Carnegie Mellon Tepper

SCHOOL OF BUSINESS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
INDUSTRIAL ADMINISTRATION
(MARKETING)

Titled
"ESSAYS ON THE IMPACT OF NEW TECHNOLOGIES ON
FIRM-CONSUMER RELATIONSHIPS"

Presented by
Francisco Orlando Cisternas Vera

Accepted by

Alan Montgomery   4/7/2017

_____   _____

Chair: Prof. Alan Montgomery   **Date**

Approved by The Dean

Robert M. Dammon   4/13/2017

_____   _____

Dean Robert M. Dammon   **Date**

ESSAYS ON THE IMPACT OF NEW TECHNOLOGIES ON FIRM-CONSUMER RELATIONSHIPS

by

FRANCISCO ORLANDO CISTERNAS VERA

Submitted to the Tepper School of Business
In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Field of Industrial Administration

at

CARNEGIE MELLON UNIVERSITY
TEPPER SCHOOL OF BUSINESS

April, 2017

Dissertation Committee:

Alan Montgomery (Chair)

Willem J. van Hoeve

Timothy P. Derdenger

Peter Boatwright

Beibei Li

To my family

Acknowledgements

Abstract

Technological developments have the potential to dramatically change the interactions between firms and their consumers. In recent years, technological improvements are occurring at a faster pace than ever before. Now companies are forced to learn quickly and react faster to remain competitive in the currently changing world. In my thesis, I explore some specific technological innovations, consider how consumers adjust, and discuss their managerial implications.

In the first chapter I study the effects of consumer adoption of mobile banking on the banking industry. The proportion of US bank customers using mobile banking has grown from 29% in 2012 to 43% in 2015. This channel of interaction is likely to keep growing due to a further increase in the adoption of smart phones, improvement in quality of mobile banking apps and channel awareness. As a consequence, some banks have reported that they may reduce their number branches by half over the next decade. The adoption of mobile banking displaces many banking functions performed through other channels like: automated teller machines (ATM), telephone banking, and online banking. Using geo-coded transaction data from a large consumer bank, a dynamic structural model to represent consumers' preferences is developed for online and physical channels. In this way changes in banking behavior due to variation in the branch network structure as well as the introduction of the mobile channel are considered. This model is used to predict the timing and type of transactions across channels. The knowledge gained with the demand model is then used to design an optimal branch network in terms of capacities, amenities, location, and number of branches. Counterfactuals allow the evaluation of different levels of channel adoption, and allows the consideration of their effect on banking transactions, and more important, the impact on customer loyalty. The model shows all channels remain relevant and, moreover, we found a strong complementarity between the physical and digital world. Therefore, instead of reducing the number of physical branches, banks should aim to adjust current branch capacities, specializing on transactions that cannot be served with digital channels. In conclusion, my findings suggest that digital channels will diminish–but never replace physical channels—and they should be redesigned correspondingly. Is important to note that this is the first time in the banking industry that

substitution of branches for digital channels are formally considered in a tool to support branch network design for the middle and long term.

Internet shopping made possible interactive displays. In the second chapter, I discuss how arranging the products on a display or retail shelf can directly influence consumer purchases by facilitating or obstructing product search. Product proximity also influences competition and the set of products that a consumer considers. We show that when products are placed closer together competition between the products increases. Thus, product display can encourage consumers to purchase products that would not have previously been purchased. The motivation for the result is that consumer search is costly and consumers focus their search on local neighborhoods that are influenced by shelf position. Since search is costly, consumers may not exhaust all possibilities, which means that position could be an important determinant of consideration. To formally model this behavior, I create a sequential consideration model. To begin the search consumers are influenced by colors, favorite brand or the closest shelf edge, where few products can be considered. Then consumers shift their focus to neighboring products, in a sequential fashion, increasing the set of products considered, to finally making the purchase decision. By doing this, the display generates spatially induced consideration sets. Using this approach, we find that demand is greatly impacted by shelf position and retailers can create plan-o-grams that can shift demand from one product to another. Our focus on using shelf design to stimulate competition contrasts with past research on shelf design that has focused mainly on cost minimization. Using shelf-experiments from a supermarket retailer, Dominick's Finer Foods, I show that re-arranging the products on the shelf can increase profits by up to 15%.

Table of Contents

Chapter 1

Chapter 1: Reshaping Bank Branch Networks due to Mobile Banking

Mobile banking has grown from 29% in 2012 to 43% in 2015 among US bank customers[1] and is likely to keep growing due to a further increase in the adoption of smart phones, improvement in the quality of mobile banking apps, and increasing awareness of these apps. Mobile banking is changing the way consumers interact with their banks, displacing many banking functions performed through other channels, such as automated teller machines (ATM), telephone banking, and online banking. In response banks are testing new branch formats, developing improved ATMs, and reducing the number of branches. Multi-channel management in the financial industry is critical to attract and retain customers. Using geo-coded transaction data from a large consumer bank, we develop a dynamic structural model that represents consumers' preferences for online and physical channels. Our demand model takes into account consumer banking behavior as a function of the branch network structure as well the mobile channel. We use this model to optimize the branch network in terms of capacities, amenities, location, and number of branches. Counterfactuals are constructed to evaluate potential levels of channel adoption and consider its effect on banking transactions and, more important, on customer loyalty. Our model shows that all channels remain relevant after mobile banking adoption; moreover, we find complementarity between the physical and digital channels. Our conjecture is that the importance of physical channels to banks are lessened in the presence of digital channels but is not replaced entirely. Our findings suggest that instead of reducing the number of branches, banks should aim to adjust current branch capacities and have physical branches specialize in those

---

[1] Consumers and mobile financial services report 2016 – Board of Governors of the Federal Reserve

transactions that cannot be served well with digital channels, such as financial advisement or resolving problems.

1.1   Introduction

Mobile banking has the potential to fundamentally change how, when, and where consumers bank. Mobile banking began in 2010 in the US with specialized access to banks' web pages for mobile devices. Since then the type and number of bank services offered through the mobile channel has increased, now ranging from simple balance inquiries or personal funds transfers to mobile payments and check deposits. As a consequence, mobile banking is changing the traditional role of physical bank branches.[2] In fact, some consumers do not visit any branches, except perhaps to open their accounts. However, this does not mean that branches should be closed, for several reasons. First, some consumers continue to be heavily reliant upon traditional branch services and ignore other channels. Second, some functions like cash withdrawals or accessing safety deposit boxes cannot be provided digitally. Third, branches provide more than just banking services, i.e., branches offer an in-person point of interaction and are strongly valued by customers, even if they are rarely used. This close interaction is also valued by banks, as an industry marketing research report about US consumers (Celent, June 2013) found that "[physical] branches are the best opportunity to cultivate strong relationships such as new customer acquisition and opening of accounts."

Currently, banks themselves are introducing new technologies that automate many functions performed by branch employees. Some options for redesigning branches include adding more sophisticated ATMs, self-service tablets, and videoconferencing services, while decreasing traditional teller services. As banks encourage their consumers to migrate transactions to self-service and less costly channels, the question bank managers are asking is: "What should the branch network look like?" The answer to this question is complex because there is interdependence between the branch network and consumer behavior. For example, if a new ATM is more conveniently located near a consumer then a weekly visit to withdraw cash at the

---

[2] We use the term bank branch to denote a physical brick-and-mortar location and use mobile or online to refer to virtual bank channels. For clarity we only use the term branch to refer to physical locations. Additionally we think of ATM's as a type of branch with extremely limited capabilities (typically only cash withdrawal).

branch may transform into two smaller withdrawals from the ATM. In turn the bank must predict demand for banking services to appropriately design the branch, so it is important for the bank to anticipate how consumer banking changes as a function of the branch network design. Our research proposes a solution to this problem through an analytical model of consumer behavior and leverages this model to predict how banks can optimize their existing branch networks.

Our framework addresses two primary objectives. The first objective is to construct a model of consumer attraction to a branch that predicts which branches a customer is likely to visit and which services are demanded at each location. This model needs to consider both choices of branches and ATMs as well as the online and mobile bank channels. To better understand consumer behavior, we use the locations of debit and credit card transactions to predict where consumers will shop and relate this to branch location usage. For example, a customer who commutes from the suburbs to downtown may be more amenable to using a branch location downtown, whereas a retired customer in the suburbs may have a strong preference for a location near their home. A challenge of this modeling process is to understand how consumer branch usage will change over time as consumers adopt online and mobile bank services. Additionally, the bank may even take steps to encourage consumers to make these changes more quickly.

The second objective is to combine our model of consumer and business demand with operational considerations of the branches to design the optimal branch network. Operational considerations include transaction capacity, customer satisfaction, branch expertise, neighborhood potential, and competitive characteristics. The goal is to provide insight into the design of an optimal branch network based on these considerations. In particular, how services and resources should be allocated across the different branches in the network? Should the number of branches be expanded or reduced in order to reach a certain level of customer satisfaction (or any other objective)? Our conjecture is that physical branches will continue to play an important role in banking, but that their role, as well as their quantity and size, will change dramatically in the next decade. We estimate our model with anonymized consumer-level data from a large US bank and construct counterfactuals that suggest that in the next decade we should observe that branches become more specialized with a limited number of services. These services will vary from branch to branch and are

determined by customers' preferences as predicted by models like the one presented in this paper. These models have become available and useful now only because of new sources of data, advances in econometric algorithms, and higher computing capacity.

1.1.1 Technological evolution of bank Industry

Modern banks can trace their origins to the rich cities of northern Italy during the early Renaissance periods14th century[3] with fractional reserve banking and banknotes appearing between the 17th and 18th centuries. The industry evolved slowly until the 20th century, when new technology fostered new types of interactions between consumers and bank. In the 1960s, the first call center and the precursors of ATMs appeared. One of the earliest call centers was created in the UK in 1964, where it was known as a "Private Automated Business Exchange" or PABX[4]. The precursor of the ATM was the Bankograph, which was installed in New York City in 1961 by the City Bank of New York[5,6]. This automated envelope deposit machine was removed after six months due to lack of customer acceptance. It is widely accepted that the first modern ATM was installed by Barclays Bank in London in 1967; since then, the technology has continued to advance with new generations featuring touch screens, video conferencing, biometrics, coin handling, scanning of individual checks without envelopes, and offering non-bank related services like dispensing movie tickets, phone cards, or traveler's checks.

Another leap in banking technology came with the advent of the Internet. The early 1980s precursors of online banking used phone lines and a keyboard to access account information; later, banks began to use the World Wide Web mainly as a way to advertise their services. In 1995, Wells Fargo was the first US bank to add account services to its website, and many banks followed. Many users were reluctant to adopt online

---

[3] Hoggson, N. F. (1926) Banking Through the Ages, New York, Dodd, Mead & Company.
[4] Science and invention in Birmingham#cite note-45
[5] Machine Accepts Cash Deposits, The New York Times, 12 April 1961
[6] "From punchcard to prestaging: 50 years of ATM innovation". ATM Marketplace. 31 July 2013. Last retrieved 15 February 2016

banking in the early years, with only 0.4% of households in the US using online banking at the end of 1999[7]. This number has grown to 31% in 2004, 47% in 2009[8], 51% by 2013 up to 64% in 2016[9].

Potentially smartphones have the ability to fundamentally alter how consumers interface with banks, being always on and ever present. Mobile banking was first introduced in 1999 through SMS[10], and later with the introduction of smart phones with WAP[11] technology that allowed consumers to access web platforms. In 2010 banks began to widely introduce special client programs (apps) for smartphones, but it was not until 2013 that they truly began to take advantage of unique mobile features, like location-based services. According to the last Survey of Consumers and Financial Services in 2016 conducted by the Board of Governors of the Federal Reserve System, the ubiquity of mobile phones is changing the way consumers access financial services: 43% of all mobile phone owners and 53% of smart phone owners with bank accounts have used mobile banking in 2015 (up from 29% in 2012, 33% in 2013 and 39% in 2014). And this rapid growth is expected to continue since 11% of phone owners with bank accounts who do not currently use mobile banking expect that they will probably or definitely use it during the next year.

The most common use of mobile banking is to make inquiries about account balances (94% of mobile banking users); the second and third most used services are money transfers and receiving alerts (58% and 56% of mobile banking users, respectively). Mobile payment is still a less common than using mobile banking with 24% of mobile phone owners reported having made a mobile payment in (28% among smartphone owners). The median frequency of use of mobile banking is five times per month. The main impediments to the adoption of mobile banking are the preference for other banking methods and security concerns: 88% of consumers who do not use mobile banking believe that their banking needs are being met without the use of mobile banking (up from 86% in 2014), whereas 73% cite concerns about security.

---

[7] Online Banking Report
[8] Survey by Gartner Group
[9] Consumers and mobile financial services report 2016 – Board of Governors of the Federal Reserve
[10] Short Message Service (SMS) is a text messaging service component of phone, web, or mobile communication systems
[11] Wireless Application Protocol (WAP) is a technical standard for accessing information over a mobile wireless network

1.1.2   Literature Review

Our research relates to other work in multichannel customer management (MCM). We follow the definition first proposed by Neslin et al. (2006): "the design, deployment, and evaluation of channels to enhance customer value through effective customer acquisition, retention, and development." In MCM, an open question is whether firms should add more channels to the traditional ones. In the banking industry, the introduction of a mobile channel is almost compulsory for medium to large banks. Blattberg, Kim, and Neslin (2008, Chapter 25) suggests that firms should encourage multiple channel adoption if the strategy increases loyalty or marketing response, but should discourage it if adoption decreases loyalty, has no impact on marketing response, or just offers customers greater convenience without increasing the firm's share of customers' wallets.

In terms of loyalty, many studies (Kumar and Venkatesan 2005; Venkatesan, Kumar, and Ravishanker 2007) show that increasing the number of channels can yield to higher customer satisfaction, and with the rapid increase in mobile adoption, banks should add this channel to prevent customer attrition. Empirical evidence suggests that multichannel availability may enhance loyalty (Shankar, Smith, and Rangaswamy 2003; Hitt and Frei 2002; Danaher, Wilson, and Davis 2003; Wallace, Giese, and Johnson 2004), although some studies suggest that increased Internet usage may erode loyalty (Ansari, Mela, and Neslin 2008).

In our paper we describe and support three new reasons banks may adopt the mobile channel and encourage their customers to do likewise. First, we show that mobile banking and physical channels are complementary and that mobile adoption can increase bank usage, which can be beneficial for the business in all the channels. When consumers choose to use the digital channels, it releases capacity in traditional channels. This allows firms to potentially reduce capacity without affecting their service level. The capacity reduction can be achieved by eliminating services from branches that are available on digital channels, allowing branches to specialize in their unique offerings which in turn may increase their efficiency. Second, the use of mobile banking affects customer loyalty by increasing switching costs, therefore reducing attrition.

Finally, mobile banking is still a differentiating factor, because not all banks fully support this channel yet and it can attract customers to do mobile banking as well as traditional banking.

The MCM literature shows that multiple channel customers are not necessarily more profitable (Kushwaha and Shankar 2013); however, a study of the banking industry (Cambra, et al 2015) shows an improvement in profit with multiple channels in cases where customers were encouraged to use high-margin channels, in dual channel combinations. We find that although the mobile channel increases the frequency of transactions, at the same time it decreases the bank's opportunity to cross-sell or up-sell products because digital channels are less efficient in this respect, which in the long term may yield lower profits.

Our research focuses on understanding how customers' channel decisions are affected by the introduction of the mobile channel in the banking industry, an issue that has not been studied previously. Channel choices became popular with the introduction of online channels, but this research is usually constrained to retailing. Chintagunta, Chu, and Cebollada (2012) are the closest to our work in terms of methodology; using a hierarchical Bayes model they found significant transaction costs to purchase in-store versus online, but they did not consider the mobile channel. Laukkanen (2007) found significant differences in value perception between mobile and online banking, which is consistent with our findings. Moreover, we find that value perception is related to location and other branch characteristics when compared with safety and awareness of the digital channels.

In order to give recommendations to branch managers considering the bank as a multichannel service, we created a mixed integer programming model (MIP) to find branch configurations that optimize the network under different objectives. The size and complexity of our problem force us to use linear formulations. The logit probabilities used in our model are non-linear in nature. We used the approach suggested by Haase and Müller (2013), based on the IIA properties of the logit form to linearize the MIP model. We adapted that approach to be able to include consideration sets in the formulation, which offers a new way to solve this optimization problem.

1.2 Modeling Consumer Financial Transactions

We obtain anonymized data from a large US bank. Our data comprises bank transaction data for a sample of more than 500,000 accounts with more than 1.7 billion transactions across all channels for over a ninety six-month period (June 2007 and Jun 2015). For each transaction, an anonymized account identifier, date, channel, amount, and type were provided along with associated customer information. Additionally, the location associated with branches is known and the locations of some debit card transactions through the merchant's postal code can be inferred (not all transactions give postal codes). **Table 1** provides a simulated example of the raw information for a consumer. The description in this example is given for illustrative purposes and is not actual data.

Date	Description	Channel	Location	+/-	Amount	Balance
11/21/15	ATM withdrawal	ATM	15213	Debit	$80.00	$125.15
11/23/15	Check deposit	Branch	15217	Credit	$250.00	$375.15
11/28/15	Salary from direct deposit	ACH	99999	Credit	$1,032.25	$1,407.40
12/01/15	Check balance	ATM	15201	-	$0	$1,407.40
12/07/15	Purchase w/ credit card	Debit	15120	Debit	$92.84	$1,314.16
12/12/15	Check Balance	Debit	99999	-	$0	$1,314.16

Table 1. Example of simulated transaction data for an individual customer.

In our dataset, withdrawals and deposits represent more than 98% of transactions that consumers perform at the bank, and as a consequence, these two types of transactions form the focus of our study[12]. Additionally, most consumers visit a branch about once a month on average. Specifically, we observe that almost 60% of consumers perform banking transactions one time per month on average. Most branch visits result in a single transaction but when more than one transaction is performed it is typically done as a single type of transaction (e.g., cash deposit and cash withdrawal).

