Summer 7-18-2014

Modeling the Dynamic Decision of a Contractual Adoption of a Continuous Innovation in B2B Market

Yingge Qu

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or publish this dissertation may be granted by the author or, in his/her absence, the professor under whose direction it was written or, in his absence, by the Dean of the Robinson College of Business. Such quoting, copying, or publishing must be solely for the scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

YINGGE QU

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

YINGGE QU
J. Mack Robinson College of Business
Georgia State University
35 Broad St., Suite 400
Atlanta, GA 30303

The director of this dissertation is:

Dr. V. KUMAR
Regents' Professor,
Richard and Susan Lenny Distinguished Chair Professor in Marketing, and
Executive Director, Center for Excellence in Brand & Customer Management
J. Mack Robinson College of Business
Georgia State University
35 Broad Street NW, Atlanta, GA, 30303

Modeling the Dynamic Decision of a Contractual Adoption of a Continuous Innovation in B2B Market

BY

YINGGE QU

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
2014

ACCEPTANCE

This dissertation was prepared under the direction of the ***Yingge Qu's*** 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 Doctoral of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

H. Fenwick Huss, Dean

DISSERTATION COMMITTEE

Dr. V. Kumar (Chair)

Dr. Yi Zhao

Dr. Denish Shah

Dr. Andrew Petersen (External, UNC at Chapel Hill)

ABSTRACT

Modeling the Dynamic Decision of a Contractual Adoption of a Continuous Innovation in B2B Market

BY

YINGGE QU

JULY. 18. 2014

Committee Chair: *Dr. V. Kumar*

Major Academic Unit: *Regents' Professor,*
Richard and Susan Lenny Distinguished Chair Professor in Marketing, and
Executive Director, Center for Excellence in Brand & Customer Management

A continuous service innovation such as Cloud Computing is highly attractive in the business-to-business world because it brings the service provider both billions of dollars in profits and superior competitive advantage. The success of such an innovation is strongly tied to a consumer's adoption decision. When dealing with a continuous service innovation, the consumer's decision process becomes complicated. Not only do consumers need to consider two different decisions of both whether to adopt and how long to adopt (contract length), but also the increasing trend of the service-related technological improvements invokes a forward-looking behavior in consumer's decision process. Moreover, consumers have to balance the benefits and costs of adoption when evaluating decision alternatives. Consumer adoption decisions come with the desire to have the latest technology and the fear of the adopted technology becoming obsolete. Non-adoption prevents consumers from being locked-in by the service provider, but buying that technology may be costly. Being bound to a longer contract forfeits the opportunity to capitalize on the technological revolution. Frequently signing shorter contracts increases the non-physical efforts such as learning, training and negotiating costs. Targeting the right consumers at the right time with the right service offer in the business-to-business context requires an efficient strategy of sales resource allocation. This is a "mission impossible" for service providers if they do not know how consumers make decisions regarding service innovation. In order to guide the resource allocation decisions, we propose a complex model that integrates the structural, dynamic, and learning approaches to understand the consumer's decision process on both whether or not to adopt, and how long to adopt a continuously updating service innovation in a B2B context.

Table of Contents

Table of Figures

Modeling the Dynamic Decision of a Contractual Adoption of a Continuous Innovation in B2B Market

Abstract

A continuous service innovation such as Cloud Computing is highly attractive in the business-to-business world because it brings the service provider both billions of dollars in profits and superior competitive advantage. The success of such an innovation is strongly tied to a consumer's adoption decision. When dealing with a continuous service innovation, the consumer's decision process becomes complicated. Not only do consumers need to consider two different decisions of both whether to adopt and how long to adopt (contract length), but also the increasing trend of the service-related technological improvements invokes a forward-looking behavior in consumer's decision process. Moreover, consumers have to balance the benefits and costs of adoption when evaluating decision alternatives. Consumer adoption decisions come with the desire to have the latest technology and the fear of the adopted technology becoming obsolete. Non-adoption prevents consumers from being locked-in by the service provider, but buying that technology may be costly. Being bound to a longer contract forfeits the opportunity to capitalize on the technological revolution. Frequently signing shorter contracts increases the non-physical efforts such as learning, training and negotiating costs. Targeting the right consumers at the right time with the right service offer in the business-to-business context requires an efficient strategy of sales resource allocation. This is a "mission impossible" for service providers if they do not know how consumers make decisions regarding service innovation. In order to guide the resource allocation decisions, we propose a complex model that integrates the structural, dynamic, and learning approaches to understand the consumer's decision process on both whether or not to adopt, and how long to adopt a continuously updating service innovation in a B2B context.

Keywords: continuous innovation, service innovation, B2B hi-tech markets, dynamic programming, bellman equation, learning process, structural model.

Introduction

We are living in an era that is continuously innovating. According to the famous Moore's law(Moore, 1965), most of the high-tech related devices, such as CPU speed, hardware size, and memory capacity etc, shows exponential improvement rate over year. In recent decades, the speed of innovation in the hi-tech markets creates turbulence in not only people's everyday life but also in the global business environment. In people's everyday life, cell phones significantly improve the way that people communicate with each other and internet dramatically change people's lives which were imaged before. In today's business world, enterprises find few opportunities to grow or even survive without applying the most advanced technology to their own networks. The inevitable consequence of rapid innovation in the high-tech market is that not only does the IT person's burden becomes heavier, but also the resources that firms have to spend on the IT department grows significantly.

Under such situation, cloud computing service walks into the B2B high-tech service market. Cloud computing is a high-tech delivery service provided through internet aiming at transferring the traditional IT burden to the service provider thus saving the business customers' internal IT cost. Cloud computing targets at including all potential IT services such as data processing and storage, employee profile management, software and hardware development, information integration, communication/networking security etc. Because of the emerging of the cloud computing services, whatever the internal IT department is doing, the B2B customers now have the option to shift to service provider for the 3 potential reasons: the specific feature of expertise; the less cost involved and relatively easier and quicker to scale up and down. Therefore, there is a lot of incentive to go through the adoption.

Although the cloud concept was introduced in 1960s, it becomes realization and popular only around late 2000s after the internet become prevalent. From the launch to present, cloud service shows remarkable market potential. Announced by IBM in 2011, more than 80% of the Fortune 500 companies become their cloud adopters, which include multiple industries such as airlines, cars, financing, insurance and supermarket (Tomasco, 2011). And the global cloud market size is predicted to reach 19.5 Billion by 2016 (Columbus, 2013).

In order to catch up with the fast changing business environment, many firms turn to purchase high-tech services from pioneering high-tech providers, which makes a continuous service innovation highly attractive in the business-to-business world because it brings the service providers both billions of dollars in profits and superior competitive advantage. Moreover, although cloud computing is a promising and attractive service in the B2B high-tech market, certain concerns, such as security, capability, privacy and integrity still exists among cloud adopters. Most of the concerns, fundamentally, are centered on the technological level of the cloud service. Therefore, the technology improvement is always the key feature to alleviate customer's concern, for example, increasing the data processing speed and storage size, introduce new methods to improve the security of the communication environment, build new data center to enhance the global connection etc. All those improvement, we define as continuous innovation or "sustained improvement" in our paper.

There is no doubt that the success of such a continuous innovation is strongly tied to a consumer's adoption decision. In such a unique and novelty hi-tech market, some potential problems require answers. First is, if the technology level keeps on updating, how the service provider should manage such continuous innovation? Specifically, should the service providers gather all the improvements and publicize them together to alert customers, or should the service providers gradually announce each step of their innovation progress to remind customers? Second is, how would the evolved technology impact B2B customers' decision? In order to answer these questions, we will have to understand customers' behaviors and decision process, which is not easy task because:

1) When dealing with a continuous service innovation, the consumer's decision process becomes complicated. Not only do consumers need to consider whether to adopt and how long to adopt (contract length), but also the increasing trend of the service-related technological improvements invokes a forward-looking behavior in the consumer's decision process.

2) Consumers have to balance the benefits and costs when evaluating the decision alternatives. Consumer adoption decisions come with the desire of having the latest technology and the fear of the adopted technology becoming obsolete. Non-adoption prevents consumers from being locked-in by the service provider, but buying that technology may be costly. When they are bound to a longer contract, they forfeit the opportunity to capitalize on the technological revolution. Frequently signing shorter contracts increases the non-physical efforts such as learning, training and negotiating costs.

Therefore, targeting the right consumers at the right time with the right service offer in the business-to-business context requires an efficient strategy of sales resource allocation. This is a "mission impossible" for service providers if they <u>do not know</u> how consumers make decisions regarding service innovation. In order to guide the resource allocation decisions, we propose a complex model that integrates the structural, dynamic, and learning approaches to understand the consumer's decision process on both whether or not to adopt, and how long to adopt a continuously updating service innovation in a B2B context. Our study offers both modeling contribution and substantive insights into an emerging, promising and continuously updating high-tech service B2B market. Our major modeling contribution is; different from adding covariates in the modeling framework; we structurally model customer decision process in a continuous innovation hi-tech B2B market. Specifically,

1) We simultaneously model customer's two layer decision: whether and how long to buy. To the best of our knowledge, this is the first study to integrate both learning and forward-looking in the customer's two layer decision, especially the continuous decision of how long to buy.

2) Using the structural model approach, we are able to quantify the underlying relationship between customer's decision and technology evolution. And we evaluate the impacts of technology evolution on customer's decision. Specifically, we are able evaluate how the two states of technology evolution could separately influence customer's decision process: one is overall technology level; the other is the technology evolution pace.

3) Finally we integrate both contract length and technology evolution into the dynamic programming process to structurally explain customer's dynamic decision process.

From substantive perspective, our model parameter estimation and policy simulation are very valuable for aiding firm's decision making. Based on our results, firms can know under different scenario, what strategy will be better choice. Our findings suggest that:

1) If a service provider expected the customers to try and purchase their new service, it will be better for the service provider to either select evenly distributed but relatively smaller steps of technology improvement or reduce the effort of signing a contract by offering customers some additional help. This way, customers show higher purchase intention although the contract length they signed is relatively shorter.

2) On the other hand, if a service provider is more willing to reap better profit from the customers, the better choice will be announcing less frequently but more influential technology improvement or providing relatively rigid policy for customers to make purchase. This way,

although customer's purchase intention is relatively low, customers tend more likely to sign a longer contract once make the purchase decision.

3) Since we also provide quantitatively evaluation on the customer's decision changes under different scenario, firms can combine with their internal profit function with our result to design the optimal strategy.

The rest of the article is organized as follows. We summarize the prior research on customer's forward-looking behaviors. We both discuss the development of modeling approach on forward-looking dynamics model and point out the research gap in the existing research that our study can contribute to. Then we propose our structural modeling approach on customer's decision process in the continuous innovation hi-tech market. Specifically, we explain how we integrate both the customer's two-layer decisions and technology evolution in the dynamic programming to account for consumer's forward looking behavior, which is one of our major modeling contribution. Then we specify the modeling estimation process including both value function simulation and likelihood estimation. Since we don't have close-form results, our modeling estimation process is based on the empirical solution. Following this, we will describe the data set we used, then discuss our model estimation result and policy simulation. Next, we conclude with our modeling contribution and managerial implication. Finally, we provide the study limitation and future research directions.

Literature Review

Our study belongs to the research stream of modeling the consumer's dynamic decision process in the high technology service market by taking into account forward looking behavior. The research stream began with the study of Guadagni and Little (1983) who utilized a multinomial logit approach to capture a consumer's choice between alternative brands. The decision process in the study is referred to as a static decision as compared to forward looking decisions because the it assumes that the consumers make choice decisions based only on their current utility. The fundamental modeling approach of the study is still rely on adding major explanatory variables to predict consumer's brand choice on consumer package goods. Therefore, this model can capture the probability of choice as a function of alternative attributes, such as price, promotion, and consumer loyalty etc. Although the study still relies on the current utility to model a consumer's choices, it lays the foundation for future dynamic modeling approaches. And the modeling idea of keeping the variable coefficients the same across different brand sizes shed light on the concept of modern structural model.

Recent studies pertain to the reality of the assumption behind the static modeling approach because consumers usually not only consider the current utility but also take into account the expectation on future utilities when making purchase decisions. For example, consumers can hold their current purchase decision when expecting an upcoming promotion season in order to get a better price cut. In this situation, consumers will show forward looking behaviors in their decision process. Moreover, consumers can learn from different sources of information or experiences and update their beliefs about the products/services, which in turn, will also impact their decision process.

The idea of "forward-looking behavior" was firstly appeared in a "regenerative optimal stopping model" introduced by Rust's study (Rust, 1987). The so-called "dynamic programming" approach was applied to determine an optimal decision for bus engine replacement. The fundamental concept of the "optimal stopping rule" is that, the decision of whether or not should the Bus maintenance replaces the bus engine at a certain time period is determined by two forces: one is the required maintenance cost; the other is the unexpected engine failure in the future. Consequently, to optimize the current maintenance decision, we will have to integrate the future discounted value into the current decision process, which is typically the "forward-looking" behavior. Similar idea and the "optimal stopping" model is applied to the new and used car market (Schiraldi, 2011) to quantify the trade-in effect on consumer's decision process.

Erdem and Keane (1996) demonstrated a structural modeling approach that embeds both forward looking dynamics and the Bayesian learning process. Different from the static model which only relates consumer's decision with current utility maximization, the "forward-looking" dynamic model considers that consumer's choice is determined by the maximization of the expected utility over a time zone. Furthermore, through the Bayesian learning process, the model can capture the influences of a consumer's past usage experience and advertising exposure on the consumer's choice between alternative brands. Through the combination of forward-looking dynamics and the Bayesian learning process, the model can evaluate the impact of a firm's marketing strategy on the consumer's brand choices in both the short and long run. In the paper, the authors explicitly define the "structure" model which aims at deriving the underlying relationship between consumer's choice and the marketing attributes based on the specification of consumer's purchase behaviors. Using structural modeling approach, pointed out by the authors, we can drive reliable policy evaluation because the parameters in the structural model are considered as the intrinsic preference of the consumers thus the parameter estimation doesn't vary with the policy change. Therefore, Erdem and Keane (1996)'s paper serves as the milestone to combine forward looking behavior and learning process to structurally model customer's decision process.

Following the distinct study in 1996, Erdem and Keane spent efforts on enrich the dynamic learning model framework. For example, the distinct study only considered the brand choice as the decision variable (Erdem and Keane, 1996). To improve the decision complexity, they integrated both brand and quantity choice into the dynamic model and aimed at finding the effect of price fluctuation on the consumer's decision process (Erdem, 2003). Although this paper didn't put the quantity as a continuous decision but divide it into discrete level of choices in the model, it signaled to us the idea that the dynamic model is not limited to the discrete choice but applicable to the continuous decision as well. Moreover, in the distinct study, consumer's utility was defined as a function of price and consumer experienced attributes (e.g. quality) which was recovered by a normal distribution (Erdem and Keane, 1996). To enrich both the utility function and the consumer learning process, Erdem et al. introduced both price and advertisement to help quantifying the consumer experienced product quality (Erdem et al., 2008). Moreover, when dealing with high-tech product brand choice and if we can obtain data related to consumer's information search, we are able to develop the "active learning" model which can quantify: to what extent that consumer gather the information can actually invoke an actual purchase decision (Erdem et al., 2005).

Using similar idea and modeling approach as the distinct study (Erdem and Keane, 1996), Ackerberg (2001, 2003) developed a dynamic learning model using data from consumer package goods (e.g. Yogurt). Different from Erdem and Keane's (1996) study that only exam the "informative" effect of advertising on consumer's decision process, this study separately evaluate three effects of advertising: informative, prestige and image. Although Ackerberg's study had limited modeling contribution, it pointed out an important feature of dynamic model which is capable of separately quantifying the multi-dimensional effects of one attribute on consumer's decision process.

Following the promising study, modeling the consumer's forward-looking dynamic decision process under different marketing contexts becomes prevalent. Most studies focus on modeling the price-initiated forward-looking behaviors in a B2C consumer's binary decision process, e.g., whether the consumer will buy or not buy a product.

Gönül and Srinivasan (1996) applied structural dynamic programming approach to model whether consumers will adjust their purchasing behaviors on consumer package good (e.g.

disposable diaper product) when they have expectation on future product promotions. The central idea of the model setup is that, consumer's forward-looking behavior is triggered by the potential cost of the product. Therefore, as rational consumers, in order to maximize their expected utility of purchase, they will have to consider the future cost in the promotion season. Although acknowledge the advantages of structural forward-looking dynamic model, the authors (Gönül and Srinivasan, 1996) also pointed out a potential limitation of the model – computation cost, which can significantly limit the state variables that we can include in the model.

Other than consumer package good, many structural dynamic models were built on consumer durable goods, especially the high-tech products. The reason is that the rapid development of product related features, such as product quality, and the obviously declined price of the existing products serve as the perfect triggers of consumer's forward looking behavior. For example, Song and Chintagunta (2003) investigated the interesting phenomena in Camera (e.g. high-tech durable good) market that the price of the new camera in the market continuously drops over time. The price declination invokes consumer's forward-looking behavior, which affects their current decisions. Reflected in the model, consumers will adjust their actual purchase time in order to maximize the expectation of both current utility and future discounted utilities. Because of the forward-looking behaviors, the present variation of the price of the product can not only influence consumer's current purchases but also alter their future decisions (Song and Chintagunta, 2003).

Using similar idea that price decline over time triggers customer's forward-looking behavior, Sriram et.al (2010) extended the model in Song and Chintagunta's study (2003) into individual level data from three categories of high-tech durable products. The major finding of the study is that, consumer's forward-looking behaviors can be interactional. Consequently, the price decline of one product category can change consumer's purchasing behavior not only within the category but also across the categories.

Ryan and Tucker's study (2012) also focus on using dynamic model to understand individual-level consumer's decision process on adopting high-tech product (e.g. video-calling technology). The interesting part of the study is that, it brings in the idea that the network evolution can play an important role in determining the consumer's dynamic decision process. Therefore, consumer's forward-looking behavior is no longer the "patent" of price or cost of the product, but can be closely related to the evolution of the technology related to the product.

Although dynamic model is a promising modeling approach, its modeling framework is more complicated and its estimation process requires more computing burden than the static models. Therefore, to dispel the suspects on the necessity of using dynamic model over static model to understand consumer's decision, many studies spent efforts on showing the advantages and benefits of dynamic model. For example, based on still the high-tech durable product (e.g. digital camcorder), Gowrisankaran and Rysman (2012) integrate both product quality elevation and product price decline in modeling consumer's dynamic demand. Other than showing the estimation and analysis of consumer's dynamic preference toward the product, the authors also confirmed that the dynamic model yield more realistic results than the static approach. Hendel and Nevo (2006) also pointed out that, although price cut, such as promotion or deals, usually is a temporary action for many products, it can create a long-term effect on consumers' decision. The estimation of the long-term effect can be significantly biased if we omit the dynamic behavior features in our modeling approach (Hendel and Nevo, 2006). Using forward-looking dynamic model to quantify consumer's brand switch behavior, Sun et al. (2003) also found that dynamic model can provide unbiased estimation on the brand-switch elasticity thus has better model-fit than the reduced-form model. And the biased estimation issue in the reduced-form model can't be solved by adding covariates. Moreover, in another study, Sun claimed that the forward-looking dynamic model is able to identify the behavioral link between the promotion and consumer's consumption decision even though the decision variable is endogenous (Sun, 2005). Using data from computer printer market, Melnikov (2012) built logit-based discrete choice model integrating both forward-looking dynamics and endogenous prices. His study not long empirically confirmed the existence of consumer's forward-looking behaviors, but also showed the better performance of dynamic model on both forecasting consumer's demand and measuring the benefits of new products than the static model.

In the high-tech product market, structural dynamic model can be applied to not only new product adoption but also other type of purchasing pattern. For example, Gordon (2009) addressed that, in product replacement (e.g. PC processor) purchases, consumer's dynamic decision process can be trigger by the obsolescence status of the product due to both the quality elevation and price reduction. Lewis (2004) utilized the discrete-choice dynamic model using data from online grocery and drug items to solve customer's sequential choices when facing loyalty program. In the paper, other than the price-related marketing mix attributes, the author claimed that the loyalty program can also incite customer's forward-looking behaviors because the benefits of the program are determined by the overall spending over time.

Different from studying consumer package goods or durable goods, Yang and Ching (2013) built a structural dynamic programming model to understand consumer's decision process on adopting and using ATM cards. Although the authors still followed the basic concept of establishing forward-looking structural model, they addressed an interesting result of identifying the impact of consumer's age on their forward-looking zones. The idea is that, compared with younger people, elder people have relatively shorter forward-looking zones, thus they receive lesser expected utility from adopting a new product which requires a certain level of adoption cost, such as learning cost. Consequently, elder people can be more reluctant to learn and adopt the product. Although this study still focus on the discrete choice of whether to adopt or use the ATM card which limit its modeling contribution, the idea of the length of forward-looking zones shed light on our study to integrate continuous decision into the forward-looking structural model.

