Carnegie Mellon University
Tepper School of Business

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
ECONOMICS

Titled

"Essays on Strategic Behavior in Entry and Exit and the Impact on Local Economy"

Presented by

I-Ta Yang

Accepted by

Rebecca Lessem	4/28/2020
Co-Chair: Prof. Rebecca Lessem	**Date**
Maryam Saeedi	4/28/2020
Co-Chair: Prof. Maryam Saeedi	**Date**

Approved by The Dean

Robert M. Dammon	5/15/2020
Dean Robert M. Dammon	**Date**

Essays on Strategic Behavior in Entry and Exit and the Impact on Local Economy

I-Ta Yang

Tepper School of Business
Carnegie Mellon University

Doctoral Committee:
Rebecca Lessem
Maryam Saeedi
Karam Kang
Beibei Li

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics

Acknowledgments

I am grateful to the members of my dissertation committee: Rebecca Lessem, Maryam Saeedi, Karam Kang, and external reader Beibei Li. They provided me valuable guidance and comments for this project. I want to thank Dennis Epple for advising me in the first two years in the program. I also want to thank my parents for being supportive throughout my student life. I thank all my colleagues and friends in Pittsburgh, especially Wenting, Leah, and Erica, for helpful discussions throughout this project. This work has been benefitted from MEA seminar and Tepper Applied Micro Lunch Seminar. Finally, I would like to thank Lawrence Rapp and Laila Lee for all the support in the past six and half years.

Abstract

Firms oversee market structures and choose the optimal strategic behavior in entry and exit to maximize profits. Understanding market structure and the role it plays in determining the extent of competition benefits both policymakers and firms. Policymakers use market structure knowledge to set regulations on industries that may hurt consumer welfare, while firms use such knowledge to estimate market capacity. My dissertation seeks to model and provide insights for three such problems in the education and banking industries.

In the first chapter, titled "Understanding the Impact of Rising Online Bank Channels on Exit of Bank Branches,“ I examine the impact of the increasing adoption of digital banking channels and of the information transferring cost on the shrinking brick-and-mortar branch network in the banking industry from 2009 to 2013. When digital financial channels become more common to consumers, banks close branches to reduce personnel and occupancy expenses; however, banks incur a cost to transfer the consumer-specific information and relationships from the closing branch to other branches. Using detailed data on consumer transactions, I estimate a model of demand for branch services and find that consumers are sensitive to the travel distance to the branch. I then use the demand and cost estimates with the network distances to predict a bank's profits under the observed and counterfactual network configurations. Using a moment inequalities approach, I find that it costs the bank between \$0.019 and \$0.065 to transfer one dollar of transactions per mile. I also demonstrate that increasing digital channel adoption reduces the demand for branches and transferring cost and accelerates the branch closing process.

In the second chapter, titled "Understanding the Effects of Bank Branch Closures on the Local Loan Market," To study the effects of bank branch closings on the local loan market, I use a unique small business loan application data set and difference-in-difference framework to estimate the post-intervention effects in both the supply and demand sides. The results show that branch closings have negative effects on the bank's business sector, especially in cities and areas with a high density of competitors' branches when loan seekers can apply for business loans in many other places. The bank tends to loosen its reviewing process to approve more loans to offset the loss of demand for business loans in the cities with branch closings. I also show that the bank's loan application processing time gets longer after branch closings, when the soft information and relationships between the local businesses and the branch are difficult to transfer and replace. Moreover, I find that the loan seekers

in the areas which already have limited availability of banking services are more difficult to get their loans approved by the bank after branch closings.

In the third chapter, titled "Understanding the Supply Side of For-Profit Colleges: Structural Analysis," I study the market entry of private for-profit colleges in the United States from 2005 to 2013. I empirically show that for-profit schools in a larger population or higher median income market often choose to enter the market earlier. Moreover, my results reveal that both the federal and state regulations regarding for-profit colleges' recruiting process have a negative impact on for-profit schools' payoff; therefore, for-profit colleges often choose to delay their timing of market entry. Furthermore, my model suggests that the competitive effects from other for-profit and community colleges in the same market are substantial and likely to push back the entry time for new for-profit colleges. As a result, I show that the proposal made by former President Obama regarding free community college has the greatest negative impact on the for-profit college industry compared to exiting federal and state regulations.

Contents

List of Tables

List of Figures

1 Chapter I: Understanding the Impact of Rising Online Bank Channels on Exit of Bank Branches

1.1 Introduction

After decades of expanding their branch networks, banks have begun reconsidering their branching strategies after the Great Recession in 2009. The number of branches in the U.S. peaked at 83,532 in 2009 and has declined annually since then, reaching 71,342 branches in 2018, as shown in Figure 1. The top three costs for a bank are personnel, occupancy expenses, and technology. Banks have been reducing personnel and occupancy and investing more money in technology to create a more efficient banking environment and to reduce costs for shareholders. Various digital channels have been created in the banking industry. The first ATM was installed in New York City in 1961. As the internet coverage became more common in the U.S., banks started to use the World Wide Web to perform account services in 1995. Mobile banking began in 2010 with specialized application for mobile devices to access the banks' web page, and banks have continually expanded their mobile banking services to include balance inquiries, funds transfers, and check deposits. Mobile banking applications and advanced ATMs create a new way for consumers to interact with banks. With the increasing adoption of smartphones, the share of customers using mobile banking has grown from 20% in 2011 to 44% in 2015 in the U.S. Instead of visiting brick-and-mortar branches, more people are using mobile applications, online platforms, and ATMs to interact with banks. In response to their customers' increasing use of technology, banks are making changes in their branch networks. For example, in 2013, a Capital One spokeswoman disclosed plans to evolve and optimize the branch network to ensure that the bank operates as efficiently and effectively as possible according to Wall Street Journal. Banks say they are carefully considering which branches to close, examining deposit levels at each branch and their customers' commute time to the nearest location (Ensign, Rexrode and Jones, 2018). The branch closing process is expected to continue and has created widespread concern regarding the negative impact on the social welfare.

While digital online platforms are becoming more efficient and convenient, brick-and-mortar branches are still important to the local economy. Therefore, not all branches should be closed. They provide not only physical access to customers but also credit sources to local businesses. Also, making large cash withdrawals, receiving certificate checks, creating wire transfers, and accessing safety deposit boxes have to be performed in the brick-and-mortar branches of a bank. Therefore, changes in the market structure of banking could widely affect the local industries. Especially the local small business owners rely on the close relationship with their local branch managers to receive the loans they need to expand and operate their businesses. In a recent paper, Nguyen (2019) shows that branch closings led to a persistent decline in credit supply to local small businesses due to the disruption of branch-specific relationships and soft information. Therefore, banks are considered as one of the most regulated industries in the U.S. The government wants to ensure local access to bank networks. For example, the FDIC now requires banks to provide a 60-day notification period before branch closings.

This paper presents a model of demand and network transferring cost in the banking deposit market that allows me to measure a bank's savings from branch closings associated with increasing digital channel adoption. In the first stage, I estimate the demand parameters from the multinomial logit model of consumers' choices on which branch to perform account transactions to pin down consumers' response to the distance. I use data on the account transactions including depository and wire transfers in City A from 2009 to 2013 in the U.S. in terms of asset size. To measure the benefits of having brick-and-mortar branches, I treat the transactions that can be made through ATMs and a mobile application by consumers regularly using digital channels as an outside option in the choice model. The results of the demand estimation suggest that regular consumers are sensitive to the distance to the branch and the branch size. They tend to choose to visit branches that are close to them, have a larger size (in square feet, and are not acquired from other banks. The percentage of consumers who use digital channels is 14.5% in 2009 and increases to 40.6 in 2013. To estimate the impact of a closing branch on consumer loyalty, I run a regression on consumer attrition level, which is defined in a similar manner as Cisternas-Vera, Montgomery, and Hoeve (2017) as the percentage of consumers with inactivity during the next six months in each branch . In City A, there is a high attrition level, 8.5%, among consumers who visited the branch at least once in six months. I find that when the bank closes a branch that is far from other branches or has a lower percentage of consumers with mobile adoption, the attrition level tends to be higher, showing that consumers could switch to other local banks due to the inconvenience caused by the branch closing. With the estimated demand function, I can predict the sum of the amount of transactions under any network configurations.

In the second stage, I quantify the cost of transferring information within the network of branches. The soft information, one of the most valuable assets in each branch, is built based on the relationship between the local branch employees and some frequent customers. Once the branch is closed, the branch needs to transfer existing customer information and relationships to other branches. The cost of these transitions depends on the distance between the closing branch and the other branches in the network. Therefore, I specify a profit function where the variable cost of transferring information from an inactive branch to the nearby branches is proportional to the net distance from the inactive branch to the other branches within the network, scaled by the sum of the amount of transactions. Similar to Holmes (2011), I assume that the observed sequence of inactive branches is optimal due to the revealed preference, so the discounted profit stream including the predicted amount of transactions, fixed costs, and variable transferring cost under the observed network configuration must be at least as large as under any alternative sequencing of branch closings. To construct the alternative network configurations, I swap the closing dates of any two events. A profit comparison of the observed and such counterfactual network configurations allows me to identify the upper and lower bounds on the inactive branches' transferring cost per dollar per mile such that the observed sequence dominates the perturbations. Following Romano, Shaikh, and Wolf (2014) and Andrews and Soares (2010), I use a moment inequalities estimator to solve the confidence set of the transferring cost without having to solve the complicated dynamic problem.

With 304 perturbations in the inequality moment estimation, the model indicates that it costs the bank between \$0.019 and \$0.065 to transfer one dollar of transactions per mile. Given the estimated demand and cost parameters, the benefit the bank receives is between \$2.278 and \$2.369 billions under the observed equilibrium. Although the sum of the amount of revenue from branches is 4.1% lower with branch closings relative to no branch closings due to the attrition, and the transferring cost appears to be between \$37.59 and \$128.6 millions, branch closings would reduce the building and staff costs. Under the observed closing sequence, the bank would surpass the profit with no branch closing between 0.53 and 1.06 years. In the longer term, the bank would generate even more profit if the attrition level could stay the same.

Next, I consider two counterfactual experiments based on the observed equilibrium. In the first counterfactual experiment, I increase the number of consumers using the digital channels to represent an increase in the adoption of digital channels as the bank is encouraging consumers to use the mobile application and the new interactive ATM. The expanding digital adoption decreases both the demand of the branches and the net transferring cost because this cost depends on the sum of the amount of transactions. The experiment implies that the bank could save between \$8.75 and \$52.94 millions from the net transferring cost depending on the percentage of the increase in the adoption of digital channels. Therefore, as more consumers use digital channels, the bank could significantly decrease the transferring cost while closing a branch. As a result, banks are expected to accelerate the branch closing process. According to FDIC data, the number of branches in the U.S. shrank by 1947 in 2018, a new record high, and the closing speed has increased year over year since 2010. The second counterfactual experiment studies the impact of relaxing the limit of digital channels. Now the bank allows consumers to perform \$5,000 deposit instead of \$2,500 thorugh a mobile phone application. However, this relaxation would only benefit consumers who are already digital channel users. With this about 3.55% of the amount of the transactions removed from the demand of the brand, the bank saves \$76.70 millions from lower transferring cost relative to the current equilibrium.

A rich body of existing research has explored the effects of network changes, but most of the previous studies focus on firms expanding their network of retail stores and distribution centers. In contrast, my paper studies network shrinking (Ishii (2005); Jia (2008); Holmes (2011); Houde, Newberry and Seim (2017)). Several papers have studied the environment of branch closings (Nguyen (2014); Cisternas-Vera, Montgomery and Hoeve (2017)). However, most of them focus on the impact on consumers' behavior and local economic outcomes. This paper is the first to study bank behaviors in the shrinking branch network in the banking industry using the moment inequality approach to specify the profit function.

The remainder of the chapter is structured as follows. In the next section, I describe the related literature. Section 1.3 provides the detail of the data sets used in the study. Section 1.4 presents the empirical framework associated with the demand estimation and the moment inequality approach. Section 1.5 shows the results of the empirical model. Section 1.6 provides the robustness check. Section 1.7 describes the current equilibrium and counterfactual experiments, and Section 1.8 concludes the chapter.

1.2 Related Literature

This study is related to recent industry organization research and economics literature on the sequence of firm exit. Takahashi (2011) estimates the impact of competition and exogenous demand decline on the exit process of movie theaters in the U.S. from 1949 to 1955 as TV bacame more common across households. In Takahashi (2011), the theaters have the incentive to choose to remain in the market and delay their exit time even if they have negative profits, because they expect to outlast their competitors. Under a similar setting as Takahashi (2011), while financial technology is becoming more common in the banking industry, the usage of branches decreases and branches are exiting the market. However, instead of each branch competing independently, in my paper the bank oversees the entire branch network and makes optimal decisions on shrinking the network.

This study is also related to recent operation research on the management of the network of brick-and-mortar stores. However, most of this body of research does not study store exit but analyze the entry of stores or distribution centers. Holmes (2011) estimates the savings in distribution costs while maintaining high store density around the distribution center. By comparing the observed store-opening sequences with the alternative network configurations with various distances from distribution centers, Holmes (2011) identifies the parameters of interest using the moment inequality approach.

To measure the cost related to the network configuration, I estimate a modified version of Houde, Newberry and Seim (2017). Their study examines the economies of density associated with Amazon's distribution center roll-out from 2006 to 2018. In the first stage, they estimate a model of demand for retail goods to predict revenues and shipping distances under the observed distribution center roll-out and under counterfactual roll-outs which swap the opening dates of any two centers. In the second stage, they also use a moment inequalities approach which compares the observed network configuration with counterfactual roll-outs to identify the shipping cost savings. As a result, they show that with Amazon's network expansion, Amazon has reduced its total shipping cost by over 50% since 2006. Instead of analyzing a network expansion like Holmes (2011) and Houde, Newberry and Seim (2017), I analyze the network reduction associated with branch closing in the banking industry. This paper adopts the moment inequality approach to estimate the information transferring cost within the branch network.

1.3 Data

In this study, I use four data sources: a U.S. bank, the American Community Survey, the County Real Estate Portal, FDIC bank data, and Glassdoor and Codigo. In this section, I describe the data set in detail.

1.3.1 Consumer Transactions

The data source for this chapter is the account transaction data sets provided by a U.S. bank. The data set includes ZIP-code-level information on where consumers live. The bank also records the following information on each transaction by each consumer: an anonymized account identifier, date, location, money amount, and channel. The transactions could be performed through branches, ATMs, online website, a mobile application, phone, debit card, and credit card. The transactions include deposit, withdraw, wire transfer, direct deposit, quick transfer, inquiry, and purchase. There are more than 1.7 billion transactions across all channels for over a ninety-six-month period from 2007 to 2013 and about 900,000 transactions performed in branches annually in a single city. Among all transaction types, the most popular transaction category is inquiries. These transactions do not need to go through branches. Most inquiries usually happen via ATMs, online website, phone, and mobile application. Table 1 provides a simulated example of raw information for transactions. The example is for illustrative purposes and is not based on actual data.

The next set of transaction types by frequency is withdrawals and deposits. Under current federal regulations, only Reg D puts a limit of six transactions per month on transfers and withdrawals from the savings or money market account. There are no other rules regarding withdrawal and deposit limits on checking accounts. However, each bank has its limit on mobile check deposit and ATM withdrawal due to the money available in the machine. For example, a large US bank allows each of its customers to deposit checks under \$2,500 per day and \$5,000 per month through a mobile phone application, and to withdraw only \$500 daily through an ATM. They will have to go to the counter in a branch to withdraw more than \$500 a day or to deposit a large amount of cash. They also have to go to a branch to perform wire transfers. In the data set, withdrawals and deposits represent more than 90% of the transactions that consumers perform at a branch. Therefore, these two transaction types form the focus of the study, because they measure the true benefit of branches.