Following the bank's nomenclature, consumers can operate with the bank through ten channels that can perform twenty types of transactions within sixteen service types. Among the transaction types, the top five account for almost 90% of the transaction types. The most popular are Inquiry and Purchase; these transactions do not use significant bank capacity. The next set of transaction types by frequency are

---

[12] There are other types of transactions, like safety deposit box access or opening an account with a sales' representative that can be important ones for the bank. For expositional purposes, we want to keep our model as simple as possible and, therefore, we ignore these types of events. However, our model can be extended for these other types of activities.

withdrawals, deposits, and credits, which are more demanding on the bank's infrastructure. **Figure 1** summarizes the channels and their usage as a fraction of the total number of transactions in our data since 2007. The most popular channels are web and the purchase with debit and credit cards. The mobile channel is only available from 2012; since then, we observe a drop in the web usage. This drop is not only due to the mobile channel, since an important shift is towards the credit card channel. This is because the bank boosted its credit cards business. The number of transactions through the Branch and Web channels decline as mobile banking increases. The decline is more prominent in the Web channel.
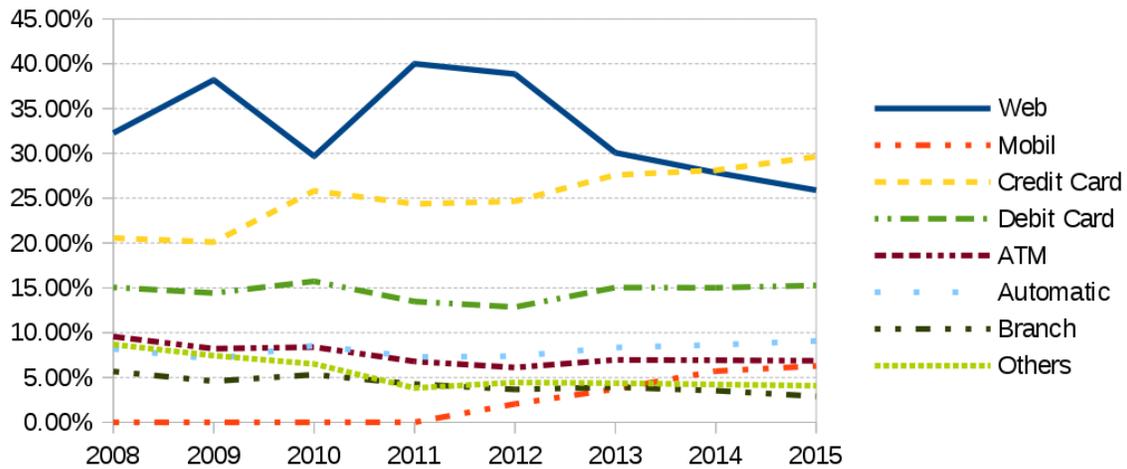
Figure 1. Channel preferences when performing current accounts and credit cards transactions.

Although this gives us an extensive set of information from which to make inferences about consumer behavior, there are important deficiencies. Clearly, consumers use cash and other payment mechanisms (e.g., credit cards from competing banks) that we do not observe. Therefore, we abstract away from our observed transactions and assume that consumers have an unobserved demand for cash that is described by a stochastic process. Additionally, consumers may receive checks at any moment, but we only observe when the checks are deposited, presumably when visiting the branch is convenient. We assume that there is an unobserved arrival of checks described by another stochastic process. Additionally, consumers experience an opportunity or holding cost when not depositing a check, which encourages them to deposit checks earlier.

Consumers are represented as rational economic agents who incur transactions costs for visiting a branch or making mobile or online transactions. These transactions costs vary by branch and time and can explain why a consumer who works close to an ATM may have a very different pattern of branch usage than one who works far away from an ATM. Every time the bank modifies its branch network by opening/closing a branch/ATM or by enabling more transaction types on its mobile channel, it is changing the transaction costs associated with the transaction, and these changes affect the frequency and intensity with which customers interact with the bank. For example, if a branch is opened in the same building where a customer works, it is likely that the consumer will visit this branch more frequently than the branch previously visited, thus decreasing the average number of transactions per trip or even the amount of cash withdrawn.

In summary, consumers make decisions in order to minimize current and future transaction and waiting costs. Therefore, we represent, these sequential decisions as the solution of a stochastic dynamic programming (SDP) model. The SDP timing is described as follows:

1. The consumers realize their needs for financial transactions (cash withdrawal and deposits).
2. Transaction costs for each alternative are realized.
3. Consumers decide whether to perform pending transactions or to wait. Their decision is made by comparing the cost of waiting one more period with the transaction cost of the "cheapest" alternative.
4. The cash balance and the amount of checks not deposited are updated according to consumers' decisions.

The timing of each period is described in **Figure 2**.
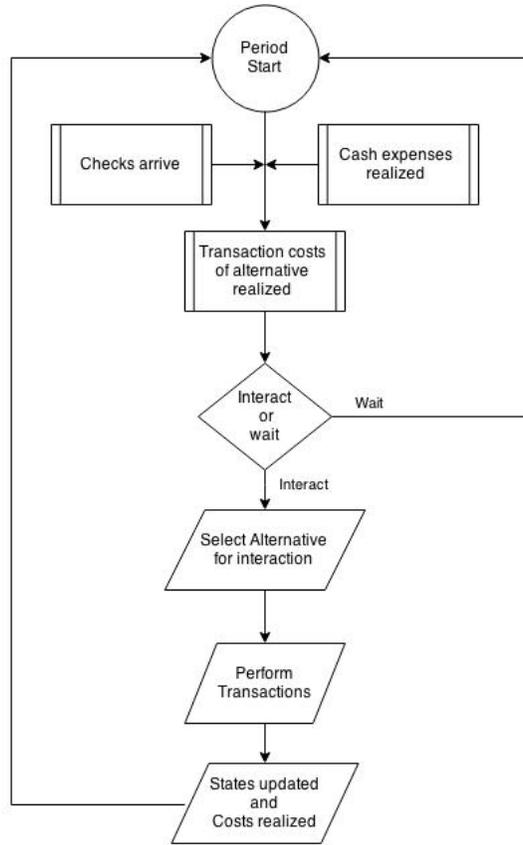
Figure 2. Model Timing

1.2.1   Model Specification

Our model has three main components: 1) waiting cost, which includes the costs associated with postponing interactions with the bank; 2) transaction costs, which are implicit costs associated with a consumer interacting with the bank; and 3) consideration set, where we assume that consumers consider only a subset of the available alternatives. We allow the set of alternatives to be updated over time which allows variation in the consideration set. At the end of this subsection we explain how we handle heterogeneity in the model.

1.2.1.1   Demand for money and its waiting costs

Cash demand. The cash demand for consumer i at period t is represented by D_{it}. Since we do not observe this demand, we model it as a random process following a Poisson distribution with rate λ_i^D:

$$D_{it} \sim Pois\left(\lambda_i^D\right) \tag{1}$$

The balance of cash held by consumer i at period t is denoted as k_{it}. If a consumer does not visit the bank to make a cash withdrawal or deposit then their cash balance is $k_{it+1} = k_{it} - D_{it}$. Alternatively, if consumers choose to visit the bank, they can decide to keep Q_{it} of cash for the next period, which is net amount of cash after withdrawals or deposits, so $k_{it+1} = Q_{it}$ in case of a bank visit.

Check deposits (h_{it}). We denote the dollar amount of checks at the beginning of period t that need to be deposited by consumer i as h_{it}. If a consumer visits the bank during period t (if a visit is made then $V_{it} = 1$ otherwise $V_{it} = 0$) then we assume the consumer deposits all checks; otherwise, the total amount of checks that are not deposited is updated by adding the checks that arrived this period:

$$h_{i,t+1} = \begin{cases} h_{it} + \sum_{s=1}^{N_{it}} A_{its} & \text{if } V_{it} = 0 \\ 0 & \text{if } V_{it} = 1 \end{cases} \tag{2}$$

N_{it} is the number of checks that arrived during the current period t, and A_{its} is the amount of check s, $s \in \{1, ..., N_{it}\}$. We do not observe the arrival of the checks, only deposits are observed. Therefore, we model these values as a compound Poisson process, where N_{it} follows Poisson arrival timing with rate λ_i^N and A_{its} is uniformly distributed within the range of checks amount deposited by the consumer Am_i from the minimum to the maximum amount deposited in previous years :

$$N_{it} \sim Pois\left(\lambda_i^N\right)$$
$$A_{its} \sim Unif\left(Am_i\right) \tag{3}$$

Holding Costs. When a consumer decides to postpone transactions with the bank for a future period, the consumer incurs a cost of not performing pending transactions. For example, not depositing a check can cause an overdraft, or not withdrawing cash might cause a suboptimal consumption. Using the notion of

opportunity costs or interest rate, we use a linear representation of waiting costs as a function of the amounts involved. The waiting cost for consumer i at period t is described by ω_{it} and defined as:

$$\omega_{it} = \underbrace{\eta_i^h \left(k_{it} - D_{it} \right) \cdot 1_{[k_{it} - D_{it} > 0]}}_{\text{Holding Costs}} + \underbrace{\eta_i^p \left(D_{it} - k_{it} \right) \cdot 1_{[k_{it} - D_{it} < 0]}}_{\text{Penalizing Costs}} + \underbrace{\gamma_i \cdot h_{it}}_{\text{Undeposited Check Costs}} \tag{4}$$

Where η_i^h is the holding cost associated with having excess cash, perhaps due to the risk of losing the cash or the opportunity cost of not having it deposited, and η_i^p is the penalty associated with not having the right amount of cash at period t which might force a consumer to forgo consumption or borrow funds. The parameter γ_i associated with not depositing a check, represents the opportunity cost or a greater risk of over drafting associated with not depositing in the period and is proportional to the amount of undeposited checks (h_{it}).

1.2.1.2   Transaction Costs

A consumer chooses a branch from a large set of branches denoted by the subscript b; one of the branches represents the mobile and online channel. We model this choice through implicit transaction costs incurred by the consumer that represent the effort and time to complete the transaction at different branches. The transaction cost is influenced by characteristics of the alternative chosen and the customers' individual characteristics and preferences. In the case of the physical branches and ATMs, our exploratory analysis showed that location is a driving factor in determining how attractive a given branch location is with respect to others. We note that while the location of the branches is fixed, customers' location is not. For example, consumers travel from home to work, or work to home, or from home to a shopping mall and these movements by the consumer may affect the branch attractiveness.

We introduce the transaction cost T_{ibt} for customer i choosing alternative b at period t as:

$$T_{ibt} = \begin{cases} \exp\left\{ -\tau_i \cdot B_{bt} - \delta_i \cdot DW_t - \pi_i \cdot 1_{[b \in \text{Used}_i]} \right\} - \varepsilon_{ibt} & \text{if } b \in CS_{it} \text{ and } b \in L \\ \exp\left\{ -\psi_i \cdot Ch_b \right\} - \varepsilon_{ibt} & \text{if } b \in CS_{it} \text{ and } b \notin L \\ \exp\left\{ -\theta_i \right\} - \varepsilon_{ibt} & \text{if } b \notin CS_{it} \end{cases} \tag{5}$$

Notice that there are three specific cases in evaluating transactions costs: alternatives that are in the consideration set (CS_{it}) and also belong to the set of physical location L (e.g., those branches or ATMs visited recently), alternatives that customer i use regularly ($\in CS_{it}$), but are not physical locations (e.g., use the computer to perform a payment), and third alternatives outside the consideration set (e.g., a consumer unexpectedly visits a branch outside their usual areas).

If alternative b is in the consideration set then the transaction cost is the sum of three components plus a random error. The first component is the inner product $\tau_i \cdot B_{bt}$ where B_{bt} is a vector of transformed attributes of alternative b at time t—for example, size or capacity of the alternative chosen. We also include in this vector the number of competitor branches in the same ZIP code, and the number of other branches of the same bank in the same ZIP code. The second component represents an individual preference for a specific day of week and is measured as the inner product of $\delta_i \cdot DW_t$, where DW_t is a vector of zeros with a one in the place of the day of the week of t and δ_i is its respective sensitivity vector of dimension seven. The third component is meant to capture persistence in usage of alternatives in the consideration set and scales an indicator, $1_{[b \in Used]}$, that detects if alternative b has been used in the past three months. In the second term, we consider alternatives that are not physical but within customer's consideration set, we estimate a fixed effect vector as an inner product of the sensitivity vector ψ_i and the vector Ch_b that is a vector of zeros with a one in the position of alternative b. Finally, if the alternative is not in the consideration set then the transaction cost is the sum of a constant θ_i and random component to give a small probability that any branch can be chosen.

We assume that consumer i chooses the alternative with the lowest transaction costs i period t:

$$v_{it} = \underset{b}{\arg\min}\, T_{ibt} \tag{6}$$

Where ν_{it} are the transaction costs of using the "cheapest" alternative. The unobserved idiosyncratic random shock ε_{ibt} in (5) is assumed to follow a Type I extreme value distribution and is i.i.d.. This assumption yields a logit form for the probability of choosing alternative b.

1.2.1.3   Consideration Sets

The data shows that most consumers visit a small number of branches, which suggests that consumers focus on branches close to areas that they live, work, or shop. The majority of branches are completely ignored when it is not in the consumer frequented areas. This is exemplified in Figure 3. This figure is a special representation of transactions at ATMs and purchases using credit/debit cards aggregated at zip code level. Each area represents a zip code area, and the filling represents the percent of transactions perform at any location within the zip code. It goes from yellow to red where yellow represents 0% and red represent 25% of transactions in a year per customer. The customer home address is indicated with a black star.
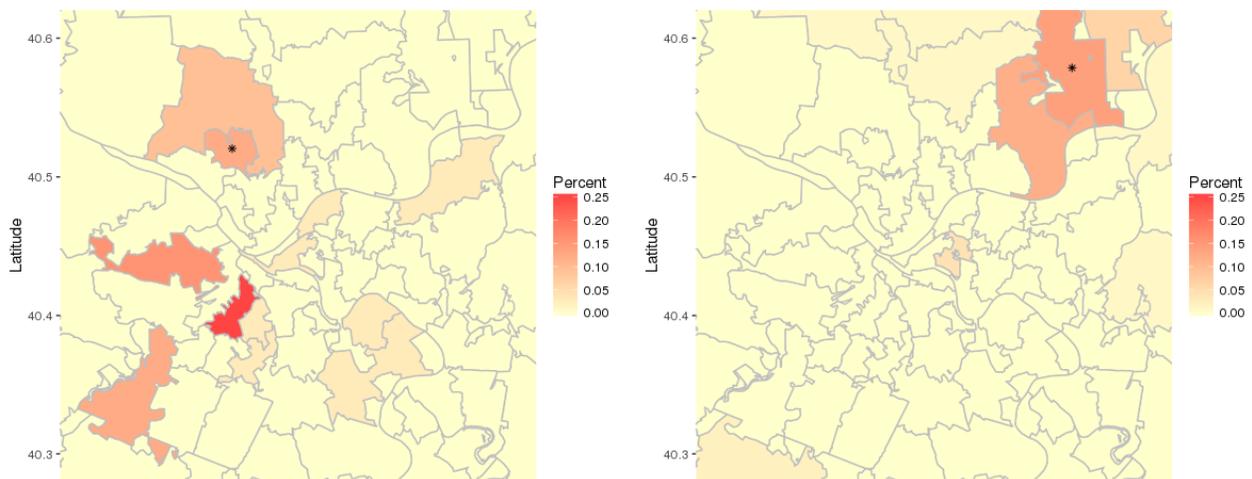
Figure 3. Two examples of consumer observed activity during 6 months

On the right example, a consumer lives in an area on the north right of this region, and performs most of her transactions near where she lives. In the case on the left, the consumer performs some transactions where she lives but the majority of them are in totally different areas. In many cases consumers focus their transactions in a few locations leaving many others with no transactions. As a consequence, most branches

are ignored regardless of their attributes. Because of this, we employ a consideration set approach (Schoker et al. 1991, Roberts and Lattin 1991). A consideration set is defined as a set of branches that a consumer will consider visiting. We later extend this to other channels as well. Consideration sets are consumer specific, since different consumers have different home addresses and frequent different areas.

An important objective is to model consumers' preferences among different channels, therefore the consideration incorporates both preferences for different branches as well as usage of mobile and web channels. Mobile channel adoption is still low, and it is expected to increase in the near future. Lack of a smart phone or being unaware of the bank's mobile banking ability are potential reasons consumers do not take advantage of the mobile channel. In our model, we choose to represent the mobile banking adoption and therefore incorporate that alternative in the consideration set, when a transaction using the channel is observed. The same procedure was used to add the Web channel. More specifically the mobile and online channels are included in the consideration set if a consumer performed a transaction on the channel in the past 12 months; otherwise these channels are not in the consideration set.

In the case of physical channels—branches and ATMs—consumers tend to concentrate their transactions in few geographic areas with some apparently random deviations. To capture the geographic influence on consideration sets we use the geographic boundaries imposed by ZIP codes and assume that consumers will form their consideration sets based upon a set of ZIP codes that they frequent. We label the most visited zip codes per customer as that customer's focal points. We assume customers build their consideration set from the branches located within their focal points plus locations connected to them.

Limiting ourselves to contiguous zones to define connected areas does not adequately reflect the visitation patterns that we observe in our exploratory analysis. The hypothesis that seems to explain the data better is that customers travel from an origin zone to a destination zone without stopping in between, presumably for working or shopping purposes. Therefore, additional data like travel distances using convenient roads or travel time does not capture the observed patterns.