The dynamic modeling approach not only can quantify the direct effects of attributes, such as price, promotion, quantify and advertising etc., on the consumer decision process, but also is able to identify the indirect effects. Gowrisankaran et al. (2011) applied the dynamic model on data from both DVD player and DVD titles markets and found that, consumers with forward-looking behaviors may form expectation on the future benefits in the DVD title market thus make multiple purchase on DVD players over time. This study pointed out a research direction of applying the dynamic model in complimentary product markets to understand consumers' decision process in both market simultaneously.

Forward-looking dynamic model can also be applied to service market. It is widely acknowledged that switching cost is a major consideration when consumers make decisions on the service providers. Using data from different types of the service market, studies also found that, switching cost can also influence consumer's decision process through invoking consumer's forward-looking behaviors. Using data from cellular service industry, Kim (2006) firstly pointed out that switching cost can be a source of consumer's forward-looking behavior. Using data from Medicare service market, Nosal (2011) addressed that switching cost significantly influence consumers' decision between advantage plan and original plan, and the amount of people choosing advantage plan could be tripled if without the switching cost. Shcherbakov (2009) specifically provided an estimation of switching costs of $109 and $186 for cable and satellite respectively in television service market. Although our study focus was not the consumer's switching behaviors, our study context belongs to the service market and the cost of signing a service contract is an important parameter that we need to estimate. The findings from these papers suggest that the switching cost in the service market, although may not be fully observable, can be identified and estimated in the dynamic model. Therefore, bringing the cost of signing a service contract in our model setup (described later) is reasonable, convincible and estimable.

Other than understanding the dynamic decision process purely from consumer's perspective, using data from video-game market, Nair (2007) developed a dynamic modeling from both consumer and firm's perspective. From consumer's perspective, consumers form a forward-looking behaviors because of the of the price drop. From the firm's perspective, firms are also forward-looking in order to consider consumer's dynamic behaviors when formulating the optimal price strategy. The findings in the study suggests that, understanding consumer's forward looking behaviors can effectively help firm to design optimal pricing strategy thus better target the right consumers at the right time with the right price. In our study, we only focus on studying consumer's decision process from consumer's perspective. Therefore, we omit reviewing the literature of dynamic models from firms' perspective which belongs to another stream of modeling stories.

Study Motivation

Although our study stems from the research stream of modeling consumer's decision process, our study possesses many unique characteristics in not only the modeling approaches but also the application contexts (Table 1):

1) We propose an improved and holistic modeling framework for a service innovation in the B2B hi-tech market. Being in the rapid developing high-tech market, a B2B consumer's concern will not only come from the cost of adopting the service but also the level and the pace of continuously developing technology that they can receive from the service provider.

2) As we pointed out previously, in this high-tech service market, there is always a contact associated with each individual purchase. This aspect suggests that, customer's adoption decision

has two layers, one is a discrete decision – whether to adopt and the other is a continuous decision which is how much to adopt, e.g. contract length. In our study, we jointly model two different decisions, e.g. a binary decision of whether to adopt and a continuous decision of how long to adopt (contract length). Although considering multiple angles of consumer decision is very common in static modeling approach (Chitaguanta 1993), up to now, binary decisions still dominate the forward looking dynamic model built on consumer's perspective. Moreover, there are some distinct features in the market which makes the customer's decision process unique in our study.

3) The first distinct feature in our study context is the physical cost of adopting the service, which can be considered as the unit price of the service. From the previous literature review, we know that price is a powerful factor that invokes a consumer's forward looking behavior that attracts most researchers' interests. In the general high-tech markets, such as camera, computer, printer, video-game etc, the price keeps on decreasing after the first release. This is one reason that most studies using price as the trigger of customer's forward looking behaviors. However, we raise another important factor of technology evolution in our study. Through iterative discussion with the managers, we realize that in our context, one distinct feature is that the physical cost of adopting the service is primarily associated with the consumer's need and will not show systematic changes within years, but the service-associated technology level keeps on updating rapidly. This distinct feature suggests that the unit price of the service is relatively consistent but the technology keep on updating. Therefore, customer still form forward looking behaviors, however, such forward looking is not triggered by price change but by the technology evolution. In another word, unlike the common contexts, the uniqueness of my study is that, the physical cost of adopting the service is not critically important in determining consumer's decision process, especially the forward-looking behavior. In fact, such a phenomenon is not novel in the field, especially in a fast developing high-tech market. For example, like I-pad, its original launch price is always the same for all generation

but the features keep on updating. Storage size increased exponentially from MB to GB to TB within a few years with the price being closer between newest models. With a similar price, consumers can always expect a better version coming in the near future.

4) Another distinct factor in our study context is non-physical cost, which we can in general define it as "contract fee". As we all know that The adoption decision in the B2B world is not an easy process but associated with remarkable non-physical efforts, such as learning the technique, training employees, cooperating multiple internal departments, negotiating with service providers and sending budget application to top manager for approval etc.. All these non-physical efforts, from modeling perspective, is latent to the researcher but will also influence consumer's decision. It is also worth mentioning that the non-physical cost is a one-time fee associated with each contract. Each time the B2B consumers decide to adopt the service and sign contract, they will have to spend the "contract fee". However, once the decision is made and the contract is signed, within the contract length, there is no additional effort of "contract fee" required. Obviously, because of this one-time contract fee, the business consumers are not willing to frequently sign the contract. The idea is that, if the consumers have already spent greater energy and resources to finalize a contract, they will intend to sign a relatively longer contract to avoid another input.

5) The last is the technology keeps on updating in the market as we illustrated previously. Moreover, once the customers make the decision of purchase, they will be bonded with the on-site technology level at the time of purchase. Clearly, from the perspective of technology keep on updating, the consumers should be willing to sign the contract frequently, e.g. sign a shorter contract, because they can capture the technology evolution more effectively.

Now, we can see that there exists a decision conflict in customer's decision process. To save the efforts on "contract fee", consumers tends more likely to sign longer contract. However, to better capture technology updating, consumers should be more inclined toward shorter contract.

Consequently, both too long and too short are not good. Customers will have to go through a "value maximization" process to select an optimal decision for the contract length decision. In order to quantify this "value maximization" process, we firstly need clarify consumer's behaviors in their decision process.

Because of the continuous technology innovation in the market, consumers will show two kinds of behaviors in their decision process. First, customer will be bonded with the on-site technology level once the contract is signed. This phenomenon suggests that customer's current decision will not only influence the current but also the future benefit associated with the purchase decision. Consequently, customers will have to consider both current utility and discounted future utility to make an optimal current adoption decision. To achieve this goal, consumers will have to form future expectation on the technology evolution according to the information set that is available to them because they are not certain about the future technology level influencing the future potential utility. From modeling perspective, this is a typical forward looking behavior.

Second, because of the technology updating, customer will have uncertainty about the speed of the technology evolution in their decision process. Therefore, consumers will form and update their belief about the future technology evolution speed. We assume customers will update their belief on future tech evolution based on the historical information. From modeling perspective, we define this as a learning behavior. In our model, we want to quantify both behaviors.

Moreover, considering the remarkable influence of technology evolution on customer's behaviors, we also want to quantify the effect of tech evolution on customer decision. Learned from the previous literature (Erdem and Keane, 1996; Ackerberg 2001 & 2003), we are able to identify multi-dimensional effects of one attribute on consumer's decision process in the forward-looking dynamic model. In our study, we define that the technology evolution will create two factors influencing consumers' decision process:

The first factor is the overall technology capability level which will encourage buyers to purchase. This means that the overall technology capability level will only influence consumers' decision of whether to buy or not. The idea is that, only if the technology level meet or exceed customer's needs, the adoption decision will be invoked. The second is the speed of technology evolution which we define as "technology evolution pace". The technology evolution pace will produce two effects: one is postponing customers' adoption decision; the other is to encourage a short term contract. This suggests that the technology evolution pace will influence consumers' both decisions of whether or not to buy and how long to buy, e.g. contract length. The intuition is that, if the customers consider the technology evolution to be very fast and expect a more advanced technology to appear in the near future, then the customers may either hold their current adoption decision to wait for the future better offer, or at most take a try on the service with a shorter contract.

Based on the previous motivations, we are interested in the following research questions: How can we structurally model customer's decision process on the high-tech service adoption? In this structural model, we want to address two issues in consumer's decision process:

1) We want to simultaneously model consumer's two-fold decisions at each decision time point: whether or not to buy and how long to buy, e.g. contract length

2) We want to evaluate the impacts of technology evolution on customer's decision.

There are several reasons that the structural model is the most appropriate modeling approach in our study. The major modeling contributions in our study are also embedded in the reasons.

1) The first reason is that we want to understand the underlying relationship between customer's decision and technology evolution; therefore, we can't rely on the approach of putting co-variates into modeling framework. Evidences from multiple previous studies show that,

especially when forward-looking behavior exists, using the reduced-form model to quantify consumer's decision process can create biased results which can't be completely alleviated by adding co-variates (Hendel and Nevo, 2006; Sun et al., 2003; Sun, 2005; Melnikov 2012). Discovering the underlying relationship between technology evolution and consumer decision process is also the first major modeling contribution in our study.

2) We will introduce in detail in the model setup section that, we will use dynamic programming approach to account for customer's forward looking behavior. Specifically, we want to integrate both consumer's <u>contract length</u> decision and the <u>technology evolution</u> (both overall technology and technology evolution pace) into the dynamic programming process to structurally explain customer's decision. In other words, we want to provide if-what answers on the impacts of technology on customer's decision. This is the second reason that we have to build a structural model. The modeling approach of integrating contribution length and technology evolution into the dynamic programming approach serves as another modeling contribution in our study.

3) The last reason is that, using structural model, we can also do the counterfactual study to show, under different scenarios of announcing the technology evolution, how the customer's decision will change. Recalled that we laid out two potential problems in the market, one is how to manage this continuous innovation and the other is how to evaluate impacts of tech evolution on customer's decision. The two unique contributions in our model setup will help us to solve the potential problems.

In summary, this dissertation offers a significant contribution in two areas: 1) we specify a model to include a consumer's multiple decisions; and 2) we develop a model estimation procedure that can provide insights on understanding a consumer's decision process in the technology evolution market by taking into account both forward-looking and learning behavior.

Insert Table 1 here

Modeling Framework

In the conceptual framework (as shown in Figure 1), we first start with modeling a consumer's current utility at each decision time point. We build the consumer's current utility as a function of the overall technology level. The current utility addresses the relationship between the "whether to adopt" decision and the current overall consumer technology need level has to be reached in order to invoke his/her purchase decision.

Next, we structure the model to account for consumer's forward looking behavior and learning behavior, which are incited by the technology evolution. The relationship between the two behaviors and the consumer's current adoption decision can be illustrated in the following manner. First, the continuous technology evolution will encourage consumers to form expectations on the future technology updating pace. Second, being in a turbulent environment, consumers will have to continuously learn from the existing information related to the technology evolution in order to update their beliefs about the technology level. Therefore, both behaviors will impact the consumers' process of determining the future value of what they expect to receive from the evolved technology, which in turn, will impact their current purchase decisions. The intuition is that, with a greater technology evolution pace, consumers should expect a better value from adopting in the future, thus either withdraw their current adoption decision or select a shorter contract to experience various aspects of the adopted technology (ie. learn the technology and apply it to the business) and decide whether or not continuing with it would be beneficial to the company.

Insert Figure 1 here

We propose to operationalize the conceptual framework by integrating three modeling approaches: dynamic programming, learning process, and structural modeling paradigm. The dynamic programming (DP) approach offers a parsimonious way to solve a sequential decision process under uncertainty (Rust, 1994) which allows us to quantify a consumer's forward looking behavior in a joint-decision model. Different from the static choice models, the DP model suggests that the consumer should generate an expected value including the utilities of both current and all future decisions to evaluate alternatives. Following the methodology proposed by Rust (1994), we will utilize the “Bellman equation” to solve the consumer’s value function. The theoretical derivation of the “Bellman equation” takes into account several unique perspectives, which is also a novel and representative contribution in our study.

First, considering both discrete and continuous decisions in the DP model elevates the difficulty in both deriving and solving the “Bellman equation.” Second, quantifying a consumer’s forward looking behaviors, in our study, requires taking into account the consumer’s learning process. The well-documented learning model has been widely applied to capture consumer’s beliefs updating (Erdem and Keane 1996). The traditional learning model (Erdem and Kean, 1996) focused on updating consumer’s belief on the mean level while assuming variance of perception to be both constant over time and known to the consumer. Moreover, the traditional learning model considered the all information that available to the consumer over time to be equally important to the consumers. Consequently, when time approaches infinity, consumer’s belief becomes a constant because consumers have no uncertainty about their belief. This idea in traditional learning model was found to be less realistic in the turbulent market especially product crisis exists (Zhao et al., 2011). To better capture consumer’s learning behavior, Zhao et al. (2011) not only relaxed the assumption of constant and known variance of perception, but also introduced an information discount factor to account for the diminishing confidence over time in consumers’ belief. We will adopt this more realistic methodology in consumer learning model to capture a consumer’s beliefs updating on both the mean level and variance of perception (Zhao et.al 2011). Finally, deriving the “Bellman equation” requires the specification of transition probability for state variables which are deterministic of consumer’s expected value. Our study includes four state variables originated from the consumer's future expectation and innate learning process. Not only has the number of state variables almost reached the upper limit in the dynamic programming model, but also the majority of the state variables are partially latent to the researcher which adds the "Partially Observed Markov dynamic model" to our modeling framework (Ricardo et.al 2010).

We will use a maximum likelihood function for model estimation. Both the latent parameters, such as the non-physical effort, and the consumer heterogeneity bring in multiple layers of integration in deriving the likelihood function. Moreover, both the Bellman equation and Likelihood function have no close-form solution and require iteratively searching for numerical solution until convergence. Basically, in order to empirically solve our model, we will have to do two loops of numerical searching: one is the searching for "value function"; the other is the searching for "value function". The process of solving "value function" is embedded in the solution of "likelihood function". This means that: 1) In order to reach the optimal solution of parameter estimations, we will have to strategically change the parameter settings and find the associated solution of "likelihood" at any given parameter settings. 2) The solution of "value function" depends on both state variables and the value of parameters in the model. Therefore, in order to get the result of "likelihood" at any one set of parameter values, we will have to solve "value function" through iteration. Therefore, these two loops of numerical searching are not parallel but two levels of nesting. Other than the two levels of nesting, to get the answer of "likelihood", we also need to solve the stochastic term at a given contract decision which also depends on numeric searching because there is no analytical solution between the stochastic term and decision variable.

It is worth mentioning that our framework is built upon the structural modeling approach. The major advantage of structural model is that the parameter estimation is invariant to the firm policy (Erdem and Kean 1996), which enables our model to detect a consumer's innate sensitivity to technology evolution.

We will estimate our model using data from a global high-tech company who continuously invests in developing an innovated high-tech service to its business partners since 2009.

Model Setup

Part 1 – current utility function

Consider in our study context where there is a set of consumers I = {i|i = 1, 2, ..., I} who make decisions on purchasing the service at different time point. Therefore, consumers' purchases are observed over the period T = {t|t = 1, 2, ..., T}, where T is the time span of the period. Let u_{it} represents the current utility that consumers (i) receive with their decision at a given purchase occasion (t). As we mentioned previously, in our study context, the technology of the service keep on updating over time. Reflected in the model, we use $Tech_t$ to represent the actual overall technology level at a given purchase occasion (t).

In our model, we assume that the utility (u_{it}) derived by consumer (i) for whether to adopt the high-tech service given the actual technology level ($Tech_t$) at current purchase occasion (t) can be represented as:

$$u_{it}|Tech_t = \gamma_{0i} + \gamma_{1i} * Tech_t + \varepsilon_{it} \quad --eqn\ (1)$$

Where γ_{0i} represents the consumer specific mean-level net-utility when the actual overall technology level =0. This mean-level net-utility represents the difference between the consumer's baseline preference toward the actual overall technology level and the potential unobserved cost associated with consumer's need. Please note that, the mean-level net-utility γ_{0i} is only consumer i-specific not a time t-specific parameter although it embraces the unobserved cost, because the unobserved physical cost is not a t-specific item in our study. As we stated previously that the physical cost of the service in our context is associated with consumer's need and has no systematic change over time. Therefore, mean-level net-utility parameter is γ_{0i} not γ_{0it}. Moreover, we don't separately specify a "physical cost" function in our model setup because the "physical cost" data has tremendous variance which blocks the valuable information that we can use to recover parameters in our model. To avoid the identification issue, we use one consumer specific parameter γ_{0i} to estimate the difference between consumer's baseline preference and the potential cost with consumer's need. We acknowledge this as a data limitation in our study. But this is not our model limitation because, once more valuable "physical cost" data is available, our model can be flexibly extended to add in another layer of "cost" function to empirically estimate the physical cost.

γ_{1i} is the consumer specific coefficient of the actual overall technology level. It captures customer's utility sensitivity of the overall technology level, e.g. to what extent the change of overall technology level will impact consumers' current utility.

$Tech_t$ indicates the actual overall technology level at a purchase occasion (t).

ε_{it} is the net-utility error for consumer (i) at purchase time (t). It captured the unobserved factors impacting consumer's net-utility, for example the service provider may provide a short-term promotion to incite consumer making purchases. We assume a normal distribution for the random component $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$. We fix the variance term at 1, e.g. $\sigma_\varepsilon = 1$ for identification issue.

We represent the consumer utility for the outside (or no adoption/purchase) option as $u_{it} = 0$ e.g. set the corresponding current utility level to zero.

This utility function setup has the following properties: a) a consumer's current utility is only associated with the current overall technology level; b) both "utility" and the "parameters" are consumer-specific in order to address the consumer heterogeneity. In our model setup, we assume that although consumers possess distinct requirements for technology, each individual consumer's need is consistent over time. Therefore, the consumer-level u_{it} also demonstrate the discrepancy of needs between consumers. Each consumer (i) will make purchase decision only if the current technology level ($Tech_t$) matches or exceeds their needs (e.g. $u_{it} > 0$).

It is worth mentioning that, the current utility (u_{it}) defined in our study is a "one-period" utility. It means that, given the technology level ($Tech_t$) at current purchase occasion (t), if the consumers decide to make the purchase and sign a contract, then on each time period within the contract length, customers will receive this utility (u_{it}). This is different from the CPG goods or durable goods that, once the customers purchase the product, they will receive an overall utility associate with the goods. Therefore, this is not a "one-time" overall utility but a "one-period" sequential utility.

Part 2 – Technology Evolution

As we mentioned previously that, in our study context, the technology level of the service keep on updating, which invokes consumer's forward looking behavior. We assume that, at current purchase occasion (t), consumers are able to fully observe the actual overall technology level ($Tech_t$). However, when consumers have forward looking behavior, they will have to consider both current utility and discounted future utilities to the current purchase occasion to make an optimal current decision. To quantify the future utilities, we will need the future overall technology level thus we need to define the technology evolution function.

We assume that the technology evolves in this way. The current actual overall technology level ($Tech_t$) is a function of previous actual overall technology level ($Tech_{t-1}$), the observed improvement of technology (IOT_t) and some unobserved technology evolution (ξ_t). (Equ (2))

$$Tech_t = Tech_{t-1} + IOT_t + \xi_t \qquad \text{— Eqn (2)}$$

Where $Tech_t$ is the actual overall technology level at current purchase occation (t), this is the same as the $Tech_t$ in the current utility function (Eqn (2)).

$Tech_{t-1}$ is the actual overall technology level at previous purchase occation (t-1).

IOT_t is defined as the "improvement of technology" at current purchase occasion (t), it was captured by the number of technology related official news released by the high-tech service provider at time (t). In our context, the high-tech service provider keeps on announcing official news to their customers. The customers can read the official news publicized on the official website every months. We collect the data of the number of news from the official website too and use it as the indicator of the "improvement of technology" in the technology evolution function.

ξ_t is the stochastic term capturing the unobserved technology evolution. We assume a normal distribution for the random component $\xi_t \sim N(0, \sigma_\xi^2)$.

In this modeling approach, we assume that the improvement of technology can be fully captured by the IOT_t, e.g. the number of technology related official news released by the high-tech service provider at current purchase occasion (t). We acknowledge that this can be a big assumption because we may have other resources that can capture the technology improvements. At current stage, we don't have other effective attributes that can help us to refine the technology evolution function. This can be another limitation of our study and suggests a direction for future study, e.g. improve the technology evolution function. We also acknowledge that, the quality of each individual news can be different, which means that the weight of each individual news representing the technology improvement can be different. However, in our model, we don't control for such quality deviation. We introduce an error term "ξ_t" to capture all aspects of the unobserved technology evolution. Therefore, the estimation of σ_ξ can inform us about the capability of official released news on capturing the improvement of technology.