The most popular channels are website and the purchase with debit and credit cards. Consumers prefer to perform inquiries and quick transfers via online websites. The mobile application, which became available in 2012, provides not only the services included on the website but also check deposit services using phone cameras. Since the mobile application became available, more consumers have moved from the website channel to the mobile application. The website transactions dropped from 39% in 2012 to 26% in 2015. At the same time, the number of transactions through ATMs and branch channels has also declined slightly.

1.3.2 Branch Characteristics

The branch characteristics data, provided by a US bank, consists of 3,461 observations of its branches in the U.S. from 2004 to 2018. It contains information about where the branch is, when the branch became inactive, and whether the branch was acquired from others. In 2015, 798 branches were inactive and 2,491 branches were acquired from others. In a large US city where is the focus of this paper, there were 69 branches in 2009 and only 55 in 2015 as shown in Figure 2. To measure the staff cost in each branch, I gather the data on the number and salary of branch employees through Codigo and Glassdoor. According to Codigo, the average number of employees in each branch is 6.5, with one branch manager and 5.5 branch sale and service associates. Glassdoor shows that the salary of a branch manager is about $67,311 and the salary of sale and service associates is $35,707.

1.3.3 County Real Estate and Geographic Characteristics

To have controls on how the geographic characteristics affect consumers' choices on visiting branches to receive banking services, I use the ACS data set. It provides annual ZIP-level geographic characteristics including age, education level, median family income level, and population. The County Real Estate Portal provides government valuations on building and land value annually. This real estate data allows me to measure the annual cost of having a brick-and-mortar branch of Bank A in City A.

1.4 Empirical Model

This section presents the demand discrete choice model and the moment inequality approach used in the study to fully estimate the model of the decision of branch closing. The empirical model setting follows Holmes (2011) and Houde, Newberry, and Seim (2017).

1.4.1 Demand Specification

In order to measure the deamnd for each branch under difference network configuration, I use a discrete choice approach. The demand side of the model follows the discrete choice literature and adapts the method of multinomial logit model. In this paper, I use both consumer-level location and demographic data. Due to the richness of the data set, I can use a model to make direct predictions on the amount of deposits made by consumers in each branch in each time period. In this environment, a consumer at a particular location l chooses one of the branches within the consideration set to make transactions in year t or the outside options including ATMs or online channels.

The transactions that made through ATMs and the bank's mobile application by consumers are catagorized as outside group. The transactions that are performed at branches are wire transfer, and transactions made by non-digital users, and deposit with over \$5,000 by digital users. Due to the limit of financial technology and cash availability, there are rules set by the bank on these transactions. To estimate the demand of deposit, I estimate the method of multinomial logit model in each individual time period. In other words, there is seperate demand estimation in every year. In this way, I can measure how fast consumers are adopting the digital channels.

Like other standard discrete choice approach, the model specifies a consumer utility function that allows for branch differentiation and heterogeneous preferences driven by consumer characteristics. I assume that the utility of individual i for branch j in time period t is

$$U_{ijlt} = U(x_{jt}, h_{it}, distance_{ij}, \epsilon_{lt}), \tag{1}$$

where x_j is a vector of branch characteristics including branch size in square feet, cost and branch merged indicator; h_i is a vector of consumer characteristics including online indicator and population; $distance_{ij}$ is the distance in miles between consumer location and branch j; and ϵ_{jt} is a logit error term which is assumed to be drawn independently and identically distributed across all consumers living in block group l in time period t. The distance term here represents the disutility of commuting from a consumer's location to each branch, which is the main factor of consumer preference. There are J branches in the model.

Each consumer chooses to perform transactions at the branch which maximizes her utility. The probability p_{jlt} that a consumer at location l deposits money at branch j in in time period t can be derived using the multinomial logit formula. The close form of p_{jlt}

relative to the probability of a consumer using an outside option can be denoted as

$$p_{jlt} = \frac{e^{Z_{jlt}}}{1 + \sum_{h=2}^{J} e^{Z_{hlt}}}, \tag{2}$$

where Z_{jlt} is the utility of the consumer in location l on choice j in time period t. The probability of a consumer using an outside option is denoted as:

$$p_{1lt} = \frac{1}{1 + \sum_{h=2}^{J} e^{Z_{hlt}}}, \tag{3}$$

Given the probability of visiting each branch, the predicted revenue for branch j is then

$$R_j = \sum_{\forall t} \sum_{\forall l} p_{jlt} \times n_{lt} \times \lambda_{lt} \times r \tag{4}$$

where λ_{lt} is the amount of deposit per consumer at location $l$at time t, n_{lt} is the number of consumers who want to deposit money at location l in time period t, and r is the rate of return respectively. There are n_l consumers at location l, and a fraction p_{jlt} of them are depositing λ_l dollars at branch j in time period t. According to a report from the Beam Bank, banks can earn 5 percent on every 1 dollar deposited. As a result, I set the rate of return to 0.05. The multinomial regression uses a maximum likelihood estimation method to measure coefficients.

In order to estimate the impact of a closing branch on consumer loyalty, I look at actual branch closures made by the bank. Even though the bank dominates the market in City A, I still see a high attrition level, 8.5%, among consumers who visited the branch at least once in six months. The attrition is defined as consumer inactivity during the next six months after a branch closing. When the bank closes a branch that is far from other branches or that has a lower percentage of consumers with mobile adoption, the attrition level tends to be higher, showing that consumers could switch to other local banks due to the inconvenience. I perform a regression for the attrition level on network distance, percentage of consumers with online channel adoption, number of competitors' branches, and acquired indication which shows whether the branch is acquired from merging.

1.4.2 Cost Specification

The cost of having a branch is included in the profit function. I consider the real estate cost and staff cost. However, because the focus of the network is within a single city, the wage and the number of employees in each branch are assumed to be the same across City A. Therefore, when I evaluate the inequality functions in later sections, the only cost included in the model is the taxable real estate value of each branch based on the City A Real Estate Portal. As the taxable land and building value are higher, both the cost to purchase and rent are more expensive. Moreover, this real estate fixed cost depends on the size and location of the branch. According to the Real Estate Portal, City A's downtown has the highest taxable real estate value in the city. Therefore, the branches in the downtown area are expensive to maintain. The fixed cost savings arise from branch closing. Once the branch is inactive, the cost becomes zero forever.

1.4.3 Network Transfer Cost

In this section, I estimate the cost of transferring information within the network of branches. Branches communicate and share information. The soft information, one of the most valuable assets in each branch, is the information that cannot be easily obtained through physical data, such as revenue, sales number, and cost. This information is built based on the relationship between the local branch employees and some frequent customers. Once a branch closes, it needs to transfer the existing customer information and relationships to the other branches. The cost of these transitions depends on the distance between the closing branch and the rest of the network. If the bank is closing a branch in an area with high-density branches, it is easier to transfer local information and customer relationships and to assist consumers to transit to other nearby branches due to the shorter commute time.

1.4.4 Profit Function

In this section, I formulate a profit function for the targeted bank that depends on the configuration of the network similar to Holmes (2011) and Houde, Newberry and Seim (2017). The profit function represents the expected stream of discounted annual profits starting from 2009. The profit function includes the annual sum of branch deposits, which depends on the location of the branch, the fixed cost of each branch, and the cost of transferring information within the network. I assume that in 2009 the bank has perfect foresight in predicting the components of the profit function and chooses the branches to close in each year to maximize its expected discounted profits. I define the set of branches that the bank opens across the city in the vector a_t, with $a_{jt} = 1$ if branch j is active in year t; a, the collection of a_t over time, represents the configuration of the network of branches. The amount of deposits in each branch depends on a_t in year t. $R_{jt}(a_t)$ is the amount of deposits in in branch j in year t.

The annual fixed cost F_{jt} for branch j in year t obtained from the County Real Estate Portal is calculated using a 30-year depreciation timeline. Once the branch becomes inactive, its annual fixed cost turns to zero. The annual staff cost S_{jt} for branch j in year t is calculated by multiplying the average number of employees and their average salary. By combining the sum of the amount of deposits, the annual fixed cost, and the annual staff cost, I denote the profit for branch j in year t as

$$\pi_{jt}(a_t) = R_{jt}(a_t) - F_{jt} - S_{jt}. \tag{5}$$

When the bank decides to close a branch, the bank needs to transfer the customer information and relationships to the other branches. The cost of transferring the information depends on the amount of deposits. I denote the variable net transferring cost by $\theta d_{jt}(a_t)R_{jt}(a_t)$, where $d_{jt}(a_t)$ is the distance from branch j to the other branches and the parameter θ represents the net transferring cost per mile per dollar of the sum of the amount of transactions. When the bank closes a branch located in a low-density branch network, the bank needs to put more effort to assist both customers and managers to become familiar with the area the inactive branch served. Therefore, the larger the distance $d_{jt}(a_t)$, the higher the transfering cost is. Moreover, if there are more relationships and information to share, the transfering cost is higher, and $R_{jt}(a_t)$ in the cost function captures this effect.

Given all these components, the discounted sum of the bank's total profit starting in 2009 is

$$\Pi(a;\theta) = \sum_{t=2009}^{2018} \beta^{t-2009} \sum_{j} \pi_{jt}(a_t) - d_{jt}(a_t - a_{t-1})R_{jt}(a_t)\theta, \tag{6}$$

where β is the discount factor, which is 0.95 by assumption. The bank is solving the profit maximization problem in 2009, in terms of the network configuration, written as

$$max_a \Pi(a;\theta). \tag{7}$$

According to the revealed preference theory, the profit under the observed network configuration (a^o) must be greater than any other counterfactual network configurations; therefore,

$$\Pi(a^o;\theta) \geq \Pi(a;\theta)\ for\ \forall a \neq a^0. \tag{8}$$

Based on this, I can construct moment inequalities by comparing the discounted profits of different network configurations to estimate the value of θ. Duo to the limitation of the data and computation power, I assume that the bank has perfect foresight in predicting for 10 years and the environment after 2013 remains the same.

1.4.5 Moment Inequalities

As in Holmes (2011) and Houde, Newberry and Seim (2017),, I construct moment inequalities by comparing the discounted profits of the observed branch network with the counterfactual configurations. The counterfactual networks I focus on are from swapping the closing date of two branches. In the data set, there are 14 branches becoming inactive. To create the counterfactual closing sequences, I not only swap the closing dates between these 14 branches but also switch the inactive branches with the active branches. In the estimation, I denote the observed branch closing sequence as a^0. In the counterfactual configurations, holding the closing dates of all remaining branches as observed in the data, I switch branch h's closing date with branch k's and denote the sequence as $a^{h,k}$. Based on the revealed preference theory, I can formulate the following linear profit difference inequality from the profit maximization question:

$$\begin{aligned}
& y^{h,k} - x^{h,k}\theta \geq 0, where \\
& y^{h,k} = \sum_{t=2009}^{2018} \beta^{t-2009}[\sum_j (R_{jt}(a_t^o) - R_{jt}(a_t^{h,k})) - (F_{jt}a_{jt}^o - F_{jt}a_{jt}^{h,k})] \\
& x^{h,k} = \sum_{t=2009}^{2018} \beta^{t-2009}[\sum_j d_{jt}(a_t^0 - a_{t-1}^0)R_{jt}(a_t^0) - d_{jt}(a_t^{h,k} - a_{t-1}^{h,k})R_{jt}(a_t^{h,k})]
\end{aligned} \quad , \tag{9}$$

where $y^{h,k}$ is the discounted amount of the transaction differential net of branch fixed cost and $x^{h,k}$ is the discounted difference in the transaction-weighted network distance from comparing the observed closing branch sequence to the one that switches the closing date from branch h to branch k. Based on the demand estimation above, I calculate both the demand from the observed sequences, a^0, and the counterfactual sequences, $a^{h,k}$. To calculate the demand under any network configuration, I assume that the status under demand estimation remains the same across any network configuration.

1.4.6 Moment Inequalities Estimation

Similar to other applied literatures using moment inequalities, Holmes (2011) and Houde, Newberry and Seim (2017), I am looking for the maximum and minimum values of θ such that all moments are satisfied. I use similar instruments to the ones employed in Holmes (2011). The instruments $z_1^{h,k}$ and $z_2^{h,k}$ separate the counterfactual configurations into two experiments. The first experiment with the dummy variable $z_1^{h,k}$ represents the counterfactual configurations which increase both the transaction-weighted network distance and result in higher discounted revenue streams, $y^{h,k} < 0$ and $x^{h,k} < 0$. By comparing the first group to the observed network configuration, I can identify the lower bound of θ. For the second experiment, $z_2^{h,k}$, I include the counterfactual configurations that decrease the weighted network distance and reduce the discounted revenue streams, $y^{h,k} > 0$ and $x^{h,k} > 0$. With the second experiment, I estimate the upper bound of θ.

Since the moment inequalities provide only the bounds of θ instead of a point estimate, I follow the approach of Romano, Shaikh, and Wolf (2014) and Andrews and Soares (2010). The null hypothesis I am testing is $H_o : \theta = \theta^o$. Given a θ, the two-step method by Romano, Shaikh, and Wolf (2014) estimates the critical value from constructing a confidence region for the moments and then using this region to obtain the information of positive moments to calculate the critical value. The modified method of moments by Andrews and Sores (2010) estimates the test statistic values and then creates a decision rule about rejecting the specific value of θ. The large value of test statistic $T(\theta^o)$ provides evidence against the null hypothes. Based on these test statistic and decision rules, the method allows me to construct 95% confidence sets which contain the true value of θ 95% of the time ($\alpha = 0.95$). To efficiently find the boundaries of the set, I use the bisection-method to guess the low and high values of the interest of the coefficient.

1.5 Results

In the first stage, I estimate the multinomial logit demand model for consumer transactions in each time period. The estimation allows me to measure the demand based on different network configuration. Table 3 shows the results of estimating the demand for consumer transactions in each time period. The results follow the expectation. There are two groups of consumers: digital users and non-digital users. The distance factor is the main factor that affects consumers' decisions on which branch to visit. The coefficient of distance variable is negative, showing that consumers tend to visit branches which are close to them. The farther the branch is from the consumer, the less likely is the consumer to visit that branch. The result also shows that consumers who are digital channel users are more sensitive to the distance factor comparing to non-digital users. The size of the branch has little negative influence on consumer choice. The magnitude of the coefficient of the size is much smaller than the distance. Moreover, according to the results, consumers tend to go to the bank's original branches instead of the branches acquired from other banks. The probability of consumers choosing digital channels increases from 14.5% to 40.6% from 2009 to 2013 as shown in Table 4. I have partial consumer data from 2015. Based on the partial data from 2015, I see the share of using digital channel remain around 40% which means that 40% is an upper bound for the probability of choosing digital channels. Therefore, the assumption regarding the environment after 2013 is valid.

In the second stage, I estimate the lower and upper bounds of the values of θ based on the moment inequalities. I present the 95% confidence sets for θ and the number of perturbations in both experiments, $z_1^{h,k}$ and $z_2^{h,k}$. There are 162 perturbations in Experiment 1 and 142 perturbations in Experiment 2. The estimates imply that the transferring cost per mile per dollar of the sum of transactions is between \$0.019 and \$0.065. The 95% confidence set does not contain zero, suggesting that the transaction cost after branch closings is significantly different from zero.