Instead, we defined the connectedness of ZIP codes by constructing a matrix of co-occurrences of pairs of ZIP codes being visited. We use the data from customer transactions from the previous year to avoid

biasing our contentedness matrix. In each cell, the number of customers who performed transactions in both ZIP codes was tallied. The diagonal measures the total number of consumers who performed transactions in a specific location. Finally, the matrix rows were normalized to reflect the percentage in each cell (i,j), which is the percentage of consumers of ZIP code i who also performed transactions in j. We call this the matrix of connectivity. Note that this matrix is not symmetric. For instance, many suburbs are well connected with downtown, but not vice versa; this is because most of the people who bank in the suburb also bank in downtown, but not a large percentage of people who bank in downtown also bank in that specific suburb. Two examples of the matrix are plotted in Figure 4, where the connectivity from the location with the black star is plotted in the map. Again, the region is split in zip codes. The density of the shading represents the degree of connectivity, where yellow is low connectivity and red is highly connected with the selected region (star).
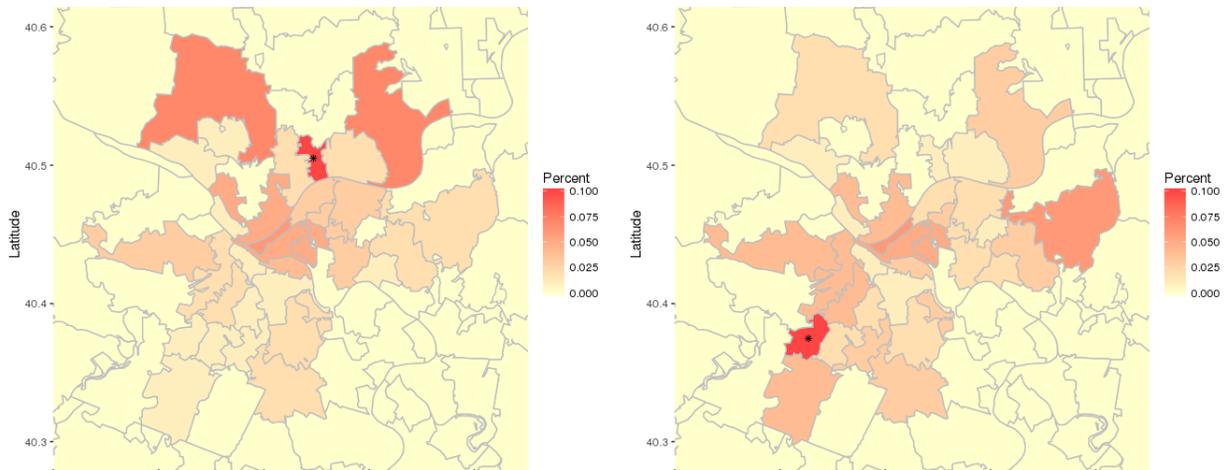
Figure 4. Two examples of connectivity regions.

An interesting aspect of this matrix is that connected ZIP locations are not necessarily contiguous. For example, this matrix shows a suburban area connected with a shopping area that are not contiguous, as depicted on the right side in Figure 4, where the zip region is marked with the star, shows high connectivity with a distant commercial zip code on the right, plus the down town region and little connectivity with places in between. On the left side in Figure 4 we can observe how a commercial district is connected with suburban

areas north east and northwest. Potentially the reason that we do not see strong patterns of connectivity for long contiguous spans is that consumers may travel by limited access highways and focus their activities near the initial starting point and terminus of their trips.

Consideration sets may change over time, perhaps due to a branch being opened or closed or even customer relocation. In the dynamic programming problem, we assume that consumers do not anticipate future consideration set changes. This assumption is needed to simplify an already complex optimization problem. However, consumers do anticipate variations in the transaction costs because of shopping or travel patterns. For example, a consumer who lives in the suburbs may have lower transaction costs associated with a downtown branch during weekdays or workweek and higher transaction costs on the weekend.

Based upon our formulation of the consideration set we can compute the probability for consumer i to visit branch b during period t to perform a banking transaction derived from equation (5):

$$
P_{ibt} = \begin{cases}
\dfrac{\exp\{T_{ibt}\}}{\exp\{-\psi_i \cdot Ch_b\} + \exp\{-\theta_i\} + \displaystyle\sum_{j \in CS_i} \exp\{T_{ijt}\}} & \text{if } b \in CS_{it} \text{ and } b \in L \\[3ex]
\dfrac{\exp\{-\psi_i \cdot Ch_b\}}{\exp\{-\psi_i \cdot Ch_b\} + \exp\{-\theta_i\} + \displaystyle\sum_{j \in CS_i} \exp\{T_{ijt}\}} & \text{if } b \in CS_{it} \text{ and } b \notin L \\[3ex]
\dfrac{\exp\{-\theta_i\}}{\exp\{-\psi_i \cdot Ch_b\} + \exp\{-\theta_i\} + \displaystyle\sum_{j \in CS_i} \exp\{T_{ijt}\}} & \text{if } b \notin CS_{it}
\end{cases}
\qquad (7)
$$

Notice that whenever a new branch is added to the consideration set the probability of the outside option is decreased, while whenever a new branch is added to the consideration set from perhaps opening a new branch there is a decrease in the probability of the outside option. Conversely closing a branch always decreases the service level for consumers that include that branch in their consideration set.

1.2.1.4 Heterogeneity

We have two sources of heterogeneity in the model. The first source is using parameter sensitivities and transaction costs parameters. We use a hierarchical Bayesian framework where for each parameter

$\phi_i \in \left\{ \eta_i^h, \eta_i^p, \gamma_i, \tau_i, \delta_i, \pi_i, \theta_i \right\}$ a normal prior distribution is defined as $\phi_i \sim N\left(\bar{\phi}, \Sigma\right)$, with $\bar{\phi}$ a linear

function of demographics. We specify a diffuse hyper-prior for these parameters. The second source of

heterogeneity is based on customer location through the consideration set. Model testing with and without

consideration sets show large improvement in fit and prediction when consideration sets are applied.

1.2.2 Consumer's Dynamic Programming Problem

In each period, consumers decide whether to visit a bank (if a visit is made then $V_{it} = 1$ otherwise it

is 0) and how much to keep for the next period $\left(Q_{it}\right)$ in order to minimize total costs of waiting,

transactions, and expected future costs. The state space is defined as $S_{it} = \left\{ k_{it}, h_{it}, D_{it}, A_{sit}, N_{it}, \varepsilon_{ibt} \right\}$. The

optimal sequence of decisions at time origin τ can be found by solving the following dynamic programming

problem with discount factor β as:

$$\min_{\{V_{it}, Q_{it}\}_{t=\tau}^{\infty}} E_{\{S_{it}\}_{t=\tau}^{\infty}} \left[C_{i\tau}\left[V_{i\tau}, Q_{i\tau}; S_{i\tau}\right] + \sum_{t=\tau+1}^{\infty} \beta^t C_{it}\left[V_{it}, Q_{it}; S_{it}\right] \middle| \{S_{it}\}_{t=0}^{\tau-1} \right] \tag{8}$$

Where

$$C_{it}\left[V_{it}, Q_{it}; S_{it}\right] = \begin{cases} \omega_{it}\left[V_{it}, Q_{it}; S_{it}\right] & \text{if } V_{it} = 0 \\ v_{it}\left[V_{it}, Q_{it}; S_{it}\right] & \text{if } V_{it} = 1 \end{cases} \tag{9}$$

We use the term $V_{it} = v_{it} + \eta_i \cdot k_{it}^*$ to represent the smallest transaction cost among all alternatives given in

(6) plus the holding cost of the optimal cash amount to be held at the end of the period, ω_{it}, as defined in (4).

We define the value function $\zeta_{i\tau}$ as follows:

$$\zeta_{i\tau}\left[S_{i\tau}\right] = \min_{\{V_{it}, Q_{it}\}_{t=\tau}^{\infty}} E_{\{S_{it}\}_{t=\tau+1}^{\infty}} \left[C_{i\tau}\left[S_{i\tau}\right] + \sum_{t=\tau+1}^{\infty} \beta^{t-\tau} C_{it}\left[S_{it}\right] \middle| \{S_{it}\}_{t=0}^{\tau} \right] \tag{10}$$

Since this is an infinite horizon dynamic problem in equilibrium the policy function is independent of time

and we can write the problem using the Bellman equation:

$$\zeta_i\left[S_{i\tau}\right] = \min_{V_{i\tau}, Q_{i\tau}} E_{S_{i,\tau+1}}\left[C_i\left[V_{i\tau}, Q_{i\tau}; S_{i\tau}\right] + \beta \; \zeta_i\left[S_{i,\tau+1}\right] | S_{i\tau}\right] \tag{11}$$

Since analytic solution is unknown for this type of equation, we rely on a numerical approach which we describe in the following section.

1.2.2.1   Identification and data limitation

The transaction frequency of consumers with the bank is observed which allows us to infer the relative tradeoffs between waiting and transaction costs. However, it is not enough to fully determine both simultaneously. Our main interest is in the transaction costs, since it drives decisions about branch choice, so we choose to fix the parameters associated with holding costs and choose to set the cost of postponing a check deposit and the cost of having excess cash to be unity for every consumer $\left(\eta^h_i = \gamma_i = 1, \forall i\right)$. Early model calibration suggested that penalty for not having enough cash was double than having excess cash. As a consequence, the penalty for not having enough cash equals two.

We calibrate the rates of consumption and check arrivals between bank visits using the first year of data which is not used for estimating the parameters. In theory, it is possible to identify these rates per consumer, but estimation tests with our data show that these parameters were hard to recover.

When computing connectivity between zip codes, to avoid simultaneity problems in the estimation, we compute them using the branch, ATM and purchase date of our customer base for a previous year. The assumption here is that connectivity remains constant from one year to year.

Finally, the discount factor β in the DPP is not possible to estimate simultaneously with other parameters without imposing additional structure in the current period utility, as shown by Rust 1994, and Magnac & Thesmar 2002. Potentially we could follow Ching and Osborne 2017 who show that is possible to estimate with state variables that affect the future costs only as a step function such as in cases of inventory problems which share similarities with our problem. However, this approach requires extra assumptions and more complexity in our parameter vector. Hence, we chose not to estimate the discount rate, but instead we

use the discount factor that was computed with similar data for the same customer set in Liu, Montgomery and Srinivasan 2014, with β value 0.998.

1.2.2.2   Understanding the Model

To illustrate our model we simulate two customers and depict their cash withdrawals, deposits and visits to branches in Figure 5. For each customer we plot three time series: the top plot represents their available cash per period; the middle plot is the amount of undeposited checks, and the bottom plot depicts which branch is visited during each period. The branch choices are indexed by an integer from one to five, and zero represents the choice to not visit any branches. The consumer on the left (panel a) depicts a customer with low holding costs relative to their transaction costs (i.e., the ratio ρ). In other words the customer is willing to wait quite a while to deposit checks and experiences a relatively high transaction cost, and as a consequence chooses to only visit the bank eight times during 100 periods. In contrast the consumer on the right (panel b), visits branches more frequently since their holding costs relative to their transactions costs are greater. In other words the consumer does not want to wait to deposit a check and visits a branch 23 times during the same 100 periods. Recall that consumers make a decision about whether to visit a branch or not based on expected future costs, so consumers are trading off the cost of holding a check against the transaction cost of visiting a branch and running out of cash.

(a) Low Holding/Transaction ratio   (b) High Holding/Transaction ratio
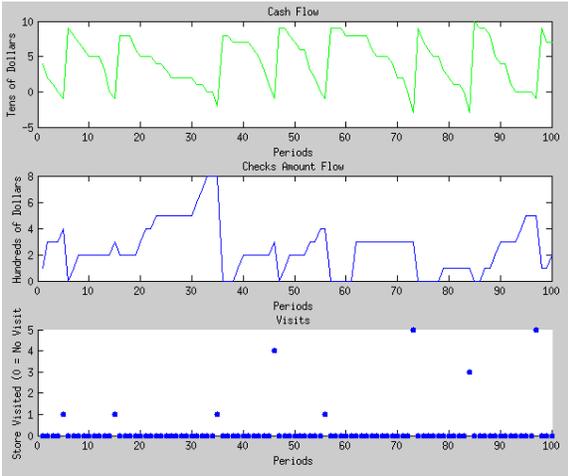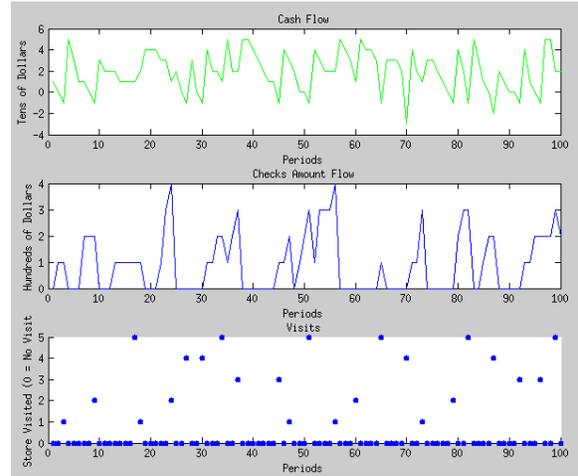
Figure 5. Model workings, bank activities vs periods.

1.2.3   Model Estimation

In order to estimate our model, it is necessary to solve a discrete choice dynamic programing (DDP) problem. In the marketing literature, many models with dynamic decisions have been estimated—for example, dynamic brand choice (Erdem and Keane 1996; Gönül and Srinivasan 1996), dynamic quantity choice (Sun 2005), or new product adoption (Song and Chintagunta 2003). However, the techniques employed in these papers cannot be used in our case due to the size and complexity of our problem.

Our problem shares many similarities with dynamic inventory problems found in the operations research literature, which is quite extensive. Harris (1913) proposed the classic model of economic order quantity (EOQ). Clark and Scarf (1960) proved the optimality of the (s, S)-policy for the stochastic demand model under very general conditions. The (s, S)-policy is a closed-form solution for the dynamic problem, where an order is placed when inventory level reaches s and the order quantity is S-s. Since 1984, efficient algorithms have provided fast computation of the policy (Ferdeguren and Zipkin). Our problem is an extension of the inventory problem to a multiple product inventory problem known as a Stochastic Joint Replenishment Problem (SJRP). Many policies have been proposed to solve this DDP, but none of them have been shown to be optimal (Ozkaya et al 2006). However, there are heuristics which can yield approximate solutions (Viswanathan 1997, Johansen, and Melchiors 2003).

To solve our DDP and simultaneously estimate the parameters of the model, we adopt the technique proposed by Imai, Jain, and Ching (2009) (which we abbreviate as IJC). This approach uses the Metropolis-Hastings algorithm to estimate the model's parameters, and within each MCMC step it iterates the value function once to improve our solution but avoid the computational cost of solving the value function. Solutions of the value function can be approximated from previous evaluations using a non-parametric approach. This method allows estimating the model without explicitly solving for the optimal policy function in every iteration. This procedure reduces our computation burden making it possible to estimate our model for a large number of consumers such as our problem.

However, even using the more efficient approach from IJC it is not possible to apply it directly to our data set due to the scale of our dataset. To speed up the computation, we employ a parallel computing approach suggested by Neiswanger, Wang and Xing (NWX 2014). Exploiting the multiplicative property of the likelihood function they propose a parallel MCMC algorithm in which subsets of the data are processed independently generating sub posterior distributions for each subset. The samples from each sub posterior can then be combined using a semiparametric mixture to make inferences about the full data posterior distribution. Because the each subset can be processed independently they can be computed in parallel thus greatly reducing the total estimation time. This combination of IJC and parallel computation was similarly used in Liu et al (2016).

The IJC method allowed obtaining estimates for our model in reasonable time for small data sets. The addition of the NWX method allowed estimating the model using the full data set. Because the demand model is going to be used as inputs in an optimization routine in a later step, we want to have estimates as precise as possible, thus using the full data set is highly valued.

In our problem we have data for over 500,000 customers. We partition the data in 250 subsets. The estimation of each one of them takes between 2.5 to 3 days. We process between 50 to 60 subsets simultaneously. Repeating this process to reach the full data set takes about two weeks. Using traditional techniques, the estimation would have taken 3.4 years of computation on a single processor.

1.3   Empirical Results

In this section, we show our estimation results and discuss their managerial implications. Using the data from a large bank in the US for a single metropolitan area, this includes the city and its surrounding suburbs. The bank in study has a couple hundred branches and a few thousand ATMs in this region. In order to understand consumer behavior, we use daily transaction data from all bank channels during one calendar year (2014). To estimate the connectivity among ZIP codes, we use the prior year of transactions (2013).

Digital channels—online and mobile—are becoming more and more important, but the vast majority of digital transactions, 85%, are inquiries to check account balances. In this analysis, however, we ignore balance inquiries and focus on banking services that involve transfers of money between accounts, check deposits and cash withdrawals. We chose these types of transactions because they represent more than 99% of all transactions performed at a branch or ATM, and they are the main reason customers choose when and how to visit the bank.

Estimation results are shown in **Table** 2. In this table, we show the results with and without heterogeneity. The models share most of the parameters, except in the preference for digital channels, where we incorporate demographics in a hierarchical fashion.

Par. Type	Channel	Homogenous		Heterogeneous	
		Est.	Var.	Est.	Var.
Waiting Parameters	Alpha_h	.013	.001	.011	.000
	Alpha_p	.063		.055	
	Gamma	1		1	
Location Parameters	Square foot	-.018	.004	-.026	.002
	Business	-.088	.004	-.089	.005
	Income level	.146	.001	.112	.003
	Medium Size	-.088	.006	-.079	.001
	Large Size	.022	.003	.012	.001
	Suburban	.259	.002	.260	.003
	Urban	-.060	.000	-.054	.002
	ATM	-.417	.001	-.230	.004
	Competition	.127	.002	.223	.003
Day of the week	Tuesday	-.076	.004	-.024	.000
	Wednesday	.107	.006	.051	.001
	Thursday	.164	.002	.141	.005
	Friday	-.082	.001	-.053	.002
	Saturday	.165	.002	.116	.002
	Sunday	.204	.001	.219	.005
Others	Visit Constant	-.122	.000	-.122	.000
	Persistence	-.099	.013	-.054	.005

	Desc. Factor	.998	.000	.998	.000
Digital Channels	Mobile	-.265	.012		
	Web	-.367	.011		
Demographic (Mobile)	Age			.142	.015
	Income Low			.089	.008
	Fixed effect			-.437	.006

Table 2. Estimates with and without heterogeneity.