As we mentioned previously, in our model setup, we assume that customers can fully observe the current technology level ($Tech_t$). However, to form forward looking structural model, we need the consumers to make prediction on future overall technology level. In order to do so, consumers will also need to predict the future technology evolution. And this requires a learning process. Specifically, consumers firstly obtain the historical information on how much official news were released in different months (e.g. $IOT_1, IOT_2 \dots IOT_t$) , then they form a belief on the future improvement of technology (IOT_{t+1}), e.g. news released speed, finally they could make a prediction on the future overall technology level ($Tech_{t+1}$).

Part 3 – Consumer learning on News releasing frequency

As pointed out previously that, it is a learning process that consumers form belief on the future improvement of technology based on the historical information of how much official news were released. And this learning process is critical for consumers to make prediction on future overall technology level. Therefore, in this section, we will specify the consumer's learning model on the "improvement of technology", e.g. the news releasing frequency(IOT_t).

In the learning model, we focus on the information role of news releasing frequency on consumer's perceptions about the improvement of technology. We assume that consumers have uncertainty about both the true mean level and the precision of the improvement of technology contained in and conveyed by the news releasing frequency. Consequently, consumers will base on the signals that they received from the news releasing frequency to form their perception on both the mean level and precision of the improvement of technology.

Therefore, we firstly specify consumers' learning of mean level and the precision of "improvement of technology" after exposure to the news releasing frequency. We assume that, the news frequency released each month provides a noisy but unbiased signal for the improvement of technology. Specifically, We defined that consumers believe the "improvement of technology (IOT_t)" at time (t) following a Poisson distribution with parameter of λ.

$$P(IOT_t) \sim Possion(\lambda) \qquad \text{—} Eqn\ (3)$$

Where λ represents the true mean level of the (IOT_t), which is unknown to the consumers. This equation suggests that the news released frequency provides imperfect information about the true mean level of the (IOT_t). As we all know that, the mean and variance are the same for Poisson distribution. Therefore, the λ not only represents the true mean level of the (IOT_t), but also captures the noisiness of information conveyed by the news released frequency.

Next we define that, consumers will learn this λ from the previous information set (I_{it}). Given the existing information set (I_{it}), a consumer has formed a prior opinion about the true mean level of (IOT_t), e.g. λ, which follows a gamma distribution:

$$\lambda | I_{it} \sim Gamma\ (\alpha_{it}, \beta_{it}) \qquad - \ Eqn\ (4)$$

Where α_{it}, β_{it} are the parameters of the prior distribution.

In the equation (4), I_{it} is defined as consumer (i)'s existing information set at time t. Several sources can contribute to the consumer information set I_{it}: 1) The first source is consumer prior knowledge about the high-tech $(\alpha_{i0}, \beta_{i0})$, e.g. the prior knowledge level before observing any official released news. The prior knowledge is an i-specific term as each consumer holds different opinions about the high-tech. 2) The second source is the official released online news (IOT_t), which is considered as the representative of the improvement of technology. Obviously, the news is a t-specific term as it is released by the high-tech service provider thus the publishing cycle has no relationship with the consumers. 3) The third source is the consumer intrinsic information process, which demonstrates how consumers continuously evaluate on the periodically released news. Each source will contribute to updating the consumer beliefs about the news described in the following equations.

After observing the actual news released at time (t), consumers begin to update their beliefs on the true mean level of (IOT_t), e.g. λ through updating the two parameters of gamma distribution $(\alpha_{it}, \beta_{it})$.

$$\begin{Bmatrix} \alpha_{it}^{*} = \alpha_{it} + IOT_t \\ \beta_{it}^{*} = \beta_{it} + 1 \end{Bmatrix} - Eqn(5)$$

Where α_{it}, β_{it} are the parameters of the prior distribution before observing the historical technology related news before time (t) and $\alpha_{it}^{*}, \beta_{it}^{*}$ are the parameters of posterior distribution of λ_t after the consumers are exposed to the number of news released at time (t).

The idea of learning model can be summarized at follow (Figure 2):

At the beginning of time t, consumers pertain a prior opinion about the true mean level of (IOT_t), which follows a gamma distribution with parameters of α_{it}, β_{it}. The true mea level of (IOT_t) is represented by λ.

During time t, the consumers receive a new information set of (IOT_t) e.g. they are exposed to a new set of official released news. According to customers belief, the number of the newly released official news (IOT_t) should follow a poisson distribution with the parameter of λ.

Combine the prior opinion and the new information set, the consumers form the posterior belief about the λ, which follow a gamma distribution with the parameter of $\alpha_{it}^{*}, \beta_{it}^{*}$.

Insert Figure 2

Finally, we consider that consumer will always place a higher weight on the more recently perceived information. We want to integrate the fact that consumer's confidence about their belief can gradually decrease by time into our model setup. To do so, we posit that consumers recall their prior evaluation on the technology related news with noise. In the model set up, we introduce the information discount factor in customer's learning. This modeling idea is originated from information processing theory in psychological study and was firstly introduced in learning model by Zhao et al. (2011).

The idea of information discount factor can be described more visually as follow. For example, we can posit ourselves as the decision makers of whether to adopt the high-tech service or not. We continuously learn from the official information that is available to us. Suppose that, we have only two information in hand, one information was a one-year old information and the other one was a new information publicized yesterday. What we need to do is to derive today's actual overall technology level to aid our decision making. The question is which information we should give more weight and which information we should give less weight when making the judgment? From a common and realistic perspective and in a rational consumer's decision process, the recent information should receive more weight, and older information is discounted. This is the concept of information discount.

Reflected in our model (Equation 6), we capture such noise in the information updating model by keeping the mean level of the news perception constant and allow the variances to increase over time (Zhao et.al 2011).

$$\begin{Bmatrix} \alpha_{i,t+1} = \delta * \alpha_{it}^{*} \\ \beta_{i,t+1} = \delta * \beta_{it}^{*} \end{Bmatrix} \qquad - Eqn(6)$$

Where δ is the parameter that accounts for the information discount process and takes the value between 0 and 1. If $\delta = 1$, it suggests that there is no information discount, e.g. consumers treat all period information equally important in their decision process.

This setting ensures that the information discount process only impacts the variance of news perception while keeping the mean level constant, as shown in Equation (7):

$$\begin{Bmatrix} E(\lambda | \alpha_{i,t+1}, \beta_{i,t+1}) = E(\lambda | \alpha_{i,t}, \beta_{i,t}) \\ Var(\lambda | \alpha_{i,t+1}, \beta_{i,t+1}) = \frac{1}{\delta} * Var(\lambda | \alpha_{i,t}, \beta_{i,t}) \end{Bmatrix} - Eqn(7)$$

Figure 3 visually displays the difference in the belief formation process between including the information discount factor (δ) and without the δ. From the plot, we can straightforwardly explain why we introduce the information discount factor in our study. On the plot, the X-axis represents the purchase occasion time (t) and the Y-axis represents the consumer's belief on the true mean level of (IOT_t), e.g. the consumer's belief on λ. It is worth mentioning that, the phenomena shown in the graph happens at infinity time condition, e.g. $t \rightarrow \infty$.

In the consumers' traditional learning process, as t approach infinity, consumers' belief gets close to a constant level regardless of the new coming information set. The reason is that, as t approach infinity, the two parameters of λ, e.g. both α and β, approach infinity at the same rate. Consequently, the variance of λ, e.g. $\mathrm{var}(\lambda) = \frac{\alpha}{\beta^2}$, will approach zero regardless of the future information set. The variance of λ approaching zero means that consumer's uncertainty about their belief will approach zero, which is equivalent to that the consumer's belief on λ becomes constant regardless of any newly released information set in the future. The underlying idea of the traditional learning process is that, when time is long enough and all the historical information were treated equally important, the consumers have already observed enough information to form their belief on λ, thus they have absolute confidence about their belief, e.g. have no uncertainty about λ. As a result, the new coming information will not alter consumers' evaluations about λ. Shown in the Figure 3, the expectation of λ, e.g. $E(\lambda)$, becomes a straight line.

However, when we introduce the "information discount" of δ, consumers' uncertainty will never approach zero. Consequently, whenever new information set shows up, consumer's belief, both mean level and uncertainty will be affected. The idea of learning process with information discount factor is that, with the information discount of δ, the historical information is no longer functional thus consumers' uncertainty will never become zero. Whenever new information are released, consumer's evaluation will be influenced. Shown in the Figure 3, consumer's expectation of λ, e.g. $E(\lambda)$, is always related to the most recent information.

Clearly, the learning process with the discount factor of δ should match better with the actual consumer's learning process (Zhao et al., 2011). In order to make our assumption more realistic, we assume the information discount exists. But our model is very flexible because we allow the discount factor to vary between 0 and 1. Our estimation will tell us whether the discount equals 1 or not. If it equals 1, this means that there is no discount, e.g. consumers treat all historical information equally important in their learning process. Clearly, the smaller the discount factor is, the heavier the historical information will be discounted.

Insert Figure 3 here

Part 4 – Consumer Dynamic Optimization on the Contract Length

So far, we have already defined the consumer's current utility function (Equation (1)); technology evolution function (Equation (2)); and consumer's learning model (Equation (3) ~ (7)). Recalled that, the reason of we develop technology evolution function and consumer learning model is because consumers have forward looking behavior in their decision process. We emphasized again that, in our model setup, we assume that customers can fully observe the current technology level ($Tech_t$). However, to form forward looking structural model, we need the consumers to make prediction on future overall technology level ($Tech_{t+1}$ ...) because they have to consider both current utility and discounted all future utilities to optimize the current decisions: whether to buy and how long to buy (contract length). That is why we need the technology evolution function.

Moreover, in order to make prediction on the future overall technology level, consumers will need to form belief on the future technology evolution, e.g. the future "improvement of technology" $(IOT_{t+1} \dots)$. And this is the reason we define learning model. Based on all the model setup from part 1 to part 3, we now can develop the consumer dynamic decision model to account for the forward looking behavior.

We model the consumers' decisions of whether to adopt the high-tech service and how long to sign a specific length of contract (CL) as a dynamic programming process. The objective of the consumer is to make a sequential decision in each of the T discrete period where T is infinite. This corresponds to the method of using infinite-horizon dynamic programming to solve consumer's markov decision process (Rust, 1994). Basically, the optimal decisions of whether to buy and how long to buy for each consumer (i) at any purchase occasion (t) is the solution to the following problem:

$$\max_{\{CL_{i1}(t_{i1}),CL_{i2}(t_{i2}),\dots\dots CL_{i\infty}(t_{i\infty})\}} E\left\{\sum_{k=1}^{\infty}\left[\sum_{\tau=t_{ik}}^{t_{ik}+CL_{ik}}\left(\rho^{\tau} * u_{it_k}\left(Tech_{t_k}\right)\right) - EC_i\right]\right\} - Eqn\ (8)$$

With the constrain of

$$0 \leq t_{i1} < t_{i1} + CL_{i1} \leq t_{i2} < t_{i2} + CL_{i2} \dots\dots t_{ik-1} + CL_{ik-1} \leq t_{ik} < t_{ik} + CL_{ik} \dots\dots$$

Where $CL_{ik}(t_{ik})$ indicates that consumer (i) make the *kth* adoption decision at time t_{ik} with a contract length of CL_{ik}.

$u_{it_k}(Tech)$ is the "current net-utility" that consumer (i) possessed from making the *kth* adoption decision at purchase occasion t_{ik}. For consumer (i), as shown in Equation (1), $u_{it_k}\left(Tech_{t_k}\right)$ is defined as follow:

$$u_{it_k}\left(Tech_{t_k}\right) = \gamma_{i0} + \gamma_{i1} * Tech_{t_k} + \varepsilon_{it_k} —— Eqn\ (9)$$

Both Equation (1) and (9) define the consumer's current utility associated with the current decision. The only difference is that: Equation (1) in general defines the current net-utility at any given purchase occasion (t); while Equation (9) specifically indicates the current net-utility at any *kth* adoption decision occasion (t_k). Therefore, we use "t_k" in Equation (9) instead of "t".

$\rho \in (0, 1)$ is the utility discount factor, which is used to discount all the future utilities to the current purchase occasion.

EC_i is the customer (i)-specific unobserved non-physical cost associated with signing a contract. EC is the short-term for "efforts of signing a new contract". As we described previously, the non-physical cost is a unique feature in the B2B world and can include searching cost, learning the technique, training employees, cooperating multiple internal departments, negotiating with service providers and sending budget application to top manager for approval etc. Considering that consumer's resources input on signing a new contract can be diversified because the consumer's needs are unique, we assume that the unobserved non-physical cost EC_i will follow a normal distribution, e.g. $EC_i \sim N(\overline{EC}, \sigma_{EC}^2)$

Equation (8) fully describes how we account for consumer's dynamic decision property, e.g. forward-looking behavior in our modeling approach. The modeling idea behind the Equation (8) is described as follow:

First, in this equation, we define that consumer's decision time span is from 0 to infinity because consumers need to consider a long-term decision. Within this decision time span, consumers can sign infinity number of contracts. As shown previously, we use the "k" as the contract index indicating each individual contracts from 1 to ∞. Obviously, "i" indicates each consumer and "CL" tells the contract length. "t_k", being more specific, is defined as the starting point of the *kth* contract.

Next, each contract can bring customer an overall value, which is represented by the expression within the bracket, e.g. $\left[\sum_{\tau=t_{ik}}^{t_{ik}+CL_{ik}}\left(\rho^{\tau} * u_{it_k}(Tech_{t_k})\right) - EC_i\right]$ in Equation (8). In this expression, the u_{it_k}, as stated before, is the current net-utility of consumer (i) making the *kth* adoption decision at decision time, also the contract starting point of (t_k). It is worth mentioning again that this net utility is a "one-period" utility, which means, within the contract, at each time period from (t_k) to $(t_{ik} + CL_{ik})$, consumer will receive this same net-utility. The expression of $\sum_{\tau=t_{ik}}^{t_{ik}+CL_{ik}}\left(\rho^{\tau} * u_{it_k}(Tech_{t_k})\right)$ in the bracket is called the "discounted cumulative utility", which means that, we discount all the periodically future utility (e.g. u_{it_k}) within the contract length "CL_{ik}" to the current decision time "t_k". "ρ" is the utility discount factor and "EC" is the unobserved non-physical cost. We mention again that "EC" is a one-time fee associated with each individual contract.

Finally, in order to understand customer's decision, we will have to maximize the long-term value that consumers obtained from signing each individual contract from 1 to infinity, e.g. the expression of $E\left\{\sum_{k=1}^{\infty}\left[\sum_{\tau=t_{ik}}^{t_{ik}+CL_{ik}}\left(\rho^{\tau} * u_{it_k}(Tech_{t_k})\right) - EC_i\right]\right\}$ in the Equation (8). That is why we see the mathematical notation of "max"at the beginning of Equation (8). It has to be pointed out that maximizing this value is a very complicated process because customers have to consider all future possibility of decisions in order to optimize the current decision. Moreover, since we don't know the future technology level $(Tech_{t_k+1} \ldots.)$ at any given decision occasion "t_k" , we have to compute the expectation of the value, which is the reason that we see the mathematical notation of "E" representing expectation at the beginning of the expression.

To illustrate this value function better, we draw a graph (Figure 4) to visually explain the modeling idea. In Figure 4, let's assume that consumer *(i)* decides to adopt the high-tech service at decision time of "t_{i1}" and sign a contract of "CL_{i1}". Then within the contract length of "CL_{i1}", like we said before that, on each time point from "t_{i1}" to "$t_{i1} + CL_{i1}$", this consumer will receive the same net-utility of "u_{it_1}" as the net-utility at the decision time "$u_{it_1}(Tech_{t_1})$" because consumer will be bonded with the on-site overall technology level once the contract is signed. In the B2B world, since the consumers will have to spend effort on signing each contract, we assume that consumers will try to solve their problems in one holistic decision thus will not make any new high-tech service purchase within the signed contract duration. And there is little change that consumers will break the contract, especially in the B2B world, due to the remarkable resources they have already spent on signing the contract and the potential penalty of breaking an existing contract. This suggests that once the consumers make the adoption decision at time "t_{ik}", the next available decision time is "$t_{ik} + CL_{ik} + 1$" .

When the contract of "CL_{ik}" ends, consumers can continue to make decision on whether to adopt again or not. At each decision occasion after "$t_{ik} + CL_{ik}$", the decision can be either yes or no. In order to make the optimal decision at each decision occasion, the consumer will have to consider all the future decision possibilities and discount the potential utilities to the current decision occasion. This dynamic decision process will move on to the infinity decision occasion "t_{∞}" and infinity number of contract length decision "$CL_{i\infty}$". That is why it is called "infinite-horizon dynamic programming". It doesn't matter whether or not there is a time gap between consecutive decisions.

Insert Figure 4

It has to be pointed out that, in our dynamic model setup, decision variable "CL_{ik}" is a "continuous decision" not a "discrete choice". When "$CL_{ik} = 0$", it indicates "no adoption decision". When "$CL_{ik} > 0$", the actual number of "CL_{ik}" indicates the contract length signed by the consumer. Ideally, "CL_{ik}" can be any number between 0 and infinity. This is a key difference between our proposed dynamic decision model and the standard dynamic decision model in the existing studies from consumer's perspective. Introducing the continuous decision in the dynamic decision model makes the derivation process of value function (described in part 5 below) becoming a contribution in the paper because the solution of the Bellman Equation (described in part 5 below) is significantly different from that in discrete dynamic decision model.

Part 5 – Derive Bellman Equation (Value Function)

As we described previously, to solve the optimal decision shown in Equation (8), we need to derive the bellman optimality equation. The fundamental idea of Bellman Equation, which was part of the dynamic programming theory, was invented by Bellman (1954, 1956) in applied mathematics studies, and later was introduced into economics field to solve discrete Markov decision process (Rust, 1994; Puterman, 1990).

When we use dynamic programming theory to solve discrete Markov decision process, we firstly need to be clear about the major components in the Markov decision process (Rust, 1994). The components, which are also the foundation for developing Bellman Equation, include:

1) The time variable (t); the time variable corresponds to the decision occasion in the market, and $t\ can\ takes\ integers\ from\ 0\ to\ \infty$

2) State variables $(\tilde{S})$; the state variables determine the outcomes of value function. Once we specify the value of state variables, we know the outcomes of value function.

3) Decision variable(s) $(\widetilde{D})$: the decision variables indicate consumers' decision at each decision occasion. In discrete Markov decision process, the decision variables can be either whether to buy the product or not, such as Song & Chintagunta's study (2003), or the choice between alternative brands, such as Erdem et al.'s study (2003). In our study, the decision variable is consumer's "contract length" decision, which is a continuous variable.

4) The transition probabilities for all state variables $\widetilde{\mathrm{P}}(\tilde{S}_{new}|\tilde{S}_t, \widetilde{D}_t)$: the transition probability is used to identify the conditional expectation of the utility over all future decision occasions. As we mentioned previously that, the essence of using dynamic structural approach to model consumer's forward looking behaviors is: in order to optimize a current purchase decision, consumers need to combine both current utility at the decision occasion and all future discounted utilities. When consumers evaluate the all future utilities, they don't know the corresponding level of the state variables. Therefore, we will have to take the expectation thus require the transition probability.

Following the theory defined in dynamic programming, we derive the Bellman Equation, e.g. the value function. The definition of "value function" is very similar as the "utility function". Defining "value function" in dynamic model, from the modeling idea perspective, is equivalent to defining "utility function" in static model. The only difference is that, the utility function refers to the current utility that consumers received from the static decision at current decision occasion; while the value function indicates the expected discounted summation of all utilities (both current and future) that consumers received from their dynamic decision under forward-looking behaviors.

Next, we derive the value function specifically for our study as follow:

$$V_i(\tilde{S}_{it}) = \max_{CL}\left\{\sum_{j=0}^{CL}\rho^j u_{it}(\tilde{S}_{it}, CL) + \rho^{CL+1}\int_{\tilde{S}_{i,t+CL+1}} V_i(\tilde{S}_{i,t+CL+1}) * dF(\tilde{S}_{i,t+CL+1}|\tilde{S}_{it}, CL) - EC_i(CL > 0)\right\} \quad --Eqn\ (10)$$

Where $V_i(\tilde{S}_{it})$ is the value received by consumer (i) at a decision occasion (t) with a given set of states $(\tilde{S}_t)$. The $V_i(\tilde{S}_{it})$ is an "optimal" value, which means that, if the consumer (i) make an optimal contract length decision at decision occasion (t) with a given set of states $(\tilde{S}_t)$, s/he should receive a value of $V_i(\tilde{S}_{it})$ which is the best value comparing with all other possible contract length decisions. This is the reason that we have "$\max_{CL}$" at the beginning of the right side of the equation.

$\tilde{S}_{it}$ represents the current state variables that consumer (i) face at the current decision occasion (t). In our model, the state variables include: $\tilde{S}_{it} = (Tech_t, \alpha_{it}, \beta_{it}, \varepsilon_{it})$. Recalled that, $Tech_t$ is the actual overall technology level; α_{it}, β_{it} are the parameters in consumer learning model determining consumer's perception on the true mean level of news released frequency; ε_{it} is the error component in the current utility function (Equation (1)) which is assumed to follow a normal distribution.