1.6 Robustness

In Table 5, I present two different sets of robustness estimates. In the profit function under network configurations, I assume that the bank has perfect foresight in predicting for 10 years and the environment after 2013 remains the same. However, this may not be the case in the real world. Therefore, I modify the number of years the bank considers while making decisions on branch closings. I change the setup to 5 and 15 years of prediction. I find a lower bound of the confidence interval for the transferring cost as low as \$0.022 and an upper bound as high as \$0.69. The confidence interval does not contain zero and provides confidence that transferring costs occur while closing branches. Moreover, Based on the partial data from 2015, I see the share of using digital channel remain around 40% which means that 40% is an upper bound for the probability of choosing digital channels. Therefore, the assumption regarding the environment after 2013 is valid.

1.7 Implications and Counterfactual Experiments

With the estimates from the model above, it is possible to evaluate the current equilibrium and consider a variety of counterfactual experiments including expanding digital adoption and extending the limitation of transactions over digital channels.

1.7.1 Current Equilibrium

In this section, I first examine the revenue and the cost of the bank under the estimated model in the current equilibrium. The transferring cost per mile per dollar of the revenue of the sum of the amount of transactions under this section falls between \$0.019 and \$0.065. Given the estimated demand and cost functions, I evaluate the bank's total sum of transactions, building costs, and staff costs over the sample period (2009–2018) under the observed network configuration and calculate the benefit the bank receives after branch closings. The results are reported in Table 6. After branch closings, the bank loses part of the total sum of transactions, but saves the building cost and the staff cost. Under the current setting, the total revenue from the sum of the deposit falls between 2.674 billions. The building and staff costs are \$132.2 and \$135.2 millions, respectively. The inactive branch needs to transfer the consumers' information, and the net transferring cost is between \$37.59 and \$128.60 millions. Combining the revenue with all the costs, the discounted profit stream generated from the branches falls between \$2.278 and \$2.369 billions.

Now I consider the discounted profit stream under the network configuration which has no branch closing at all during the same sample period as shown in Table 6. Although there is no reduction in the total amount of transactions, the bank needs to pay the building and staff cost to maintain all the branch. The total revenue from sum of the transactions is 2.728 billions. The building and staff costs are \$269.7 and \$169.6 millions. However, since all banks are active, there is a one-time transferring cost for each branch closing. the bank can save the building and staff cost in the long run. The total cost under the environment with no branch closing at all is 42.67% higher than under the current network configuration. The discounted profit stream with no branch closing becomes \$2.289 billions. The results show that between 0.53 and 1.06 years after the closing progress in 2013, the profit associated with the current network configuration would be higher than the profit with no branch closings. Besides branch closings, the bank is developing more advanced digital channels and expanding the online banking service. The following implication follows the digital adoption experiment.

1.7.2 Expanding Digital Adoption

In the current equilibrium, I assume that the consumers who are not using alternative digital channels including ATMs, website, and mobile applications will not be aware of those possibilities. Therefore, the demand function includes the non-digital users' transactions, and the digital users' transactions that can only be made at a branch in this setting. However, the bank is building the new interactive ATM and better mobile applications and helping consumers to explore the new services on their phone. To represent an increase in the adoption of digital channels, I increase the number of consumers using digital channels.

In the current equilibrium, the percentage of consumers who use digital channels is 14.5% in 2009 and increases to 40.6 in 2013. There are two different setups in this experiment. In the first setup, I reduce the number of non-digital transactions by 10%, 20%, and 30% from all branches. In the data set, I randomly select non-digital transactions which is under \$2,500 and make them perform over digital channel. The reason I only choose transactions under \$2,500 is that Bank A allows each consumer to deposit checks under \$2,500 per day and \$5,000 per month through a mobile phone application. Their transactions which can be performed over an ATM or mobile application are now considered as outside option. In the second setup, I randomly choose non-digital transactions only from the inactive branches.

The results in the first setup are shown in Table 10. The multinomial estimation for different level of adoptions are shown in Table 7, 8, and 9. The net transferring cost is between \$28.84 and \$105.4 millions for a 10% increase in digital adoption and between \$20.87 and \$75.66 for a 30% increase in digital adoption. The bank's savings from the net transferring cost are between \$8.75 and \$52.94 millions depending on the percentage of the increase in digital adoption. When this percentage increases, both the amount of revenue from branches and the transferring cost decreases. However, if we combine the revenue from the branches with revenue from the digital channel, the net gain is positive. The results in the second setup are shown in Table 11. The net transferring cost falls between \$16.91 and \$48.87 millions for a 10% increase in digital adoption and between \$7.42 and \$23.74 for a 30% increase in digital adoption. If the bank focuses on helping more consumers in inactive branch to adopt the digital channels, the transferring cost would lower significantly in the estimation. Instead of spending time and money educating consumers in every branch, the bank could benefit more by just educating the consumers in branches that they are planning to close in the near future to lower the transferring cost.

All in all, this experiment indicates that as more consumers use digital channels, both the demand and the transferring cost for the branches decrease significantly; therefore, banks are expected to accelerate the branch closing process. According to FDIC data, the number of branches in the U.S. shrank by 1947 in 2018, a new record high, and the closing speed has increased year over year since 2010.

1.7.3 Extending the Limitation of Transactions over Digital Channels

Although there are no regulations on banking services, each bank has its limit on mobile check deposits and ATM withdrawals due to the money available in the machine. Also, some banks do not allow online wire transfers. In this study, Bank A allows each consumer to deposit checks under \$2,500 per day and \$5,000 per month through a mobile phone application and withdraw only \$500 daily through ATMs. They will have to go directly to the counter in the branch to withdraw more than \$500 a day or to deposit a large amount of cash. Moreover, there is no online wire transfer service in Bank A. Currently, the banking industry is investing a lot of money into financial technology to create a more efficient environment for both the bank and consumers. For example, Chase Bank allows consumers to request wire transfers online up to \$100,000 daily, and Bank of America allows \$1,500 cash withdrawals from ATMs. With more advanced machines and technology, the bank can relax the limitation over digital channels and decrease the demand of the branch.

In this experiment, I assume that Bank A allows consumers to perform \$5,000 cash deposit instead of \$2,500 through a mobile phone application. However, this relaxation would only benefit consumers who are already digital channel users. Therefore, I select only the transactions from digital users with the amount of money between \$2,500 and \$5,000. Under this scenario, only about 3.55% of the transactions in the data set would be removed from the demand of the branch. The results are provided in Table 12. After extending the limit, the net transferring cost falls between 36.09 and 117.3 millions which is about 6.4 millions reduction of the net transferring cost. Even though it only benefits a small number of consumers and the marginal transferring cost stays the same, it still decreases the transferring cost while closing a branch.

1.8 Conclusion

I examine the impact of rising adoption of digital bank channels and information transferring cost on the shrinking brick-and-mortar branch network in the banking industry from 2009 to 2013. As digital channels become more common to consumers, banks are closing branches to reduce personnel and occupancy expenses; however, the bank has to incur the cost to transfer the customer accounts and the soft information including consumer-specific relationships to other nearby branches. In the estimation of the demand, the consumers are sensitive to the travel distance to a branch, as expected. The bank carefully considers the network configuration, the transferring cost, the branch demand, and the potential duplicate serving areas to decide which branch to close. In the model, the net transferring cost depends on the distance between the closing branch and the other branches in the network and the magnitude of the demand; therefore, it captures the pattern that the bank tends to close the branches that serve the same areas as other branches and have a lower magnitude of demand due to the lower net transferring cost.

The study shows that expanding digital channel adoption could accelerate the branch closing process. In the experiment, the digital adoption decreases the demand of the branches. Moreover, it also reduces the net transferring cost if the branch is closed. As a result, it creates an incentive for the bank to close even more branches because it could save even more with lower transferring costs. However, branch closings harm the relationship between consumers and the branches, especially small business owners. The business loan seekers usually rely on the soft information which is built by the close relationship with the branch to receive credit supply. I examine the negative impact on loan industry in Chapter 2.

This is the first study to model the shrinking network of branches in the banking industry. As digital usage continues to grow, the model would be useful to determine the banks' savings from branch closings. This study has some potential extensions. With more data, I can measure consumer behavior with other competitive branches to determine the competitive effects and potential delay in the exit process which is commonly seen in a competitive environment.

Figure 1: Number of Branches from 1990 to 2018

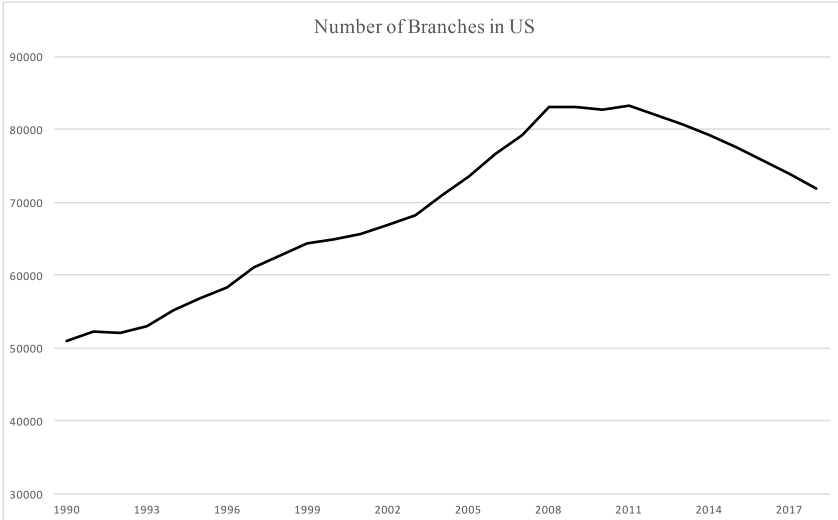


Table 1: Example of Transaction Data

Date	Description	Channel	Amount
04/30/13	Check Deposit	Branch	\$150.00
05/12/13	ATM Deposit	ATM	\$88.00
05/13/13	Check Balance	Mobile	\$0
05/13/13	ATM Deposit	ATM	\$200.00
06/01/13	Cash Deposit	Branch	\$32.24

Figure 2: The Numer of Branches in City A from 2009 to 2015

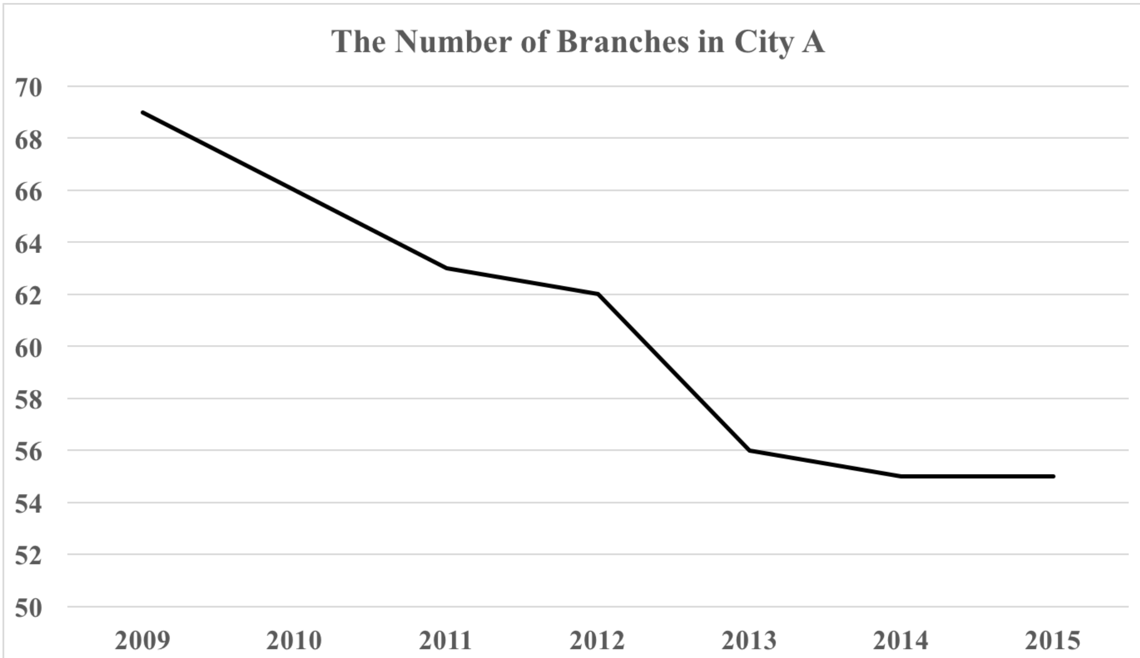


Table 2: Summary Statistics of Branch-Level Data

	Variable	Mean	Std. Dev.	Min	Max
Inactive	Branch Real Estate Cost	901240.93	1059946.03	108000	4116941
Active	Branch Real Estate Cost	603467.34	422330.85	43500	1671700
Inactive	Network Distance	4.58	2.37	1.56	10.73
Active	Network Distnace	4.00	2.80	0.31	12.24
Inactive	Amount of Deposit	2022.32	686.11	1115.21	3187.28
Active	Amount of Deposit	2885.07	1450.81	750.40	7732.44
Inactive	Number of Transactions	7221.33	9664.18	996	32691
Active	Number of Transactions	17342.36	15681.88	1149	86496

Table 3: Multinomial Logit Demand Model

	2009	2010	2011	2012	2013
Distance (Mile)	0.018***	-0.012***	-0.002	-0.033***	-0.032***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)
Distance*Online (Mile)	-0.205***	-0.309***	-0.346***	-0.414***	-0.361***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Acquired Branch	-0.570***	-0.491***	-0.569***	-0.599***	-0.485***
	(0.013)	(0.010)	(0.013)	(0.014)	(0.014)
Branch Cost (Dollar $\times$.001)	-0.009***	-0.005***	-0.001***	-0.004***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	50000	50000	50000	50000	50000
Number of Branches	69	67	65	62	56

Standard errors clustered at the zip code level level. *** p<0.01, ** p<0.05, * p<0.1

Figure 3: Probability of Choosing Digital Channel

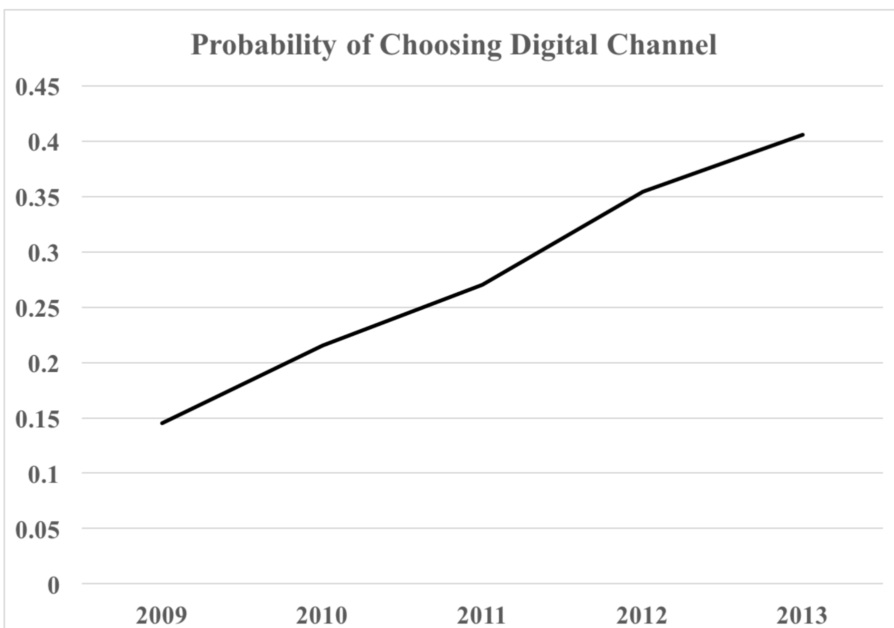


Table 4: Probability of Choosing Digital Channel

	Predicted Probability	# Branches
2009	0.145	69
2010	0.215	67
2011	0.270	65
2012	0.354	62
2013	0.406	56

