersereau (2007), "A Learning Approach for Interactive Marketing to a Customer Segment," *Operations Research*, 55 (6), 1120-35.

Cao, H. Henry and Baohong Sun (2007), "Value of Learning and Acting On Customer Information," *working paper*.

Cassandra, Anthony R. (1998), "A Survey of POMDP Applications," *Working Notes of AAAI 1998 Fal l Symposium on Planning with Partial ly Observable Markov Decision Processes, 17-24*.

Dunbar, R. I. M. (1992), "Neocortex size as a constraint on group size in primates," *Journal of Human Evolution*, 22 (6), 469-93.

Fader, Peter S. and Bruce G. S. Hardie (2009), "Probability Models for Customer-Base Analysis," *Journal of Interactive Marketing*, 23 (1), 61-69.

Garc, xed, M. a-Murillo, and H. Annabi (2002), "Customer Knowledge Management," *The Journal of the Operational Research Society*, 53 (8), 875-84.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V. Kumar, Nathaniel Lin, and Nalini Ravishanker S. Sriram (2006), "Modeling Customer Lifetime Value," *Journal of Service Research*, 9 (2), 139-55.

Hauser, John R., Glen L. Urban, Guilherme Liberali, and Michael Braun (2009), "Website Morphing," *Marketing Science*, 28 (2), 202-23.

Jain, Dipak and Siddhartha S. Singh (2002), "Customer lifetime value research in marketing: A review and future directions," *Journal of Interactive Marketing*, 16 (2), 34-46.

Kaelbling, Leslie Pack, Michael L. Littman, and Anthony R. Cassandra (1998), "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, 101 (1–2), 99-134.

Keane, Michael P. and Kenneth I. Wolpin (1994a), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76 (4), 648-72.

---- (1994b), "The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence," Federal Reserve Bank of Minneapolis.

Kumar, V., Girish Ramani, and Timothy Bohling (2004), "Customer lifetime value approaches and best practice applications," *Journal of Interactive Marketing*, 18 (3), 60-72.

Kumar, V., S. Sriram, Anita Luo, and Pradeep K. Chintagunta (2011), "Assessing the Effect of Marketing Investments in a Business Marketing Context," *Marketing Science*.

Luo, Anita and V. Kumar (2013), "Recovering Hidden Buyer–Seller Relationship States to Measure the Return on Marketing Investment in Business-to-Business Markets," *Journal of Marketing Research*, 50 (1), 143-60.

Montoya, Ricardo, Oded Netzer, and Kamel Jedidi (2010), "Dynamic Allocation of Pharmaceutical Detailing and Sampling for Long-Term Profitability," *Marketing Science*, 29 (5), 909-24.

Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27 (2), 185-204.

Porta, Josep M., Nikos A. Vlassis, Matthijs T. J. Spaan, and Pascal Poupart (2006a), "Point-Based Value Iteration for Continuous POMDPs," *Journal of Machine Learning Research*, 7, 2329-67.

Porta, Josep M., Nikos Vlassis, Matthijs T. J. Spaan, and Pascal Poupart (2006b), "Point-Based Value Iteration for Continuous POMDPs," *JOURNAL OF MACHINE LEARNING RESEARCH*.

Poupart, Pascal, Nikos Vlassis, Jesse Hoey, and Kevin Regan (2006), "An analytic solution to discrete Bayesian reinforcement learning," in Proceedings of the 23rd international conference on Machine learning. Pittsburgh, Pennsylvania: ACM.

Ramani, Girish and V. Kumar (2008), "Interaction Orientation and Firm Performance," *Journal of Marketing*, 72 (1), 27-45.

Reinartz, Werner J. and V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (1), 77-99.

---- (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.

Rust, Roland T. and Tuck Siong Chung (2006), "Marketing Models of Service and Relationships," *Marketing Science*, 25 (6), 560-80.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting your customers: who are they and what will they do next?," *Manage. Sci.*, 33 (1), 1-24.

Schweidel, David A., Eric T. Bradlow, and Peter S. Fader (2011), "Portfolio Dynamics for Customers of a Multiservice Provider," *Management Science*, 57 (3), 471-86.

Scott, Steven L. (2002), "Bayesian Methods for Hidden Markov Models: Recursive Computing in the 21st Century," *Journal of the American Statistical Association*, 97 (457), 337-51.

Sun, Baohong and Shibo Li (2011), "Learning and Acting on Customer Information: A Simulation-Based Demonstration on Service Allocations with Offshore Centers," *Journal of Marketing Research*, 48 (1), 72-86.