$\tilde{S}_{i,t+CL+1}$ represents the transition states, meaning the states at the beginning of next available decision occasion, e.g. $(t + CL + 1)$, e.g. after the current chosen contract duration ends. The state variables include: $\tilde{S}_{i,t+CL+1} = (Tech_{t+CL+1}, \alpha_{it+CL+1}, \beta_{it+CL+1}, \varepsilon_{it+CL+1})$.

To better understand the value function, we need to view it in two parts:

1) The first part is the expression of $\sum_{j=0}^{CL} \rho^j u_{it}(\tilde{S}_{it}, CL)$, which represents the cumulative discounted utilities that consumer (i) received from the contract length decision (CL) that s/he made at the decision occasion (t) with a given set of state $(\tilde{S}_{it})$. Within the contract length(CL), at each time period, consumer (i) received the same utility as the one at the decision time (t), therefore, we are able to compute the exact value of the cumulative discounted utilities within the contract length.

2) The second part is the expression of $\rho^{CL+1} \int_{S_{i,t+CL+1}} V_i(\tilde{S}_{i,t+CL+1}) * dF(\tilde{S}_{i,t+CL+1} | \tilde{S}_{it}, CL)$, which represents the discounted expected value of the all future utilities after the contract length (CL) ends. Consumers should face the new set of states $(\tilde{S}_{i,t+CL+1})$ when the contract length (CL) at decision occasion (t) ends, consequently, the value of the all future utilities should be $V_i(\tilde{S}_{i,t+CL+1})$. Because we don't observe the actual value of the future new states $\tilde{S}_{i,t+CL+1}$ when

consumers are trying to optimize the decision at current occasion (t), we need to take the expectation, e.g. integrate the new states out in the expression, thus need the transition probability of $dF\left(\tilde{S}_{i,t+CL+1}\middle|\tilde{S}_{it}, CL\right)$.

It is worth mentioning that, the derivation of the Bellman Equation is based on the assumption of stationary Markov decision process, meaning that, consumer's decision rule is consistent at each decision occasion (t) (Rust, 1994). That is the reason that we see the value function only depends on the state variable but not depends on the decision occasion (t).

Part 6 – Likelihood Function

After we specify the value function (Equation (10)) to quantify consumer's decision process, the next step is to empirically estimate the model, which requires the likelihood function.

In our model setup, we can categorize the parameters into four groups: 1) the parameters in the utility function: γ_{0i}, γ_{1i}; 2) the parameters in the technology evolution function: σ_{ξ}; 3) the parameters in the consumer learning model: $\alpha_{i0}; \beta_{i0}$; δ; where $\alpha_{i0}; \beta_{i0}$ represents the initial value of the $\alpha_i; \beta_i$; and 4) efforts of signing the contract: ECi. Please note that, since the $\gamma_{0i}, \gamma_{1i}, ECi$ are consumer-specific parameters, therefore, the ultimate parameters estimated by the MLE become $\overline{\gamma_0}, \sigma_{\gamma_0}; \overline{\gamma_1}, \sigma_{\gamma_1}; \overline{EC}, \sigma^2_{EC}$.

Next, we need to derive the probability distribution of decision variable (CL) given all the parameters. This requires us to know the expression of stochastic terms as a function of the decision variable (CL), which requires the value function we defined before (Equation (10)).

According to the value function (Equation (10)), we know that

$$CL_{it} = G_V(\gamma_{0i}, \gamma_{1i}, Tech_t, \alpha_{it}, \beta_{it},\ \varepsilon_{it}, EC_i) \qquad -- \text{Eqn (11)}$$

To show the expression more obviously, here we use "G_V" to represent the value function.

Based on our model setup of technology evolution function (Equation (2)) and consumer learning model (Equation (5) & (6)), we have the following functional forms:

$$Tech_t = \mathcal{G}_{tech}(\,IOT_1, \ldots..IOT_t, \xi_1, \ldots.\xi_t); \qquad --Eqn\ (12-1)$$

$$\alpha_{it} = \mathcal{G}_{\alpha}(\alpha_{i0}, IOT_1, \ldots..IOT_t, \delta\,) \qquad --Eqn\ (12-2)$$

$$\beta_{it} = \mathcal{G}_{\beta}(\beta_{i0}, t, \delta) \qquad --\mathrm{Eqn}\ (12-3)$$

Similarly, we use the "$\mathcal{G}_{tech},\ \mathcal{G}_{\alpha},\ and\ \mathcal{G}_{\beta}$" to represent the function of technology evolution and the two parameters in learning model respectively.

Combine equation (11) and equation (12), we have:

$$CL_{it} = G_V(\gamma_{0i}, \gamma_{1i}, t, IOT_1, \ldots..IOT_t, \xi_1, \ldots.\xi_t, \alpha_{i0}, \beta_{i0}, \varepsilon_{it},\ \delta,\ EC_i) \ --Eqn\ (13)$$

As shown in Equation (13), we have two sets of stochastic terms in our model: one is ε_{it}; the other include $\xi_1, \ldots.\xi_t,$. To simplify the expression and the estimation process of our likelihood function, we want to derive the relationship between decision variable "CL_{it}" and "ε_{it}" only. To do so, we denote the contract length CL_{it} as a function of ε_{it} conditional on both the parameters and stochastic term of $\xi_1 \ldots \xi_t$, e.g.

$$CL_{it} = G_V(\varepsilon_{it}|\widetilde{\Lambda}, \{IOT_t\}, t) \qquad --Eqn\ (14)$$

Where $\widetilde{\Lambda}$ represents the set of all parameters, e.g. $\widetilde{\Lambda} = \{\gamma_{0i}, \gamma_{1i}, EC_i, \delta,\ \xi_1, \ldots.\xi_t, \alpha_{i0}, \beta_{i0}\}$

Please note that the $\{\mathrm{IOT_t}\}, \mathrm{t}$ are not parameters but the "news frequency per month" and "decision time" in our data. The Equation (14) suggests that, we don't treat $\xi_1, \ldots.\xi_t$ as stochastic term but as time-specific parameters in our likelihood function.

Based on the assumption of $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and $\xi_t \sim N(0, \sigma_\xi^2)$, we can write the conditional density function of observed data CL_{it} by using the Change-of-Variable theorem:

$$P(CL_{it}|\tilde{\Lambda}) = f\left(G_V{}^{-1}\left(\varepsilon_{it}|\tilde{\Lambda},\{IOT_t\},t\right)\right) * \|J(\varepsilon_{it} \to CL_{it})\| \quad -- Eqn\ (15)$$

Next, what we need to do is integrating out all the heterogeneity terms and the unobserved terms including $(\gamma_{0i}, \gamma_{1i}, EC_i, \xi_1, \dots . \xi_t, \alpha_{i0}, \beta_{i0})$. Finally, we could derive the Likelihood function as:

$$L\left(\tilde{\Theta}\middle|\{CL_{it}\},\{IOT_t\},t\right) = \oint_{\tilde{\Lambda}} f\left(G_V{}^{-1}\left(\varepsilon_{it}|\tilde{\Lambda},\{IOT_t\}\right)\right) * \|J(\varepsilon_{it} \to CL_{it})\| * \tilde{\pi}(\tilde{\Lambda}|\tilde{\Theta}) \quad -- \quad Eqn\ (16)$$

Where $\tilde{\Theta} = \left\{\, \overline{\gamma_0}, \sigma_{\gamma_0}; \overline{\gamma_1}, \sigma_{\gamma_1}; \overline{EC}, \sigma_{EC}^2; \delta;\ \sigma_\xi; \sigma_\varepsilon;\ \overline{\alpha_0}, \sigma_{\alpha_0}; \overline{\beta_0}, \sigma_{\beta_0} \,\right\}$

Obviously that, in order to find the Likelihood for parameter estimation, we need the function of $\varepsilon = f(CL)$, which is actually the $G_V{}^{-1}$ function in Equation (15) and (16). Since we still don't have the close-form solution for $\varepsilon = f(CL)$, we will need to empirically compute the following items: 1) Given a "CL", find the corresponding "ε"; and 2) the $\|J(\varepsilon_{it} \to CL_{it})\|$ by iteration.

We can draw a diagram to visually show the logical link between the six parts of the model setup (Figure 5). Basically, when consumers optimize their dynamic decisions, they will consider two parts of the benefits associated with the current decision: One is the current utility of the decision, e.g. the utility function we established in Part 1; the other is the future utilities of decisions. In static model, consumers only have to consider the first part of the benefit. The second part of benefits, which involves the evaluation of future utilities of current decision, is the major difference from the static model and is the major reason why we call it forward-looking dynamic model.

Consumers don't know the future states to quantify the future utilities thus they will have to form expectation on the future states. Reflected in our model, we need to build the technology evolution function in Part 2 and the learning model in Part 3 to understand how consumers form their expectation on the future states. Finally, we put both current utility and future utilities together to understand consumers' dynamic decision optimization process, which lead to the dynamic model in Part 4. Next, to solve the maximization process in the dynamic model, we utilize the bellman equation to derive the Value Function shown in Part 5. And to estimate our model, we need to derive the Likelihood function discussed in Part 6.

Insert Figure 5 here

Model Estimation

We estimate the model parameters using a "Maximum Likelihood Estimation" approach. Specifically, we use the "Simplex Method" to identify the optimal parameter estimation result (Nelder and Mead, 1965). The detailed steps of performing "Simplex Method" are described in Appendix C. As we mentioned previously in Modeling Framework section, our model estimation process includes two loops of numerical searching, e.g. the simulation of "value function" and the searching for maximum likelihood function. We described the detailed steps of value function simulation in Appendix A. Moreover, in order to compute the likelihood function, we also need to find the ε at a given contract decision of CL and solve the Jacobian of $\|J(\varepsilon_{it} \rightarrow CL_{it})\|$, whose detailed steps are described in Appendix B and D respectively. We can draw a diagram to visually show the estimation process (Figure 6) and summarize the detail steps as follow:

Step 1: At a given set of parameters and pre-defined boundary and grids of state variables, we can compute the corresponding value function (Appendix C)

Step 2: Combined the simulated value function and the actual data of purchase time, contract length and news frequencies; we can calculate the likelihood function for the given set of parameters. Here we need to utilize Appendix A, B and D to identify the result of likelihood.

Step 3: We integrate the likelihood computation into Simplex Process to find the optimal parameter estimation, e.g. the set of parameters corresponding to maximum likelihood. If the results pass the searching criteria of Simplex Process, then we stop and report the results; otherwise, we go back to step 1 and do the computation again.

Insert Figure 6 Here

Data Description

Our data comes from a global leading high-tech company (service provider). Its products include almost all kinds of high-tech products and infrastructures, such as hardware, software, and personal computers, and covers both B2B and B2C settings. From 2007, they began to build a cloud service for their B2B consumers.

In total our data includes 218 business-to-business buyers. All of them have had at least one historical purchase with the service provider before purchasing the hi-tech service. The characteristic suggests that the 218 buyers should consider the service provider in our study as the first choice over other competitors when adopting the high-tech service, not only because the service provider holds the superior power in the market, but also because the buyers are existing B2B consumers who retain better knowledge about having a relationship with the service provider.

Our data includes both the service transaction and the official news items. Both data were collected at the monthly level. The time frame of service transaction data ranges from January. 2009 to September 2011 and the data include both consumer adoption decision time (year & month) and the decision of contract length (in unit of months). The official news-count data time-frame ranges from October 2008 to January 2012 and the data were recorded as number of news items per month.

We present both a graphical illustration and summary statistics of consumer's contract length decision and officially released news-counts in Figure 7. On average, consumers purchase the service for 14.5 months, which is slightly longer than one year, but the range can cover from 1 month to 60 month, e.g. 5 years (Figure 7, left table). Additionally, the average number of official news-counts that the service provider released every month is 3.95, with the minimum of 1 and maximum of 11 (Figure 7, right table). Moreover, the patterns of the contract length and the news-counts (Figure 7) suggest that, in general when the news-count is high, the contract length tends to be low. Recall that news-count is used to capture the improvement of technology. We previously mentioned that when consumers consider the tech evolution to be fast, they may either postpone their adoption decision or take a try with a shorter contract; our data pattern shows that the concept is empirically true.

Insert Figure 7 here

Structure of Value Function

Before illustrating the model estimation results and policy simulation, it is worth showing the structure of the value function because the essence of using the dynamic programming approach to model consumer's forward looking behavior is the construction of value function for capturing consumer's decision process. Since our primary interest is to explore the underlying relationship between technology evolution and consumer's decision process, we will illustrate the pattern of consumer's value as a function of contract length and the two drivers of technology evolution, e.g. the overall technology level and technology evolution pace.

Please note that the "value" we compute and plot on Figure 8 and Figure 9 is the $v\left(\tilde{S}_t, CL\right)$ (Equation (A2)) thus it is a function of both state variable "$\tilde{S}_t$" and the decision variable "CL". Since the state variable $\tilde{S}_t$, after simplification (Appendix A), includes ($Tech,\ \frac{\alpha}{\beta}, \varepsilon$). To show the relationship between $v\left(\tilde{S}_t, CL\right)$ and "CL" at different states of "$Tech$" and "$\frac{\alpha}{\beta}$" respectively, we compute the $v\left(\tilde{S}_t, CL\right)$ at fix value of "$\varepsilon = 0$".

Figure 8 shows the pattern of value function as a function of both contract length and the overall technology level, e.g. "$Tech$". When overall the technology level is low (Figure 8, left panel), the shape of consumer's value indicates an optimal contract length at $CL = 0$. This suggests that, when technology level is very low, consumers don't think about making a purchase. When the overall technology level is higher (Figure 8, right panel), the optimal contract length is greater than zero, meaning consumers purchase the service. The relationship between the consumer's decision and the overall technology level implies that, only if the overall technology level elevates to exceed the consumer's need, then the consumer will consider making the purchase.

Insert Figure 8 here

Next we will show that, how the pattern of value functions as a function of both contract length and the technology evolution pace, e.g. "$\frac{\alpha}{\beta}$" (Figure 9). From the Figure 9 – Panel 1, we can see that when the technology evolution speed is low, the shape of consumer's value indicates that the optimal contract length decision appears at very high "CL" value. This suggests that, when technology evolution pace is low, if the buyer decides to buy, they are more likely to buy with a longer contract because they may not expect a significant technology improvement to happen in the near future. The comparison of the 4 panels in Figure 9 shows a clear relationship between technology evolution pace and consumer's contract length decision, e.g. consumer's contract length decreases as the technology evolution pace increases. The relationship matches with our prior expectation that, consumers will tend to have a try on approach the cloud service when expecting a fast technology innovation pace.

Insert Figure 9 here

Simulation

In this section, we demonstrate the ability of our model to recover parameters and ensure empirical identification. Our simulation scheme is as follow: First, we simulate the news frequency data, based on a Possion distribution, $X \sim Pois(\lambda)$ and we generate the news frequency for 36 time periods. Then, we simulate the two random error term: $\varepsilon_{it} \sim N(0,1);\ \ and\ \xi_t \sim N(0,\sigma_\xi)$. Now we can compute the "$Tech_t$" using the technology evolution function (Equation (2)). Next, we simulate the purchasing behavior of 200 individual consumers. Then reasons that we select 36 time periods and 200 individuals are: first, they match with our actual data size; second, the simulation and estimation process take heavy computation burden. Therefore, to both satisfy the requirements of simulation and reduce the computation burden, we select the 36 time periods and 200 individuals. In order to simulate consumer's purchasing behavior, it is obvious that we need the value function. First, we generate the three heterogeneous terms: $\gamma_{0i} \sim N\left(\overline{\gamma_0}, \sigma_{\gamma_0}\right);\ \gamma_{1i} \sim N\left(\overline{\gamma_1}, \sigma_{\gamma_1}\right);$ and $EC_i \sim N(\overline{EC}, \sigma_{EC})$. Then combine both the heterogeneous parameters of $\gamma_{0i}, \gamma_{1i}\ and\ EC_i$ and the fix parameters of $\delta\ and\ \sigma_\xi$, we can generate the value function $V^*\left(\tilde{S}^*\right)$ (Appendix A). Based on the computed value function, we can decide both the timing of purchase and the length of contract for each consumer at each purchase occasion. Please note that, we only consider parameters of $\{\overline{\gamma_0}, \sigma_{\gamma_0}; \overline{\gamma_1}, \sigma_{\gamma_1}; \overline{EC}, \sigma_{EC}^2; \delta;\ \sigma_\xi\}$ and in our simulation process with the following reasons:

1) Since ε is the random error in current utility function (Equation (1)) which determines consumer's decision of whether to buy or not, e.g. binary decisions. It is acknowledged that in choice model, the variance of error term is not identifiable. Therefore, we fix $\sigma_\varepsilon = 1$ for identification issue.

2) Also, we fix the both the initial parameter of belief updating, e.g. $\alpha_{i0}\ and\ \beta_{i0}$ for two reasons. First, the cloud service is a novel technology in the market thus it is reasonable to assume that consumers don't possess knowledge about the specification of the technology at the beginning time. Second, from our empirical findings, the absolute values of these two initial parameters don't contribute significantly to the parameter estimation but add a computation burden. Therefore, we fix $\alpha_{i0}\ and\ \beta_{i0}$ at zero for all consumers.

As shown in Table 2, all the parameters are reasonably estimated. This result suggests that we are able to reasonably identify and recover the parameters.

Insert Table 2 Here

Model Comparison and Validation

We compare our model with the static models (OLS and Tobit) for model comparison and validation. In the Tobit model, we use both the news frequency and the cumulative news-count as the independent variables. The model fit statistics are shown in Table 3. The Bayesian information criterion (BIC) result shows that our proposed model outperforms both OLS and Tobit model in quantifying consumer's purchasing decision. The superiority of our model over the static model suggests the importance of including consumer's forward-looking behavior in understanding consumer's dynamic purchase behaviors. We will focus on discussing the full model specification in the next section.

Insert Table 3 Here

Empirical Results and Discussion

In this section, we first discuss some of the estimation results of the proposed model and explain the meaning of the estimation results. Based on the parameter estimation results, we will show how consumer's value is shaped and how the consumers' contract length decision will change as a function of the focal state variables (Equation (10)). Then we will illustrate the policy simulation outcomes and propose the associated managerial implication.

<u>Parameter Estimation</u>

The parameter estimation appears in Table 4. Standard error of the estimation suggests that all estimated parameters are significant at the 95% significance level. We take into account the consumer heterogeneity for three parameters in the model, e.g. γ_0 which is the mean-level net-utility; γ_1, which is the coefficient of technology in the utility function (Equation (1)), and EC, which is the parameter to capture the efforts of signing a contract. Therefore, their parameter estimations include both mean-level and standard deviation. Also, we ignore the estimation of the two initial parameters α_{i0} and β_{i0} and fix them at zero for all consumers with the reason described previously in simulation section.

Insert Table 4 here

From Table 4, we can see that, $\overline{\gamma_0}$ is estimated to be negative. Recall that γ_0 indicates the mean-level net-utility when the actual overall technology level =0 and represents the difference between the consumer's baseline preference toward the actual overall technology level and the potential unobserved cost associated with consumer's need. A negative $\overline{\gamma_0}$ suggests that, on average, consumer's baseline preference toward the technology can't outperform the cost of the technology thus is not strong enough to incite a purchase.

Next, we can see that, the $\overline{\gamma_1}$ is a positive number suggesting that, on average, consumer's utility should increase as the overall technology level increase, which make sense in the real world. And we observe that the estimated absolute value of $\overline{\gamma_1}$ appears to be small, but the resulting contribution of technology on the net-utility is remarkable. We can use some rough calculation to show. In our data, if we only compute the overall technology level by the summation of news in each month, e.g. we assume the unobserved part of $\xi_t = 0$ in Equation (2); within our data time frame (e.g. from Jan. 2009.1 to Sep. 2011), we can reach an approximate result of overall technology from 3 to 147. Then, using Equation (1), we can compute the impact of the overall technology level on the consumer's net-utility, which is from 0.0417 to 2.0433. Since we fix the variance of the net-utility error $\sigma_\varepsilon = 1$ due to the identification issue, we can see that given the parameter estimation of $\overline{\gamma_1}$, the contribution of the overall technology level on consumer's net-utility is significant.

δ is the information discount factor, which is used to capture the process of consumer retrieving information. As we mentioned previously that, information processing theory suggests that consumers put more weight on the more recent information. The estimated information discount is 0.829 suggesting a relatively heavy information discount. It is worth mentioning again that, the smaller the value of information discount is, the less weight that historical information will stay in consumer's mind when consumers making prediction on future technology level. With a discount factor of 0.829, more than two-year old information is no longer considered in consumer's evaluation of future technology level. This also suggests that, consumers may concentrate more on the most recent information on evaluating the technology evolution.