Table 5: Robustness: Modifying Rate of Return

	Transferring Cost	# of Perturbations
5 Years	[0.022, 0.062]	Group 1: 186; Group 2: 163
10 Years	[0.019, 0.065]	Group 1: 162; Group 2: 142
15 Years	[0.042, 0.069]	Group 1: 198; Group 2: 167

Table 6: The Comparison between Current Equilibrium and Network without Branch Closings

	Current Equilibrium	No Closing
Revenue (Billions)	2.674	2.728
Building Cost (Millions)	132.2	269.7
Staff Cost (Millions)	135.2	169.6
Marginal Transferring Cost (Dollar/Mile)	[0.019, 0.065]	0
Transferring Cost (Millions)	[37.59, 128.6]	0
Discounted Profit (Billions)	[2.278 2.369]	2.289

Table 7: Counterfactual Experiment 1: Multinomial Logit Demand Model (10%)

	2009	2010	2011	2012	2013
Distance (Mile)	0.030***	-0.015***	-0.010***	-0.034***	-0.025***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Distance*Online (Mile)	-0.287***	-0.265***	-0.280***	-0.346***	-0.321***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
Acquired Branch	-1.126***	-0.643***	-0.598***	-0.654***	-0.628***
	(0.015)	(0.013)	(0.014)	(0.010)	(0.015)
Branch Cost (Dollar $\times$.001)	-0.013***	-0.007***	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	50000	50000	50000	50000	50000
Number of Branches	69	67	65	62	56

Standard errors clustered at the zip code level level. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Counterfactual Experiment 1: Multinomial Logit Demand Model (20%)

	2009	2010	2011	2012	2013
Distance (Mile)	0.000	-0.003***	-0.005*	-0.010***	0.017***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Distance*Online (Mile)	-0.221***	-0.319***	-0.304***	-0.374***	-0.355***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
Acquired Branch	-1.126***	-0.656***	-0.640***	-0.713***	-0.660***
	(0.014)	(0.013)	(0.014)	(0.015)	(0.015)
Branch Cost (Dollar $\times$.001)	-0.014***	-0.009***	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	50000	50000	50000	50000	50000
Number of Branches	69	67	65	62	56

Standard errors clustered at the zip code level level. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Counterfactual Experiment 1: Multinomial Logit Demand Model (30%)

	2009	2010	2011	2012	2013
Distance (Mile)	0.018***	-0.033***	-0.004***	0.013***	0.055***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)
Distance*Online (Mile)	-0.240***	-0.332***	-0.339***	-0.423***	-0.447***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.0034)
Acquired Branch	-1.338***	-0.780***	-0.713***	-0.716***	-0.6275***
	(0.155)	(0.014)	(0.015)	(0.02)	(0.016)
Branch Cost (Dollar $\times$.001)	-0.016***	-0.010***	-0.001***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	50000	50000	50000	50000	50000
Number of Branches	69	67	65	62	56

Standard errors clustered at the zip code level level. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Counterfactual Experiment 1: Expanding Digital Channel Adoption

	Current Equilibrium	10% Adoption	20% Adoption	30% Adoption
Revenue (Billions)	2.674	2.523	2.305	2.104
Marginal Transferring Cost (Dollar/Mile)	[0.019, 0.065]	[0.020, 0.063]	[0.017, 0.059]	[0.016, 0.056]
Transferring Cost (Millions)	[37.59, 128.6]	[28.84, 105.4]	[25.23, 92.18]	[20.87, 75.66]
Discounted Profit from Branches (Billions)	[2.278 2.369]	[2.076, 2.152]	[1.946, 2.012]	[1.761, 1.816]

Table 11: Counterfactual Experiment 1: Expanding Digital Channel Adoption in Inactive Branches

	10% Adoption	20% Adoption	30% Adoption
Revenue (Billions)	2.664	2.646	2.627
Marginal Transferring Cost (Dollar/Mile)	[0.009, 0.026]	[0.007, 0.020]	[0.005, 0.016]
Transferring Cost (Millions)	[16.91, 48.87]	[11.77, 31.96]	[7.42, 23.74]

Table 12: Counterfactual Experiment 2: Extending the Limitation of Transactions

	Current Equilibrium	Counterfactual Experiment 2
Revenue (Billions)	2.674	2.585
Marginal Transferring Cost (Dollar/Mile)	[0.019, 0.065]	[0.020, 0.065]
Transferring Cost (Millions)	[37.59, 128.6]	[36.09, 117.3]
Discounted Profit from Branches (Billions)	[2.278 2.369]	[2.205, 2.277]

2 Chapter II: Understanding the Effects of Bank Branch Closures on the Local Loan Market

2.1 Introduction

The objective of this paper is to study the effects of bank branch closings on individual business borrowers. After decades of expanding their branch networks, banks have begun reconsidering their branching strategies due to changes in technology and associated changes in consumer behavior. Online banking platforms and advanced automated teller machines have been widely developed. Instead of visiting brick-and-mortar bank branches, more people are using mobile applications or online platforms to use banking services, including making inquiries, transfers, and deposits. The banking system has made huge improvements in technology. In response to their customers' increasing use of technology, banks are making changes in their branch networks. For example, in 2013, a Capital One spokeswoman disclosed plans to evolve and optimize the bank's branch network to ensure that the bank operates as efficiently and effectively as possible according to Wall Street Journal. Banks say they are carefully considering which branches to close, examining deposit levels at each branch and the commute time to the nearest location (Ensign, Rexrode and Jones, 2018). Figure 4 shows that after the recession in 2009, the number of branches for all banks in the U.S. has been decreasing. The branch closing process is expected to continue and has created widespread concern regarding the negative impact on the social welfare.

Bank branches are important to the local economy. They provide not only physical access to customers but also the credit sources to local businesses. Changes in the market structure of banking could widely affect the local industries. Therefore, banks are one of the most regulated industries in the U.S. The government wants to ensure local access to the bank networks. For example, the FDIC implements a policy that requires banks provide a 90-day notice before any branch closure. However, there is still a lot of criticism about the negative influence on the local economy of current branch closures in both rural and urban neighborhoods.

This paper evaluates the impact of branch closings on the local credit supply and demand as measured by the volume of small business lending. Having access to a unique data set of individual small business loan applications, I am able to examine the impact of branch closures from many more aspects than other papers. First, I study how branch closures affect local consumers' behavior when searching for loans. The loan seekers might look for an alternative loan provider depending on the availability of loan access in the area. Second, I measure the damage to loan seekers caused by the loss of personnel-specific soft information. The soft information is the information that cannot be easily obtained through physical data, such as revenue, sales number, and cost. This information is usually built based on the relationship between the local branch loan managers and the loan seekers, which depends on the existence of nearby brick-and-mortar bank branches. The damage to loan seekers is measured by the approval rate and the length of processing time. Third, using the locations of loan seekers and bank branches, I estimate the geographic spillover of the impacts of branch closings. Consumers living in urban areas could face different effects from those living in rural areas. Moreover, using the individual-level data, I compare the different types of applicants including industries and qualities.

To estimate the impact of branch closings on the local credit market, I compare the control and treatment groups using a difference-in-difference framework. Loan seekers who are located within 15 miles from an inactive branch are considered as the treatment group. However, one of the difficulties of such a study is that both the decision of branch closure and the tendency of loan seekers could be related to economic shocks; thus, the endogeneity problem could break down the framework and create bias in the estimation. To overcome this problem, I consider only the inactive branches which were acquired from other banks or merged with the branches of other banks. Yang (2019) shows that a bank tends to close the acquired branches due to the redundancy with the original branches. As a result, the decision of the acquired branch closure is not driven by the local economic factors, especially the local business loan shocks.

A rich body of existing literature has studied how the local bank networks matter for local economies (Kerr and Nanda (2009); Gilje (2012); Gilje, Loutskina and Strahan (2013); Townsend and Zhorin (2014); Nguyen (2019)). However, different from exiting literature, this paper uses a unique individual-level data set to study the local effects of branch closings on both the demand and supply sides. This paper performs the analysis of local loan markets at a much detailed scale. The authors of these papers are concerned that large banks are using universal platforms on evaluating loan applications and that relationship-intensive lendings are ignored in such platforms. Therefore, when large banks acquire smaller ones, the small business owners who are seeking extra money supply could be damaged. Other authors estimated the effects of negative local credit supply shocks on the state or county level (Peek and Rosengren (2000); Ashcraft (2005); Greenstone, Mas and Nguyen (2019)). However, with the unique date set, I study the branch-level shocks on the credit supply.

This study yielded several interesting results. First, branch closings have a negative impact on the local credit demand. The areas with branch closures lose 563.1 loan applications relative to the areas with no branch closures. The average requested loan amount is \$172,353. Therefore, the bank loses about \$97 millions a year from small business loans due to branch closures. The result indicates that local business owners may search for new alternatives to submit their loan applications due to the loss of brick-and-mortar branch access.

Second, the loan applications in the areas with branch closings have a higher approval rate but a longer reviewing process. The approval rates in the areas with branch closings are 8.11% higher than in the areas with no branch closings. In this study, I use the data set provided by a large U.S. bank. One of the reasons the approval rate increases after a branch closing is that the bank adopts the strategy to approve more business loans in order to offset the loss of loan demand. Moreover, the bank needs more time to approve or decline an application in the areas with branch closings. The reviewing process of loan applications could take longer because of the lack of labor force or the difficulties in collecting and transferring information.

Finally, branch closings have more negative effects on the volume of loan applications in the areas closer to the inactive branch and with a higher density of competitors' branches. The branches in the areas with more than 20 competitors' branches lose 1,322 applications, and those in the areas with less than 20 competitors' branches lose only 579.1 relative to the areas with no branch closures. Cities usually have a high density of branches; therefore, the result shows the bank could lose profit from business loans by closing a branch located in a city.

The remainder of the chapter is structured as follows. In the next section, I describe the detail of the data sets used in the study. Section 2.3 providesthe empirical framework associated with the identification. Section 2.4 shows the results of the empirical model. Section 2.5 concludes the paper.

2.2 Data

In this study, I use three data sources: a U.S. bank (Bank A), the American Community Survey, and the FDIC's Summary of Deposits. In this section, I describe the different aspects of the data set in detail. The summary statistics are shown in Table 13.

2.2.1 Loan Applications

The unique anonymized data set I analyze consists of over 100 thousands loan applications from a large U.S. bank (Bank A) on a small business industry in the U.S. from 2010 to 2018. The applicants are located accross US. For each application, an anonymized account identifier, date, industry, income level, loan requested amount, and approval result were provided by Bank A along with the anonymized customer information.

The rich loan application data set makes the studies on both the demand and supply sides feasible. From the demand perspective, the applicants are located across US. The applicants own businesses with 1,174 different NAICS codes, operating in 8 industry segments. For example, the applicants own dental offices, landscaping services, full-service restaurants, religious organizations, etc. The mean and median amounts of the requested loans are about 170 thousands and 7 thousands, respectively. About one-third of the applicants are labeled as low-moderate income level, which is below 80% of the median income for the area they reside in. From the supply perspective, the approval rate from Bank A is between 50 to 60 percent. The mean and median amounts of the approved loans are about 100 thousands and 25 thousands, respectively. The largest approved amount is over 7 millions. Considering all the applications, the bank takes on average 4 weeks to approve or decline a loan application.

2.2.2 Branch Characteristics

The branch characteristics data, provided by Bank A, consists of 3,461 observations of branches in the U.S. from 2004 to 2018. It contains information about where the branch is, when the branch became inactive, and whether the branch was acquired from others. The branches are located across US. In 2015, 798 branches were inactive and 2,491 branches were acquired from others. By combining the loan application and branch data set, I can analyze how close the applicants are located to the branches. The applicants are about 5 miles away from the closest branch on average. Surprisingly, the furthest applicant is over 4 thousands miles away from the branch. The mean and median distance between the applicants and inactive branches are about 2.7 and 1.7 miles, respectively.

2.2.3 Competition and Geographic Characteristics

With more credit access availability, local business owners have more flexibility to submit loan applications. In order to study the relationship between credit access and the demand for business loans, I use the FDIC's Summary of Deposits, which provides annual information of all branches belonging to FDIC-insured institutions. The information consists of the status, address, and amount of deposits of the branch. After pinpointing their locations using latitude and longitude, I can determine the distance and number of competitors' branches in the applicants' neighborhoods. On average, within an area with a 2-mile radius from a Bank A branch, the applicants can access about 6.9 competitors' branches. The largest number of competitors' branches in the same size area is 126. In terms of distance, a Bank A branch is on average 0.53 miles away from the closest competitor's branch. The furthest competitor's branch is 16.67 miles from a Bank A branch.

To have controls on the real economic effects of branch closures on loan applications, I use the U.S. gross domestic product (GDP) from the Bureau of Economic Analysis and the American Community Survey (ACS) data from the United States Census Bureau. The ACS data set provides annual ZIP-level geographic characteristics including education level, median family income level, and population.

2.3 Identification and Empirical Framework

This section presents the different versions of difference-in-difference frameworks and the robustness checks of the model.

2.3.1 Model

This paper evaluates the impact of branch closures on the local credit supply and demand. I use a difference-in-difference (DID) estimation to measure this impact similar to Nguyen (2019). The estimation is feasible since there is data on the number of small business applications from both pre-intervention and post-intervention periods. However, the empirical challenge is that the bank chooses which branches to close, and the decisions of branch closings might be related to the local economy including small business loan demand. In other words, the bank has the incentive to deactivate the branches whose future profits are expected to be low. This creates an endogeneity problem, which makes the estimation of the effects of branch closing subject to bias.

Based on the findings of Yang (2019) and Nguyen (2019), the bank tends to close the branches acquired from other banks or merged with the branches of other banks because the acquired branches serve the same areas as the existing branches. Therefore, the bank's decision of closing the acquired branches is not driven by local factors, especially small business credit condition. As a result, I select only the inactive branches that are acquired from other banks from the data and create the treatment group based on these branches to solve the endogeneity problem.

Tha main focus of the research is to study the effect of a branch closure on local lending. The primary specification of the DID framework is

$$y_{mt} = \alpha_t + X_i\beta + M_m\gamma + Competition_i\rho + Post_i\theta + Treat_i\omega + (Post_i \times Treat_i)\delta + \epsilon, \quad (10)$$

where α_t is the year fixed effect; X_i is a vector of applicant i's characteristics; M_m is a vector of geographic information in the area m; $Post_i$ is a dummy equal to 1 if the date of application i is in the post-intervention time period; and $Treat_i$ is a dummy equal to 1 if the applicant i is in the treatment group. An applicant is in the treatment group if there is at least one branch closing within 8 miles from the location of the applicant. The geographic information includes population, median family income level, and median age. The coefficient δ is the main focus of this paper, and it measures the difference, conditional on controls, in the outcome y_{mt} between the control and treatment groups in the period following a branch closure.