The numbers in **Table 2** represent the sensitivity of different attributes to the transaction costs. The negative numbers represent a reduction in the transaction costs. We can use this numbers to compute percent variation in the probabilities.

For example visiting a branch on a Saturday is 15%[13] less likely than visiting the exactly same branch on a Monday. We find that consumer's favorite day to bank is Friday, followed by Tuesday and then Monday. This strong desire to perform transactions on given days can be explained by the time compression caused by the limited weekend hours. Because of this lack of availability customers tend to use more branches located near where they work, as opposed to branches near where they live.

We conjecture that consumers are using branches near their home, work, or shopping locations that are captured in the consideration sets. Using the model we measure state dependence on customers, within alternatives that customer considers. Using alternatives that were used in the previous six months, raised the probability of using it again between 6% and 10% with respect to an identical alternative but not visited in the previous six months.

Branches in high income ZIP codes are visited almost 13% less than other branches in the other ZIP codes. Further, we found that consumers prefer branches in suburban regions. Urban branches are, in fact, chosen 6% more than rural branches, and surprisingly suburban branches are the least preferred.

The finding that high income regions are visited almost 15% less than other branches in other regions suggests that high income consumers perform fewer transactions at a branch, which is consistent with the lack of urgency to perform transactions based on need for cash or to deposit a check. Wealthier customers tend to use more branches and more channels to perform transactions, thus alleviating the burden

---

[13] Using the properties of the logistic model.

for branches in high-income areas. Another effect that explains this behavior is that high income locations are populated with high branch density; thus, increased competition among branches reduces the attractiveness of individual branches. Similarly, a preference for suburban branches can be a consequence of branch density, because urban regions tend to have higher branch density than suburban and rural areas; therefore, the competition for consumers is stronger in those regions.

Our model shows that consumers prefer large branches. We were surprised to find that medium sized branches are far more preferred than small branches. Specifically their attractiveness increased by more than 8%. In addition to branch size, we use the log of the square footage of the branch as a proxy for a branch's capacity. The size of a branch is a proxy for its capability to perform more types of transactions, for its capacity to perform each service (e.g., more tellers), and for its increased amenities. Customers show that they prefer these attributes in the branches but, surprisingly, not in high proportions. We further determined that an increase of 10% in the surface area makes the branch only between 2% to 3% more attractive.

The choice among channels is influenced by attributes and number of other alternatives in each consideration set, in addition to the preferred locations of each customer.

We found that average transaction costs for ATMs are smaller than branches, which is consistent with our hypothesis, but to a lesser extent than expected. We found that ATMs are approximately 8% more attractive than branches. The mild consumer preference for ATMs was counter-intuitive to us, but can be explained by three major effects. First, the number of ATMs is more than ten times higher than that of branches, so each ATM receives less attention. Second, at ATMs, consumers cannot perform all the type of transactions that are possible at branches. Third, ATMs present more security concerns than branches, so people tend to perform high value transactions at branches. Although ATMs can be more convenient, many customers still prefer branches.

We find that few consumers use mobile channels, but more than 70% of those who try mobile banking at least once continue using it later on. Since the mobile channel is relatively new, we consider consumers to be aware of this channel after they try it once, thus incorporating it within their consideration set. When the

mobile channel is within the consideration set, the alternative is 12% more attractive than the average branch, whereas the online channel is 22% more attractive than the average branch.

1.3.1 Model performance

We used the proposed model to predict customer usage of branches and other channels in a given time horizon. The model is compared to what actually happened and also with three benchmarks models commonly used in the industry. First benchmark used we call "Distance" model, in this model demand shift based on proximity. Customers of closing branches switch to branches near the closing branch, with this approach attrition cannot be predicted, although average attrition levels can be added to generate better performance. The second benchmark is another commonly used practice that we call "Usage"; in this practice consumers are expected to transfer their transactions to the most used alternative other than the closing branch. As a third benchmark we incorporate a traditional discrete choice model. We use a logit model with the attributes of the branches as explanatory variables. In **Table** 3 we show the predictions of branch usage for a set of eight branches before and after a branch at location A is closed.

| | Data | | Proposed | | Bench Mark | |
Branch	Before	After	Model	Distance	Usage	Logit
Closing Location A	26.9%	0.0%	0.0%	0.0%	0.0%	0.0%
Location B	7.5%	20.3%	18.0%	35.1%	17.7%	10.5%
Location C	3.5%	5.6%	6.4%	2.7%	7.9%	4.9%
Location D	2.1%	2.2%	2.9%	2.2%	5.6%	3.4%
Location E	2.7%	2.1%	2.4%	2.4%	6.3%	2.8%
Location F	1.0%	1.9%	2.0%	0.3%	0.9%	1.1%
Location G	0.8%	3.7%	3.6%	1.1%	0.9%	1.6%
Location H	1.3%	1.2%	1.6%	1.0%	1.2%	1.8%
Attrition		8.1%	**6.8%**	**3.8[14]%**	**1.2%**	**3.6%**
RMSE[15]			**1.4%**	**5.2%**	**4.5%**	**2.9%**

Table 3. Prediction of customers switching to other locations when location A is closed[16].

In **Table** 3 we show the percent distribution of transactions performed at location A to H[17]. Only customers that performed at least one transaction at location A were considered. Among all the transactions performed by those consumers almost 27% was performed at that location and the rest of transactions were performed in multiple other locations, some of them were very close to A but in some cases not even in contiguous zip codes, but with high *connectivity,* for example two branches connected by a highways. The column "After" shows the percent distribution of transaction on branches three months after the branch at location A was closed.

---

14   Average attrition that year after a branch was closed.
15   Root mean squared error
16   Numbers were masked to protect data source.
17   We limited the number of branches to 8 for exposition purposes.

Location A is in a commercial area, and after its closure some branches nearby were highly impacted but to a lesser extent than anticipated using a rule of distance (fifth column in **Table** 3), where the closer the branch was to A the higher the impact. Instead location G was highly impacted, which is in a suburban area.

The proposed model performs better when we test the fit using root squared mean error using each prediction with respect to the actual values, computed for the top twenty most affected branches. When closing multiple branches the model shows an even larger gain in performance with respect to the benchmark alternatives. We used the fact that during a month period the bank closed three branches. We show the shift in demand from three months before and after the closures in the same way as shown in the previous example.

	Data	Proposed	Bench Mark		
Branch	**After**	**Model**	**Distance**	**Usage**	**Logit**
RMSE[18]		**2.9%**	**10.5%**	**6.5%**	**5.3%**
Δ RMSE		**+1.5%**	**+4.3%**	**+2.0%**	**+1.5%**

Table 4. Model comparison after the simultaneous closure of three branches.

In this example we show 2 branches located within 4 blocks of each other and a third branch located 6 miles away. The first two branches were located in the downtown, a commercial area with a high branch density. The third branch is located in a suburban area near a commercial district. Given the increase in the number of decisions that were made all the models show a decrease in their performance, the proposed model clearly dominates the alternatives and the improvement with respect to the alternative models is even stronger, which suggest the model will be still be the best under multiple closing decisions. The greatest gain was in mid affected branches that were located in different areas from the closing branches but were highly connected to the areas of the closing branches.

---

18 Computed using the top 30 most affected branches

1.4   Branch Optimization

To find the optimal branch network we have to make decisions about opening and closing branches within our target area. The number of permutations of potential locations leads to a combinatorial explosion, so an efficient optimization tool is required. To make our problem more tractable we assume that only a single new branch can be opened in each ZIP code. Obviously there are many possible locations within each ZIP code to locate a branch, but we lack detailed information about location availability, building costs, rental and leasing costs, government regulations, and which locations the bank is actually considering for expansion. Therefore, we have restricted our simulations to new "average" branches with a medium size and average capacity. However, this not a limitation of our technique but rather of our information set[19].

Given the complexity of the decisions, we choose to use a mixed integer programing (MIP) formulation to explore the feasible solutions in short amount of time, and suggest optimal branch configurations. This technique not only allows to solve this problem efficiently, but also to extend the model to accommodate additional constrains that would greatly improve the usability of the model. Some extensions are discussed in chapter 1.4.2.

In order to use a MIP formulation is required to have a closed form for the customer choice probabilities conditional on a branch network. The branch network design enters the probabilities through the consideration sets in the transaction costs. Using the posterior of the parameters associated with our demand model we simulate transaction costs T_{ijt} for a single month for a sample of customers[20] S , using all the current branches, available channels, and the outside option represented as set of alternatives B . The transaction cost depends on the time and the realization of the uncertainty, so using the simulation we approximate an expected transaction cost which is independent of time. We set these values as the expected costs for customer i of using alternative j as μ_{ij}. We let μ_{i0} be the average of the transaction costs of the

---

[19] The number of branches per ZIP code can be extended. However, the complexity of the optimization grows multiplicatively with the number of branches per ZIP code.

[20] We generate a random sample of 10,000 customers to speed up these computations.

alternatives outside the consideration set, and let $\mu_{i,-1}$ be the transaction cost of the outside option. We define x_j as binary variable, taking value one when we decide to keep branch j open and zero if the branch is closed, G_{ij} is binary parameter matrix that takes the value one if the alternative j is in the consideration set of i and zero otherwise.

Using μ_{ij} to compute the choice probabilities by plugging its value in the logit form of the probabilities is an approximation. This simplification forced us to check the solution using the true consumer probabilities using the full demand model specification. Once we got the solutions using the optimization tool, we compare the attrition level of this solution, with the neighboring solutions[21], to confirm or reject the proposed solution. In case it was rejected we eliminate that solution and repeat the process. In the example exposed here, the solution found using the MIP in all cases was better than the neighboring solutions; therefore we are, at least, using a local optimal solution.

The logistic form of our demand model yields a non-linear optimization problem. Given the size and complexity of our model, a direct optimization of our model is not feasible. Instead, we re-formulate the problem as a linear programming model. Taking advantage of the IIA property of the logistic probabilities we notice that the ratio between branches in the consideration set should remain constant for each customer.

$$p_{ij} \le \frac{\exp(\mu_{ij})}{\exp(\mu_{ik})} p_{ik} + (1 - G_{ik}) \qquad \begin{array}{l} \forall i \in S \\ \forall j, k \in B \\ j \ne k \end{array} \qquad (12)$$

We need to impose a constraint to ensure that probabilities add to one:

$$\sum_j p_{ij} = 1 \qquad \forall i \qquad (13)$$

We also impose that consumers can only choose alternatives that are available:

---

[21] A neighbor solution is a solution that considers all the solutions that the same status (open/close) for all the branches, with only one exception (one branch that is open in target solution is closed or vice versa)

$$p_{ij} \leq x_j \qquad \begin{aligned} &\forall i \in S \\ &\forall j \in B \end{aligned} \tag{14}$$

Additionally, alternatives outside the consideration set have equal probability:

$$p_{ij} \leq p_{ik} + G_{ik} \qquad \begin{aligned} &\forall i \in S \\ &\forall j, k \in B \\ &j \neq k \end{aligned} \tag{15}$$

This linearization is similar to the approach shown in Haase and Müller (2014), but we adapted the idea to include considerations sets.

Following this formulation we can now use different objective functions based on the customer's service level, which we define as $\left(1 - p_{i,-1}\right)$. This measures the probability that a consumer visits a branch. Our objective is to choose a branch format such that the probability that consumers choose to visit our bank is maximized:

$$\max_{x_j} \sum_i \left(1 - p_{i,-1}\right) \tag{16}$$

In this simulation we assume that the number of branches desired by the bank is fixed at the value *NB*, which yields the constraint:

$$\sum_j x_j = NB \tag{17}$$

In other words, our objective is to maximize the service level. Assuming customers are more likely to leave the bank with lower service levels, this objective reduces customer attrition.

1.4.1 Potential extensions

The use of a mixed integer programming technique enhances the flexibility of the model. It allows the addition of further considerations that would be complex to perform using econometric techniques. Here we show two examples.

1.4.1.1 Number of branches within a range

In some areas the bank would like to set a minimum (**m**) and or a maximum (**M**) number of branches that a zip code or a collection of zip codes (**Z**) can allocate. This can be useful when regulations mandate certain limits in the number of branches per area.

$$m \leq \sum_{j \in Z} x_j \leq M \tag{18}$$

Note that this constraint can be used with list of branches not necessarily grouped by zip code, but per type, format, size, revenue or any other filter the decision maker desires.

1.4.1.2 Limit the number of changes

Another option is to penalize modifications to the network, or constrain the number of alterations to the network. This can be motivated in order to make smooth adjustments to the network to prevent a big disruption in the bank network, or to satisfy budgetary constraints.

This can be done by adding a term in the objective function to penalize the difference with the current solution. The term that needs to be added to the objective function is:

$$\Gamma(x_j - x_j)^2 \tag{19}$$

Where gamma is the penalty weight, and x_j is the original solution vector. The above constrain is non-linear, so we need to add an auxiliary variable y_j and the following set of constrains.

$$\begin{aligned} x_j - x_j \leq y_j \quad \forall j \\ x_j - x_j \leq y_j \quad \forall j \end{aligned} \tag{20}$$

And replace the term in (19) by

$$\Gamma \sum_j y_j \tag{21}$$

Another way to reduce the modifications would be through a hard restriction, limiting the number of modifications (**L**) over the entire network, or a subset Z. The constraint would be:

$$\sum_{j \in Z} \left(x_j - x_j \right)^2 \leq L \qquad (22)$$

Using the same auxiliary variable defined in (20) we replace the constraint in (22) by:

$$\sum_{j \in Z} y_j \leq L \qquad (23)$$

Other applications can include optimizing a subgroup of branches and fixing the rest, budget constraints, and human resource management, among many others.

1.4.2   Counterfactuals

Using our model estimates and the linear programming approximation just described we perform three counterfactual simulations. First, we analyzed customer attrition when the bank closes a branch. Second, we analyzed locations for opening new branches. Third, we analyzed the impact of increased mobile channel adoption by consumers.

1.4.2.1   Branch Closure

When the bank decides to close a branch, customer utility decreases, as customers need to replace the branch with another alternative that has higher transaction costs. Therefore after closure, the customer's probability of attrition increases or remains the same. In order to determine which branch to close, we can use our model to find branches to close that minimizes the negative impact on customers. To determine optimal branch closure we use the following objective function:

$$\min_b \sum_{i,b} \Delta P_{ib} \cdot CV_i \qquad (24)$$

Where ΔP_{ib} is the variation on the probability of attrition of customer i when closing branch b. CV_i is the value of customer i. The customer value is computed as the average balance in the account for the prior year[22].

---

[22] An example of the optimization problem can be found in Appendix 1.

In order to test the impact of making decisions using this approach, we look at actual branch closures made by the bank. During the period of analysis, the bank closed several branches in order to comply with competitive regulations. One closure was for a moderate size, medium-income level branch located in a suburban region that is near a commercial district. In the data, we observed a high attrition level[23] among customers who visited the branch at least once in the six months before the branch closed. The data showed close to 8.5% attrition, while the model predicted an expected attrition level of just over 6.5% among the same customers during the next six months. One reason for this discrepancy is that when banks close a branch, they tend to close more than one at a time; this combined effect might cause an increase in the attrition level.

The ZIP code connectivity has an impact not only in the number of branches but also on the type of branches it should have. Using the model we tested different levels of degree of connectivity between ZIP codes, although qualitative sets of open/close branches remain similar, in some cases it showed the same number of branches, but with different type of branches opened. This is due to the change in the number and composition customer influenced by the branch.

We expected the attrition level to be 3.2%, and the expected weighted value loss ($\Delta P_{ib} \cdot V_i$) was 27% of the value loss computed closing the bank choice this implies a 73% reduction in value loss. For reference, we can convert attrition level difference into dollars, using customer lifetime value (LTV). We assume that LTV can be computed using the following formula:

$$LTV = \sum_{t=1}^{T} m_i \frac{r_i^t}{(1+d)^t} \tag{24}$$

Based on our data we used the following parameters: average profitability per consumer $300 per year, and using a conservative valued for retention rate of 98.5% annually, and an annual discount factor of 3% with an infinite time horizon we can compute customer LTV. According to these parameters, the difference between closing the branch suggested by our model, and the branch the bank actually closed, the bank would have

---

[23] Attrition is defined as customer inactivity during the next six months.

saved $530,000 with a single decision. During that year, the bank closed more than 20 branches in different locations. Thus, if we expect similar savings per each closure, in one year the bank would have saved more than $10 million using this tool. The objective function is not minimizing attrition, but attrition weighted by customer value. If we were to minimize attrition without weights and using the same numbers as before for the same branch, the annual decision would have saved the bank more than $12 million in lifetime value during that year.

1.4.2.2   Branch Opening

When a bank opens a new branch, it is important to determine the type and location of the branch. We can use our model to predict by how much the probability of attrition decreases when a branch is added. The model was used to rank regions in terms of potential attrition gain when adding a branch. We found that locations with low connectivity and low branch density are good candidates for a branch opening. Alternatively, we found that in locations where branch presence is low, mobile channel adoption should be encouraged. Conversely, in regions with high competition, a large branch with more amenities is recommended.