The σ_ξ is the standard deviation term for the random error in technology evolution function (Equation (2)). It is used to capture all the unobserved impacts on technology level other than the news-count. According to the estimation result, the estimated $\sigma_\xi = 1.2764$ suggests a relatively large variance. This result implies that the news-count is not able to fully capture the technology improvement thus there some errors exist in the technology evolution setting.

Value Function under Parameter Estimation Results

Based on the estimated parameters, we can compute the value function specifically for our study. Figure 10 shows the 3D plot between the cloud adopter's value and the two focal state variables of the technology evolution pace and the overall technology level, represented by $\frac{\alpha}{\beta}$ and $Tech$ respectively. The "ε" state variable has been integrated out. Therefore, the "value" we plot in Figure 10 is the $V^*(\tilde{S}_t^*)$, which is independent of the decision variable "CL".

The 3D value function (Figure 10) shows that, the increase of both overall technology and technology evolution will increase the consumer's value. In the reality, consumers should benefit from evolutions in technology. Therefore, the value function result matches with the real world situation in that consumers gain value from both overall technology elevation and increases in the speed of innovation .

Insert Figure 10 here

<u>Relationship between Contract Length Decision and Technology Evolution</u>

Figure 11 is a 3D plot showing the relationship between the optimal contract length decision and the two focal state variables of "$\frac{\alpha}{\beta}$" and "$Tech$". The optimal contract length is the "decision" variable of "CL" in the value function (Equation (10)). Since optimal contract length (CL) is a function of all state variables of $\frac{\alpha}{\beta}$ $, Tech\ and\ \varepsilon$, in order to show how consumer's optimal contract length decision changes by the two focal variables of $\frac{\alpha}{\beta}$ $, Tech$, we calculate the optimal contract length shown in Figure 11 at $\varepsilon = 0$.

We find several observations from Figure 11. First, from the direction of "overall technology", we can see that the "optimal contract length" stays at "zero" until the "overall technology" reach a certain level. This finding suggests that, only if the overall technology level, e.g. "$Tech$", beyond consumer's need, the purchase decision will be invoked. Second, from the direction of "technology evolution pace", we can see that the "optimal contract length" shows a clear downhill shape. This observation indicates that, the contract length decision is impacted by the technology evolution pace, e.g. "$\frac{\alpha}{\beta}$" and the relationship is clearly negative; e.g. the optimal contract length decreases as technology evolution becomes faster. Third, still from the direction of "technology evolution pace", we observe that, when "technology evolution pace" increase to a certain level, the shape of "optimal contract length" shows a clear "indentation" toward larger “overall technology level”. The phenomenon tells that, when the pace of technology evolution increases to a certain level, the consumer’s purchase decision is also delayed.

In summary, the results shown in Figure 11 suggest that, our model supports our previous assumption of: when technology evolution is fast, consumers will expect a better technology to show in the near future thus they may either postpone their purchase decision or sign a shorter contract to try the new technology.

Insert Figure 11 here

Policy Simulation

Given the parameter estimation results, we will conduct two settings of policy simulation. The first setting is that we will compare consumer’s adoption decisions at two different paces of technology evolution. The first is that the technology evolution happens more frequently but each step of evolution is relatively smaller, which is defined as “jogging pace” evolution. The “jogging pace” evolution is represented by an evenly distributed news-count per month. The second is that the evolution happens less frequently but each step of evolution is relatively larger, which is defined as “leaping pace” evolution. The “leaping pace” is represented by a large amount of news released simultaneously in one month, and then no news for several months. To make the two paces of evolution comparable, we will keep the total news-count released constant. The second setting is that we will compare the consumer’s adoption decisions at two different levels of the “effort of signing a contract”: benchmark efforts vs. half of the benchmark efforts.

The reason of selecting the two settings is that, both settings can be linked with service provider’s strategic decision-making, meaning that the service provider, at least to some degree, can adjust the levels of the two settings. As we all know that in hi-tech market, both releasing new technology and update existing tech are very strategic. It is a very common phenomenon in hi-tech market that firms may research and develop a new tech this year, but release it 5 years later. The results of our first setting policy simulation can aid firms in strategically managing the evolution in their technology. Although the efforts of signing a contract is majorly under the control of business-to-business consumers, service providers can still implement some plans to partially alter it. For example, the service provider could proactively offer to help reduce their consumers' efforts in signing the contract, such as help the consumers train their internal IT employees. On the other hand, the service provider could increase the effort involved in signing a contract by acting less friendly toward the consumer. Therefore, both of the two settings of policy simulation can provide valuable strategic implications.

The policy simulation of the first setting shows the following results. First, comparing the "optimal decision time" between the two paces of evolution (Figure 12), we find that the jogging pace has "longer" decision time. Specifically, from jogging pace to leaping pace, we see a 54.18% drop in the optimal decision time. It is worth mentioning that, the "longer" decision time indicates a "later" purchase decision thus actually suggests a "smaller" purchase probability. The comparison of optimal decision time between the jogging and leaping evolution paces suggests that, when facing more frequent but smaller steps in technology evolution, consumer's purchase probability tends to be low, meaning consumers are less interested in making the purchase. On the contrary, although less frequent, the larger step evolution in the leaping pace can create a stronger incitement on consumer's purchase decisions, e.g. consumers are more likely to make purchase. Therefore, from the perspective of optimal decision time, the leaping pace is better than the jogging pace because consumers show higher purchase intention (probability).

Next, comparing the "optimal contract length" between the two paces of evolution (Figure 13), we find that the jogging pace has "longer" contract length. Specifically, from the jogging pace to leaping pace, we observe a 16.93% drop in the optimal contract length decision. Intuitively, a longer contract length should give higher profitability. Therefore, from the perspective of contract length decision, the jogging pace is better than the leaping pace because consumers tend more likely to give long-term profits. In summary, the result of the first setting policy simulation (Figure 12 and 13) tells us that, when facing more frequent but smaller steps in evolution, although the purchase probability is relatively low, once the consumer makes the purchase decision, they tend to sign a longer contract. On the contrary, although the larger step in evolution tends to incite consumers to make purchases, consumers are less likely to sign a profitable contract.

Insert Figure 12 and 13 here

The policy simulation of the second setting also shows very interesting results. First, we find a 14.02% drop in the optimal decision time when the efforts of signing contract reduce 50% (Figure 14). As we mentioned previously that a "longer" decision time suggests a "lower" purchase probability, while a "shorter" decision time suggests a "higher" purchase probability. The finding indicates that, when the efforts of signing a contract are reduced, consumers should be more willing to make a purchase, which makes sense in the actual world. Therefore, from the perspective of optimal decision time, lower effort involved in signing a contract is better because consumers are more likely to make purchase.

Next, we observe a 9.71% drop in the optimal contract length decision when the efforts of signing a contract is reduced 50% (Figure 15). Considering that a longer contract is more profitable, from the perspective of contract length decision, a higher effort of signing a contract is better because consumers are more likely to give higher profits. This finding maybe contradicts to some common sense because we may easily think why consumers should even make the purchase if the efforts of signing a contract are high? It need to be clarified that, the case of "how long consumers will sign the contract" that we discuss here is conditional on "consumers make the purchase decision". The underlying idea behind our result here is that, when consumers decide to purchase and they have already spend great efforts on finalizing the contract, they are more likely to sign a longer contract to avoid another input; which make perfect sign in B2B world.

In summary, the result of the second setting policy simulation (Figure 14 and 15) reminds us that, although lowering the efforts of signing a contract tends to positively stimulate a consumer's purchase decisions, the firm may not reach desirable profitability because consumers tend to sign a shorter contract in order to better capture future technology evolutions. On the other hand, consumers are more reluctant to make the purchase when facing relatively higher efforts of signing a contract; however, once they make the purchase decision, they tends more likely to sign a profitable contract.

Insert Figure 14 and 15 here

Managerial Implication

Our study contributes substantively in two ways. First, our model demonstrate that it is important to consider both consumer heterogeneity and forward-looking in consumer's decision process especially in the rapidly developed high-tech market. Both researchers and managers should recognize that the necessary to account for forward-looking behavior in modeling consumer's purchase decision. Moreover, based upon our model estimation results, we can improve a firm's understanding on consumer's behaviors in the turbulent market. For example, the parameter estimations of $\overline{\gamma_0}$ and $\overline{\gamma_1}$ tell us that, although at baseline condition (e.g. technology level =0), consumer's preference is not strong enough to suppress the cost, consumers' sensitivity toward the technology improvement is still remarkable. The negative mean-level net-utility may reflect the properties of the emerging technology which consumers don't want to take the risk of being the first adopters, especially in B2B world. But, managers should be confident on the impacts of technology improvement thus be consistent in the innovation to attract consumer's purchase intention. Furthermore, the estimation of information discount factor δ tells us that, when evaluating the technology evolution, consumers are far from treating all available information equally but place relatively more weight on the more recent information. This suggests that managers can strategically alter consumer's decision process by altering the way of how to publicize the information.

Second, based on our policy simulation findings, firms can know what strategy will be the better choice under different scenarios. Moreover, to be more specific on how the result can aid firm's decision making, we draw a managerial decision making diagram to show it (Figure 16). The findings in our policy simulation study shows that, each policy has both strength and weakness from managerial point of view in continuously developing high-tech B2B market. For example, the "leaping" pace technology evolution and "reducing efforts of signing contract (e.g. reducing non-physical cost)" can incite consumers to make purchase, but consumers are more likely to sign a shorter contract, e.g. relatively less profitable. In fact, both the "leaping pace" technology evolution and the "reducing non-physical cost" strategies are very common in high-tech B2C market. The reason is that, in B2C world, consumers only decide on whether to buy or not which also determines firms' profit, therefore these two strategies perfectly matches firms' need of inciting consumers to buy. We can easily find examples in B2C world that belongs to the domain of these two strategies, such as Apple who announces a new model of product only once every year and keeps silence in between. Although consumers don't have "contract efforts" in B2C world, they need to spend "searching and comparing etc." efforts on their decisions which also belongs to the "non-physical cost". The phenomena of opening new stores, providing online services and exhibiting products for consumers to try and compare etc. show how the firms put in efforts on helping consumers to reduce the "non-physical cost". However, in B2B contractual market, service providers have to consider consumer's both decisions because 1) their profit is mainly from the contract length decision and 2) the contract length is conditional on consumer's purchase decision. Therefore, service providers make trade-off between the two policies within each scenario (Figure 16) depending on which consumer's decision they put on more weight in the current environment. For instance, if a service provider is relatively new and aims to seize a share in the market, they may be more willing to incite consumers to buy and try their new service. Then suggested by our findings, it will be better for the firm to select the "leaping evaluation pace" or "reduce contract signing effort" by offering consumers additional aid, because these two strategies can help elevate the consumers' purchase probability. On the other hand, if the firm is already very dominant in the market and are primarily concerned about profit (e.g. contract length), then it will be better to select the "jogging evaluation pace" or "increase the contract signing effort" by initiating strict and inflexible contract policy to the consumer. Then according to our results, although consumers are less inclined to make the purchase decision, if they still decide to buy or there is no other choices, consumers tend to sign a longer contract.

Finally, since our policy simulation also provides a quantitative evaluation of the consumer's decision changes under different scenarios, service providers can combine our results with their internal profit function to design the optimal strategy for profit maximization.

It is very straightforward to apply our decision making framework (Figure 16) to help service providers on selecting a better strategy. First, we only need the consumer's purchase data, e.g. purchase time and purchase quantity (e.g. contract length) and the news released frequency data to build the consumer dynamic model. If the news-count data is not available, the model can also be flexibly revised for other type of "count" data representing technology evolution. Next, we can perform the policy simulation by putting in different types of the policies, such as the pattern of announcing technology evolution and the possible degree of "contract effort" decrease/increase from service provider's side, and quantitatively compute the impact of each type of policy on consumers' two decisions: whether to buy and how long to buy. Then we can entail the policy simulation findings to optimize service provider's strategy plan and potential profitability.

Conclusion and Future Directions

In this study, we develop a new dynamic modeling approach to understand the impact of technology evolution on consumer's purchasing behaviors. In the model, we account for both consumer heterogeneity and consumer's forward-looking behaviors. Our model extends the current dynamic model by introducing consumer's continuous decision into the dynamic decision process. This study makes significant contribution to the marketing literatures from both methodological and substantive perspectives. From the modeling perspective, we develop an estimable dynamic model to understand how consumers make the purchase decisions in a turbulent environment where the technology is continuously developing. From the substantive domain, we provide insights into consumer's decision process under different scenario that can aid in manager's decision making. We show that, consumer's sensitivity toward the technology evolution is remarkable thus improving technology is an effective strategy on inciting consumer to make purchase. We address that, at current stage of the high-tech service, consumers still rely more on the most recent information to evaluate the technology evolution. We also empirically demonstrate that, both the pattern of technology improvement and the efforts of signing a contract can affect consumer's forward-looking behaviors thus the consumer's final decisions on both purchasing time and contract length.

Finally, our study still has several development spaces for future research. First, we only have the news frequencies in our data to help us quantify the technology evolution function (Equation (2)). If more information were available, we may be able to incorporate them into the technology evolution function thus better quantify the technology improvement. Second, we assume a linear relationship in the technology evolution function. Although the technology evolution model setup is still applicable in our study context after discussing and confirming with the managers, it may not be applicable to other study context. It would be a good avenue for future research to select a more flexible relationship in the technology evolution function, such as log-linear or multi-nomial etc.. Last, we assume that the news-count follows a Possion distribution in our study. Although Poisson distribution is a well-utilized discrete distribution for “count” data, its lack of flexibility in many applications is also well-documented, such as the over dispersion problem. Future research can consider select a more flexible distribution to better quantify the news-count distribution, for example, the negative binomial distribution etc.

Tables and Figures

Figure 1. Modeling Framework of Consumers' Dynamic Adoption Decision Process

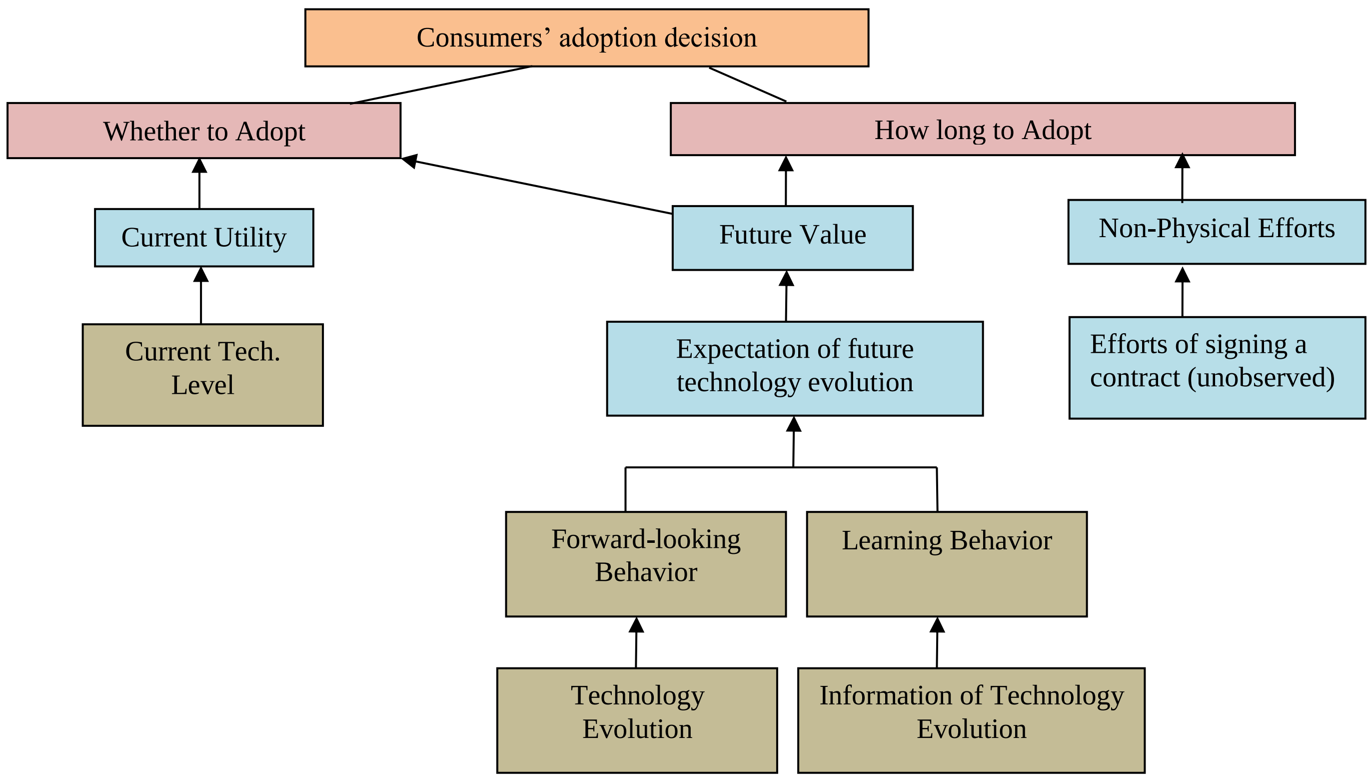

Figure 2. Graphical illustration of the learning model

Consumer's prior belief about λ at time (t)	$\lambda \| I_{it} \sim Gamma\ (\alpha_{it}, \beta_{it})$
During time (t), consumers receive new set information IOT_t	$P(IOT_t) \sim Possion(\lambda)$
Combine the prior opinion and the new information set, the consumers form the posterior belief about the λ at time (t)	$\lambda \| I^*_{it} \sim Gamm\ (\alpha^*_{it}, \beta^*_{it})$ $\begin{Bmatrix} \alpha^*_{it} = \alpha_{it} + IOT_t \\ \beta^*_{it} = \beta_{it} + 1 \end{Bmatrix}$

Figure 3. Graphical illustration of the discounted information process

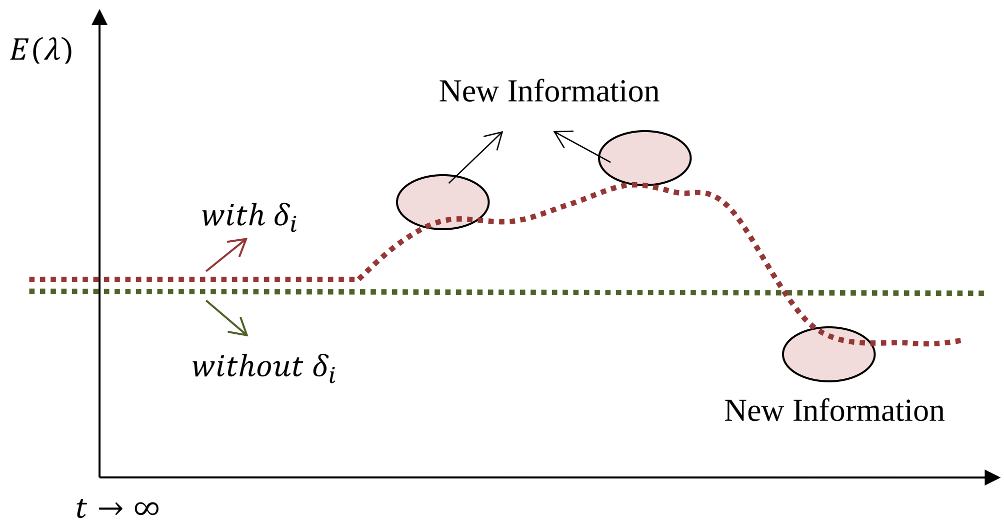

Figure 4. Graphical illustration of idea of dynamic decision model framework

$$\max_{\{CL_{i1}(t_{i1}),CL_{i2}(t_{i2}),\ldots\ldots CL_{i\infty}(t_{i\infty})\}} E\left\{\sum_{k=1}^{\infty}\left[\sum_{\tau=t_{ik}}^{t_{ik}+CL_{ik}}\left(\rho^{\tau} * u_{it_k}\left(Tech_{t_k}\right)\right) - EC_i\right]\right\} - Eqn\ (8)$$