Sun, Baohong, Shibo Li, and Catherine Zhou (2006), "“Adaptive” learning and “proactive” customer relationship management," *Journal of Interactive Marketing*, 20 (3-4), 82-96.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (4), 106-25.

Venkatesan, Rajkumar, V. Kumar, and Timothy Bohling (2007), "Optimal Customer Relationship Management Using Bayesian Decision Theory: An Application for Customer Selection," *Journal of Marketing Research (JMR)*, 44 (4), 579-94.

APPENDIX A. ALGORITHM FOR SOLVING POMDP & HMM

This Appendix provides the details on how we solve the Dynamic Programming problem we described in section 2.4. Since the difference between POMDP and HMM is the belief update function, we will focus on how to solve POMDP first. From the data, we found the maximum spending on marketing contact on a customer is \$150, therefore, we assume that our marketing spending budget for a customer is \$150 per months.

$$DCV_i^{POMDP} = V_i^*(b_{i,t})$$

$$= Max_{MKT_{it}} \left\{ \sum_{s=1}^{2} b_{it}(s) * \left[E_{b_{it}}\left[Y_{ijlt} * \overline{REV_{il}} \,|b_{it}\right] - MKT_{it} \right] \right.$$

$$\left. + \rho \sum_{j=1}^{2} \sum_{l=0}^{4} b_{it}(s=j) P\left(Y_{ijlt} | b_{it}(s), MKT_{it}\right) \left[V_i^* \left(b_{i,t+1}\left(s' \;\; = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right)\right)\right] \right\}$$

Subject to

$$b_{i,t+1}\left(s' \;\; = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) = \frac{\sum_{s=1}^{J} b_{it}(s) q_{itjj'} \;\; P_{ijlt}}{\sum_{s=1}^{J} \sum_{l=1}^{J} b_{it}(s) q_{itjj'} \;\; P_{ijlt}}$$

$$tslp_{i,t+1} = \begin{cases} 1 & if\ Y_{ijlt} > 0 \\ tslp_{i,t} + 1 & if\ Y_{ijlt} = 0 \end{cases}$$

$$MKT_{it} \leq 150$$

The general idea is to use point-based value iteration with value interpolation(Keane and Wolpin 1994b). The dynamic programming corresponding to the bellman equation above is

$$(TV)DCV_i^{POMDP} = V_i^*(b_{i,t})$$

$$= Max_{MKT_{it}} \left\{ \sum_{s=1}^{2} b_{it}(s) * \left[E_{b_{it}}\left[Y_{ijlt} * \overline{REV_{il}} \mid b_{it} \right] - MKT_{it} \right] \right.$$

$$\left. + \rho \sum_{j=1}^{2} \sum_{l=0}^{4} b_{it}(s=j) P\left(Y_{ijlt} | b_{it}(s), MKT_{it} \right) \left[V_i^* \left(b_{i,t+1}\left(s' = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) \right) \right] \right\}$$

T is the operator we apply to the value function repeatedly, and the value function will converge to the optimal value (Bertsekas 2000).

The details procedure is shown below.

Algorithm for Solving POMDP

1. Load the emission and transition function parameters from the estimation results.

2. Simulate two state variables: relationship state belief and time since last purchase, Belief state is between 0 and 1, we divide the state space into 40 grid points between 0 and 1. From the data, the longest inter-purchase time is 24 months. Therefore, the state variable *TSLP* ranges from 1 to 24.

3. Starting at period 1, calculate one time expected profit at each marketing expenditure level. Find the maximum profit and optimal marketing level for each relationship state grid point. Save the maximum profit as the initial Value for value iteration.

4. Run a linear regression with the Value as the dependent variable and transformations of the state variables as independent variables. Specifically,

$$V^*(b, TSLP) = \alpha_0 + \alpha_1 b + \alpha_2 b^2 + \alpha_3 \ln(b) + \alpha_4 + \alpha_5 TSLP + \alpha_6 TSLP^2 + \alpha_7 \ln(TSLP)$$

Obtain the parameters for value function approximation in the value iterations(Keane and Wolpin 1994b).

5. Start the value iteration by setting the initial value as the one-time expected profit calculated in 3. Then calculate the belief update functions by:

   1). Calculate the transition probabilities according to equations in section 2.2.3.

   2). Calculate the state dependent choice probabilities according to equations in section 2.2.2.