The model can be used to generate a map to show the locations in need of new branches as an example shown in Figure 6
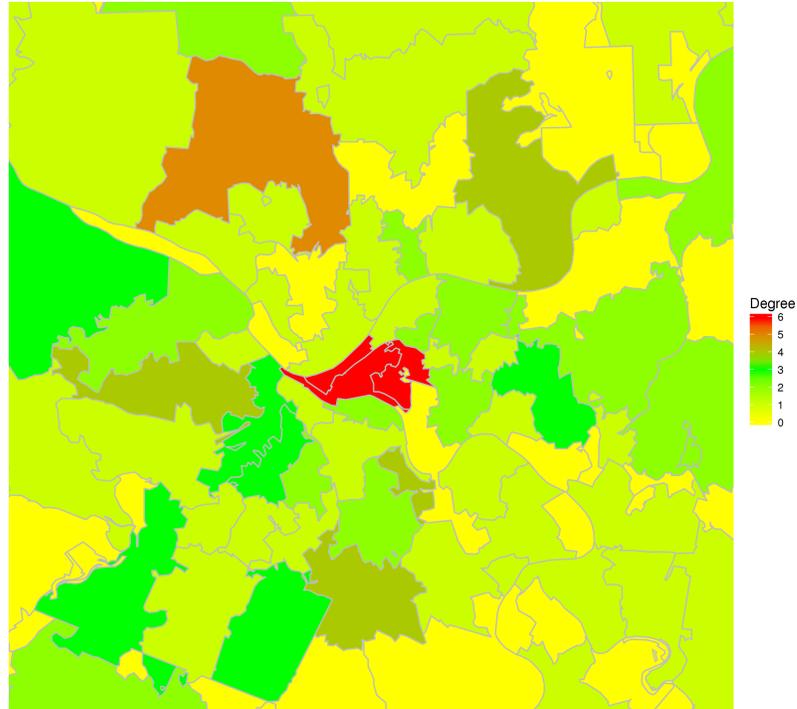
Figure 6. Example of branch requirement

In Figure 6 the branch requirements[24] are plotted. When the shaded region is closer to red, the need for branch capacity is higher. The plot in this figure is just to demonstrate a possible usage of our methodology to determine location for new branches.

1.4.2.3   Mobile Adoption

The model assumes that consumers who have never used a mobile channel are not aware of that possibility; therefore, it is not included in the consideration set. To represent an increase in the adoption of mobile channels, we increase the mobile channel adoption among consumers who were aware of the option but who had never used it before. In other words, we include this alternative in the consideration set, so consumers who find this channel more attractive than other alternatives will begin to use it. When we add a

---

24   Real branch requirement have been mask with random noise. The plot is just a demonstration of the a usage of the map tool.

mobile channel to the consideration set, customers' utility can potentially increase, since they now can substitute an alternative channel for this one if it gives them higher utility.

In our data set, the percentage of consumers who use a mobile channel is 18%. We ran counterfactuals for mobile adoption at 30%, 50%, and 70% levels of adoption. Specifically we random chose individuals in our sample and made them aware of the channel to complete the required adoption level. This can be interpreted as a campaign to educate consumers about mobile banking. Many banks are already using these campaigns to encourage and accelerate the transition of customers to the digital world. For each level of adoption, we then reevaluate the attrition probabilities. We look at small regions between one and three ZIP codes, and evaluate whether it is possible to eliminate a branch without increasing the attrition level above the level where it was before increasing the mobile adoption. When consumers adopt a mobile channel, the probability of attrition decreases more than 37% on average. For example, if a consumer had an attrition probability of 6%, then after the mobile channel is adopted this probability drops to 3.8%.

As we expected an increase in mobile adoption generates a substitution effect with transactions performed at branches, causing an excess capacity in that channel. However, even with high levels of mobile adoption, branches are always needed. This is not only because there are services that cannot be done through other channels, but also because for many customers' transactions performed using digital channels are far less preferred and lack of branches might make these customers switch to competitors.

It was surprising to find that an increase in mobile adoption can lead to an increase in demand for branches. We found two mechanisms that yield this effect. First, there is an attraction effect. Mobile channels may create the perception that the bank as a whole is more attractive. In response customers respond by switching transactions from competitors to the bank through mobile and branch channels. Second, there is a branch switching effect. Mobile channels create a distortion within the bank, switching transactions from one branch to another. Consider, for instance, a consumer who deposits large checks at a branch every Saturday. For convenience, the consumer also simultaneously performs other transactions at this branch. When a mobile channel becomes available, the consumer can perform the urgent deposit using their smartphone on Saturday but postpone the rest of their transactions until Monday when they go to work

and are near to a physical branch location. In this case, the consumer's preferred branch is different than the one that is typically used on Saturdays. In this way, mobile channels alter the location of some transactions.

The overall branch usage depends on the summation of these effects. The substitution effect seems to dominate in most scenarios. Attraction and branch switching effects tend to be weaker when mobile adoption very high or very low. However, banks are in a transition period now, since mobile adoption is not too low, nor too high, and in this scenario, these effects can have a substantial impact. Using the knowledge gleaned from our methodology, we can summarize our findings in the following table:

		Branch Density	
		High	**Low**
Competition	**Leader**	Mobile ↑ => Branch Usage ↓	Mobile ↑ => Branch Usage ↓
	Follower	Mobile ↑ => Branch Usage ↓	Mobile ↑ => Branch Usage ↑

If the bank is dominant and there is high branch density, then the substitution effect dominates due to the bank's role as leader. Also, since the market share of competitors is small there are not many new customers to attract. Conversely, when the bank is a follower and does not have many branches, there are many customers who can migrate to the bank from competitors. Because branches are sparse, the mobile channel makes the bank very attractive to new customers, so an increase in mobile channel adoption can lead to an increase in branch usage. In the other quadrants, both effects coexist, and the substitution effect tends to dominate.

1.5   Conclusion

Our analysis suggests that it is essential to make branch opening and closure decisions using the predictions of the adoption of online and mobile channels. Making decisions without considering these other channels can lead to expensive decisions and market share loses. From the model, we conclude that rural areas and low density regions can be well served with mobile channels. When facing strong competition, the

mobile channel is not enough and branches are needed. Given the current industry trend to reduce the number of branches we would recommend against closing branches in highly competitive regions.

An unexpected finding was that an increase in mobile adoption may lead to an increased demand for branch transactions. In regions where the bank is not a leader and there is low branch density, we recommend that the bank be prepared for an increase in demand at branch locations when mobile adoption increases. Depending on the situation, the bank should consider opening a branch in the region. This can be a good strategy to grow, attracting customers from competitors. As mobile usage continues to grow, our model will be increasingly useful for banks that need to determine optimal branch configurations. And with small changes, our model can be used to make decisions about other existing and new channels.

1.6   References

Ansari A., Mela C., & Neslin, S. (*2008*) Customer Channel Migration. Journal of Marketing Research: February 2008, Vol. 45, No. 1, pp. 60-76.

Arrow, K, Harris. T., & Marschak, J. (1951). Optimal Inventory Policy. Econometrica.

Blattberg, R., Kim, B., & Neslin, S. (2008). Database Marketing: Analyzing and Managing Customers. Springer-Verlag New York, eBook ISBN 978-0-387-72579-6.

Board of Governors of the Federal Reserve System. (2014) "Consumers and mobile Financial Services". *Federal Reserve Board Publications.*

Cambra-Fierroa, J., Wagner, K., Melero-Poloc, I., & Sesec, F. (2016). "Are multichannel customers really more valuable? An analysis of banking services" IJRM 1140.

Capgemini (2015), "World Retail Banking Report", http://www.worldretailbankingreport.com.

Clark, A. & Scarf, H., (1960). Optimal Policies for a Multi-Echelon Inventory Problem. Management Science.

Accenture Consulting (2013). "Celent Model Bank 2013: Case Studies of Effective use of Technology in Banking", web: https://www.accenture.com/us-en/insight-celent-model-bank-2013-banking-summary.aspx

Chintagunta, P. K., Chu, J., & Cebollada, J. (2012). Quantifying transaction costs in online/off-line grocery channel choice. Marketing Science, 31(1), 96-114.

Danaher, P. J., Wilson, I. W., & Davis, R. A. (2003). A comparison of online and offline consumer brand loyalty. *Marketing Science*, *22*(4), 461-476.

Erdem, T. & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. Marketing science, 15(1), 1-20.

Federguren, A. & Zipkin, P. (1984). A Combined Vehicle Routing and Inventory Allocation Problem. Operations Research.

Gönül, F. & Srinivasan, K. (1996). Estimating the impact of consumer expectations of coupons on purchase behavior: A dynamic structural model. Marketing Science, 15(3), 262-279.

Haase, Knut, and Sven Müller. (2014). "A comparison of linear reformulations for multinomial logit choice probabilities in facility location models." European Journal of Operational Research 232, no. 3 689-691.

Harris, F. (1913). How many parts to make at once. The Magazine of Management.

Hitt, L. M. & Frei, F. X. (2002). Do better customers utilize electronic distribution channels? The case of PC banking. Management Science, 48(6), 732-748.

Hoehle, H., & Venkatesh, V. (2015). Mobile application usability: conceptualization and instrument development. Mis Quarterly, 39(2), 435-472.

Imai, S., Jain, N., & Ching, A. (2009). Bayesian estimation of dynamic discrete choice models. Econometrica.

Johansen, S.G. & Melchiors, P. (2003). Can-order policy for the periodic-review joint replenishment problem. Journal of the Operational Research Society.

Kumar, V. & Venkatesan, R. (2005). Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior. Journal of Interactive Marketing, 19(2), 44-62.

Kushwaha, T. & Shankar, V. (2013) Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics. Journal of Marketing: July 2013, Vol. 77, No. 4, pp. 67-85.

Liu, X., A.L. Montgomery and K. Srinivasan (2016). Optimizing Bank Overdraft Fees with Big Data. Carnegie Mellon University, Working Paper.

Magnac, T., & Thesmar, D. (2002). Identifying dynamic discrete decision processes. Econometrica, 70(2), 801-816.

Neiswanger, W., Wang, C., & Xing, E. (2013). Asymptotically exact, embarrassingly parallel MCMC. arXiv preprint arXiv:1311.4780.

Neslin, S., Grewal, D., Leghorn, R., Shankar, V., Teerling, M., Thomas, J., & Verhoef, P. (2006). Challenges and Opportunities in Multichannel Customer Management. Journal of Service Research, Volume 9, No. 2.

Laukkanen, T. (2007),"Internet vs mobile banking: comparing customer value perceptions", Business Process Management Journal, Vol. 13 Iss 6 pp. 788 – 797.

Ozkaya, B.Y., Gurler, U., & Berk, E. (2006). The stochastic joint replenishment problem: A new policy, analysis, and insights. Naval Research Logistics.

Roberts, J. H. & Lattin, J. M. (1991). Development and testing of a model of consideration set composition. Journal of Marketing Research, 429-440.

Rust, J. (1994). Structural estimation of Markov decision processes. Handbook of econometrics, 4, 3081-3143.

Shankar, V., Smith, A. K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. International journal of research in marketing, 20(2), 153-175.

Shocker, A. D., Ben-Akiva, M., Boccara, B., & Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. Marketing letters, 2(3), 181-197.

Song, I., & Chintagunta, P. K. (2003). A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. Quantitative Marketing and Economics, 1(4), 371-407.

Sun, B. (2005). Promotion Effect on Endogenous Consumption. Marketing Science 24(3) 430–443.

Taha, A., Jahed, D. H., Ahmad, M. N., & Zakaria, N. H. (2013, November). Antecedents of customer satisfaction in mobile commerce: A systematic literature review. In 2013 International Conference on Research and Innovation in Information Systems (ICRIIS) (pp. 554-558). IEEE.

Takeyasu, K. (2008). Mobile marketing. Information Systems Research Methods, Epistemology, and Applications, 328.

Valentini, S., Montaguti, E., Neslin, S. (2011) Decision Process Evolution in Customer Channel Choice. Journal of Marketing: November 2011, Vol. 75, No. 6, pp. 72-86.

Venkatesan, R., Kumar, V., & Ravishanker, N. (2007). Multichannel shopping: causes and consequences. Journal of Marketing, 71(2), 114-132.

Viswanathan, S. (1997) Periodic review (s, S) policies for joint replenishment inventory systems. Management Science.

Wallace, D. W., Giese, J. L., & Johnson, J. L. (2004). Customer retailer loyalty in the context of multiple channel strategies. Journal of retailing, 80(4), 249-263.

Yang, B., & Ching, A. T. (2013). Dynamics of consumer adoption of financial innovation: The case of ATM cards. Management Science, 60(4), 903-922.

Chapter 2

Chapter 2: Influencing Product Competition through Shelf Design

Arranging products on a retail shelf can directly influence consumer purchases by making products easy to find, and also indirectly by influencing the competitive set of products that a consumer considers. In this research we develop a model of product choice that accounts for shelf position as well as price and promotion. Using this model we show that when products are placed closer together competition between products increase. Thus shelf design can encourage consumers to purchase products that would not have previously been purchased. The motivation for this result is that consumer search is costly and consumers focus their search on local neighborhoods that are influenced by shelf position. Since search is costly, consumers may not exhaust all possibilities, this in turn means that position is an important determinant of consideration. Formally we model this behavior as a sequential consideration set. To begin the search consumers initial starting points are influenced by colors, favorite brand or the closest shelf edge. Consumers shift their focus to neighboring products, in a sequential fashion, which increases the set of products considered, until finally making the purchase decision. In other words the shelf arrangement spatially influences consideration sets. Applying our results to shelf experiments conducted in a supermarket retailer we find that demand is greatly impacted by shelf position and retailers can create plan-o-grams that can shift demand from one product to another which can potentially increases profits by up to 15% through improved plan-o-grams. Our focus on using shelf design to stimulate competition contrasts with past research on shelf design that has focused mainly on cost minimization.

2.1   Introduction

Experimental work in retail shelf design has a long history in marketing research (Pauli and Hoecker 1952). Early studies were largely experiment and focused on establishing space to movement (Cox 1970, Anderson 1979). Before the 1980's most research attention focused on reducing shelving operation costs, with little consideration for the demand side. In the 1980s models that incorporated demand into shelf design was first proposed by Corstjens and Doyle (1981, 1983) who included demand elasticities within a space allocation optimization problem, Bultez and Naert (1988) developed a decision tool (S.H.A.R.P.) to help managers optimize the shelf space allocation. In the 1990's technological developments like electronic scanners, digital data bases, and the increase in computing capacity allowed researchers to study other aspects of shelf design. More complex demand structures were used (Allenby 1989) and aspects like assortment and variety were studied (Borin, Farris, and Freeland 1994; Hoch, Bradlow and Wanskin 1999). The improvement in technology and research shifted the attention from minimizing costs to maximizing profits (Van Ryzin and Mahajan, 1999).

In the late 1990's, new technology was used to detect where consumers where looking when facing a shelf. This allowed researchers to understand how products positioning affects its demand (Pieters and Warlop 1998). Theoretical and empirical work by Sayman, Hoch and Raju (2002) showed that the location of store brands was strongly influenced by national brand location. Chandon (2007) showed that shelf positions have important demand implications and argued that "Unless you're Coca-Cola, it's important to be visible on the shelves," concluding that shelf position can be even more important than the product's brand name.

Stüttgen, Boatwright and Monroe (2011) used eye-tracking data to note that consumers tend to compare products in neighborhoods defined by spatial proximity. Consulting firms like Usercentric and GMO JMI have found similar effects also using eye-tracking technology and create heat-maps like in Figure 7. They observe consumers focusing on a particular product, and then looking at products that are located nearby, before making the purchase decision.

Figure 7. Eye tracking heat map for the cereal shelf as measured by Usercenter (User Centric[25]).

A primary contribution of our research is to show how a physical, spatial search of the shelf can influence consideration. As such our research contributes to the extensive research on consideration set formation. Although others have mentioned shelf-position as having an impact on consideration, to our knowledge no formal models that estimate this effect have been made. A challenge that we address in our research is that we estimate the consideration effects using aggregate, weekly, store-level data that managers readily have available. A secondary contribution of our research is to develop implications for shelf-design based upon our model. Specifically, we can explain why products within a brand family would wish to be blocked together while store-brands would want to be located near their targeted national brand.

2.2   Model

To model consumer purchase we assume that consumers break their purchase decision into two steps. The first step is the construction of a consideration set (Hauser and Wernerfelt 1990, Roberts and Lattin 1991). The primary motivation is that consumers have a large variety of choices and use screening rules to reduce the number of alternatives that must be evaluated (Hoyer 1984). The second step is the choice of the best item (or potentially no choice if an outside good is permitted) given the consideration set. Mehta,

---

[25] User Centric is a digital consulting company. It was named among top100 digital companies in Chicago. Usercentric.com

Rajiv and Srinivasan (2003), Pancras (2010), Draganska and Klapper (2010), and Kim, Albuquerque and Bronnenberg (2010) have consider the effect of the consideration set on demand estimation and managerial decision making.

Our modeling framework follows Berry, Levinsohn and Pakes (1995) except for our inclusion of consideration sets. There have been a number of recent extensions that have included consideration sets into this framework as well (Bruno and Vilcassim 2008, Draganska and Klapper 2010, Kim, Albuquerque and Bronnenberg 2010, Seiler 2012). Our point of departure with these papers is that we assume that the consideration set is determined by a search of the product shelf. For example, Mehta, Rajiv, and Srinivasan (2003) model the consideration set as a function of advertising and past consumption on choice. We conjecture that previous researchers have not included shelf position in the consideration set due to the lack of observation and variation in shelf design. Our dataset has a series of controlled shelving experiments that is uniquely able to address this weakness. Hence, although our model specification is unique in the literature, our primary purpose is to consider the substantive issue of shelf position.

2.2.1   An aggregate model of market share with latent consideration sets

To predict purchase probability of product *j* by consumer *i* during week *t* for store *s*, we use a random utility model:

$$U_{ijts} = \mathbf{z}_j'\boldsymbol{\alpha} + \mathbf{x}_{jts}'\boldsymbol{\beta}_{is} + \xi_{jts} + \varepsilon_{ijts} \, , \tag{1}$$

Where $\mathbf{z}_j$ is a vector of product specific dummies (e.g., brand, size, color, or design), $\mathbf{x}_{jts}$ is a vector of *a* observed product attributes (e.g., price, promotion), ξ_{jts} is a random shock that is common across consumers for a given product, week, and store combination, and ε_{ijts} is an idiosyncratic error which follows an extreme value distribution. There are *M* products (*j*=0,…, *M*) in our category, *S* stores (*s* = 1,…,*S*), and *T* weeks (*t* = 1,…,*T*). We include an outside good, which we denote by *j*=0, to capture the utility from not purchasing in the category. To identify our model we set the utility of the outside option to be zero (U_{i0ts}=0).