In this paper, because of the rich unique data set available, it is possible to study the effects of branch closings on both the supply and demand aspects. Therefore, the dependent variable y_{mt} varies according to the focus of the effects, that is, the demand or the supply side. For the study of the demand side, y_{mt} is the number of loan applications in a given area m and year t. Small businesses in the area where there is a branch closing might search for alternative credit providers. Consumer behavior could be affected by the number of alternative choices, which is represented as *Competition* in the model.

For the study of the supply side, there are two different dependent variables. First, y_{mt} is the approval rate by Bank A in a given area m and year t. The approval rate could be impacted by the branch closing because the soft information between the branch and the local industry is deteriorated. In this case, the model can measure the effect of branch closings on the credit supply to the local businesses. Second, y_i is the number of loan application processing days by Bank A of a given applicant i. The length of processing time includes both approved and declined applications. Therefore, this model determines the impact of branch closings on the pace of processing rates by Bank A for both approved and declined loan applications.

2.3.2 Robustness

As the treatment and control groups differ in the number of loan applications, the internal validity of the DID framework requires a parallel trend assumption. As shown in Figure 5, before time period 0, both the control and treatment groups are increasing and close to the parallel trend. In order to have a robustness check on the parallel trend assumption of the DID framework, I estimate a multi-period version of the DID which includes both leads and lags of the treatment. In this manner, the model checks if the treatment and control groups behave similarly before the intervention happens. In other words, if the leads of the treatment matter in the estimation, the DID framework is problematic since the intervention occurs before the expected time, and this could create bias in the estimation.

The primary specification of the multi-period DID framework is

$$y_{mt} = \alpha_t + X_i\beta + M_m\gamma + Competition_i\rho + Treat_i\omega + \sum_{\tau}(D_{i\tau} \times Treat_i)\delta_\tau + \epsilon, \quad (11)$$

where α_t is the year fixed effect; X_i is a vector of applicant i's characteristics; M_m is a vector of geographic information in the area m; $Treat_i$ is a dummy equal to 1 if the applicant i is in the treatment group; $D_{im\tau}$ is a dummy equal to 1 if year t is τ years after the branch closing in the area m. Here τ ranges from -5 to 5. The geographic information includes population, median family income level, and median age. The coefficients of leads and lags, δ_τ, show the difference between the control and treatment groups in outcome y before and after the branch closing, respectively. As shown in Figure 6, 7 and 8, the coefficients of lead are close to zero meaning that the parallel trend assumption is valid.

2.4 Results

The results in this chapter are shown in 95% confidence interval instead of the exact number due to the agreement of the usage of the data source.

2.4.1 Branch Closing and Credit Demand

In this section, I show that branch closings have a negative impact on the local credit demand. The dependent variable in this section is the number of loan applications in a given year and a given area from Bank A. The result of the model is shown in Figure 6. The lead variables in the pre-intervention periods are not statistically significant, meaning that the difference between the control and treatment groups is not significantly different from zero. As a robustness check, the lead variables show that the assumption of the parallel trend before the intervention effects is valid in the study.

The coefficients of lags show that in a setting with branch closings, there is a decline in the volume of new small business loan applications that Bank A receives. The absolute value of the difference between the control and treatment groups increases in a range between 380 and 600 from the first period to the fourth periodm and all of them are statistically significant; therefore, the negative impact of branch closings on credit demand intensifies with time.

To better express the magnitude of the effects of branch closings, I estimate the less flexible two-period version of the DID. The specification of the framework is

$$y_{mt} = \alpha_t + X_i\beta + M_m\gamma + Competition_i\rho + Post_i\theta + Treat_i\omega + (Post_i \times Treat_i)\delta + \epsilon, \quad (12)$$

where $Post_i$ is a dummy equal to 1 if the date of application i is in the post-intervention time period and $Treat_i$ is a dummy equal to 1 if the applicant i is in the treatment group. In Table 14, the interest of the coefficient, δ, is in the confidence interval between -563.30 and -562.90, indicating that Bank A in the areas with branch closings loses about 563 loan applications relative to the areas with no branch closings. The average requested loan amount is about 170 thounds dollar. Therefore, Bank A loses about $97 millions in a year from small business loans due to branch closures.

In addition, branch closings reduce the local customers' access to the brick-and-mortar branches of Bank A. As a result, local businesses refuse to continue to seek credit supply through Bank A. It is possible that they might search for new alternatives to submit their loan applications. The competition effects will be addressed later in this paper. The result here shows that branch closings reduce the local industries' access to credit, which could not only hurt the local economy but also damage the bank's own credit business by making it lose potential loan seekers in these areas.

2.4.2 Branch Closing and Credit Supply

Branch closings may destroy the relationship between banks and loan seekers. Some papers in the existing literature (Stein (2002); Berger, Miller, Petersen, Rajan and Stein (2005)) argue that soft information is lost when big banks take over the local banks and control the local area. Although banks transfer information between branches after merging, the soft information is difficult to transfer. Therefore, as this information is lost after branch closures, the credit access channel could be stricter, which makes it more difficult for business loan seekers to obtain successful loan contracts. To address this question, one of the dependent variables is the approval rate of a given year and area; another one is the number of days it takes to process a loan application.

First, in the study of the approval rate, the estimation shows that the branch closings have positive effects on the approval rate of small business loans. In the multi-period version of the DID, the coefficients of leads in Figure 7 are not statistically significant. Therefore, the pre-trend robustness test indicates that the assumption of a parallel trend is valid. In the two-period version of the DID in Table 14, the focus of the coefficient δ is in the confidence interval between 8.11 and 8.12, indicating that the approval rate in the areas with branch closings is higher than in the areas with no branch closings. It is surprising that the credit supply channel is actually wider after removing a branch in the area, contradicting the past findings of losing soft information. However, the difference between this paper and others is that my data set allows me to focus on only one bank strategy instead of looking at the aggregate data. Therefore, Bank A might adopt the strategy of approving more business loans in order to offset the loss of the loan demand in the areas with branch closings. Nevertheless, the strategy of Bank A cannot represent the whole picture of the bank industry.

Second, in the study of the loan application processing rate by Bank A, the model indicates that the length of the reviewing process in the areas with branch closings increases by less than a week. In the multi-period version of the DID, the coefficients of leads in Figure 8 indicate that the parallel trend test is valid. The results of a two-period version of the

DID in Table 14 show that it takes a little longer for Bank A to process loan applications, meaning that the closings have a negative impact on the loan seekers and the bank. To analyze this question deeper, I separate the data set into two groups based on whether the application was approved (the approved group)or declined (the declined group). The processing time of the declined group is longer in the treatment group after the intervention than the decline group. As the demand for business loan decreases and the approval rate increases, the lack of labor force and the difficulties of transferring and collecting the soft information from the local industries could be the reasons for longer processing time.

2.4.3 Geographic Spillover

Branch closures may impact nearby customers differently than the customers located further away. To examine how local the effects of a branch closing are, I vary the distance of the treatment groups from their closest inactive branch and compare the difference-in-difference coefficients in both the demand and supply sides. To simplify the comparison, I use the two-period version of the DID instead of the multi-period. The main variable of interest is $Post \times Treated$. The independent variable for the demand side is the volume of applications, and for the supply side, the approval rate.

First, Table 15 shows the coefficients of the treatment group for the demand side with varying distances from the closest inactive branch: [-711.7, -711.0] for less than 3 miles, [-21.26, -21.01] for 3–6 miles, and [373.9, 374.8] for 6–10 miles. The demand for a business loan from Bank A decreases the most within the areas with a 3-mile radius from an inactive branch. As a result, the branch closing has a greater negative impact on the local credit demand in the nearby areas. The customers farther away from the branch are affected the least from a branch closing.

Second, Table 16 shows the coefficients of the treatment group for the supply side with varying distances from the closest inactive branch: [2.52%, 2.54%] for less than 3 miles, [0.46%, 0.47%] for 3–6 miles, and [-1.8%, -1.79%] for 6–10 miles. A comparison of the demand andsupply sides reveals that as the volume of the applications decreases, the approval rate increases. The reason could be that Bank A has an incentive to approve more applications in the areas with lower demand to offset the loss of applications due to branch closings.

2.4.4 Competition and Urban vs. Rural Areas

In this section, I compare the impacts of branch closings in areas with a different number of competitors' branches. As urban areas have a high density of bank branches, the analysis of the impacts of the strength of competition is the same as the analysis of the competition inurban and rural areas. I separate the applications into two groups using the threshold of 20 competitors' branches within an area with a 3-mile radius from an inactive branch. Then, I run two-period versions of the DID using two independent variables, the volume of applications and the approval rate.

First, as shown in Table 15, the estimations of the volume of applications indicate that Bank A loses more applications after branch closures in the areas with a high density of competitors' branches. The interest of the coefficient is in the confidence interval between -1,327 and -1,318 in the areas with more than 20 competitors' branches, and between -579.5 and -578.6 in the other areas. The cities with more branches have various credit access options available to business loan seekers. Therefore, when a nearby branch is closed, the loan seekers are likely to choose competitors' branches.

Second, the results of the supply side show that branch closures have positive effects on the approval rate in the areas with a high density of competitors' branches. As seen in Table 16, the interest of coefficient is in the confidence interval between 0.114 and 0.117 in the areas with more than 20 competitors' branches and between -0.813 and -0.811 in the other areas. After its branch closings in urban areas, Bank A has less demand for business loans due to strong competition from other banks in these areas; therefore, the strategy of Bank A might be to loosen the rules of the loan application reviewing process and increase the approval rate.

The estimation shows that Bank A has a lower approval rate in the areas with less competition after branch closings. Loan seekers in the areas which already have less availability of banking services have even more difficulties getting their business loan applications approved. As a result, branch closings could have larger negative effects on the local economy in these relatively disadvantaged areas.

2.4.5 Heterogeneity Across Applicants

In this section, I study the impact of branch closings on different qualities of applicants depending on their income level. I separate the applicants into two different groups: low-to-medium-income and high-income groups. The estimations of the volume of applications show that Bank A loses more applications from the high-income group than from the low-to-medium-income group in the areas with branch closings. The coefficients of the applicants in the low-to-medium-income and high–income groups are in the confidence interval between -692.1 and -688.5 and between -1,095 and -1,085, respectively (see Table 15). The reason could be that high-income-level loan seekers are higher-quality borrowers for credit institutions; therefore, they would have access to more credit options. When a nearby branch is closed, the high-income-level loan seekers have a higher incentive to choose an alternative option for their business loans.

The analysis for the supply side in Table 16 shows that the applications from the high-income-level group have a lower approval rate in the areas with branch closings. The estimations for the low-to-medium-income and high-income groups are in the confidence interval between 2.62% and 2.66% and between -3.50% and -3.48%, respectively (see Table 16). Bank A chooses to approve the loans of fewer high-income-level applicants after branch closures while at the same time the demand of high-income-level applicants is decreasing.

2.5 Conclusion

This paper uses a unique small business loan application data set and difference-in-difference framework to estimate the impacts of branch closings on both the demand and supply sides of business loan markets. I show that closings have negative effects on the bank's business sector, especially in cities and areas with a high density of competitors' branches when loan seekers have more alternative options to apply for business loans. The bank tends to loosen its loan application reviewing process to approve more loans to offset the loss of demand of business loans in the cities with branch closings. However, the study also shows that the length of the bank's application processing time gets longer after branch closings when the soft information and relationships between the local businesses and the branch are difficult to transfer and replace. Moreover, I find that these effects have a negative relationship with the distance between the applicants and branches, meaning the effects are localized.

Altogether, branch closings hurt not only the local loan business market but also the bank itself in the loan sector. The bank loses more high-income-level applicants in cities when they have a higher chance to be approved by other credit institutions. While the bank loosens its reviewing process in the areas with more competitors, it also creates more risk to the bank and also the whole banking system. In addition, loan seekers in the areas which already have limited availability of banking services are more difficult to get their loan applications approved by the bank after branch closings. In order to increase the credit supply in these relatively disadvantaged areas, the government needs to put more regulations on the bank industry regarding the branch closing and loan reviewing process.

Figure 4: Number of Branches from 2010 to 2018

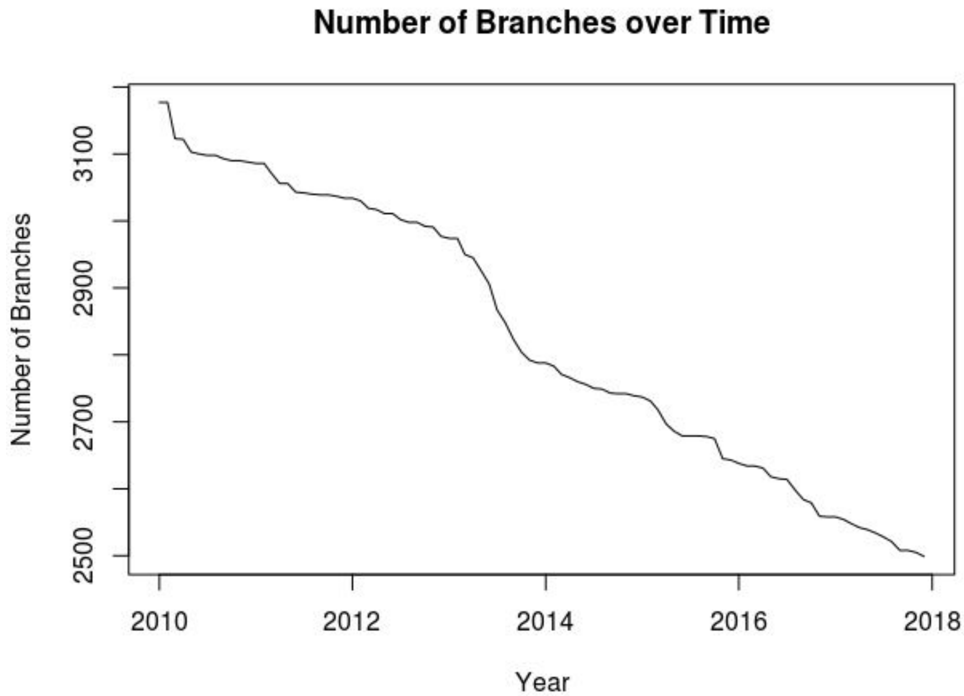


Figure 5: Credit Demand between Treatment (Black) and Control Group (Red)