   3). Calculate belief update function by

$$b_{i,t+1}\left(s^{'} = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) = \frac{\sum_{s=1}^{J} b_{it}(s) q_{itjj^{'}} \ P_{ijlt}}{\sum_{s=1}^{J} \sum_{l=1}^{J} b_{it}(s) q_{itjj^{'}} \ P_{ijlt}}$$

$$tslp_{i,t+1} = \begin{cases} 1 & if\ Y_{ijlt} > 0 \\ tslp_{i,t} + 1 & if\ Y_{ijlt} = 0 \end{cases}$$

6. Calculate the approximated expected future value function by

$$V\left(b_{i,t+1}\right) = \sum_{j=0}^{1} \sum_{l=0}^{4} b\left(s_{i,t}\right) * P_{ijlt} \left[\sum_{\iota=0}^{7} \alpha_{\iota} * b_{i,t+1}\left(s^{'} = 1 | b_{it}, Y_{ijlt}, MKT_{it}\right) \right]$$

   where, the $\alpha_{\iota}{}^{s}$ were the parameters from linear regression in step 4.

7. Update the value function by,

$$V^{*}\left(b_{i,t+1}\right) = V^{*}\left(b_{i,t}\right) + \rho * V\left(b_{i,t+1}\right)$$

With the updated $V^{*}\left(b_{i,t+1}\right)$, go back to step 1 and start over again. Iterate the process until the process converges to a fixed point(Bertsekas 2000).

The dynamic programming problem for HMM is

$$DCV_i^{HMM} = V_i^*\left(b_{i,t}\right)$$

$$= Max_{MKT_{it}} \left\{ \sum_{s=1}^{NS} b_{it}(s) * \left[E_{b_{it}}\left[Y_{ijlt} * \overline{REV_{il}} \mid b_{it}\right] - MKT_{it}\right] \right.$$

$$\left. + \rho \sum_{s=1}^{2} b_{it}(s) \left[V_i^*\left(b_{i,t+1}\left(s' \mid b_{it}, MKT_{it}\right)\right)\right] \right\}$$

where, $b_{i,t+1}\left(s' \mid b_{it}, MKT_{it}\right) = b_{it} * q_{it11} + (1 - b_{it}) * q_{it01}$

$$tslp_{i,t+1} = \begin{cases} 1 & if\ Y_{ijlt} > 0 \\ tslp_{i,t} + 1 & if\ Y_{ijlt} = 0 \end{cases}$$

To solve this problem, replace the 5 3) in the algorithm for POMDP above.

APPENDIX B. SIMULATION PROCEDURE FOR POMDP & HMM

After solving the dynamic programming problem, we obtained optimal polices for each customer. The optimal policy is a mapping of belief state and *TSLP* to marketing actions. To show how to manage relationship proactively through adaptive learning and conditional planning, we simulate the multiple customer and firm interactions for each customer to see how marketing actions will respond accordingly. This appendix provides details about the simulation procedure.

Simulation Procedure for Optimization with Adaptive Learning (POMDP)

1. Obtain the optimal marketing policy from the optimization results.

2. At t=1, initial relationship belief state *s* as the initial state distribution calculated by the pre-purchase marketing actions based on Section 2.2.1, set the initial *TSLP* as 1. Obtain the optimal marketing action from the marketing policy based on the two state variables.

3. Simulate state dependent purchase utilities for both state by the emission function as in Section 2.2.2 with the optimal marketing action and state variables plus a GEV random error.

4. Simulate state variable for this period as a random draw from

$$u \sim Uniform[0,1], Simulated\ State\ Sims = \begin{cases} 0 & u < s \\ 1 & u \geq s \end{cases}$$

5. Simulate state dependent choice based on the simulated relationship state from step 4 and the state dependent purchase random utility from step 3.

6. Simulate purchase amount based on step 5 and the average purchase quantity of the purchase level.

7. When there was a purchase, simulate promotion indicator using the average promotion probability for the customer.

8. Simulate transition force at the optimal marketing level and *TSLP* based on section 2.2.3.

9. Update the belief based on the simulated purchase level, optimal marketing and *TSLP*.

10. Save the state, optimal marketing action, simulated state, simulated state dependent choice and profit for time t. Then move to the next period with the updated state variables.

For the simulation of the pure exploitation HMM strategy, replace the belief update function in step 8 with

$$b_{i,t+1}(s' \ |b_{it}, MKT_{it}) = b_{it} * q_{it11} + (1 - b_{it}) * q_{it01}$$

APPENDIX REFERENCES

Bertsekas, Dimitri P. (2000), Dynamic Programming and Optimal Control: Athena Scientific.

Keane, Michael P. and Kenneth I. Wolpin (1994a), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76 (4), 648-72.