We assume that consumers consider only a subset of products, which we refer to as a consideration set C_{its}. We do not directly observe the consideration set, but infer some probability that a consumer uses this

consideration set, which we denote as $p(C_{its})$. The construction of this probability is discussed in the following subsection. We can show that for consumer i the probability of selecting an item in a given consideration set C_{its}, follows the usual multinomial logit choice model, if the item belongs to C_{its}, and otherwise is zero.

$$\Pr\left(U_{ijts} \geq \max\left(\{U_{ikts} \mid \forall k \in C_{its}\}\right)\right) = \frac{\exp\left\{\mathbf{z}_j{}'\boldsymbol{\alpha} + \mathbf{x}_{jts}{}'\boldsymbol{\beta}_{is} + \xi_{jts}\right\}}{\sum_{k \in C_{its}} \exp\left\{\mathbf{z}_k{}'\boldsymbol{\alpha} + \mathbf{x}_{kts}{}'\boldsymbol{\beta}_{is} + \xi_{kts}\right\}} \tag{2}$$

We assume that the outside good, denoted as $j = 0$, is always included in the consideration sets: $p(0 \in C_{its}) = 1 \quad \forall C_{its}$.

Consumer price and attribute preferences are captured by the parameters $\boldsymbol{\beta}_{is}$ which we assume to be normally distributed with its mean linearly related to the observed vector of consumer demographics ($\mathbf{d}_s$) of the store's trading area[26]:

$$\boldsymbol{\beta}_{is} \sim N(\boldsymbol{\Gamma}\mathbf{d}_s, \boldsymbol{\Omega}) \tag{3}$$

We do not observe individual consumers, so we must integrate over the individuals to predict market share. Marginalizing over consideration sets and also the common consumer shocks yields the predictions of market share (w_{jts}):

$$w_{jts} = \sum_{C_{its}} \int_{\boldsymbol{\beta}_{is}} \frac{\exp\left\{\mathbf{z}_j{}'\boldsymbol{\alpha} + \mathbf{x}_{jts}{}'\boldsymbol{\beta}_{is} + \xi_{jts}\right\} \cdot I(j \in C_{its})}{\sum_{k \in C_{its}} \exp\left\{\mathbf{z}_k{}'\boldsymbol{\alpha} + \mathbf{x}_{kts}{}'\boldsymbol{\beta}_{is} + \xi_{kts}\right\}} p(\boldsymbol{\beta}_{is}) \, p(C_{its}) \, d\boldsymbol{\beta}_{is} \ . \tag{4}$$

However, in our case this integral is not analytically known, so we approximate it with P random draws from (3), which we denote as $\boldsymbol{\beta}_{is}^d$, where $d=1,\ldots,D$:

---

[26] A store's trading area in our dataset is determined by a market research firm which looks for a contiguous area around the store which it predicts most consumers of the store are drawn from. These trading areas are approximately a zip code and are discussed by Hoch et al. (1995).

$$w_{jts} \approx \frac{1}{P} \sum_{C_{its}} \sum_{d=1}^{D} \frac{\exp\left\{ \mathbf{z}_j{}'\boldsymbol{\alpha} + \mathbf{x}_{jts}{}'\boldsymbol{\beta}_{is}^d + \xi_{jts} \right\} I(j \in C_{its}^u)}{\sum_{k \in C_{its}} \exp\left\{ \mathbf{z}_k{}'\boldsymbol{\alpha} + \mathbf{x}_{kts}{}'\boldsymbol{\beta}_{is}^d + \xi_{kts} \right\}} p\left(C_{its} \right) .$$

(5)

The estimation procedure is discussed in the Appendix.

2.2.2 Consideration Sets Based Upon Shelf Position

Unfortunately we do not directly observe consumer movements[27], but our construction of the consideration set is motivated by the physical and eye movement of an individual consumer through the aisle to construct a consideration set. Chandon et al (2006) show that eye movements are known to correlate with product consideration, and such movements can be modeled as a Markov model (Stüttgen, Boatwright and Monroe 2011). Although we do not directly observe the consideration sets, we propose the probability that a consumer considers a product can be modeled as composition of three processes: (1) the consumer's approach to the shelf, (2) information gathering in a limited physical neighborhood of products and finally (3) a transition process between different neighborhoods in the shelf.

1. Approach to the shelf. To initialize the category search we assume that consumers approach to the shelf to particular product (**initial focal product**) in the shelf. We assume that a product that was purchased in a previous occasion or a familiar product is very likely to be used as starting point to explore the category. We represent this process as probability distribution over the entire set of products per consumer and store $\boldsymbol{\pi}_{is0}$, initializing these parameters with market shares update them using individual predicted market shares from (5).

2. Information gathering process. When consumers look at the focal product they gather information relevant for the purchase decision, like price, and then they compare it with nearby products. Studies using eye tracking like Stüttgen et al (2011) show that consumers gather information about many products in a category before making the purchase decision, even if they begin the search in the product they

---

[27] In current retail environments it is not possible to routinely collect this data, although in the future we expect improvements in eye tracking cameras and lower costs may make it possible to collect this data. If this new data becomes available then instead of thinking of the considerations as unobserved it would allow us to treat them as partially observed, which can be directly incorporated into our model.

will finally purchase. We represent the neighborhood of products considered simultaneously with the focal product j by the set $\mathcal{BC}_{jts}$ that depends on the vertical and horizontal distance from the focal product, i.e $i \in \mathcal{BC}_{jts}$ if $d_{ij} < D$. Where d_{ij} is a distance function between products i and j. Several specifications for the distance functions were tested, but the difference was negligible in the estimation. Therefore as a parsimonious representation we decided to use the rectilinear distance which is the sum the horizontal and vertical distance. We also penalize the vertical distance in order to make the horizontal distance more important to be consistent with Pieters and Warlop (1998):

$$d_{ij} = \Delta x + \varphi \cdot \Delta y \qquad (6)$$

3. Shelf Transition process. After finishing comparisons within the consideration set, consumers may want to consider additional products by shifting to another shelf location. The new location is represented by a new focal product and its neighborhood. The resulting consideration set is the union of the current set and the new basic consideration set. The probability of moving from focal product j to j' for store s at week t is represented by the transition matrix $\mathcal{T}_{ts}$ that depends on the product location in the shelf. Each element of the matrix is calculated using the logit form as follows:

$$\mathcal{T}_{ts}(j,j') = \frac{\exp\left\{d_{jj'} + \phi \cdot \text{Display}_{j'}\right\}}{\sum_{k \neq j} \exp\left\{d_{jk} + \phi \cdot \text{Display}_{k}\right\}} \quad \forall j \neq j'$$

$$\text{where}: d_{jj'} = \log^{-1}\left(d_{jj'}\right) \text{ and } \mathcal{T}_{ts}(j,j) = 0 \qquad (7)$$

In this way incorporate the assumption that is more likely to move to a near location in the shelf and that is also more likely to move toward a product with display.

These transitions from one focal product to another create a set of products resulting from the union of all the previous basic consideration set associated to the focal products visited. Thus the products considered after λ_i steps by consumer i at store s is:

$$C_{ist} = \bigcup_{k=1}^{\lambda_{is}} \mathcal{BC}_{j(k),ts} \qquad (8)$$

Where $j_{(k)}$ is the focal product visited at step k. Again, we do not observe the number of steps, so we use a random approach assuming a truncated Poisson distribution:

$$\lambda_{is} \sim \text{Poisson}(\overline{\lambda}_s)\,\text{I}[\lambda_{is} \leq K] \tag{9}$$

Where K is the maximum number of steps that a consumer performs.

To compute the distribution of reaching any consideration set used in (5), we use a recursive equation to compute the probability of increasing the consideration set C_{ist} with focal product j by moving to j' in one step is given by:

$$p(C_{ist}^{'}) = \mathcal{T}_{ts}\left(j_{(\lambda)}, j_{(\lambda+1)}^{'}\right)p(C_{ist}) \tag{10}$$

Where $C_{ist}^{'} = C_{ist} \cup B_{j'}$. For a given path $\mathcal{K}_{\lambda_{is}}$ of λ_{is} steps, defined for the sequence of products $\left\{j_{(k)}\right\}_{k=1}^{\lambda_{is}}$ that generates the consideration set $C_{its}^{*} = \overset{\lambda_{is}}{\underset{k=1}{\bigcup}}BC_{j(k)}$, we can define its probability as:

$$p\left(\mathcal{K}_{\lambda_{is}}\right) = \prod_{k=1}^{\lambda_{is}}\left[\mathcal{T}_{ts}\left(j_{(k)}, j_{(k+1)}^{'}\right)\right]\pi_{is0\,j_{(1)}} \tag{11}$$

Now, to compute the probability of any consideration set C_{ist}^{*} we have to sum over all the paths that result in this consideration set:

$$p(C_{ist}^{*}\mid\lambda_{is}) = \sum_{\mathcal{K}_{\lambda_{is}}\,:\,C_{its}^{*}\,=\,\overset{\lambda_{is}}{\underset{k=1}{\bigcup}}BC_{j(k)}}p\left(\mathcal{K}_{\lambda_{is}}\right) \tag{12}$$

Finally to compute the probability of reaching any consideration set we integrate over the distribution of λ_{is} in (12) to give the marginal probability of the consideration set:

$$p\left(C_{its}\right) = \sum_{\lambda_{is}=1}^{K}p\left(C_{its}\mid\lambda_{is}\right)\cdot p\left(\lambda_{is}\right) \tag{13}$$

A problem that arises in the estimation procedure is the number of possible paths that can lead to the same consideration set grows combinatorically with the number of products. In order to solve this issue

we approximate (5) with simulation paths to draw $C^u_{ist} \sim p(C_{ist})$ using (13), and then use that result to

compute (5) numerically as follows:

$$w_{jts} \approx \frac{1}{U} \sum_{C^u_{its}} \frac{1}{P} \sum_{d=1}^{D} \frac{\exp\{\mathbf{z}_j{'}\boldsymbol{\alpha}+\mathbf{x}_{jts}{'}\boldsymbol{\beta}^d_{is}+\xi_{jts}\} I(j \in C^u_{its})}{\sum_{k \in C^u_{its}} \exp\{\mathbf{z}_k{'}\boldsymbol{\alpha}+\mathbf{x}_{kts}{'}\boldsymbol{\beta}^d_{is}+\xi_{kts}\}} \tag{14}$$

We discuss each of these components in more detail, and then illustrate this model with an example in the

next subsection.

Initial Probability. We propose that consumers have some probability of starting at an initial product

which is proportional to the market share of the product. There are two reasons for this specification. The

first is that consumers are likely to start at commonly purchased products, which is the product's market

share. For example, in the soft drink category it is likely that a consumer starts at Coca-Cola. A second

reason is that before search has begun consumers do not have any information about where to begin, hence

they would set rational expectations about what is the best product in the category, which again would be the

product's market share.

There are many possible alternative approaches for choosing the initial probability vector. The initial

probability could be influenced by location. For example, half consumers likely start at the left end of the

shelf and the other half likely start at the right end of the shelf. Alternatively, consumers may start at their

favorite product, last product purchased, randomly start their search at any position in the shelf, or start at

those items that have in-store displays. We tested these various rules and found that found that the best rule

for initializing position is proportional to the past market share. In other words consumers probabilistically

start their search at the last product that they purchased.

Transition Probability. The transition probability matrix captures how consumers move their attention

from one area to the next. In our construction of the transition matrix we favor consumers making more

horizontal movements than vertical movements (Pieters and Warlop 1998). Our intuition is that there is a

greater probability that consumers move to an adjacent shelf location than a shelf location farther away.

Therefore our transition matrix is constructed by assuming that the probability of moving to a new shelf

location is proportional to the inverse of the distance between the current and the future location. To illustrate the transition matrix consider a shelf with three products, we define the distance between i and j as d_{ij}, where $d_{ij} = 1$ and $d_{ik} = 2$. The probability of moving from i to j and k are 2/3 and 1/3 respectively. To define the transition probability matrix, first we need to define the distance function and distance D that will define our Basic Consideration Sets. Second, we compute the distances among all physical centers of the Basic Consideration Sets.

Stopping Rule. As we mentioned before, the number of movements follows a truncated Poisson distribution. The number of movements can also be thought of as a stopping rule for search in the category. We assume that there are segments of consumers that are very impatient and spend very little time searching for products in the shelf, and there are other kinds of consumers that have more time or finding a better product gives them high utility and performs a more detailed search in the shelf. We don't observe these segments, or individual data, but we can estimate the proportion of consumers that behave in one way or another by estimating the probability of belonging to these latent segments.

We assume that consumers decide a priori how long to search and not to adjust it based upon expected price. This formulation is consistent with De Los Santos, Hortacsu, and Wildenbeest (2012), but contrasts with Kim, Albuquerque and Bronnenberg (2010), who develop an optimal search pattern. We are sympathetic to endogenous stopping rules, but do not implement them because of their complexity and likelihood that consumers use heuristics in deciding how long to shop. The decision to stop is effortful, especially if incomplete information about prices is assumed at each step of the decision. Therefore it is logical to think of heuristics that consumers may employ to balance the tradeoff between the accuracy and effort in deciding the appropriate heuristic (Bettman, Payne and Johnson 1993). One could view a fixed stopping criterion as an optimal search heuristic when cognitive effort is present. Given the lack of individual search information we believe it is difficult to properly test this proposition and as a consequence assume that stopping length is randomly chosen.

In **Figure 8**, we can see an example of a consumer search. In this example the search begin at the edge of the shelf eye level on a focal product. Immediately a consumer can compare this product with its basic consideration set.

Figure 8. Focal product and its basic consideration set

The search can finish there, in which case the consumer will but the product, within the consideration set with higher random utility ($\lambda_{is} = 1$), or continue the search. Now the set of products that will be included will depend on their proximity to the beginning of the search and how long is the search, as shown in **Figure 9**
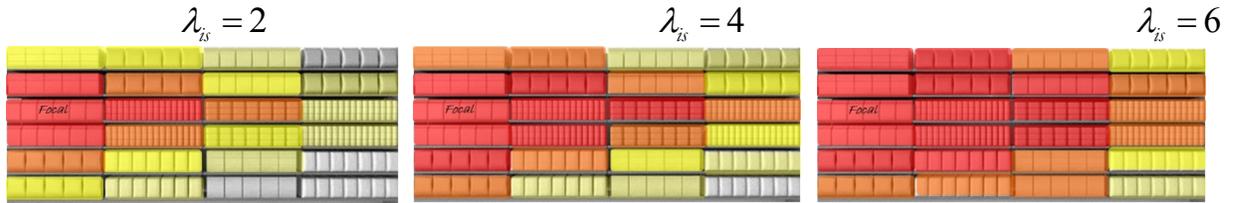
Figure 9. Consideration set with $\lambda_{is} = 2, 4$ and 6.

In this way, the model induce a covariance matrix that relays on product relative location to other products, which has allows a better understanding the influence of product display in the demand.

2.2.3 Alternative Models with Fixed Positions and Nested Logit Models

Fixed Positions. An alternative model formulation that follows Dreze, Hoch, and Purk (1994) is to assume that there are fixed shelf position effects. Specifically, we use the following attribute vector in our model:

$$U_{ijts} = \alpha_{js} + \mathbf{x}_{jts}' \boldsymbol{\beta}_{is} + \phi_1 xpos_{jts} + \phi_2 xpos_{jts}^2 + \phi_3 ypos_{jts} + \phi_4 ypos_{jts}^2 + \xi_{jts} + \varepsilon_{ijts} , \qquad (15)$$

where *xpos* and *ypos* measure the position of the product on the shelf from the left and bottom edge of the category respectively. We point out that often one cannot estimate this model since the inclusion of product intercepts and the lack of variability in shelf positions means that the shelf position effects are indistinguishable from the shelf position. Hence having some variation in shelf-position is critical for estimation.
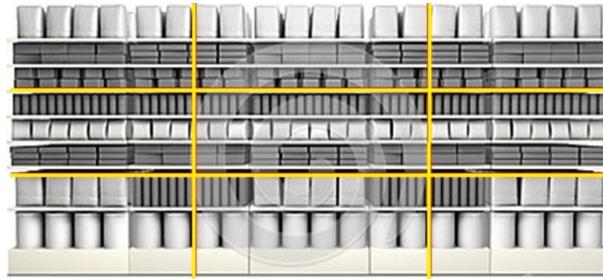
Figure 10. Location based consideration sets of the nested logit model.

Nested Logit Model. Another alternative model is the nested logit, in which the nests are based upon shelf position. We propose that shelves can be divided into three vertical and three horizontal locations to yield nine areas as illustrated in **Figure 10.** Intuitively this model is similar to ours in the sense that consumers first decide upon an area of the shelf, and then deliberate amongst the items within this location.

2.3 Model Discussion

In this chapter we describe the data used to estimate our model, present the result for the category Paper Towel, and discuss the price elasticities inferred from the model.

2.3.1 Data Description

We begin our study in the category *Paper Towel.* This category is particularly interesting to test our hypothesis. The size of the items is physically large and it requires consumers to move through the aisle and not just their eyes, therefore we conjecture a high search cost. In addition this is a mature category with high retailer inventory cost, it is a frequently purchase category (approximately once every two weeks). Quality tiers exist (low to medium vs premium) and little difference in demand exist vs Suburban and urban areas. Brands have multipack or single product versions.

Variation in product position is key in our analysis; in this regard, thee shelf designs were designed, in addition to the control design, which kept the original design at the stores.

Control. The stores originally group the products by brand. Within each brand block, all formats were together, and premium labels were scattered as in **Figure 11**

Hi-Dri	Other	Gala
Brawny	Scott	
Bounty	Private Label	Viva

Figure 11. Control design –grouped by brand.