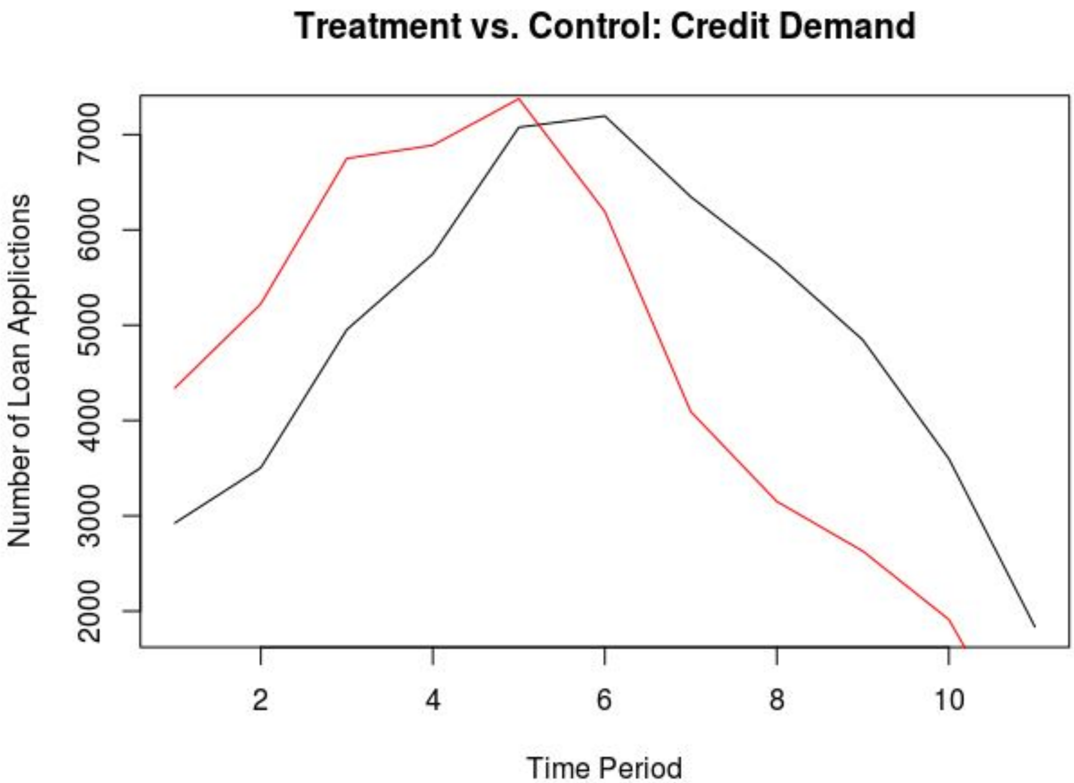


Table 13: Summary Statistics of Application Data

Variables	Mean	Median	Std. Dev.
Number of Applications	773.20	540.00	765.72
Number of Requested Amount	172,353	70,000	395,572.3
Number of Approved Amount	102,635	25,000	233,827.4
Population	25,423	23,599	16,251.72
Income	27,946	27,488	11,425.17
Number of Competitors' Branches	6.879	5.00	9.1917
The Distance: Applicants and the Closest Active Branch	4.963	1.357	54.276
The Distance: Applicants and the Closest Inactive Branch	11.303	6.527	55.359
Number of Applicants in Control Group	53,672		
Number of Applicants in Treatment Group	48,879		
Number of Applicants in Post-intervention	34,379		

Figure 6: Multi-period DID: Number of Loan Applications

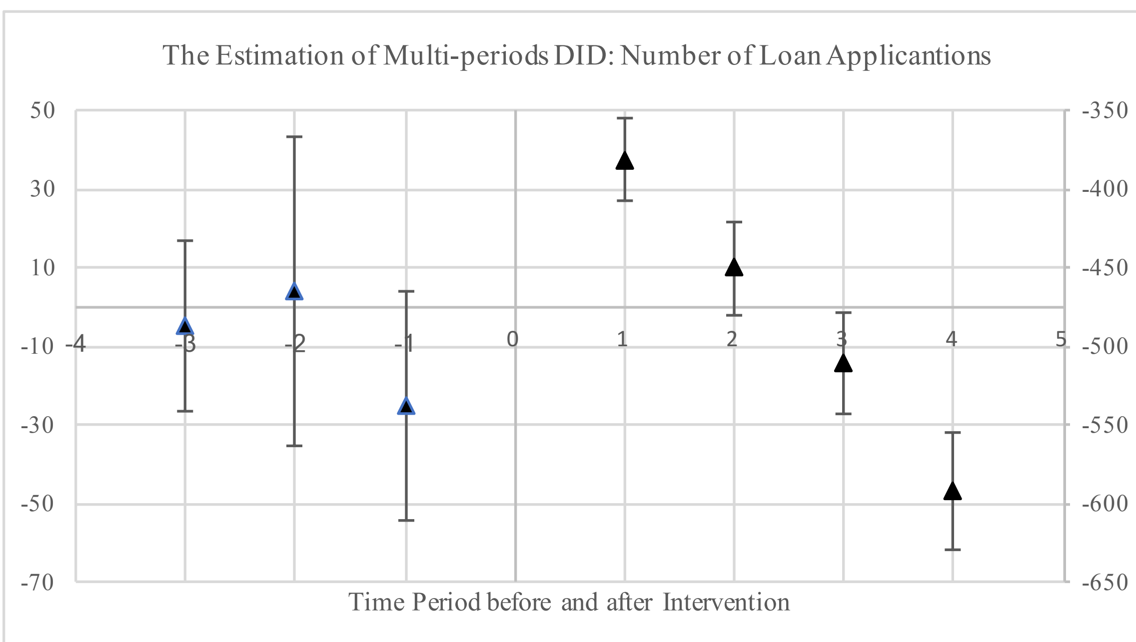


Figure 7: Multi-period DID: Approval Rate

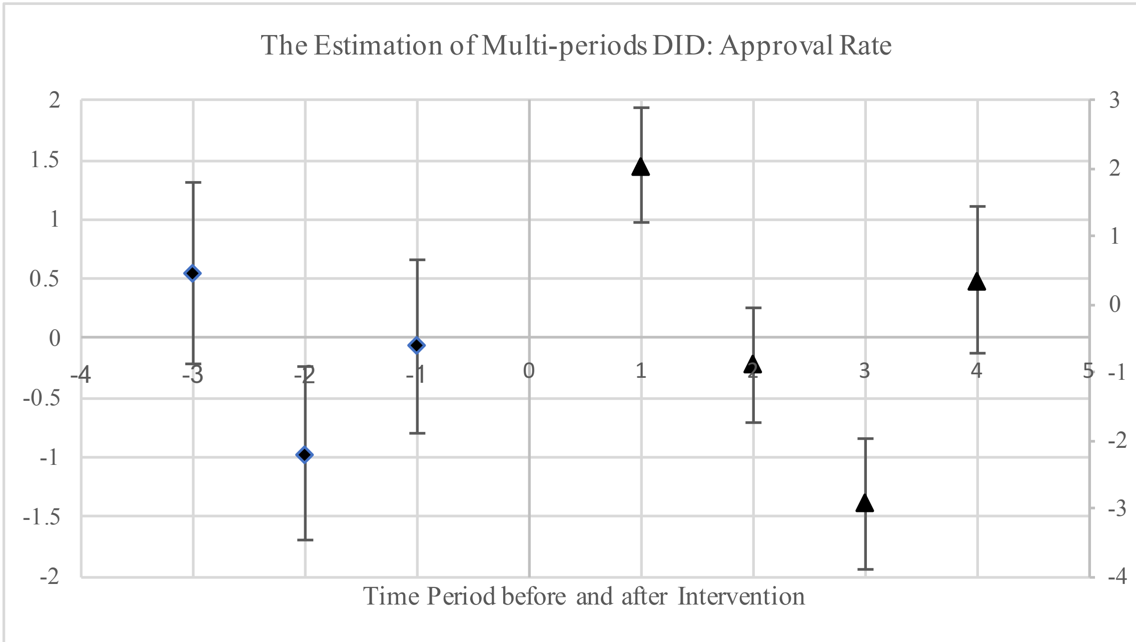


Figure 8: Multi-period DID: The Length of Processing Time

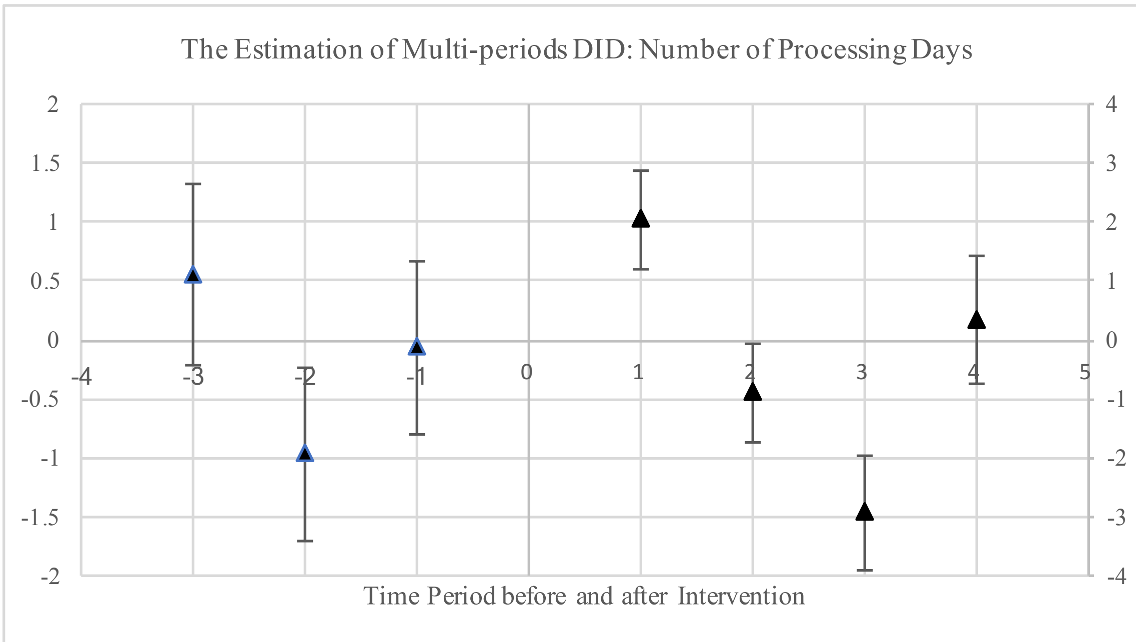


Table 14: The Estimation of δ on Both Supply and Demand Sides (95% CI)

	The Estimation of Coefficient δ under 95% CI	# Control	# Treatment
Number of Applications	[-563.30, -562.90]	62,847	58,028
Approval Rates	[8.108, 8.118]	62,847	58,028
Processing Days	[3.086, 3.130]	62,847	58,028

Table 16: The Estimation of δ on the Approval Rate: Different Groups of Interests (95% CI)

Categories	**The Estimation of Coefficient** δ (95% CI)	**Number of Data Points**
The Distance from the Closest Inactive Branch:		
Less than 3 mile	[2.521, 2.540]	71,419
3 mile to 6 mile	[0.460, 0.469]	68,632
6 mile to 10 mile	[-1.832, -1.795]	69,844
The Number of Competitors within 3 mile**:**		
Less than 20	[-0.813, -0.811]	64,505
Greater than 20	[0.114, 0.117]	38,046
The Quality of Applicants: Low-Medium Income Level		
High Income	[-3.513, -3.481]	71,172
Low-Medium Income	[2.621, 2.662]	31,379
The Quality of Applicants: Requested Amount		
Less than 40,000	[-0.455, -0.437]	31,656
40,000 to 100,000	[1.130, 1.140]	35,160
Greater than 100,000	[0.088, 0.096]	35,735

Table 15: The Estimation of δ on the Number of Applications: Different Groups of Interests (95% CI)

Categories	**The Estimation of Coefficient** δ (95% CI)	**Number of Data Points**
The Distance from the Closest Inactive Branch:		
Less than 3 mile	[-711.7, -711.0]	71,419
3 mile to 6 mile	[-21.26, -21.01]	68,632
6 mile to 10 mile	[373.9, 374.8]	69,844
The Number of Competitors within 3 mile**:**		
Less than 20	[-579.5, -578.6]	64,505
Greater than 20	[-1327, -1318]	38,046
The Quality of Applicants: Low-Medium Income Level		
High Income	[-1095, -1085]	71,172
Low-Medium Income	[-692.1, -688.5]	31,379
The Quality of Applicants: Requested Amount		
Less than 40,000	[-606.4, -604.1]	31,656
40,000 to 100,000	[-649.4, -646.1]	35,160
Greater than 100,000	[-604.5, -598.4]	35,735

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

3 Chapter III: Understanding the Supply Side of For-Profit Colleges: Structural Analysis

3.1 Introduction

The objective of this paper is to estimate the impact of recent government regulation and competition between colleges on market entry timing. In the current policy debate about American post-secondary education, private for-profit colleges are at the center of discussion. Their goal is to provide education while generating a positive profit. Therefore, many quality issues have been raised for for-profit colleges. Although there are some literature analyzing the quality and demand of for-profit school sector, little is known about the supply decisions of private for-profit colleges. Thus, this paper will focus on studying the supply decision made by the owners of private for-profits.

According to Deming, Goldin, and Katz (2012), students in for-profit colleges are less satisfied with their education and classroom experience compared to others who go to public and private nonprofit institutions. For-profit colleges tend to target independent students and those from low-income families who are eligible for Pell grants or federal student loans. They accounted for 24 percent of Pell grant disbursements and 26 percent of federal student loan disbursement in 2008-2009 even though they only enrolled 12 percent out of the students of the total college students. According to the 90/10 rule, for-profits can receive no more than 90 percent of their revenues from Title IV federal student aid including grants, loans, and work-study. However, for-profits are still able to receive a major fraction of their revenues from the government. Starting at 2010, state and federal government have begun focusing on the problems in the for-profit sector and released regulations to control the quality of education in for-profit schools. The detail of the regulations is described in Section 3.2.

In a strategic environment, a school's decision on when to enter the market depends on not only the cost and the market situation, but also the government regulation and the competition effect from other similar type of schools within the area. In the article "For-Profit Colleges" written by Deming, Goldin, and Katz (2013), they compared for-profits with community universities since they claim that both sectors have similar local serving property. Thus, this paper found the strong competition effect from other local for-profit and community colleges. The growth of for-profit sector has raised concerns at both the state and federal levels about the quality of education in for-profits, the strategy they use to attract students, the graduation rate, and the success of the graduates in finding a job. To show the effect of competition effect and influence of the government regulation, I develop and estimate a dynamic structural model of school market entry.

In the model setup, preemption motive effect is captured among for-profit schools. Schools are able to enter the market early to discourage other competitors to enter the market. Whoever decides to enter the market will lower the payoff for the competitors. The other competitors may decide to wait until the cost is low enough for them to earn profit before activating the schools. In this paper, I treat the markets independently and each market characteristics including the size of population, and people's wealth level will affect the activation time for for-profit schools.

By assuming schools make their own decision sequentially in every period, the setup provides me a unique subgame perfect equilibrium. After specifying the payoff and cost function for all schools, backward recursive algorithm can be used to solve the equilibrium play and Method of Simulated Moments (MSM) Estimator (McFadden (1989)) can be applied to estimate the parameters in the structural model. From the estimation, the competitive effect from other for-profit schools and community colleges and the effect from market variables can be revealed.

Estimates indicate that for-profit schools in a larger population or higher medium income market enter the market earlier. Both federal regulation and state regulation have negative effect on for-profit schools payoff and will delay the timing of market entry for for-profits; however, state regulation has a much less significant impact compared to federal regulation. The competitive effects from other for-profit and community colleges in the market are strong and likely to push back the activation time for for-profits. On the cost side, the real cost of opening the school declines 14% per year.

There are three counterfactual experiments performed in this paper. Each of them quantify the effects of different regulation on for-profit school industry. In the first experiment, I apply the state regulation to all the markets. The state regulation has insignificant influence on payoff of for-profits; therefore, there is no change in the decision of the timing of for-profit market entry in any market. In the second experiment, I allow the federal regulation to take action during the early stage of the model. In this experiment, early federal regulation does delay a small portion of the entry timing of for-profit schools. The third experiment is related to the recent proposal made by President Obama. The proposal is about making community college free for everyone and easier access to people who want to attend higher education. Although this proposal still needs to pass the legislation process, it has strong impact on for-profit industry. In this experiment, I double the number of community colleges in each market. This could be interpreted as the influence of community colleges on for-profit industry is stronger under the proposal. As a result, the proposal does push back the market entry time of for-profits and has the strongest effect among all three counterfactual experiments.