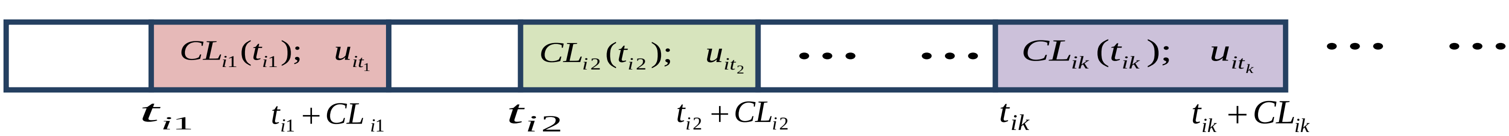

Figure 5. Graphically show the logical link between 6 parts of model setup

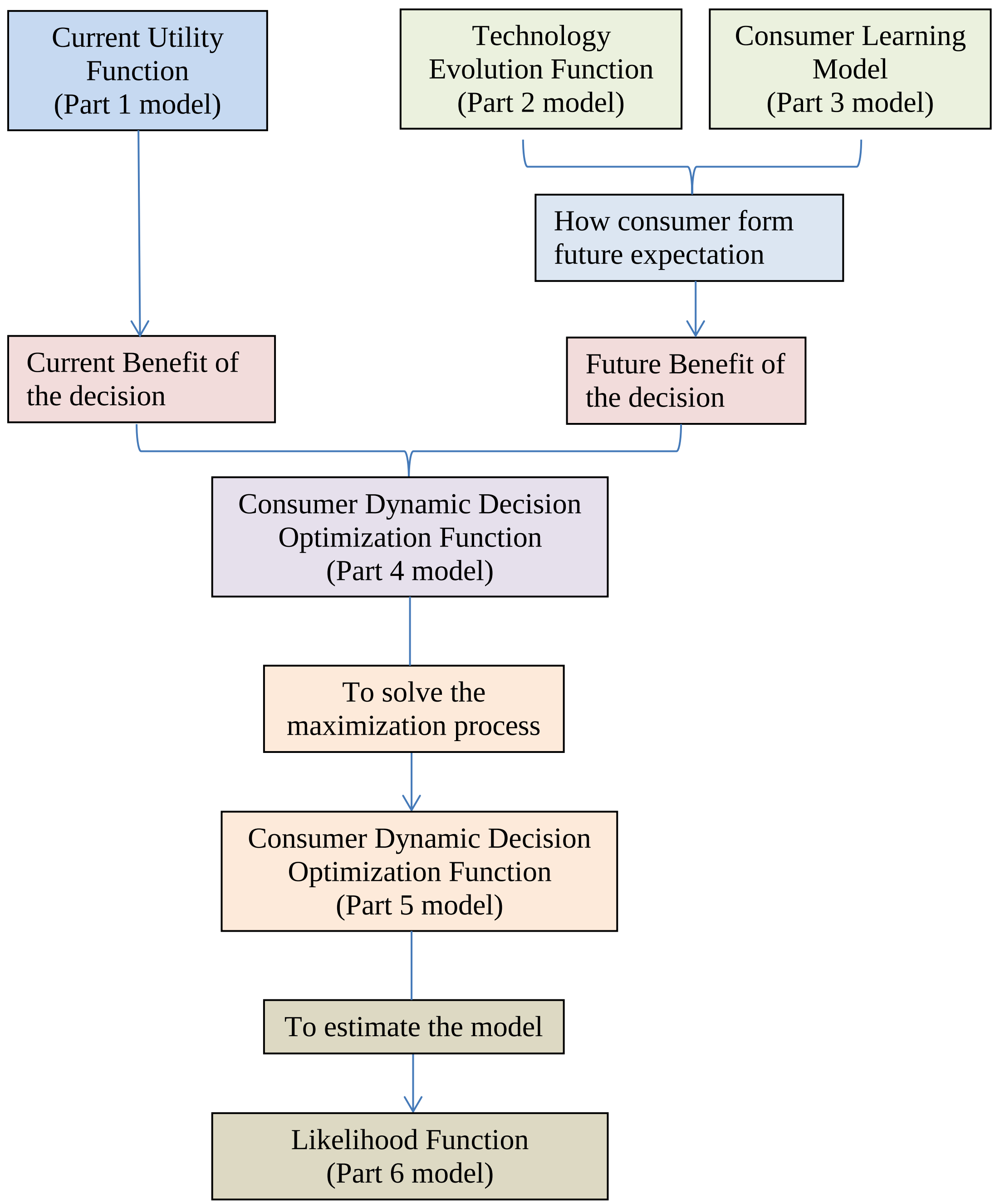

Figure 6. Graphically illustration of the parameter estimation process

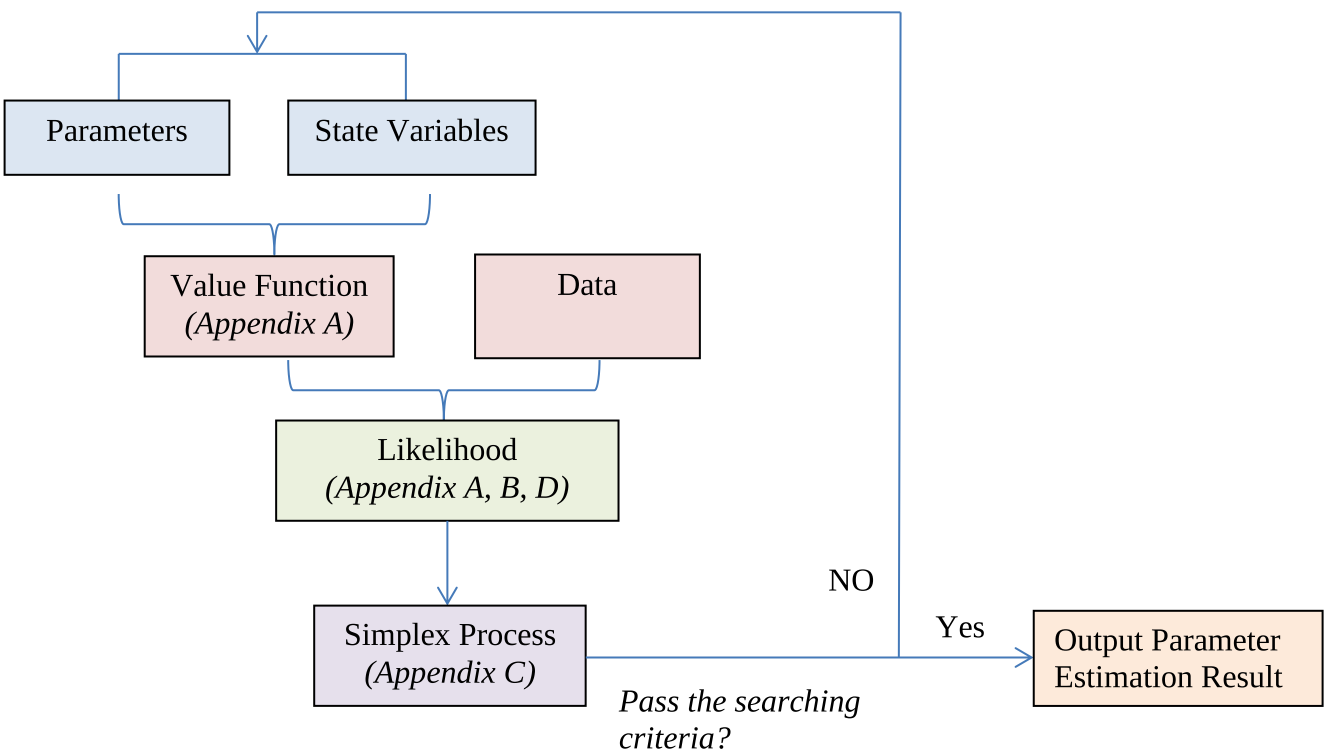

Figure 7. Graphical illustration of the monthly basis contract length and news-count per month

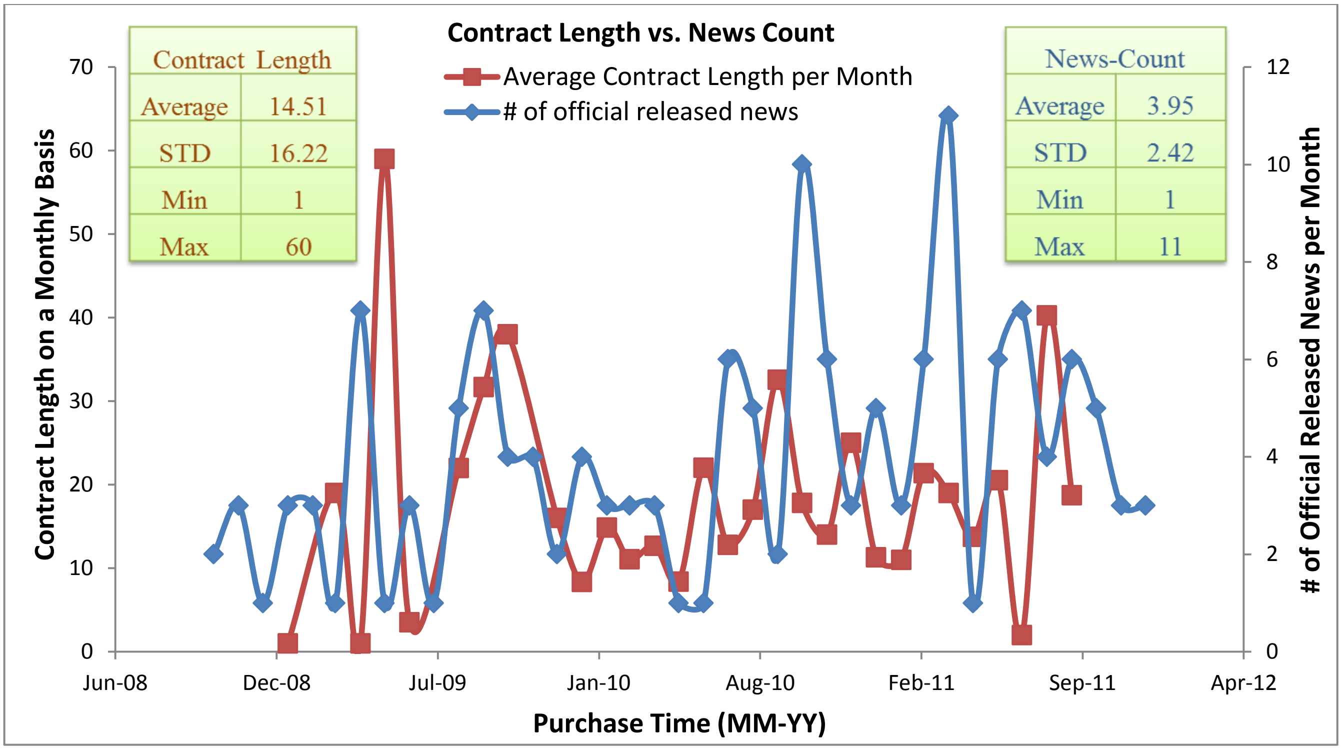

Figure 8. Comparison the patterns of consumer's value between different levels of overall technology

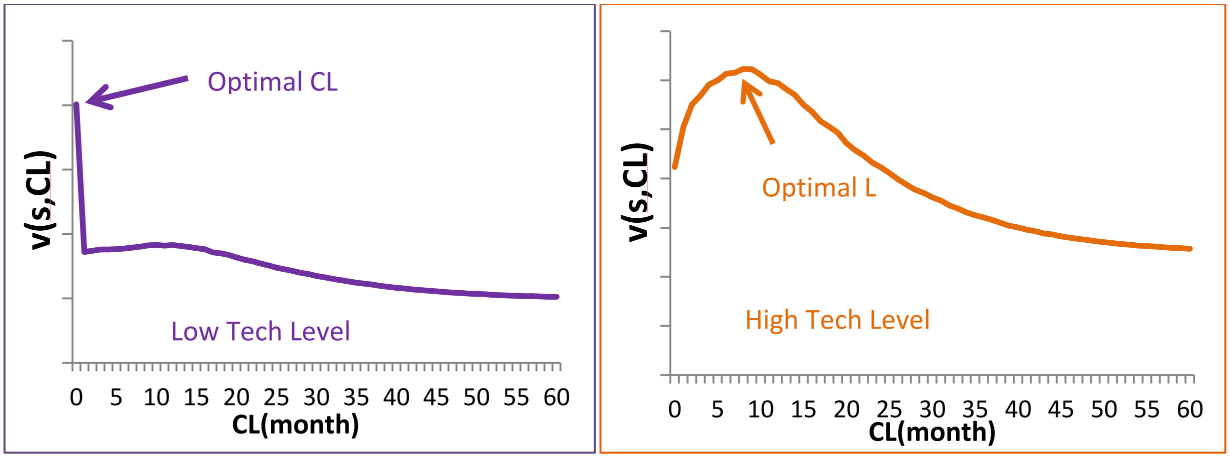

Figure 9. Comparison the patterns of consumer's value between different levels of technology evolution page

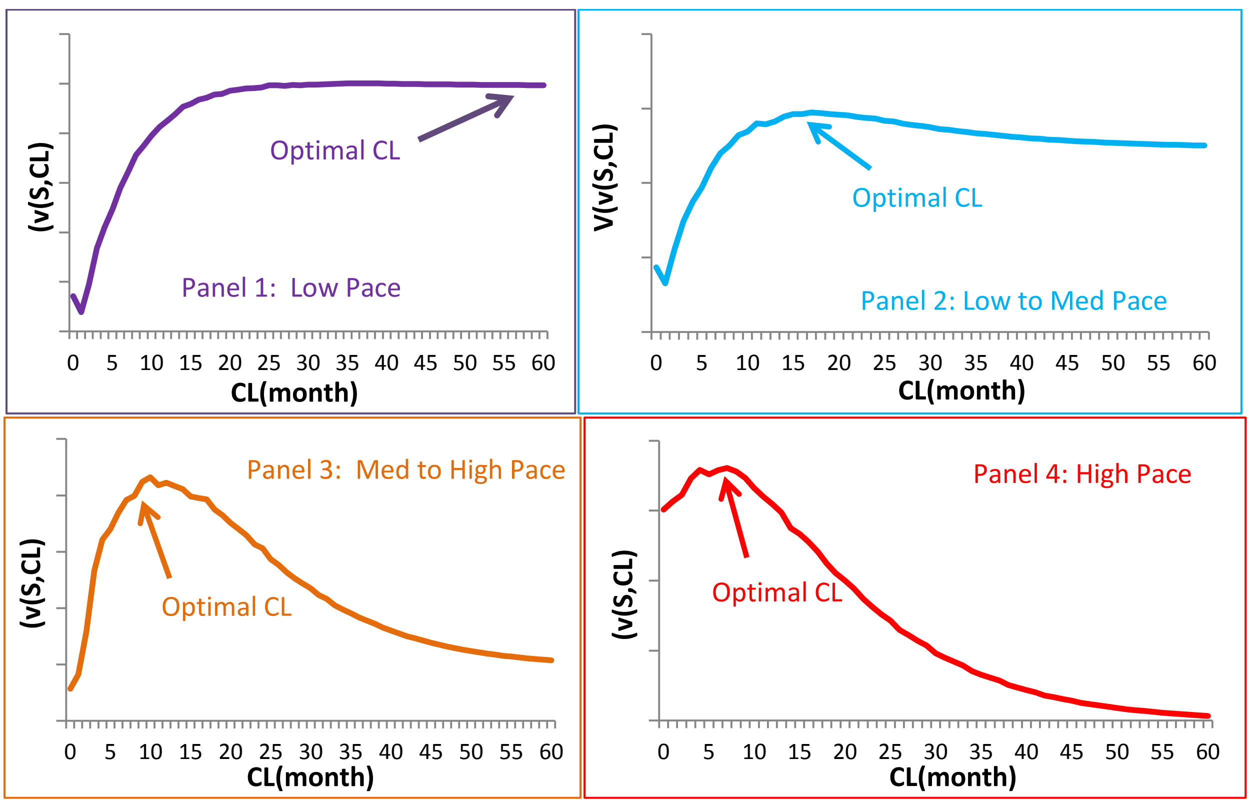

Figure 10. Relationship between cloud adopter's value and two focal state variables

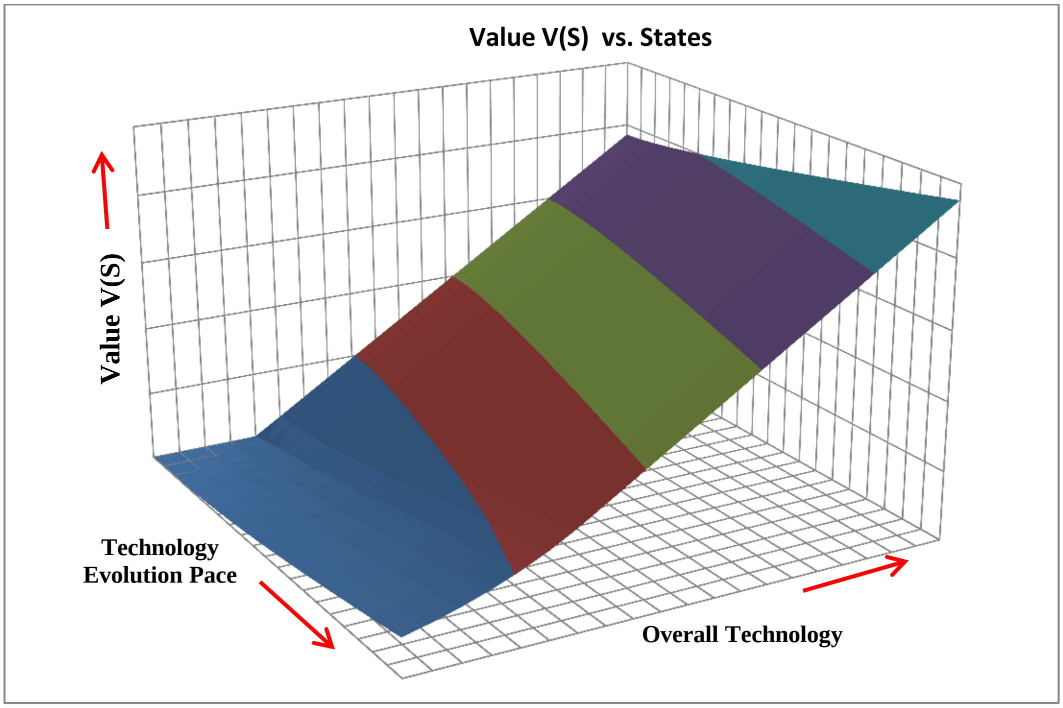

Figure 11. Relationship between contract length decision and two focal state variables

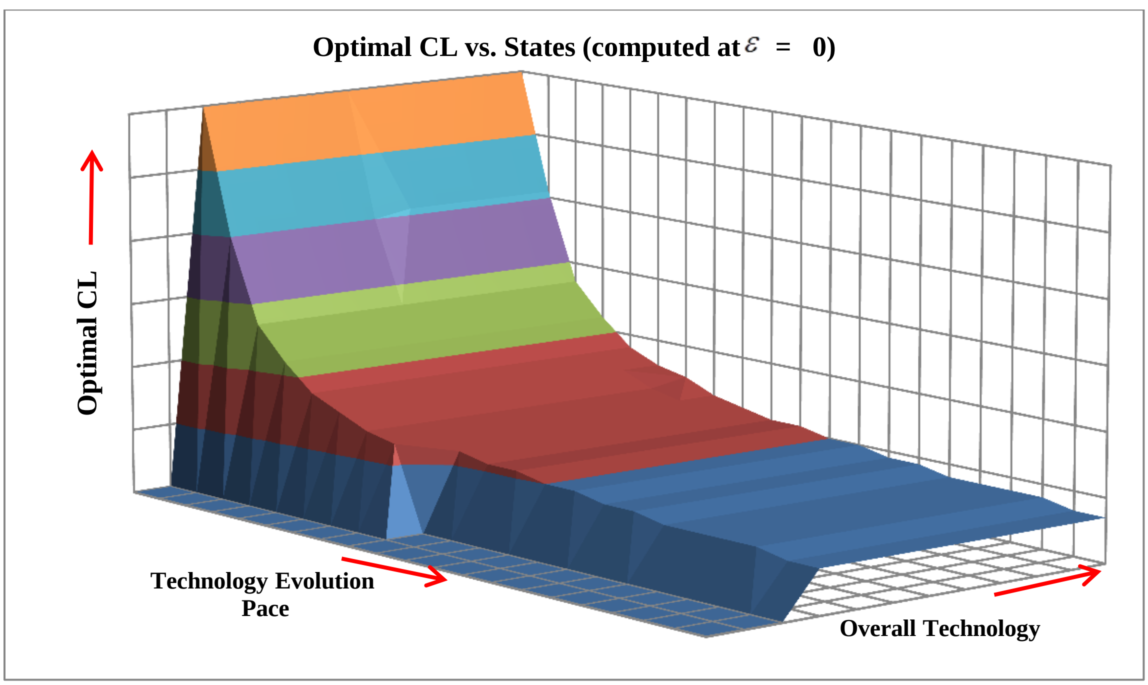

Figure 12. Comparison of optimal decision time between jogging and leaping pace

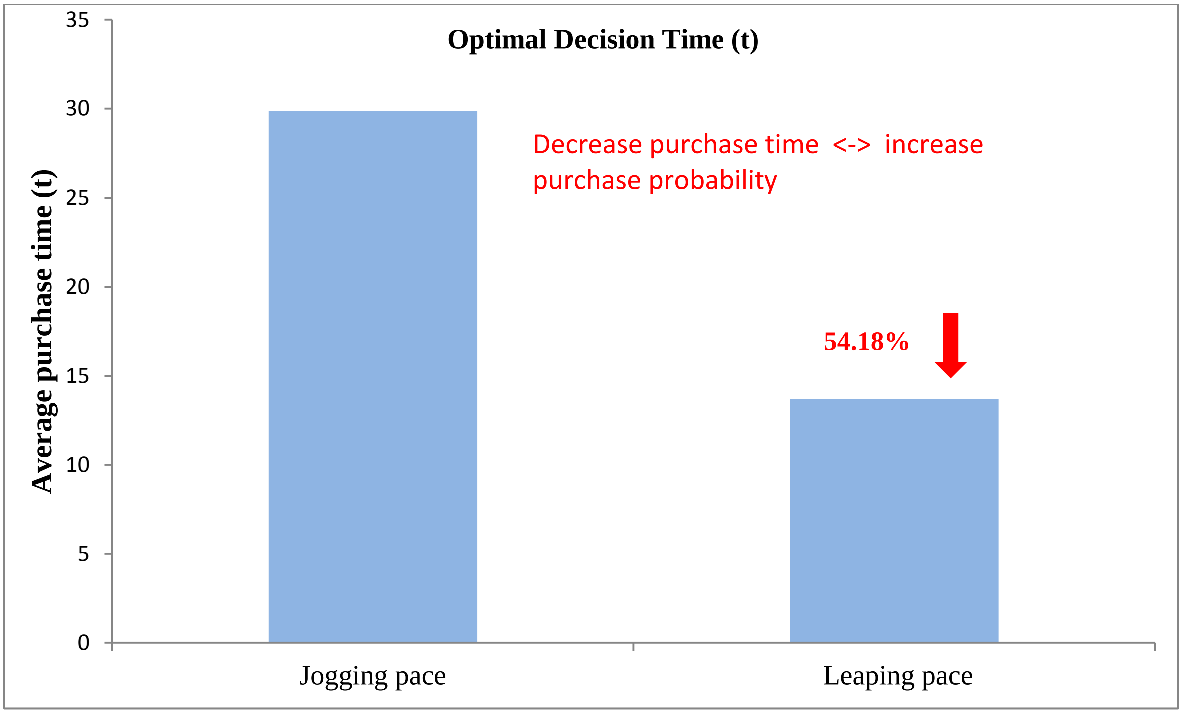

Figure 13. Comparison of optimal contract length decision between jogging and leaping pace