Treatment 1 –Quality tier 1– organizing by quality and size by placing single rolls on top shelf, multipacks in the middle and price brands on the bottom as in **Figure 12**

Premium
Multi-pack/Big roll
Price/Mid price

Figure 12. Shelf treatment 1

Treatment 2 –Quality tier 2– Organizing by quality and size placing price/mid-price and single rolls on the top shelf, multipacks in the middle and premium on the bottom as in **Figure 13**

Price/Mid price
Multi-pack/Big roll
Premium

Figure 13. Shelf treatment 2

Treatment 3 –Difficult to shop– Separating the single jumbo size towels from the multipacks, the jumbo size organized by quality from top to bottom and the multipacks organized by size from top to bottom, as in **Figure 14**.

	Jumbo	Multi-pack/Big roll
Premium		
Mid-price		
Price		

Figure 14. Shelf treatment 3.

In this experiment 60 stores participated, 15 stores were randomly assigned to each shelf design. The experiment included 17 weeks of data before and 16 after the treatment was implemented. So we have all 60 stores with the control design for 17 weeks, then we 15 stores in each design for 16 weeks.

To avoid confusion with size of stores, small adjustments were made to accommodate all designs in all type of stores and each 15 store group were balanced in terms of size.

2.3.2   Estimation results

To test the performance of our model and how it compares with the alternatives we use three designs to estimate the model and then validate the prediction of our model with respect to the fourth design that was left out.

In **Table 5** we estimate the models leaving the treatment one for validation, in **Table 6** we left the treatment 2 for validation and similarly in **Table 7** we left treatment 3 for validation. We also estimate the validation leaving the control out, as show in **Table 8**.

Parameter	Simulated Consideration Sets			Nested Logit			Logit with fix position		
	Est.	Std.		Est.	Std.		Est.	Std.	
Hi-Dri	0.047	0.180		0.420	0.158	***	-0.204	0.196	
Bounty	0.276	0.179		0.607	0.162	***	-0.138	0.189	
Dominick	0.688	0.250	***	1.240	0.222	***	0.547	0.260	**
Brawny	-0.596	0.257	**	-0.510	0.226	**	-0.819	0.265	***
Gala	-1.424	0.964		-1.576	0.920	*	-2.310	0.952	**
Scott	-0.034	0.154		-0.083	0.136		-0.311	0.165	*
Viva	0.151	0.183		0.401	0.162	**	-0.370	0.192	*
Mardi Gras	0.096	0.264		0.623	0.238	***	-0.054	0.270	
Green Forest	0.325	0.297		0.203	0.262		0.091	0.310	
Bolt	-1.129	0.421	***	-1.085	0.391	***	-1.714	0.421	***
Multi Count	0.486	0.377		0.858	0.463	*	0.683	0.386	*
Design	0.441	0.058	***	0.516	0.054	***	0.623	0.054	***
Big Format	0.076	0.095		0.239	0.091	***	0.416	0.094	***
Feature	0.324	0.030	***	0.743	0.026	***	0.198	0.031	***
Display	-0.420	0.006	***	-0.786	0.006	***	-0.354	0.006	***
Price	-0.559	0.125	***	-0.967	0.159	***	-0.637	0.132	***
Time(h)	6.7			0.00			0.4		
Loglikelihood	**1080.04**			**1148.46**			**1115.70**		
LLH on Hold out	**121.48**			**134.80**			**132.09**		
Mean Square Error	**7.34E-06**			**6.11E-05**			**5.60E-05**		

Table 5. Estimates with Treatment 1 held out

Parameter	Simulated Consideration Sets			Nested Logit			Logit with fix position		
	Est.	Std.		Est.	Std.		Est.	Std.	
Hi-Dri	0.033	0.185		0.500	0.165	***	-0.099	0.207	
Bounty	0.294	0.183		0.626	0.168	***	0.175	0.207	
Dominick	0.546	0.258	**	1.344	0.230	***	0.085	0.289	
Brawny	-0.543	0.259	**	-0.492	0.230	**	-0.137	0.300	
Gala	-1.366	0.980		-1.538	0.934	*	-0.927	1.046	
Scott	-0.086	0.156		-0.064	0.140		-0.067	0.175	
Viva	0.136	0.186		0.447	0.167	***	0.107	0.209	
Mardi Gras	0.025	0.272		0.681	0.247	***	0.021	0.295	
Green Forest	0.224	0.298		0.329	0.264		-0.244	0.329	
Bolt	-0.994	0.424	**	-1.063	0.398	***	-0.058	0.493	
Multi Count	0.477	0.389		0.814	0.467	*	0.542	0.391	
Design	0.368	0.059	***	0.524	0.055	***	0.105	0.067	
Big Format	-0.072	0.095		0.242	0.092	***	-0.103	0.099	
Feature	0.307	0.030	***	0.714	0.027	***	0.068	0.032	**
Display	-0.406	0.006	***	-0.801	0.006	***	-0.302	0.007	***
Price	-0.552	0.131	***	-0.914	0.161	***	-0.388	0.123	***
Time(h)	6.1			0.00			0.4		
Loglikelihood	**1078.26**			**1145.68**			**1094.34**		
LLH on Hold out	**123.41**			**137.67**			**135.62**		
Mean square error	**7.96E-06**			**8.48E-05**			**9.18E-05**		

Table 6. Estimates with Treatment 2 held out

Parameter	Simulated Consideration Sets			Nested Logit			Logit with fix position		
	Est.	Std.		Est.	Std.		Est.	Std.	
Hi-Dri	0.023	0.179		0.435	0.156	***	-0.129	0.199	
Bounty	0.248	0.180		0.604	0.163	***	0.113	0.204	
Dominick	0.557	0.248	**	1.297	0.220	***	0.089	0.277	
Brawny	-0.590	0.259	**	-0.536	0.226	**	-0.149	0.298	
Gala	-1.327	0.992		-1.618	0.939	*	-0.532	1.104	
Scott	-0.112	0.153		-0.118	0.135		-0.080	0.172	
Viva	0.124	0.183		0.397	0.162	**	0.095	0.205	
Mardi Gras	0.012	0.264		0.622	0.238	***	0.006	0.285	
Green Forest	0.185	0.296		0.193	0.262		-0.259	0.327	
Bolt	-1.065	0.420	**	-1.112	0.391	***	-0.086	0.488	
Multi Count	0.468	0.374		0.847	0.456	*	0.514	0.380	
Design	0.380	0.058	***	0.531	0.055	***	0.116	0.065	*
Big Format	-0.026	0.093		0.260	0.091	***	-0.079	0.097	
Feature	0.320	0.029	***	0.756	0.026	***	0.067	0.032	**
Display	-0.409	0.006	***	-0.792	0.006	***	-0.304	0.006	***
Price	-0.532	0.124	***	-0.944	0.156	***	-0.366	0.118	***
Time(h)	5.3			0.00			0.5		
Loglikelihood	**1180.73**			**1148.74**			**1096.11**		
LLH on Hold out	**128.81**			**134.49**			**129.23**		
Mean square error	**4.10E-05**			**6.00E-05**			**5.22E-05**		

Table 7. Estimates with Treatment 3 held out

Parameter	Simulated Consideration Sets			Nested Logit			Logit with fix position		
	Est.	Std.		Est.	Std.		Est.	Std.	
Hi-Dri	-0.432	0.481		0.095	0.419		-0.551	0.540	
Bounty	0.178	0.558		0.774	0.496		0.073	0.615	
Dominick	0.589	0.683		1.494	0.607	**	0.246	0.744	
Brawny	-1.053	0.784		-0.805	0.715		-0.408	0.928	
Gala	-1.588	3.112		-1.782	2.960		-1.679	3.205	
Scott	-0.408	0.447		-0.074	0.388		-0.369	0.503	
Viva	0.376	0.584		0.632	0.499		0.256	0.656	
Mardi Gras	0.376	0.830		0.717	0.760		0.373	0.888	
Green Forest	-0.149	0.895		0.329	0.801		-0.305	0.977	
Bolt	-1.948	1.939		-1.744	1.853		-0.526	2.342	
Multi Count	1.164	1.076		0.769	1.236		1.303	1.092	**
Design	0.501	0.163	***	0.549	0.167	***	0.343	0.169	
Big Format	0.074	0.241		0.290	0.240		-0.039	0.248	**
Feature	0.450	0.080	***	1.008	0.070	***	0.187	0.086	***
Display	-0.429	0.018	***	-0.826	0.016	***	-0.334	0.019	*
Price	-0.632	0.334	*	-0.890	0.393	**	-0.537	0.324	
Time(h)	2.6			0.00			0.3		
Loglikelihood	**-370.52**			**-404.88**			**-397.43**		
LLH on holdout	**-849.06**			**-883.31**			**-861.31**		
Mean square error	**4.86E-05**			**5.97E-05**			**6.02E-05**		

Table 8. Estimates with Control group held out

In all the analysis the consideration our approach fits the data better, and it is particularly superior when predicting a new design, suggesting that consumers behave closer to the proposed theory.

Logit model with position behave better than nested logit, but the difference is minor when predicting new design.

As a difference with other studies like Chandon et al 2009, this model does not put enfaces on a specific location on the shelf; instead, it uses product location with respect to its competition.

Another reason to use the logit with product location as controls is to test how good a model with shelf location performs. According to our findings, our theory fits the data significantly better.

2.3.3   Elasticity Structure

To better understand our model we consider an analysis of the implied price elasticities. We derive the properties of this model by analyzing the properties of w_{jts} in (14). This function can be interpreted as the market share of product j at a given purchase occasion. We observe that that if the price sensitivity is

negative, $\dfrac{\partial w_{jts}}{\partial p_l} < 0$ for $l \neq j$ with equality only when $l \notin C_{its}^u$. Also that $\dfrac{\partial w_{jts}}{\partial p_j} < 0 \ \forall j$. These properties

are intuitive and common in the literature. In addition to those, this model has new properties:

1. The cross-elasticities are affected by the number of consideration sets they share, which in turn in affected by the products proximity. If these products never share a consideration set, the effect of a price change in one, will not affect the demand in the other. On the other hand, the more consideration sets they share, the greater is the impact in each other demands

2. Free riders, products can benefit by being placed next to attractive products. In many categories there are brands leader that draw a lot of attention, in some cases products positioned near a brand leader enjoys more attention, and increase their sales.

3. Attractiveness effect, there are some locations in the shelf that are more attractive than others.

The second effect explains why commonly observed own brands close to national brands (Sayman, Hoch and Raju 2002). Given our results, we also argue that competition increases for neighboring products,

damaging the profitability of the products. Notice that both effects coexist and to understand what effect dominates is another good reason to build the optimization model. In particular, our model incorporates the demand free rider effect in the starting position, where consumers begin the search. In some cases consumers may begin the search in their favorite product, but then change their mind and purchase a neighboring product.

In the third effect, studies about product position, based on eye tracking data (Chandon, Hutchinson, Bradlow and Young, 2007 - 2009), show that there are positions that are more attractive than others, and this effect can be very strong. In our model this effect is represented in two parts, first is in the starting position, where more attractive locations are more likely to be selected to start the search, the second is in the movement from stage to stage in different consideration sets. We want to emphasize that this effect is in consequence of the physical shelf and how consumers search through it. The search for products on a retail shelf has physical constraints. For example if a consumer wants to go from the right edge to the left edge, or if they are looking at the top part of the shelf and they want to look the bottom part, they are likely to go through the middle part. In consequence, when we represent the shelf movement as a function of the distance between consideration sets, this effect is implicitly included in the model.

These three properties are not found in the standard demand models. This model allows us to answer relevant questions about positioning, competition, shelf arrangement and adding/eliminating products in shelves. Another important property of this model is the correlation of utilities for similar products. For example, we can think about subcategories like diet beverages. We believe the impact of an increase in the price of *Diet Coke* to have greater impact on *Diet Pepsi* than for a regular drink. Since we have individual level parameters, the consumers that enjoy *Diet Coke* (we presume this is the most preferred in the utility ranking and therefore buy it) are likely to also enjoy *Diet Pepsi*. When the price of diet coke rises, the consumers that decide not to purchase *Diet Coke* now are more likely to buy a similar product, like *Diet Pepsi*. This is the main reason we decided to use random coefficients in our model, given that we do not observe individual level data directly.

Our model can converge to the standard logit model in two ways, first by increasing the size of the consideration sets, such all the products belong to every consideration set, then for every stopping distribution and for every starting position, the model simplifies to the standard logit. Second, by allowing a large search in the shelf, then the final set of products will include all of them.

2.3.4 Shelf design comparison

When blocking products by quality, it also separates the products by price. The model predicts small difference between treatments 1 and 2. In both cases the products were grouped in terms of quality, in case 1 giving the premium products the tested more attractive place in the shelf, whereas in the second treatment the best shelf location was given to the price products. Our model did not presume better locations, just better neighborhoods, and in both cases the consideration sets looked similar. This observation vas validated with the results, giving similar sales to both shelf types. Bounty, the expensive leading brand was not found in many consideration sets as the price brands, the model predicted an increase in its demand and a lower demand for non-leading premium brands, which was also confirmed with the data, confirming our theory.

The model predicted an average decrease in profitability of 3% in both treatment 1 and 2. In the data we found in increase of 8% and 11% respectively, which compare with an increase of 14% in the control group, giving an adjusted loss of 6% and 3% respectively. In treatment 3 the model predicted a loss of profit of 5%, which compared with an increase of 7%, which compare with the control increase of 14% yielding an adjusted loss of 7%.

2.4 A Cross-Category Empirical Study

In this study we use a sales data set from Dominick's Finer Foods (DFF). DFF is a retail chain with more than 80 stores located in the Chicago Area. This data set contains weekly sales for consumer packaged goods in 12 categories for 5 years between 1989 and 1994. Besides the information about product attributes like price, we have the weekly position of each product on the shelf in bi-dimensional coordinates. This data set includes shelf experimentation conducted during the Micro-marketing project at the University of Chicago (Dreze et al).

We observe weekly market shares and location of products for different categories in retail stores. Most categories have two or three shelf design changes induced by experimental changes. Ideally there would be even more, but practically a reason that shelf experimentation is not common is because it is expensive to conduct since it requires a good deal of labor to rearrange SKUs and insure the pan-o-gram is followed properly.

When we contrast the data set we had available with the ideal one, we have to consider that prices are not randomly assigned to the products. It is reasonable to think that firms set prices in order to maximize profits, and they do this considering information that is not observable from the researcher perspective. This behavior influences the error term causing an endogeneity problem. This problem can be solved, in theory, by incorporating instrumental variables.

Ideally we would observe individual level data. However, only observe weekly sales data at the store-level. Therefore we must aggregate our individual model to the store-level. Although we lack individual data, we do know the demographic profile of each store's trading area, which may be helpful in understanding the distribution of customers that each store faces.

2.5   Designing an Optimal Shelf

To find optimal shelf displays we used a mixed integer programing model (MIP). In the first part of this chapter we describe the main model and in the second part we discuss its benefits and limitations.

2.5.1   Optimization problem

From the perspective of the retailer, the objective function is to maximize profit, considering the individual product demand at any given time w_j [28] and their margin m_j. We define the position of product j in the shelf by the horizontal coordinate of its lower left corner x_j, L_{Lj} binary value that takes value 1 is product j is in level L and y_{jk} binary variable that takes value 1 if product j is located to the left of product k.

---

[28] Subscript consumer and time to simplify notation in the exposition of the optimization model

Each product has a utility estimated with the demand model, but its probability of being chosen depends on its location and the products belonging to the same consideration set. Using the traditional logistic approach to compute the probabilities yields a non-linear formulation, instead we compute the probability of being chosen using a new approach that yields a linear formulation for the probabilities and simultaneously accounts for the consideration set formulation. We define the probability that consumer i choose product j as p_{ij}. In order for these probabilities to represent preference based on the product utility u_j and the consideration sets we use the following constrains

Consumer can only choose a product that is in the consideration set

$$p_{ij} \leq C_i \quad \forall i,j \tag{15}$$

Logit formulation of the probabilities

$$p_{ij} \leq \frac{\exp(u_{ij})}{\exp(u_{ik})} p_{ik} + (1 - C_{ik}) \quad \forall i,j,k \neq j$$
$$\sum_j p_{ij} = 1 \quad \forall i \tag{15}$$

Avoid overlapping products, where hd_j define the horizontal dimension of product j, and TSL define the total length of the shelf.

$$x_j + hd_j \leq x_k + y_{kj} \cdot TSL \quad \forall j,k$$
$$y_{jk} + y_{kj} = 1 \quad \forall j \neq k \tag{15}$$
$$y_{jj} = 1 \quad \forall j$$

The previous formulation assumes that all products are in a one level long shelf. In order to compute the initial consideration sets we need to compute the horizontal h_j and vertical v_j coordinates of the center of the products, using the length of the shelf SL and the shelf height SH that has N levels. This can be computed as follows:

$$h_j = x_j + \frac{hd_j}{2} - \sum_{l=2}^{N}(l-1)\cdot SL\cdot L_{lj}$$

$$v_j = \frac{vd_j}{2} + \sum_{l=2}^{N}(l-1)\cdot SH\cdot L_{lj}$$

(15)

In addition we need an indicator of precedence in the horizontal and in the vertical direction, so if j is located to the left of k no matter the level, zh_{jk} takes the value one, and if j is located below k, zv_{jk} takes the value one.

$$\frac{h_j - h_k}{SL} \le zh_{jk}$$

$$\frac{v_j - v_k}{3SH} \le zv_{jk}$$

(15)

Now we can compute the basic consideration BC_j of the focal product j as the products within a distance D from the center of the product j. In the model BC_{jk} take value one is k belong to BC_j and zero otherwise. We can compute the consideration set with the following set of constrains where depending on the relative location one constrain holds.

$$BC_{jk} - \left(zh_{kj} + zv_{kj}\right) \le 1 - \frac{\left(h_j - h_s + v_j - v_k - D\right)}{D + SL + N\cdot SH}$$

$$BC_{jk} - \left(zh_{kj} + zv_{jk}\right) \le 1 - \frac{\left(h_j - h_s + v_k - v_j - D\right)}{D + SL + N\cdot SH}$$

(15)

$$\frac{D - \left(h_j - h_s + v_j - v_k\right)}{D + SL + N\cdot SH} \le BC_{jk} + zh_{kj}\cdot(D + SL + N\cdot SH) + zv_{kj}\cdot(D + SL + N\cdot SH)$$

$$\frac{D - \left(h_j - h_s + v_k - v_j\right)}{D + SL + N\cdot SH} \le BC_{jk} + zh_{kj}\cdot(D + SL + N\cdot SH) + zv_{jk}\cdot(D + SL + N\cdot SH) \quad \forall j \ne k$$

Now we can come back and compute C_{ist} using this representation of BC_j and (8)

2.5.2 Model benefits and limitations

The number of possible shelf designs is factorial in the number of products. Even with as few as 15 products the number of designs exceeds the trillion, which emphasize the need for a tool to quickly select

designs. The optimization model proposed solve this problem for categories with about 26 products, using the addition of redundant constrains and using the standard ILOG CPLEX 12.5.1.