The remainder of the chapter is structured as followed. Section 3.2 provides a summary of college environment and regulations by state and federal government in the United States. Section 3.3 describes the model setup, assumptions and solution algorithm. Section 3.4 summarizes the data sources and the characteristics of the data set including school and market level data. In Section 3.5, the specific forms of payoff and cost functions are provided. The detail of estimation method (MSM) is also described in this section. Moreover, Section 3.5 presents the estimates of the model. Section 3.6 shows the results of three counterfactual experiments related to three different regulation levels. The last section, Section 3.7, concludes the whole chapter and provides a description of potential future work.

3.2 Summary of College Environment

3.2.1 College Entry

In this section, I first describe the big picture of college entry in 1980 and 2010 and then I will focus on the period from 2005 to 2013 which this paper will be targeting. By comparing the situation in 2010 to 1980, the college supply and cost have increased much. As more people are receiving post-secondary education, the tuition and Pell-grants per student are getting higher. As in Table 17, the number of for-profit schools and full-time students in for-profit colleges increase in a much faster pace than the other two sectors. Moreover, unlike public and not-for-profit schools, most of the supply expansion of for-profits comes from new colleges.

Table 18 summarizes the percentage of all full-time college students served by public, private not-for-profits, and for-profits in each college size group between 1980 and 2010. The table shows that the share of full-time students served by for-profits served increased from 1.28 percent in 1980 to 10.83 percent in 2010 while the share served by public decreased from 69.46 percent to 62.35 percent and the share served by not-for-profit decreased from

29.26 percent to 26.85 percent.

Table 19 shows the number of for-profit colleges in 248 Core Based statistics Areas (CBSA) from 2005 to 2013. The number of for-profit schools increased dramatically from 105 to 1309 in these nine years. Figure 10 and 11 show the distribution of active for-profit schools in 2004 and 2013. Figure 12 shows the number of school entry between 2005 and 2013 in each CBSA. From these distribution maps, we see that the areas with the largest number of for-profits in 2004 are located in New York, California, and Florida and these markets still remain the largest in 2013. The structure of post-secondary education in the United States has changed due to the enlarging demand and supply side, especially in for-profit sector. All in all, IPEDS shows that for-profit sector has expanded in a much faster speed than public and not-for-profit sectors over the past thirty years.

3.2.2 Government Regulation

The growth in for-profit sector has raised concerns on the amount of money in loans and scholarships for-profits receive, the graduation rate, and the success of the graduates in finding jobs. In addition, the federal government worries about the large subsidies that for-profits receives. According to National Conference of State Legislatures (NCSL), for-profits received $32 billion in federal grants and loans during the 2009-2010 school year. In order to have better control of education quality in the for-profit sector, lawmakers in both state and federal governments are adding or modifying the regulation of for-profit colleges. The following paragraph will discuss state and federal actions on for-profits in detail.

The details of the regulations regarding the for-profit colleges are presented in the NCSL website. In California, the government revised the Cal-Grant, California's student financial aid program, in 2011. After updating the restriction for the Cal-Grant, for-profit colleges must have a graduation rate of at least 30 percent and a maximum loan default rate of 15.5 percent. According to NCSL, nearly 80 percent of for-profit colleges in California will no longer be eligible to receive Cal-Grant funds. In April 2011, SB695 was passed in Maryland. The state government not only requires for-profit institutions to provide data on enrollment, graduation and retention rates to the Maryland Higher Education Commission, but also prohibits school from paying commission to recruiters, or using false tactics to attract students. Moreover, starting from July 1, 2013, Maryland state government will no longer provide financial assistance for for-profit colleges.

In 2013, Connecticut passed HB 5500 which requires all higher eduction institutions including for-profit colleges to provide uniform financial aid information created by the Consumer Financial Protection Bureau and the U.S. Department of Education to every prospective student. In 2010, Michigan's governor passed the Proprietary Schools Act. This restricts proprietary schools which are defined as for-profit institutions teaching a trade or vocation that do not have authority to grant degrees to sell goods and services produced and provided by students if the schools meet the criteria specified in the law: schools need to disclose that the products are produced or served by students and make the production or service process as part of the instruction.

In addition to four state regulations above, federal government has also released some regulations to protect student and taxpayer investments. In 2011, the Department of Education passed the gainful employment regulations to ensure students are having enough returns from investing in for-profit schools. The government look closely at the students loan repayment rates and debt-to-income ratio to decide whether the schools are eligible to use federal student aid. According to the Department of Education, programs in for-profit schools receiving federal student aid must meet at least one of three benchmarks: on average, students must have loan repayment rate higher than 35 percent, debt-to-total income ratio less than 12 percent, or debt-to-discretionary income ratio less than 30 percent. The government predicts that more than 5 percent of programs in the for-profit sector will lose eligibility for financial aid due to the new gainful employment regulation.

In 2015, President Obama made an announcement that he wants to make the first two years of community college free. This will create large accessibility for everyone to attend higher education. According to the proposal, the federal government would pay for 75 percent of the cost of community college and the participating states would be required to cover the remaining tuition balance. Although this is still a proposal and need to pass legislation to accomplish, free community colleges will have a strong impact on for-profit college industry. This will cause the for-profit schools difficulty in recruiting students.

3.3 Model

3.3.1 Basic Setup

The model in Schmidt-Dengler (2006) is used in this paper. This is an entry model which captures the strategic interaction in the timing of school activation and variation in market structure and regulations.

Time is discrete and finite with $t = 1, 2, ..., 9$. There are L independent markets and I^m schools in market m. Every period, the owners of the schools decide whether to enter the market or not. Let $a_i^t \in \{0, 1\}$ denote school i's action at time t where 0 is not activated and 1 is activated and a^t be an I-vector denoting the action profile chosen by all schools within one market at time t. Once the school is activated, it will remain activated in the remaining of the period. Let h_t^i be school i's history of actions which contains zeros until school i decides to enter the market. In this model I only focus on pure strategies. A pure activation strategy from school i is a function mapping the history to an element of the action set: $s_t^i : h_t \to A_t^i(h_t)\ \forall h_t \in H_t$ where H_t is the set of all possible action histories at time t and the action set $A_t^i(h_t) = \{not\, activated,\, activated\} = \{0,\, 1\}$.

A school receives 0 profit per period before activating and $\pi_1^i(n_t^m)$ thereafter. When the owner of the school decides to activate the school at periodt^i, there is a sunk cost of $C(t^i)$. As a result, a school's discounted intertemporal profits are

$$\Pi^i = \sum_{t=t^i}^{9} \beta^t \cdot \pi_1^i(n_t^m) - \beta^{t^i} \cdot C(t^i)$$

where β is the discounted factor and the owners of the schools would choose when to enter the market to maximize the discounted intertemporal profits.

3.3.2 Assumptions

The assumptions regarding the payoff and cost functions are similar to the assumptions in Schmidt-Dengler(2006). There are two assumptions for the payoff function:

1. (monotonicity) $\pi_1^i(n_t^m - 1) \geq \pi_1^i(n_t^m)$ and $n \geq 1,\ i \leq I$
2. (positive return) $\pi_1^i(n_t^m) \geq 0;\ \forall\, 1 \leq i,\ n \leq I$

The monotonicity assumption states that schools are competing in the markets. As the number of the schools increases in the market, the payoffs declines. The positive return assumption shows that schools can always earn positive revenue by activating.

The cost function is assumed to be decreasing, convex and bounded. The cost function is falling exogenously over time at a decreasing rate and will fall to a level such that the payoff from activating is higher than cost. In addition, by assuming all schools moving sequentially will avoid the potential multiplicity in a simultaneous move discrete time game (Berry 1992). Because time is finite, the setup is a finite horizon game of perfect information in the setup. This is a unique subgame perfect equilibrium, since schools will never be indifferent between activating and not activating.

3.3.3 Solution Algorithm

By using a backward recursive algorithm, the unique subgame perfect equilibrium can be computed. First, the profit of each school is calculated in the last period and ordered to get the rank of profitability. At period $\bar{t}$, all schools are activated. Therefore, at $\bar{t}-1$, school I which has the least profitability knows that all schools will be activated next period without knowing the history of the play. The least profitability school I chooses to enter the market if the increase in period payoffs when activating current period is larger than the cost saving when activating next period which means that the school will have more profit by activating now since the payoff outweighs the cost saving. After achieving the decision of school I, I compute the continuation value of activating versus not activating and get the decision result of school $I-1$. Similarly, I calculate the continuation value for all other schools solving backward recursively and repeat the process until the first period. By using this solution algorithm, the equilibrium decision from all schools can be reached.

3.4 Data

Sources of data. In this paper, we used a data set consisting of observations on 1309 for-profit colleges and 316 community colleges in US from 2005 to 2013. These colleges are located in 248 Core Based Statistics Areas (CBSA). The data were drawn from National Center for Education Statistics (NCES), Integrated Postesecondary Education Data System (IPEDS), and American Community Survey (ACS). Demographic information in each metropolitan and micropolitan statistical area is drawn from American Community Survey (ACS).

3.4.1 School Information

The data from IPEDS includes the information on location (state, county, and CBSA), enrollment, cost, SAT/ACT scores, grants, and tuition. For-profit schools have been focusing on online education recently. To account for this issue, I am able to determine whether the schools focus on in-class or online education by the 2012 distance educational data provided by IPEDS which consists of the number of the total full-time undergraduates enrollment and the number of students who only take online courses. By using this ratio, we can estimate what percentage of total college students in each school are taking online courses only. After removing schools serving more than 75 percent of their students through internet, I still have 1309 for-profit colleges remain in the data set.

I cluster the college location by CBSAs. The market with largest number 57 of for-profit colleges is in New York, Northern New Jersey, and Long Island metropolitan area. In Figure 9, the histogram shows that most of the markets are small. There are 188 markets having less than five for-profit colleges and only 32 markets with more than 10 for-profits which is shown in Table 21. Most of the large markets are located in New York, California, and Florida. IPEDS also provides the information on community colleges. New York, Northern New Jersey, and Long Island metropolitan area also has the largest number 19 of community colleges.

3.4.2 Market Information

There are total 917 CBSAs in the United States of America. After removing missing data from ACS and IPEDS, the data set consists 248 markets. The data from ACS includes the size of population, the racial/sexual composition of population, median income, median age and the number of housing units, etc. every year. The summary statistics is provided in Table 20 and Table 21. In the data set, there are 41 markets with state regulation that puts restriction on for-profit industry and these markets are located in four states: California, Connecticut, Maryland, and Michigan.

3.5 Estimation

3.5.1 Specification

This section will now provide the specific functional form of payoff and cost functions used in this paper. The payoff function will depend on the market characteristics and the random shocks in market and individual school. There are M independent markets, with

I^m schools operating in market m. I observe the market level variables Z^m and W^m in market m and each school i's activation year t^i.

To account for unobserved factor in the setup, there is an additive random component ϵ^i in the payoff function. This profitability shock is drawn independently across markets from a strictly monotonic and continuous distribution function F and is observed by all schools, but not by the econometrician. Similar to Berry (1992) and Schmidt-Dengler (2006), the unobserved term can be correlated across schools within a market and also let unobserved market characteristics to affect the profitability shocks of for-profit schools. I made a similar assumption on shock formulation as Berry (1992) which is $\epsilon^i = exp(\sqrt{(1-\rho^2)}\nu^i + \rho\nu^m)$ where ν^i is the school specific profitability shock and ν^m is the market specific profitability shock. Both school and market shocks are drawn from standard normal distribution and $\rho \in [0,1]$. Here are the payoff function and cost function for school $i = 1^m, ..., I^m$ where n_t^m is the number of schools in the market m in time period t and $\theta = (\alpha_1, \gamma_1, \gamma_2, \delta_1, c, \lambda, \rho)$

$$
\begin{aligned}
\pi_0^i(n_t^m, \theta) &= 0 \\
\pi_1^i(n_t^m, \theta) &= \alpha_1 + Z^m \cdot \gamma_1 + \delta_1 \cdot log(n_t^m) + \epsilon^i \\
C(t^i, \theta) &= (c + W^m \cdot \gamma_2) \cdot \lambda^{t^i}
\end{aligned}
$$

Note that the cost function depends on some market level variables and converges to zero at a decreasing rate $\lambda \in (0,1)$. In the payoff function, the market level variables Z^m include population, median income, number of community college,and indicators for federal and state regulation in given market m. Although the content of each state regulation is different, we assume they all have the same effect on schools' profit. In the cost function, the market level variables W^m includes the size of population in market m. The cost function includes the size of the population in the market because the cost in bigger and crowded market is usually higher due to the land value. Under this estimation specification, there are 11 parameters to estimate.

3.5.2 Identification

The indentification follows Schmidt-Dengler (2006). The parameter vector of interest θ is point identified if two parametric specifications are not observationally equivalent. When school i decides to enter the market, the school receives payoff $\pi_1^i(n_t^m, \theta)$. If school i enters the market last among all the schools at period t^i, the payoff and cost functions must hold that

$$\begin{aligned}
\pi_1^i(I,\theta) - (c + W^m \cdot \gamma_2) \cdot \lambda^{t^i}(1-\beta\lambda) &\geq 0 \\
\pi_1^i(I,\theta) - (c + W^m \cdot \gamma_2) \cdot \lambda^{t^i-1}(1-\beta\lambda) &< 0
\end{aligned}$$

where I is the total number of schools in the market m. From the condition above, I can estimate $(\alpha_1, \gamma_1, \delta_1)$. In the model, the discounted factor is fixed at .94. There are two explanations for cost falling over time. The first one is that the development of the technology allows school to spend less to incorporate new computer and other technology on campus. The other one is that other similar schools have already planned the recruiting strategy and schools which decide to enter the market later can modify or improve the business model and strategy of the exiting one.