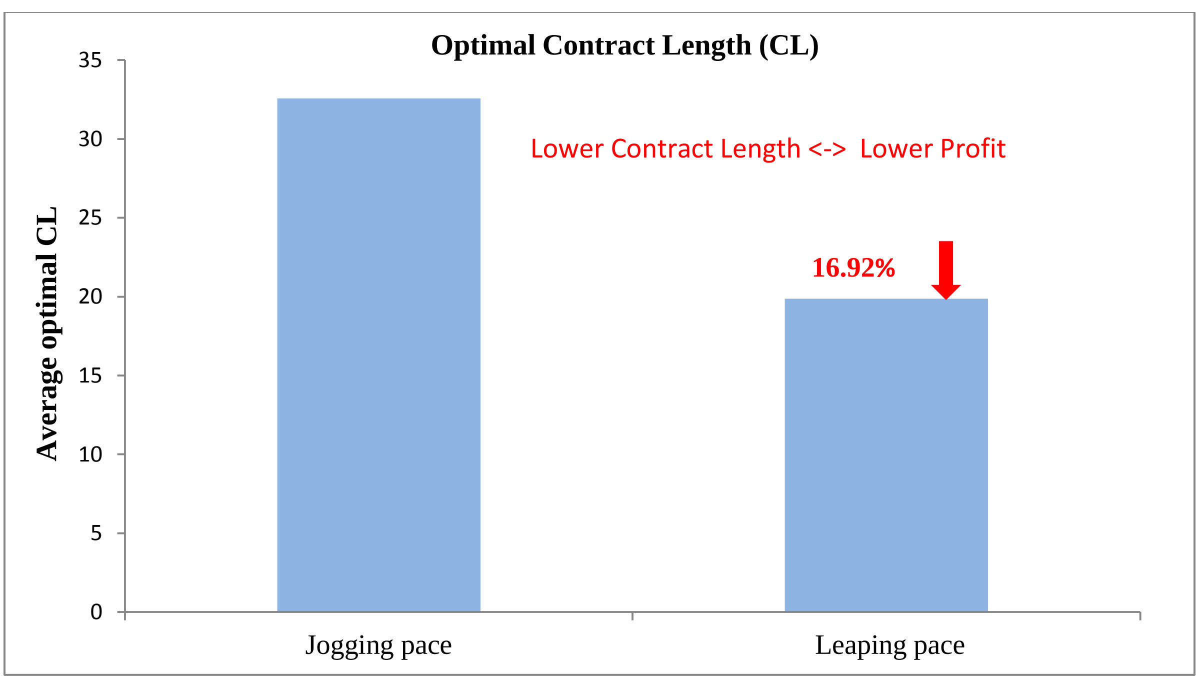

Figure 14. Comparison of optimal decision time between two levels of “efforts of signing contract”

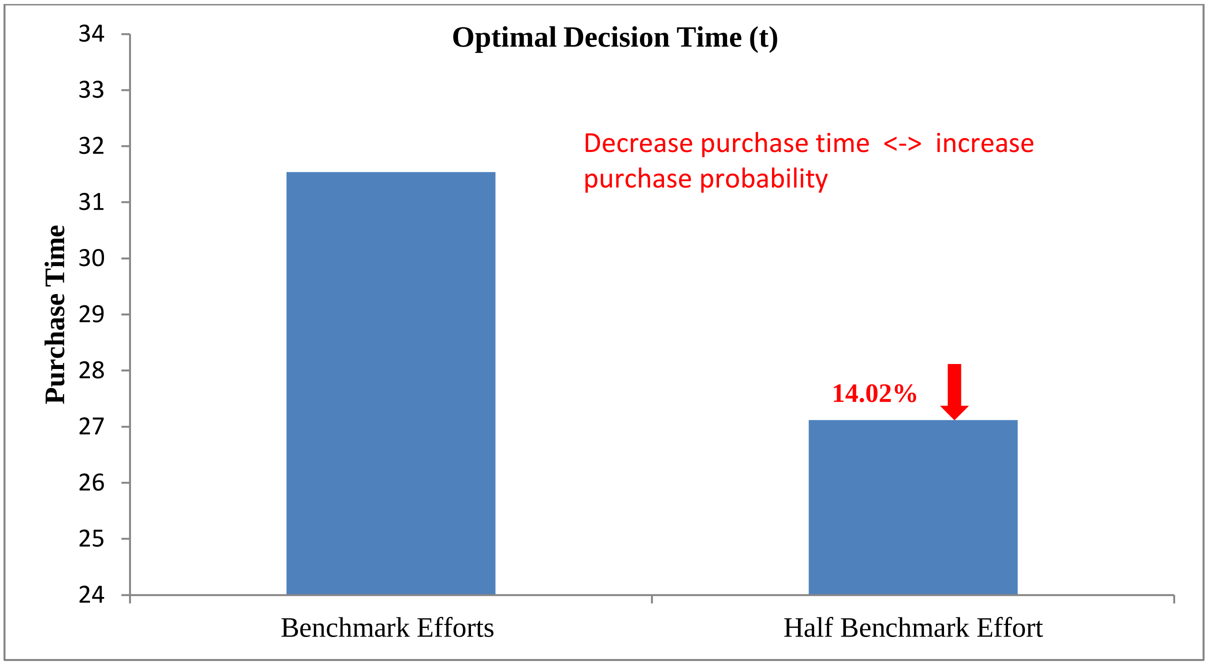

Figure 15. Comparison of optimal contract length decision between two levels of “efforts of signing contract”

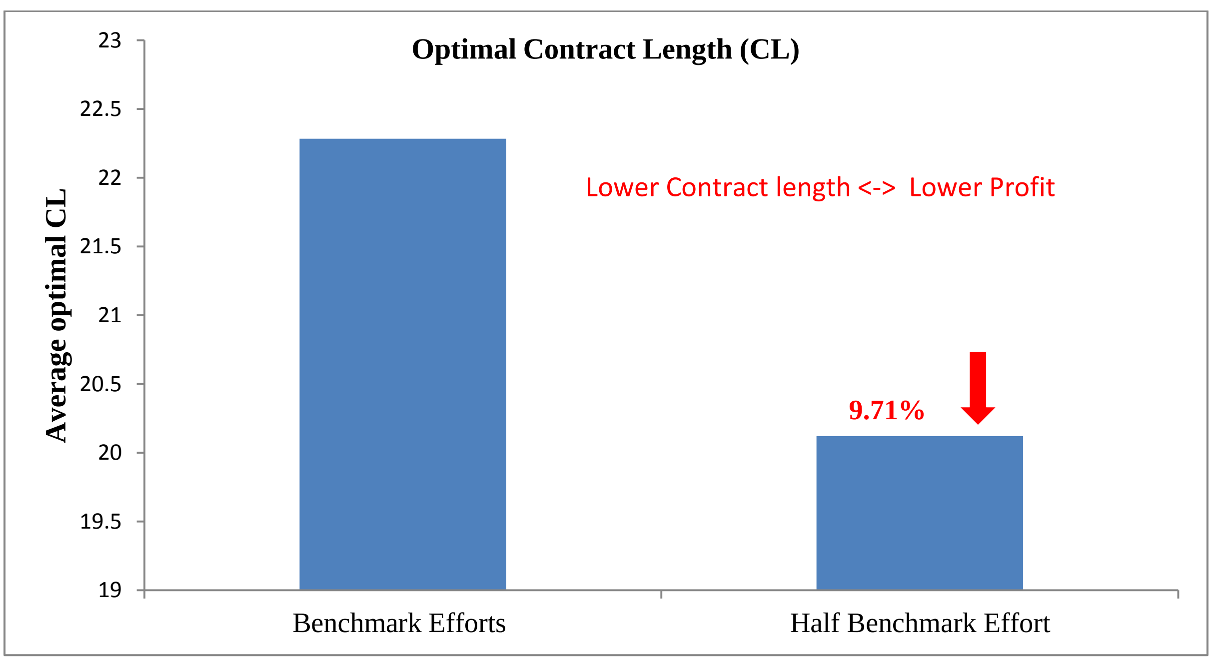

Figure 16. A Managerial Decision Making Diagram of Consumer's Dynamic Purchasing Behavior

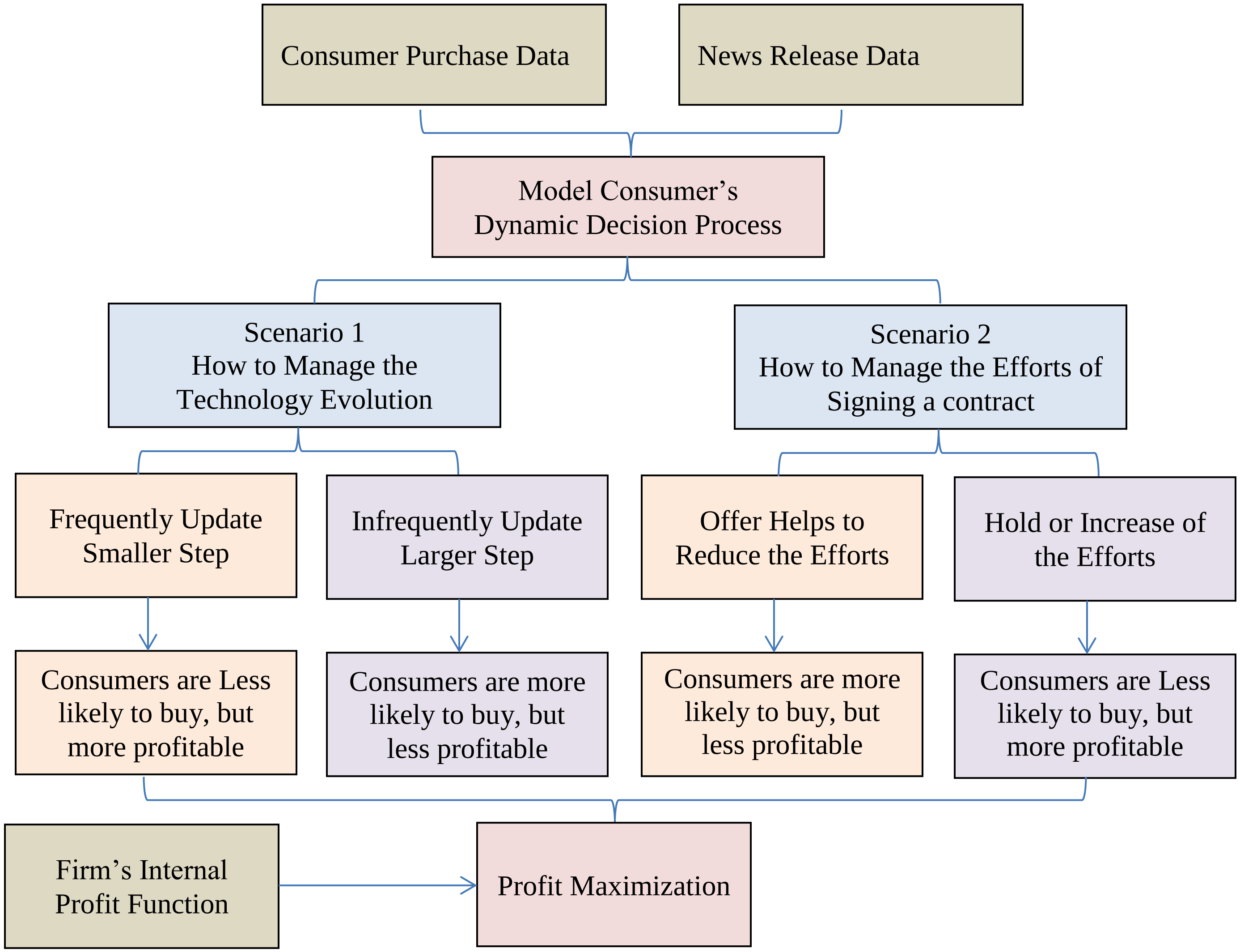

Table 1. Relative Contribution of this Study

Consumer Decision Process Studies	Offers in sights for consumer's choice between alternative?	Account for consumer's forward looking dynamic behavior?	Consider consumer's learning behavior?	Consider both binary and continuous decision?	Consider factors invoking forward looking behavior other than price?
Guadagni &Little (1983)	√				
Erdem & Keane (1996)	√	√	√		√
Gönül & Srinivasan (1996)	√	√			
Ackerberg (2001, 2003)	√	√	√		√
Erdem (2003)	√	√	√		√
Song & Chintagunta (2003)	√	√			
Sun, et al. (2003)	√	√			
Lewis (2004)	√	√			√
Sun, et al. (2005)	√	√			
Erdem et al., (2005)	√	√	√		√
Hendel and Nevo (2006)	√	√			
Kim (2006)	√	√			√
Nair (2007) *(equilibrium)*	√	√			
Erdem et al., (2008)	√	√	√		√
Gordon (2009)	√	√			√
Sriram, et al. (2010)	√	√			

Nosal (2011)	√	√			√
Gowrisankaran, et al. (2011)	√	√			√
Schiraldi, 2011	√	√			
Gowrisankaran and Rysman (2012) *(equilibrium)*	√	√			√
Ryan and Tucker (2012)	√	√			√
Shcherbakov, 2009	√	√			√
Yang and Ching (2013)	√	√			√
Current Study	√	√	√	√	√

Table 2. Simulation Results

Parameters			True Value	Estimation	SE
Mean-level Net-Utility	γ_{0i}	$\overline{\gamma}_0$	**-1**	**-0.9852**	**0.0246**
		σ_{γ_0}	**0.1**	**0.1213**	**0.0359**
Technology Coefficient	γ_{1i}	$\overline{\gamma}_1$	**1.5**	**1.5345**	**0.3157**
		σ_{γ_1}	**0.1**	**0.0982**	**0.0348**
Effort of Contract	EC_i	$\overline{EC}$	**10**	**11.753**	**1.7248**
		σ_{EC}	**1**	**1.2001**	**0.4948**
Information Discount	δ		**0.95**	**0.9392**	**0.0941**
Technology Evolution Error	ξ_t	σ_ξ	**1**	**0.9893**	**0.1339**

Table 3. Model Comparison Results

	BIC
OLS	**12596**
Tobit	**2809**
Proposed Model	**1131**

Table 4. Parameter Estimations

Parameters			**Estimation**	**SE**
Mean-level Net-Utility	γ_{0i}	$\overline{\gamma_0}$	**-1.7711**	**0.0396**
		σ_{γ_0}	**0.0576**	**0.0177**
Technology Coefficient	γ_{1i}	$\overline{\gamma_1}$	**0.0139**	**0.0055**
		σ_{γ_1}	**2.2824E-04**	**9.2499E-05**
Effort of Contract	EC_i	$\overline{EC}$	**14.9406**	**2.0719**
		σ_{EC}	**1.0472**	**0.5136**
Information Discount	δ		**0.8291**	**0.0822**
Technology Evolution Error	ξ_t	σ_{ξ}	**1.2764**	**0.3004**
Initial Parameter of Belief Updating		α_0	**Fix at 0**	**N.A.**
Initial Parameter of Belief Updating		β_0	**Fix at 0**	**N.A.**
Utility Time Discount Factor		ρ	**Fix at 0.90**	

Appendix A – Computation of the value function

In this section, we describe the details of how we empirically solve the value function. To simplify the expression, we neglect the consumer (i) subscripts. As we mentioned previously, in our model, the consumers solve an infinite-horizon stationary Markov decision process in order to optimize their current contract length decision. Ideally, we should iteratively compute the value function until it converges at all decision time points. However, to reduce the computation burden, we follow the "backsolving" method introduced by Erdem and Keane (2003). We select a terminal period T and assume that: when the transit period beyond the terminal period T, consumer's value become zero at all state points (Erdem and Keane, 2003), e.g. $V(S, t \geq T) = 0$. The underlying idea of this method is that, when consumers make current purchase decision, the utilities at far away future is almost discounted to zero although consumers have forward looking behavior.

After we assume $V(S, t \geq T) = 0$, we can backwardly calculate the $V(S, T-1), V(S, T-2) \dots V(S, 1)$ using the value function in Equation (10). As addressed in Erdem and Keane's study (2003), in the "backsolving" computation procedure, we need to select an enough large terminal period T in order to ensure that the value functions become stable. We follow the criteria specified in Erdem and Keane's study (1996) to find the terminal period T, e.g. if we fix the utility discount factor at 0.9, a terminal period $T \geq 100$ is enough for finding a converged value function. In our study, we select a terminal period $T = 150$ with the utility discount factor fixed at 0.9.

It is worth mentioning that, in our dynamic programming model, we have in total four state variables ($Tech_t, \alpha_{it}, \beta_{it}, \varepsilon_{it}$). All of the state variables are continuous. Therefore, as pointed out in Erdem and Keane's study (2003) that, we can't solve value function for all potential state points. Therefore, we need to use approximation method (Erdem and Keane, 2003). First, we define "region" and "grid" for each state variable. Then, we can compute the exact result of value function at each "grid" points (Note: the grid points are defined by all state variables). Next, if the states of the value function that we need don't fall on the "grid" points, we will use the Kernel non-parametric regression (Altman, 1992) to find the approximate solution.

Then, we want to reduce the dimension of the state variables to both simplify the value function computation process and ease the computation burden. We can use a simple example to show why reducing the dimension of the state variables offers us the computation advantages. Suppose we select 10 grids for each state for calculating the value function, and then we need to compute value function for 10^4 grid points occupying a 4-dimensional matrix. If we can reduce the dimension of state variables, the number of value function that we need to compute can be reduced exponentially. We finally are able to reduce the dimension of state variables to 2-dimension and the process is shown as follow:

First, we can integrate out the state variable of "ε" in the value function because it is an "iid" random component (Equation (1)), which means that "ε_t" doesn't depend on the value of "$\varepsilon_{t-1}, \varepsilon_{t-2} \dots \varepsilon_1$". Then the value function shown in Equation (10) can be re-defined as follow:

$$V^*(\tilde{S}_t^*) = \int_\varepsilon V(\tilde{S}_t)\, d\varepsilon$$

$$= \int_\varepsilon \max_{CL} \left\{ \sum_{j=0}^{CL} \rho^j u_t(\tilde{S}_t, CL) + \rho^{CL+1} \int_{\check{S}_{t+CL+1}} V(\tilde{S}_{t+CL+1}) * dF(\tilde{S}_{t+CL+1} | \tilde{S}_t, CL) - EC_i(CL > 0) \right\} d\varepsilon$$

$$= \int_\varepsilon \max_{CL} \left\{ \sum_{j=0}^{CL} \rho^j u_t(\tilde{S}_t, CL) + \rho^{CL+1} \int_{\tilde{S}^*_{t+CL+1}} V^*(\tilde{S}^*_{t+CL+1}) * dF(\tilde{S}^*_{t+CL+1} | \tilde{S}^*_t, CL) - EC_i(CL > 0) \right\} d\varepsilon$$

$$--(A1)$$

Where $\tilde{S}_t^* = (Tech_t, \alpha_t, \beta_t)$ and $\tilde{S}^*_{t+CL+1} = (Tech_{t+CL+1}, \alpha_{t+CL+1}, \beta_{t+CL+1})$; and we now can reduce the state variable space from 4-dimensional to 3-dimensional.