Although many categories fall into this category, there are other categories with more products. For those categories we use a heuristics that divide the self in smaller pieces and optimize the products within those locations, then we evaluate the best trades of products pairwise and proceed to re-optimize until no more pairwise trades improve the profit of the shelf.

2.6   Conclusions

Starting with Hauser (1978) it has long been realized that explaining consideration is more important than choice conditional upon the consideration set. Therefore, our research contributes to this idea with an explicit representation of how consideration sets are influence by shelf design. Our approach extends the marketing shelf literature to induce demand correlation based on products proximity with their competitors in the shelf, and not only their relative strengths. We find that shelf location by itself can influence the product demand, but the addition of its surrounding products adds information to make significant improvements in fit, prediction power and decision support. We conclude that products are more affected by the relative location of competitors than location alone. It is possible to influence competition in the shelf by closing or enlarging the gap between competitors. The difference in profitability of different designs can be up to 7%. The mechanism exploited to improve shelf profit involves influencing competition through shelf design to favor more profitable products. In order to find better shelf designs, traditionally different designs were tested and adjusted. This task can be very expensive and only evaluate a handful of different designs. In this paper we propose a new technique to explore many shelf designs and in many cases finding optimal designs with a reduced cost.

2.7   References

Allenby, Greg M. (1989), "A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data", *Marketing Science*, 8 (3): 265-280.

Berry S. (1994) "Estimating Discrete-Choice Models of Product Differentiation", *The RAND Journal of Economics*, 25 (2): 242-262

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium,"*Econometrica*, 63, 841–890.

Boatwright, Peter and Joseph C. Nunes (2001), "Reducing Assortment: An Attribute-Based Approach", *Journal of Marketing*, 65 (3): 50-63.

Borin, N., P. Farris, and J. Freeland (1994), "A Model for Determining Retail Product Category Assortment and Shelf Space Allocation", *Decision Science*, 25 (3): 359-384.

Bradlow, Eric T. and Vithala R. Rao (2000), "A Hierarchical Bayes Model for Assortment Choice", *Journal of Marketing Research*, 37 (May); 259-268.

Brown, J. and Wildt, A. (1992), "Consideration Set Measurement", *Journal of the Academy of Marketing Science*, (20); 235-246.

Bruno, Hernan A. and Naufel J. Vilcassim (2008), "Structural Demand Estimation with Varying Product Availability", *Marketing Science*, 27 (6): 1126-1131.

Bultez, A. and P. Naert (1988), "SHARP: Shelf Space Allocation for Retailers Profit", *Marketing Science*, 7 (3): 211-231.

Cardell N.S. (1997). "Variance component structure for the extreme value and logistic distributions with application to models of heterogeneity," Econometric theory, first presented at IMEO, WSU 1991.

Chandon, Pierre, J. Wesley Hutchinson, Eric T. Bradlow, and Scott Young (2007), "Measuring the Value of Point-of-Purchase Marketing with Commercial Eye-Tracking Data", in *Visual Marketing: From Attention to Action*, Michel Wedel and Rik Pieters, eds. Mahwah, NJ: Lawrence Erlbaum Associates, 225-58.

Chandon, Pierre, J. Wesley Hutchinson, Eric T. Bradlow, and Scott H. Young (2009), "Does In - Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase", *Journal of Marketing*, 73, 1-17.

Corstjens, Marcel and Peter Doyle (1981), "A Model for Optimizing Retail Space Allocations", *Management Science*, 27 (7): 822-833.

Corstjens, Marcel and Peter Doyle (1983), "A Dynamic Model for Strategically Allocating Retail Space", *Journal of the Operational Research Society*, 34 (10): 943-951.

Cox, Keith K. (1970), "The Effect of Shelf Space upon Sales of Branded Products", *Journal of Marketing Research*, 7 (1): 55-58.

Curseu, A., T. van Woensel, J. Fransoo, K. van Donselaar, and R. Broekmeulen (2009), "Modeling handling operations in grocery retail stores: an empirical analysis", *Journal of the Operational Research Society*, 60 (2), 200-214.

De Los Santos, Barbur, Ali Hortaçsu, and Matthijs R. Wildenbeest (2012), "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior", *American Economic Review*.

Draganska, Michaela and Daniel Klapper (2010), "Choice Set Heterogeneity and the Role of Advertising: An Analysis with Micro and Macro Data", Stanford Graduate School of Business, Stanford University.

Dreze, Xavier, Stephen J. Hoch, and Mary Purk (1994), "Shelf management and space elasticity", *Journal of Retailing*.

Guadagni P and Little J, (1983). "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, 22 (3): 203-238

Hariga, Moncer, Abdulrahman Al-Ahmari, and Abdel-Rahman A. Mohamed (2007), "A joint optimization model for inventory replenishment, product as- sortment, shelf space and display area allocation decisions", *European Journal of Operational Research*, 181: 239-251.

Hauser, John R. and Wernerfelt, Birger (1990), "An Evaluation Cost Model of Consideration Sets", *Journal of Consumer Research* (16): 393-408

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

Irion, Al-Khayyal, and Lu (2004), "A Piecewise Linearization Framework for Retail Shelf Space Management Models".

Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg (2010), "Online Demand Under Limited Consumer Search", *Marketing Science*, 29 (6): 1001-1023.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003), "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation", *Marketing Science*, 22 (1): 58-84.

Miller, G.A. (1956), "The magic number seven plus or minus two: some limits on our capacity to process information", *Psychological Review* 63 (2): 81–97

Paas, F., Renkel, A., & Sweller, J. (2004), "Cognitive Load Theory: Instruc- tional Implications of the Interaction between Information Structures and Cognitive Architecture", *Instructional Science* 32: 1–8.

Pauli, Hans and Raymond William Hoecker (1952), "Better utilization of selling space in food stores", U.S. Department of Agriculture, Production and Marketing Adminstration, November 1952.

Payne, Bettman, Johnson (1993), *The Adaptive Decision Maker*, Cambridge University Press, ISBN 0-521-41505-5

Pieters, Rik and Warlop, Luk (1998), "Visual attention during brand choice: The impact of time pressure and task motivation", *International Journal of Research in Marketing*, 16: 1-16.

Roberts, John H. and Lattin, James M. (1991), "Development and Testing of a Model of Consideration Set Composition", *Journal of Marketing Research*, (28); 429-440.

Serdar Sayman, Stephen J. Hoch and Jagmohan S. Raju (2002), "Positioning of store brands", *Marketing Science*

Stuttgen, Peter, Peter Boatwright, and Robert T. Monroe (2011), "A Satisficing Choice Model", *Marketing Science*.

Sweller, J. (1988), "Cognitive load during problem solving: Effects on learning", *Cognitive Science*, 12 (2): 257–285

Van Nierop E, Bronnenberg B, Paap R, Wedel W, and Hans Franses P (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data", *Journal of Marketing Research*, Feb, 63-74.

Van Ryzin, G., S. Mahajan (1999), "On the relationship between inventory costs and variety benefits in retail assortments", *Management Science*, 45 (11); 1496-1509.

Zufryden, Fred, (1986) "A Dynamic Programming Approach for Product Selection and Supermarket Shelf-Space Allocation", *Journal of Operational Research Society*, 37 (4).

Appendix A

Appendix Chapter 1

3.1 Other Optimization procedures

In this paper, we want to answer questions about proper branch network structure and how the bank should react to the introduction of the mobile channel and continue growth of alternative channels. There are many objectives we could use to improve the branch network; here, we use consumer attrition for two reasons. First, we observe increased levels of customer attrition when closing branches. Second, it is an objective measure that can be computed directly from observed data.

We compute regional maps with consumers' activity with current configurations and shifts in consumer activities due to branch network configurations or channel adoption. Based on this shifting and variation in transaction/holding costs caused by this shift, we compute what variation generates the minimal attrition level or lower decrease in service level. Based on the current branch network configuration, the optimization model can suggest what branches could be closed without increasing consumer attrition more than a certain level. We also can recommend openings or increase capacity in a region that shows high demand for branch activity. With these tools we can run counterfactuals with different levels of adoption of alternative channels, what branches would be not necessary and in what regions would be profitable to encourage consumers to move to these alternative channels. We define consumer attrition percentage as the proportion of inactive consumers for more than 3 months. A consumer is considered inactive when she doesn't perform any interaction with the bank in a given period.

## 3.2 Branch closure optimization example

The output of the consumer demand model can be used to compute customers' probabilities of using other alternatives or quitting the bank. This optimization can be done with any number of branches in the analysis—for example a neighborhood, a ZIP code, or a city. To demonstrate the optimization workings, we use an example of five branches. We first use the model to compute all the transition probabilities for a given consumer when the bank closes each branch, as depicted in **Table 1**.

| S(b,b')	B1	B2	B3	B4	B5	Attrition
B1	-	0.15	0.25	0.30	0.25	0.05
B2	0.20	-	0.35	0.12	0.30	0.03
B3	0.40	0.15	-	0.19	0.20	0.06
B4	0.20	0.30	0.29	-	0.17	0.04
B5	0.25	0.35	0.18	0.20	-	0.02

Table 1. Example of transition probability after closing a branch

In this example, the customer would be least likely to quit the bank if B5 were closed, because its probability of attrition is the smallest (0.02). Following the same procedure for multiple customers and weighting the attrition probability by their value, it is possible to choose the optimal branch to close.

3.3   Example of Usage with fixed scale
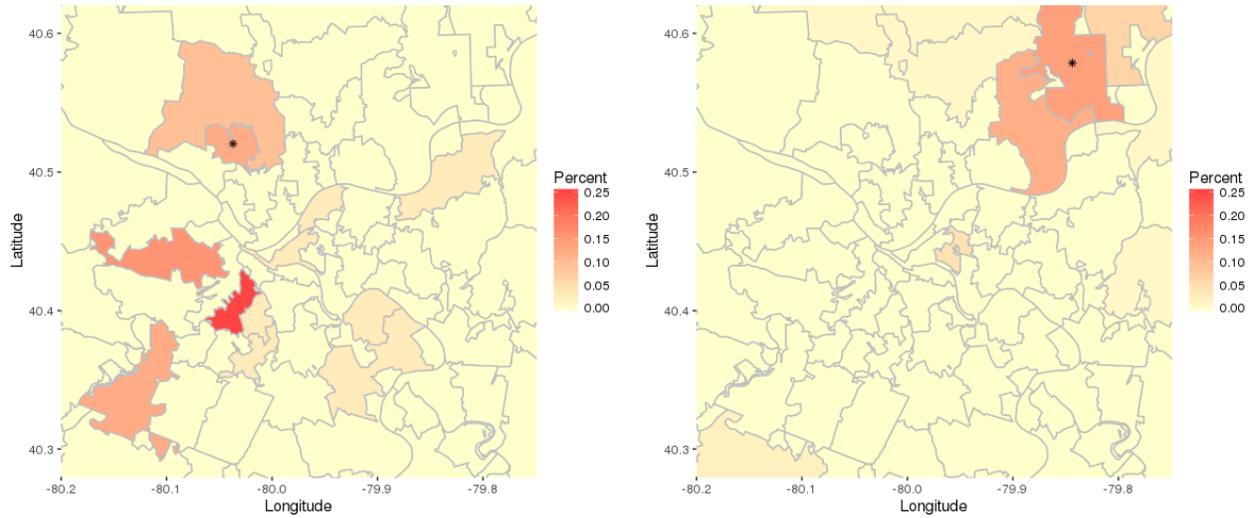
3.4   Example of Map plots with geographical data
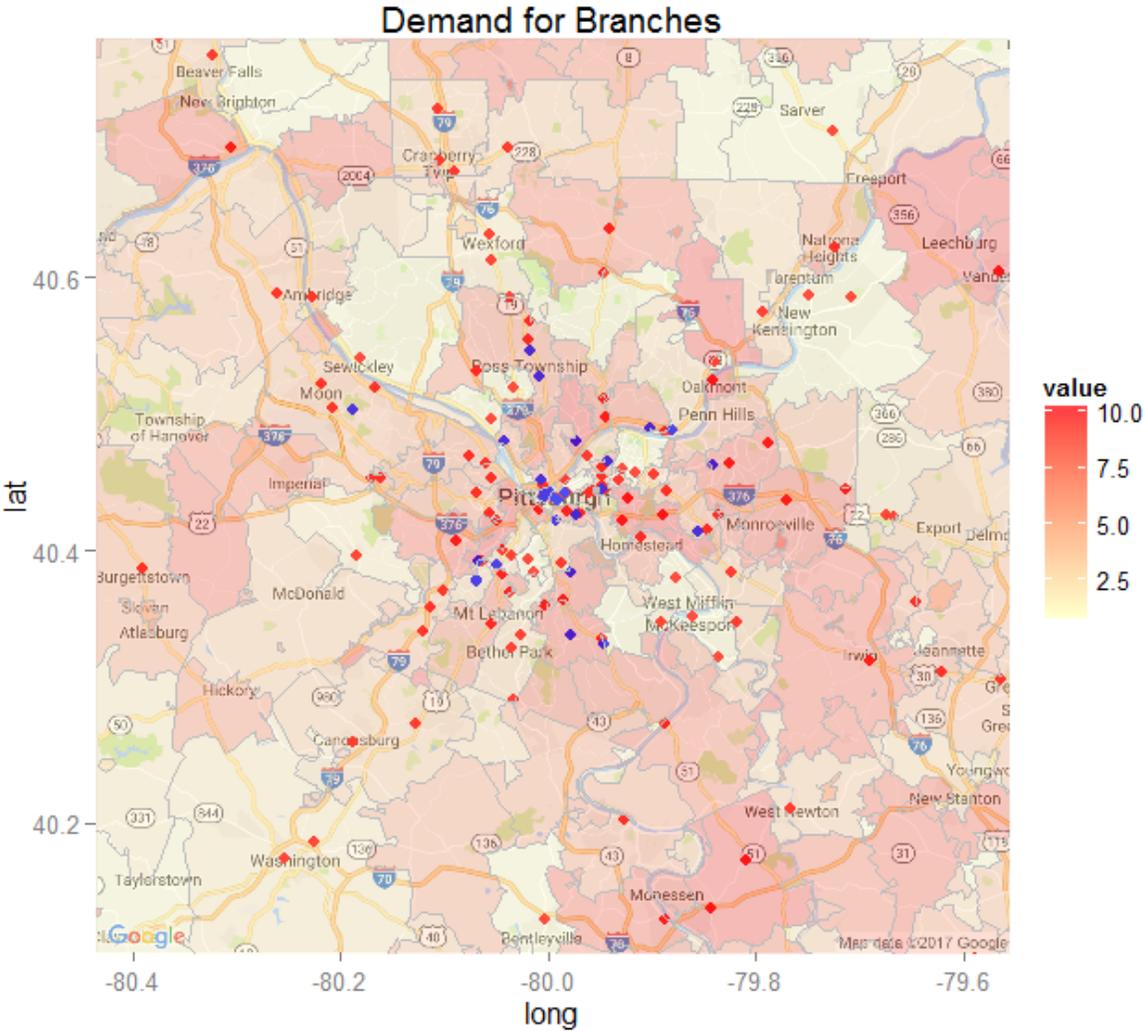
Appendix B

Appendix Chapter 2

4.1   Data Description: Can Soup, RTE Cereals and Toilet Tissue

In this study we wanted to use the model using more than one category, to observe the how generalizable it is. We chose to estimate these three models for three categories: Can Soup, ready to eat (RTE) Cereals, and Bathroom Tissue.

Can Soup: This is a mature category with 5 top sellers (like chicken noodle and tomato), 10 medium sellers (like vegetables) and 50 lower volume sellers (like Oyster stew). Consumers have difficulty shopping in this category because of 70 different flavors, each dominated by red and white Campbell's package. In this category previous studies try to organize the shelf by flavor, which did not succeed apparently because they could not match the consumers' mental representation. The study conducted in this category organized products alphabetically. This data set is interesting given that consumers need to get many details about products before making the final purchase, which we hypothesize, will be shown in the sequential search.

RTE Cereals: This category contains around 200 different products, with 125 heavily marketed. The market by manufacturer is relatively concentrated with 40% of sales for Kelloggs, 30% for General Mills and the rest distribute among Ralston, Post, Quaker and Nabisco. This category is always blocked by manufacturer. The shelf study in this category used the "bull's eye" setting, where Kelloggs and General Mills where positioned in the center to encourage consumers to shop at more than one end of the aisle. This design is interesting from the perspective of this study, since it tests our conjecture.

Bathroom Tissue: The products are segmented by size: single rolls, 4 packs, and multipacks. Before the shelf experiments products were grouped on the shelf by brand. The shelf experiments aimed to inhibit the price comparisons among sizes and proposed two treatments that made comparisons between sizes more difficult, as depicted in **Figure 15**.

Figure 15. Shelf designs for Bath Tissue

We choose this data set because it is a category with items that are physically large and it requires consumers to move through the aisle and not just their eyes. In order to achieve a better fit for all the models, we grouped the brands with low market share, and we also reduced the weeks and stored analyzed to allow us to try many different parameter configuration and reasonable amount of time.

4.2 Experimental Shelf design

Shelf design for the base case (D1):

Design D 1 Category ptw