Second, we know that, unlike the state variables of "$Tech_t$" and "α_t", the parameter "β_t" is a deterministic term and can be exactly computed by time because it has no relationship with the unobserved or random components of $(IOT_{t+1} \ldots IOT_{t+CL};\ \lambda;\ \xi_{t+1} \ldots \xi_{t+CL})$. Therefore, we can create a new state variable by combining both "α_t" and "β_t", e.g. "$\frac{\alpha}{\beta}_t$". This new state variable is not a pure mathematical expression but have physical meaning. Recalled that in consumer learning model (Equation (4)); "$\frac{\alpha}{\beta}_t$" indicates consumer's mean-level perception about "λ", while "λ" is the true-mean level of news released frequency. Therefore, "$\frac{\alpha}{\beta}_t$" actually suggests consumer's perception on the mean-level of the technology evolution pace. It is worth mentioning that, reducing state space by replacing "α_t" and "β_t" with "$\frac{\alpha}{\beta}_t$" makes our value function non-stationary. The reason is that, "α_t" and "β_t" can create both "$\frac{\alpha}{\beta}_t$" and "$\frac{\alpha}{\beta^2}_t$" (e.g. consumer's uncertainty about "λ") influencing the value function at different period. Although we reduce the state space by introducing "$\frac{\alpha}{\beta}_t$" as the state variable, we lose control on the "$\frac{\alpha}{\beta^2}_t$" thus both our transition probability and our value function becomes t-specific, e.g. $F_t\left(\tilde{S}^*_{t+CL+1} \middle| \tilde{S}^*_t, CL\right), V_t\left(\tilde{S}_t\right)\ and\ V^*_t\left(\tilde{S}^*_t\right)$ instead of $F\left(\tilde{S}^*_{t+CL+1} \middle| \tilde{S}^*_t, CL\right), V\left(\tilde{S}_t\right)\ and\ V^*\left(\tilde{S}^*_t\right)$ respectively. Begin from here, we denote all value function and transition probability as t-specific at a given the state space.

Last, we explain how we compute the maximization process in the value function. Let's first define another format of value function without the maximization notation, e.g.

$$v_t\left(\tilde{S}_t, CL\right) = \sum_{j=0}^{CL} \rho^j u_t\left(\tilde{S}_t, CL\right) + \rho^{CL+1} \int_{\tilde{S}^*_{t+CL+1}} V^*_t\left(\tilde{S}^*_{t+CL+1}\right) * dF_t\left(\tilde{S}^*_{t+CL+1} \middle| \tilde{S}^*_t, CL\right) - EC_i(CL > 0) \quad --(A2)$$

Note, the difference between the value function $v_t(\tilde{S}_t, CL)$ defined here and those we defined previously (either $V_t^*(\tilde{S}_t^*)$ or or $V_t(\tilde{S}_t)$) is that, $v_t(\tilde{S}_t, CL)$ is the function of both state variables $\left(Tech_t, \frac{\alpha}{\beta_t}, \varepsilon_t\right)$ and decision variable (contract length, $CL_t \geq 0$). In discrete dynamic model, consumer's decision space only contains two values: 0 = not purchase and 1 = purchase. Therefore, we can identify consumer's optimal decision only by compare the value function between $v_t(\tilde{S}_t, 0)$ and $v_t(\tilde{S}_t, 1)$. However, in our study, the decision variable of contract length is a continuous variable. Therefore, unlike the discrete dynamic model, we will need to iteratively searching for the optimal decision point, and we use "Golden Search" method to reach the optimal decision at a given state, e.g. $CL_{t(optimal)}(\tilde{S}_t) = CL_t |\max\left(v_t(\tilde{S}_t, CL_t)\right)$ and the associated value is $V_t^*(\tilde{S}_t^*) = \int_\varepsilon \ V_t(\tilde{S}_t) \ = \int_\varepsilon \ v_t(\tilde{S}_t, CL_{t(optimal)})$. Please note that the "Golden Search" method can only help us to find the max $\left(v_t(\tilde{S}_t, CL)\right)$ for $CL_t \geq 1$ because it is only applicable to finding the extremum of a strictly unimodal function. Therefore, we also need to compare the "Golden Search" result with the $v_t(\tilde{S}_t, 0)$ to find the value function: $V_t(\tilde{S}_t) = \max(v_t(\tilde{S}_t, 0), max(v_t(\tilde{S}_t, CL)|CL_t \geq 1))$.

Now, we can simulate the value function as follow:

Step 1: Define the "region" and "grid" for the state variables. Now we have three groups state variables ($Tech,\ \frac{\alpha}{\beta}, \varepsilon$) and ε can be separately integrated out like the unobserved components. Therefore, we only need to define "region" and "grid" for ($Tech,\ \frac{\alpha}{\beta}$). Because of the existence of unobserved components, we can't identify the exact "region" for ($Tech,\ \frac{\alpha}{\beta}$). The ideal region for "$Tech$" and "$\frac{\alpha}{\beta}$" can be $(-\infty, \infty)$ and $(0, \infty)$. But according to their physical meanings, e.g. "$Tech$" is the overall technology level and "$\frac{\alpha}{\beta}$" is the technology evolution pace, we can approximately use the real data to help us define the “region” of them in the value function computation. Next, suppose that we select M “grid” for both state variables. We can discretize the “region” of the state variables by selecting “M” evenly distributed points, e.g. $(Tech_1 \ldots Tech_M)$ and $(\frac{\alpha}{\beta}_1 \ldots \frac{\alpha}{\beta}_M)$.

Step 2: Assume terminal period T=150, and $V_t^*(\tilde{S}_t^*, T \geq 150) = 0$. Please note that, the value function we used here is $V_t^*(\tilde{S}_t^*); where\ \tilde{S}_t^* = (Tech_t,\ \frac{\alpha}{\beta}_t)$.

Step 3: At T=149, use equation (A2), we can compute the $v_t(\tilde{S}_t, CL_t)$. Here, we also need to specify the “region” of CL that we use to compute the value function. Ideally, the region of CL_t should be $(0, \infty)$. To reduce the computation burden, we will use the max(CL_t) in the data as the upper limit for computing the value function of $v_t(\tilde{S}_t, CL_t)$.

Step 4: We use the iterative computation of $v_t(\tilde{S}_t, CL_t)$ to help us identify the $V_t^*(\tilde{S}_t^*)$ for "$T = 149$". The detailed steps are: first, we find the max $\left(v_t(\tilde{S}_t, CL_t)\right)$ for $CL \in [1, \max(CL_t)]$ using “Goldern Search” method; then we compare the result with $v_t(\tilde{S}_t, CL_t = 0)$ to get $V_t(\tilde{S}_t)$; finally, we can update the value function for $V_t^*(\tilde{S}_t^*, T = 149)$ using $V_t^*(\tilde{S}_t^*) = \int_\varepsilon V_t(\tilde{S}_t) d\varepsilon$ for all $M * M$ combination of grid points. For any value that the “region” of the state variables doesn’t belong to, we can use the “Kernel non-parametric regression” to find the approximated solution. At the same time, we can also find the consumer’s decision rule at any given state, e.g. $CL_{t(optimal)}(\tilde{S}_t)$.

Step 5: Finalize the solution of value function by computing both $V_t^*(\tilde{S}_t^*)$ and $CL_{t(optimal)}(\tilde{S}_t)$ to the time period of $T = 1$, e.g. iteratively redo step 3 and 4 for $T = 148, T = 147 \ldots . T = 1$.

Appendix B – Empirical finding "ε" at a given "CL"

This appendix describes how we empirically compute the "ε" at a given "CL" since we don't have close-form function for $\varepsilon = f(CL)$. We empirically observed that there is a monotonic increasing relationship between "ε" and "$CL|CL > 0$", which means that, as "ε" increases, the "CL" also increases for all "CL" when "$CL > 0$". Therefore, we will use the bisection method to iteratively find the solution of "ε" at a given "CL". The detailed steps are described as follow:

Step 1: We need to compute the Value Function, e.g. $V^*(\tilde{S}^*)$ at a given set of parameters, e.g. $\{\gamma_{0i}, \gamma_{1i}, EC_i, \delta, \sigma_\varepsilon, \sigma_\xi, \alpha_{i0}, \beta_{i0}\}$ following the steps shown in Appendix (A).

Step 2: Assume an initial upper and lower bound of "ε", e.g. "ε_l and ε_u" for computing the corresponding "CL", e.g. e.g. "CL_l and CL_u". For example, we can select the "$\varepsilon_l = -5$ and $\varepsilon_u = 5$".

Step 3: We use the finalized $V^*(\tilde{S}^*)$ to compute the "CL_l and CL_u" corresponding to "ε_l and ε_u" respectively. The process is very similar as the "step 3 and 4" defined in Equation (A2): at a given "ε", first, we find the $\max\left(v(\tilde{S}_t, CL)\right)$ for $CL \in [1, \max(CL)]$ using "Goldern Search" method; then we compare the result with $v(\tilde{S}_t, CL = 0)$ to find the $CL_{optimal}(\tilde{S}_t)$ where $CL_{optimal}(\tilde{S}_t) = CL|\max\left(v(\tilde{S}_t, 0), \max\left(v(\tilde{S}_t, CL)|CL \geq 1\right)\right)$.

Step 4: We can use the bisection method to find the exact "ε" for "$CL > 0$". If "$CL = 0$", we can find the upper limit of "ε_{ul}", e.g. "$CL = 0$" for all values of "$\varepsilon \leq \varepsilon_{ul}$"

Appendix C – Simplex Method for finding MLE

In this section, we describe the details of how we empirically solve the MLE using Nelder–Mead Simplex Method (Nelder and Mead, 1965), which is a well-established numeric method for searching for the extremum when the first-order derivative of the function can't be analytically solved. Since this method targets at the "minimum" of the non-linear function with multi-dimensional parameter space, we revise the process to find the "maximum" of "$likelihood$" function in our study. The detailed steps are described as follow:

Step 0: We select initial values for the m-dimension parameters in the likelihood function: e.g. $\tilde{\theta}^0 = \{\theta_1^0, \theta_2^0, \theta_3^0 \ldots . \theta_m^0\}$. Then for each individual parameter, we add in a certain step-length $\Delta\theta$ and keep other parameters unchanged so that we obtain the following vector space: $\{\tilde{\theta}^0, \tilde{\theta}^1, \tilde{\theta}^2, \ldots \tilde{\theta}^m\}$, which we name as the "simplex" (Nelder and Mead, 1965). To be clear of the notation, we use "i" to denote the set of parameters; and use "j" to denote values of each individual parameter.

Step 1: For each of the (m+1) set of parameters, we can compute the corresponding "$likelihood$". To simplify the exposition, we use "f" to denote the "$likelihood$". Therefore, we obtain "$\tilde{f} = \{f^0, f^1, \ldots f^m\}$" where "$f^i$" is the "$likelihood$" outcome of parameter set "$\tilde{\theta}^i$" where $i = 0 \ to \ m$.

Step 2: We find the maximum and minimum in "$\tilde{f}$". We denote that: $f^h = \max_i(f^i)$; $\tilde{\theta}^h = \tilde{\theta}^i | \max_i(f^i)$; $f^l = \min_i(f^i)$; $\tilde{\theta}^l = \tilde{\theta}^i | \min_i(f^i)$.

Step 3: We compute the centroid for each parameter, e.g. the $\bar{\theta}_j$ where $i \neq l$. We use $\overline{\Theta}$ to denote the centroid for all parameters: $\overline{\Theta} = \{\bar{\theta}_1, \bar{\theta}_2 \ldots \bar{\theta}_m\}$

Step 4: We will use three operators: reflection, contraction and expansion to replace one set of the parameters in the "simplex". The detailed steps are shown as below:

Step 5 (reflection): We compute the "reflection" as $\tilde{\theta}^r = (1+\alpha) * \overline{\Theta} - \alpha * \tilde{\theta}^l$ and the corresponding "$likelihood$" as "f^r". The "α" is an arbitrarily selected positive constant.

Step 5-1: If $f^l \leq f^r < f^h$, then we replace the "f^l" with "f^r" and the associated "$\tilde{\theta}^l$" with "$\tilde{\theta}^r$".

Step 5-2 (Expansion): If $f^r \geq f^h$, then we compute the "expansion" as $\tilde{\theta}^e = \gamma * \tilde{\theta}^r + (1-\gamma) * \overline{\Theta}$ and the corresponding "$likelihood$" as "f^e". The "γ" is an arbitrarily selected positive constant. We replace the "f^l" with "f^e" and the associated "$\tilde{\theta}^l$" with "$\tilde{\theta}^e$".

Step 5-3 (contraction): If $f^r < f^i \; for \; all \; i \neq l$, then we compute the "contraction" as $\tilde{\theta}^c = \beta * (centroid + \tilde{\theta}^l) + (1-\beta) * \overline{\Theta}$ and the corresponding "$likelihood$" as "f^c". The "β" is a number between 0 and 1. If $f^c > f^l$, then we replace the "f^l" with "f^c" and the associated "$\tilde{\theta}^l$" with "$\tilde{\theta}^c$". If $f^c \leq f^l$, then we compute a new (m+1) set of m-number parameters as: $\theta^h_{j_{new}} = \theta^h_j \; and \; \theta^i_{j_{new}} = \frac{(\theta^i_j + \theta^h_j)}{2} \; for \; all \; i \neq h$; and the associated new (m+1) results of "$likelihood$", e.g. "$\tilde{f}$".

Step 6: If the $std(\tilde{f}) > \varepsilon$, then we go to *Step 2*, otherwise, we finalize the results.

Appendix D – Computation of Jacobian in Likelihood Function

In this section, we describe the details of how we empirically compute the Jacobian, e.g. $\|J(\varepsilon_{it} \rightarrow CL_{it})\|$ in the Likelihood function. As described previously, there is no close-form solution for $\varepsilon = f(CL)$, we will need to empirically compute the Jacobian by iteration. Fundamentally, the Jacobian of $\|J(\varepsilon_{it} \rightarrow CL_{it})\|$ is the partial derivative of $\frac{\partial \varepsilon}{\partial CL}$. We follow fundamental concept of partial derivative, e.g. $\frac{\partial f}{\partial x_i}$ =$\lim_{h \rightarrow \infty} \frac{f(x_1, \dots x_i + h, \dots x_n) - f(x_1, \dots x_i, \dots x_n)}{h}$ and we derive the detailed steps to empirically find the solution of $\frac{\partial \varepsilon}{\partial CL}$ as follow:

Step 1: For a given set of parameters, we find the value function following the steps described in Appendix A.

Step 2: At a given number of CL and $CL + \Delta CL$, where ΔCL is a very small value, we empirically compute the corresponding ε and $\varepsilon + \Delta\varepsilon$ following the steps described in Appendix B.

Step 3: We can approximately compute the Jacobian as: $\frac{\partial \varepsilon}{\partial CL} = \frac{\Delta \varepsilon}{\Delta CL}$

Selected References

Ackerberg, Daniel A. (2001), "Empirically Distinguishing Informative and Prestige Effects of Advertising," *The RAND Journal of Economics*, 32(2), pp. 316-333

Ackerberg, Daniel A. (2003), "Advertising, Information, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 44(3), 1007-1040

Altman, N. S. (1992), "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician* 46(3), 175–185.

Bellman, Richard (1954), "The Theory of Dynamic Programming," *Bulletin of the American Mathematical Society*, 60, 503-516,

Bellman, Richard (1956), "Dynamic programming and Lagrange multipliers," *Proceedings of the National Academy of Sciences of the United States of America,* 42(10), 767-769.

Chintagunta, Pradeep K. (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, 12(2), 184-208.

Columbus, Louis (2013), "Predicting Enterprise Cloud Computing Growth," *http://www.forbes.com/sites/louiscolumbus/2013/09/04/predicting-enterprise-cloud-computing-growth/*

Erdem, Tülin and Michael P. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing* Science, 15(1), 1-20.

Erdem, Tülin (2003), "Brand and Quantity Choice Dynamics under Price Uncertainty," *Quantitative Marketing and Economics*, 1, 5–64

Erdem, Tülin, Michael P. Kean, T. S. Öncü and J. Strebel (2005), "Learning about computers: An analysis of information search and technology choice," *Quantitative Marketing and Economics*, 3(3) 207–246.

Erdem, Tülin, Michael P. Keane and Baohong Sun (2008), "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 27(6), 1111-1125

Gönül, Füsun and Kannan Srinivasan (1996), "Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model," *Marketing Science*, 15(3), 262-279.

Gordon, BR (2009), "A Dynamic Model of Consumer Replacement Cycles in the PC Processor Industry," *Marketing Science*, 28(5), 846–867.

Gowrisankaran, G., M. Park, and M. Rysman (2011), "Measuring Network Effects in a Dynamic Environment," Unpublished Manuscript, Boston University, 2011.

Gowrisankaran, G. and M., Rysman (2012), “Dynamics of Consumer Demand for New Durable Goods,” *Journal of Political Economics*, 120(6), 1173–1219.

Guadagni, Peter M., and John D. C. Little (1983), “A Logit Model of Brand Choice Calibrated on Scanner Data,” *Marketing Science*, 2(3), 203-238.

Hendel, I., and A. Nevo (2006), “Measuring the Implications of Sales and Consumer Stockpiling Behavior,” *Econometrica*, 74, 2006, 1637–73.

Kim, J. (2006), “Consumers' dynamic switching decisions in the cellular service industry,” Working Paper 06-24, University of Wisconsin-Madison, The Networks, Electronic Commerce, and Telecommunications Institute.

Lewis, Michael (2004), “The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention,” *Journal of Marketing Research*, 41(3), 281-292

Montoya, Ricardo, Oded Netzer, and Kamel Jedidi (2010), “Dynamic Allocation of Pharmaceutical Detailing and Sampling for Long-Term Profitability,” *Marketing Science*, 29(5), 909-924.

Moore, Gordon E. (1965), “Cramming more components onto integrated circuits,” *Electronics, 38(8)*, 1-4

Nair, H. (2007), “Intertemporal Price Discrimination with Forward-Looking Consumers: Application to the US Market for Console Video-Games,” *Quantitative Marketing and Economics*, 5(3), 239–292.

Nelder, John A., and Roger Mead (1965), "A simplex method for function minimization," *The computer journal*, 7(4), 308-313.

Nosal, K (2011), “Estimating Switching Costs for Medicare Advantage Plans.” Unpublished Manuscript, University of Arizona, 2011.

Puterman, M.L. (1990) “Markov Decision Processes,” in: D.P. Heyman and M.J. Sobel, eds., *Handbooks in Operations Research and Management Science,* Volume 2. North-Holland/Elsevier: Amsterdam.

Rust, John (1994), “Structural Estimation of Markov Decision Processes,” *Handbook of Econometrics*, Vol.4, Chapter 51, 3082-3139.

Rust, John (1987), “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55(5), 999–1033.

Ryan, Stephen P. and Catherine Tucker (2012), “Heterogeneity and the Dynamics of Technology Adoption,” *Quantitative Marketing and Economics*, 10, 63–109

Shcherbakov, O. (2009), “Measuring Switching Costs in the Television Industry.” Unpublished Manuscript, University of Mannheim, 2009.

Song, Inseong and Pradeep K. Chintagunta (2003), "A Micromodel of New Product Adoption with Heterogeneous and Forward-Looking Consumers: Application to the Digital Camera Category," *Quantitative Marketing and Economics*, 1, 371-407.

Sriram, S., Pradeep K. Chintagunta and Manoj K. Agarwal (2010), "Investigating Consumer Purchase Behavior in Related Technology Product Categories," *Marketing Science*, 29(2), 291-314

Sun, Baohong, Scott A. Neslin and Kannan Srinivasan (2003), "Measuring the Impact of Promotions on Brand Switching When Consumers Are Forward Looking," *Journal of Marketing Research*, 40(4), 389-405

Sun, Baohong (2005), "Promotion Effect on Endogenous Conumption," *Marketing Science*, 24(3), 430-443.

Tomasco, Steven (Apr. 07, 2011) "Industry Leaders Worldwide Embrace IBM Clouds to Transform Business Processes," *http://www-03.ibm.com/press/us/en/pressrelease/34201.wss*

Yang, Botao and Andrew T. Ching (2013), "Dynamics of Consumer Adoption of Financial Innovation: The Case of ATM Cards," *Management Science*, 60 (4), April 2014, 903–922

Zhao, Yi., Ying Zhao, and Kristiaan Helsen (2011), "Consumer Learning in a Turbulent Market Environment: Modeling Consumer Choice Dynamics After a Product-Harm Crisis," *Journal of Marketing Research*, 48 (April 2011), 255-267.