To show how the model setup is sufficient to identify the parameters, let's look at the simplified model:

$$\begin{aligned}
\pi_0^i(n_t^m,\theta) &= 0 \\
\pi_1^i(n_t^m,\theta) &= \alpha_1 + \delta_1 \cdot log(n_t^m) + \epsilon^i \\
C(t^i,\theta) &= c \cdot \lambda^{t^i}
\end{aligned}$$

In a market with only one for-profit school, the school chooses to be activated in year t if

$$\begin{aligned}
\pi_1^i(1,\theta) - c \cdot \lambda^t(1-\beta\lambda) &\geq 0 \\
\alpha_1 + \epsilon^i - c \cdot \lambda^t(1-\beta\lambda) &\geq 0
\end{aligned}$$

therefore, the probability that school i is activated in year t is $Pr(\epsilon^i \geq -\alpha_1 + c \cdot \lambda^t(1-\beta\lambda))$. The fraction of schools which has not enter the market by year t is

$$S(t) = F(-\alpha_1 + c \cdot \lambda^t(1-\beta\lambda))$$

and from the data, I can get $F^{-1}(S(t)) = -\alpha_1 + c \cdot \lambda^t(1-\beta\lambda)$. The fraction of activated schools in single school markets can be observed in three different periods: t, $t+1$, $t+2$. Then I can solve for λ by forming:

$$\begin{aligned}
\frac{F^{-1}(\hat{S}(t)) - F^{-1}(\hat{S}(t+1))}{F^{-1}(\hat{S}(t)) - F^{-1}(\hat{S}(t+2))} &= \frac{c \cdot \lambda^t(1-\beta\lambda) - c \cdot \lambda^{t+1}(1-\beta\lambda)}{c \cdot \lambda^t(1-\beta\lambda) - c \cdot \lambda^{t+2}(1-\beta\lambda)} \\
&= \frac{1}{1+\lambda}
\end{aligned}$$

After knowing λ, c can be solved using

$$F^{-1}(\hat{S}(t)) - F^{-1}(\hat{S}(t+1)) \quad = \quad c \cdot \lambda^t(1-\beta\lambda) - c \cdot \lambda^{t+1}(1-\beta\lambda)$$

After knowing λ and c, α_1 is solved using

$$F^{-1}(S(t)) = -\alpha_1 + c \cdot \lambda^t(1-\beta\lambda)$$

In markets with more than one school, we have competition effect δ_1 to identify. The second school in a duopoly market chooses to enter the market in year t if

$$\begin{aligned} \pi_1^2(1,\theta) - c \cdot \lambda^t(1-\beta\lambda) &\geq 0 \\ \alpha_1 + \delta_1 \cdot log(2) + \epsilon^2 - c \cdot \lambda^t(1-\beta\lambda) &\geq 0 \end{aligned}$$

In the model setup, the second school in the market has less profitability than the first school. Thus, the distribution of ϵ^2 is the second order statistic $F_{(2)}$. By observing the fraction of firms that has not entered the market in duopoly markets and knowing α_1, c, and λ, δ_1 can be found from

$$F_{(2)}^{-1}(S(t)) = -\alpha_1 - \delta_1 + c \cdot \lambda^t(1-\beta\lambda)$$

3.5.3 Method of Simulated Moments

A Method of Simulated Moments (MSM) Estimator (McFadden (1989)) is used in this paper. Unlike GMM, I will not need to have closed form solution for moments and equilibrium activation time. This method estimates parameters of the structural model by simulating the model and updates the model parameters to have the model data moments that match the actual data moments as closely as possible. In this way, I can make sure that model matches key features of actual data. I will now introduce the steps for executing MSM in this paper:

Let $\theta = (\alpha_1, \gamma_1, \gamma_2, \delta_1, c, \lambda, \rho)$ be the vector of estimated parameters. For identification, I need at least as many moments as parameters to identify the model.

Step 1: Compute a J-dimensional vector of moments $\hat{\psi}_T^d$ from actual data and $J \geq Dimention\, of\, \theta$.

Step 2: Obtain a draw $v^m$of market specific profitability shock and S draws of an I^m-dimensional vector of individual profitability shocks $[v^1, v^2, \ldots v^{I^m}]$ for every market $m \in M$ where I^m is the total number of schools in the last time period.

Step 3: For a given parameter vector θ, compute the period payoff and cost for every draw of shocks from Step 2 in every market m. Then determine the profitability ranking among schools in each market.

Step 4: Compute the last activation time $\bar{t}$ using the decision rule on the least profitable school in each market and solve the model recursively from $\bar{t}$.

Step 5: From Step 4, I have $S \cdot L$ of equilibrium play. For each simulated equilibrium play from each draw, we compute the simulated moments $\psi^s(\theta)$. Then I calculate the average simulated moments

$\hat{\psi}^s(\theta) = \frac{1}{S}\sum_{s=1}^{S} \psi^s(\theta)$.

Step 6: Compare the average moments from the model simulations $\hat{\psi}^s$ to those from the actual data $\hat{\psi}^d$. Update the parameters θ based on the weighted distance between observed and simulated moments: $\hat{\theta}_{S,T}(W) = argmin_\theta[\hat{\psi}^d - \hat{\psi}^s(\theta)]\, W\, [\hat{\psi}^d - \hat{\psi}^s(\theta)]$ where W is the weighting matrix.

Step 7: Repeat Steps 3 to 6 until the data and model moments are as close as possible, i.e. until a vector θ is found that minimizes the weighted distance between observed and simulated moments.

The estimator $\hat{\theta}_{S,T}(W)$ is consistent and $\sqrt{T}(\hat{\theta} - \theta_0)$ is asymptotically normally distributed with zero mean and covariance matrix

$Q_S(W) = (1 + \frac{1}{S})[E\frac{\partial b(\theta_0)}{\partial \theta}^{'}\, W\, E\frac{\partial b(\theta_0)}{\partial \theta}]^{-1}$

where $b(\theta) = \hat{\psi}^d - \hat{\psi}^s(\theta)$ and $\frac{\partial b(\theta_0)}{\partial \theta}$ is the derivative of the vector of moments with respect to the parameter vector. The optimal weight matrix $W = (Ebb')^{-1}$ weighs moments more heavily if they are better identified to improve the efficiency of the estimator.

3.5.4 Moments

I need at least as many moments as parameters to identify the model. In the model, there are 11 parameters to estimate where $\theta = (\alpha_1, \gamma_1, \gamma_2, \delta_1, c, \lambda, \rho)$. The market level variables γ_1 in payoff function include population, median income, number of community colleges, indicator for federal and state regulation and γ_2 in cost function includes the population. Therefore, I first select the number of activated schools in each year starting from 2005 till 2012 which produces me eight moments. In order to capture the market size effect, I group the number of activated schools in 2010 with the size of population in the markets: large, medium and small. This will produce me another three moments. To capture the effect of state regulation, I interact the number of activated schools in 2007 and 2011 with the indicator of state regulation which means we have two additional moments. Moreover, I also include the number of activated schools in markets with small and medium number of for-profit schools that will be activated in 2013 in order to have better control on markets with smaller scope of for-profit industry. [1]This results in a total of 15 moments for our model.

3.5.5 Parameter Estimates

Table 22 shows the parameter estimates for the model which includes the market level variables, regulation variables, and competitive effects from both for-profit and community colleges. The number of simulation draws is twenty per market. In the payoff function, the coefficients of population and median income have positive estimations and both are significantly different from zero indicating that population and income in a market have positive effect on the payoff of for-profits after activating. Therefore, for-profit schools are likely to enter the market earlier with higher population and median income.

The estimated coefficient of the effect of community colleges is -8.4450 and is significantly different from zero indicating that for-profits in a market with more community colleges are likely to delay the timing of market entry. Similarly, the estimated coefficient of the competitive effect from other for-profit colleges in the same market δ_1 is -3.6357 which is less than half of the competitive effect from community colleges, but it is still significantly different from zero. This means that for-profits colleges have less profit when markets have more other activated for-profit colleges. The estimated coefficients of federal regulation and state regulation are both small (-0.0857, -0.0158) and only the coefficient of federal regulation is significantly different from zero. In the cost function, the sunk cost declines at a rate of about 16 percent per year which gives for-profit schools incentive to delay the timing of market entry. As in Berry (1992) model setup, I allow unobserved market characteristics to affect the profitability shocks of for-profit schools. The estimated coefficient term ρ is 0.3348 representing that the correlation between market shock and school shocks. This correlation is significantly different from zero. This means that a high draw of a common market shock indicates that for-profit colleges in a market gain more from activation and a low draw of a common market shock indicates that for-profit colleges gain less from activation.

[1]Markets with more than 5 and less than 20 for-profit schools in 2013 are considered as medium number of for-profits market; Markets with less than 5 are small number.

3.5.6 Goodness of Fit

To show the fit of the model, I simulate the model by drawing shocks and compare the entry timing of for-profit colleges from simulations with data. The results are presented in Figure 14. Figure 14 compares the observed market entry timing with the simulation result from the model by different groups of markets including markets with state regulation and markets with small, medium and large numbers of for-profit schools[2]. In the result given by the model, for-profit schools in markets with smaller number of for-profit schools tend to enter the market later comparing to the observed market entry time. In other words, the model underpredict the timing of activation for smaller and medium markets in terms of number of for-profit schools. In markets with small and medium number of for-profits, the model fits the observation well after period four.

More for-profit colleges in markets with large number of for-profit schools tend to enter the market in the early stages compared to the observation, but the market entry starts to slow down in middle stage of the model. Therefore, the model overpredict the timing of activation in large market in terms of number of for-profit schools in the early stages of the game, and underpredict the timing of activation in middle stage of the game after period 3. Comparing to the result for Schmidt-Dengler (2006), the fitness of the model is a bit worse in this study; however, we do have more for-profits schools in each market creating more variation and also more markets in the data set making the model harder to simulate from the observation. This will become to the future work to improve the fitness of the model for such a large data set with many competition and markets.

3.6 Counterfactual

The structural model setup allows us to do counterfactual experiments on three different policies happening recently for controlling for-profit industry. From these experiments, I can determine how each policy affects the entry decision of each for-profit college. In the first experiment, I expand the state regulation to every state in the United States. In the second experiment, I extend the federal regulation by having earlier executive year. In the last experiment, I increase the number of community colleges in each market. This can be

[2]Markets with more than 20 and less than 30 for-profit schools in 2013 are considered as large number of for-profits market; Markets with more than 5 and less than 20 for-profit schools in 2013 are considered as medium number; Markets with less than 5 are small number.

related to recent proposal made by President Obama which is making community college more appealing to people. The result of each experiment will be presented in this section.

3.6.1 Expanding State Regulation

I first examine the effect of state regulation after allowing all the markets to adopt the current state regulation. There are four states: California, Connecticut, Maryland, and Michigan, having regulation to control the for-profit colleges industry. They all have different regulations, but I treat them having equal effect in the model. Under the model setup, all the markets are independent. The entry decisions between markets do not affect each other.

The result for expanding state regulation is shown in Table 23. After applying the state regulation to all markets, there is no change in the timing of activation for for-profit colleges. The simulation result stays the same as the result with only four states having the regulation. In Table 22, the estimated coefficient for state regulation in payoff function is small comparing to size of population, median income and number of community colleges and is not significantly different from zero. As a result, it is not surprising that there is no difference in state regulation counterfactual experiment.

3.6.2 Extending Federal Regulation

In this experiment, I extend the federal regulation period by moving the starting executing date from 2011 to 2005. The federal regulation is applied to every independent market. The result for extending federal regulation is shown in Table 24. The schools do delay the timing of market entry starting from 2007; however, the effect on delaying the entry timing is small. The estimated coefficient of federal regulation in payoff function is small and significantly different from zero. Therefore, federal regulation has more influence on the decision of for-profits than state regulation.

3.6.3 President Obama's Proposal: Expanding Community College Industry

This experiment is related to the recent proposal made by President Obama. President Obama wants to make the first two years of community college free as is with high school. This will make community colleges more appealing to people who want to receive higher education and have strong competitive impact on for-profit school industry.

In this experiment, I double the number of community colleges in each market to represent the increase of competitive effect. Under the model setup, markets are independent. Therefore, community colleges in each market only affect the entry decision of the for-profit colleges in the same market. The result of expanding community college industry is shown in Table 25. In this experiment, the timing of activation for for-profit schools delays much starting from the beginning 2005 till the end 2012 of the period comparing to the original simulation and two previous experiments. In the early stage before 2007, about 35 percent of schools less than original simulation choose to enter the market when community college industry expands. After 2007, about 40 percent of schools less than original simulation choose to enter the market. All in all, President Obama's proposal about making community colleges free creates strong competitive impact on for-profit industry and reduces the incentive for for-profit schools to be activated.

3.7 Conclusion

In this study, I developed a structural model to study the entry timing of for-profit colleges from 2005 to 2013. The model can be estimated by using method of simulated moments which does not require any closed form solution. The result shows that population size and median income of markets have positive effect on for-profit colleges. Moreover, the competitive effects from for-profit and community colleges on profit is strong.

Because of the model setup, I am able to perform counterfactual experiments for state and federal regulations and President Obama's proposal. I found out that there is no effect on the timing of market entry after applying state regulation to all markets and only little delay on the timing of market entry after extending the period of federal regulation. However, the proposal made by President Obama about making community colleges free to people who want to receive higher education has strong impact on the for-profit industry. For-profit colleges have less incentive to be activated when community colleges are more appealing to the public.

The analysis in this paper can be extended to examine government policies on restricting for-profit college industry. The model can measure the effect of the policies on the decision of the entry of for-profit colleges. Although the model does a great job on fitting the data on smaller markets, the model tends to overestimate in large markets. Comparing to the result shown in Schmidt-Dengler (2006), the goodness of fit in this paper is worse since there are more markets and more for-profit colleges in each market which make the dynamic game more complicated. Future work may include improving the model to have better fit in larger and more complicated environment. Moreover, in the current framework, there is no uncertainty about future payoffs and costs which is unrealistic and also the model does not allow the owners to choose what the size and type of colleges they want to build. The extension to incorporate school owner's choices besides entry timing is left for future research.

Figure 9: Histogram for Number of For-Profit Schools within the Market in 2013

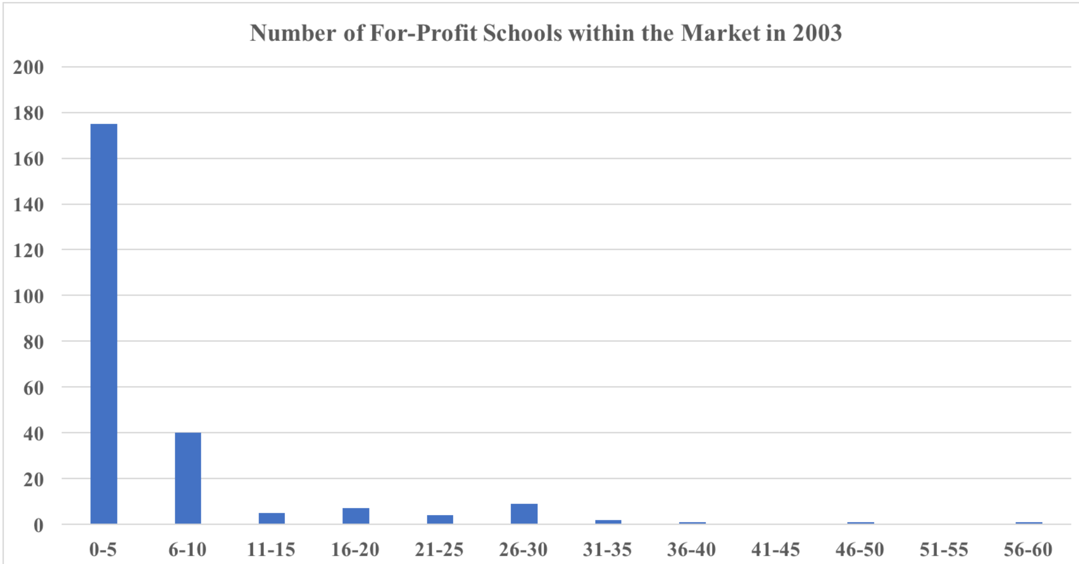


Table 17: Supply Change in Each Type of Colleges

Type of Colleges	**Year**	**Number of Schools**	**Enrolled Full-time Students**	**Supply Increases from Existed Colleges**	**Supply Increases from New Colleges**
Public	2010	677	5,131,561 (54.80%)	5,061,358	70,203
	1980	632	3,315,400		
Not-For-ProfitNot-For-Profit	2010	1361	2,210,010 (58.22%)	2,120,394	89,616
	1980	1045	1,396,755		
For-Profit	2010	691	889,011 (1359.00%)	231,444	657,567
	1980	77	60,933		

Table 18: Number of Students Served by Colleges in Each Size Group (percentage is the fraction over total number of students)

Type of Colleges	Year	Under 1000	1000-4999	5000-9999	10000-19999	20000 above	Students by Sectors	Total Students
Public	2010	43,251 (0.53%)	703,786 (8.55%)	1,191,837 (14.48%)	1,798,627 (21.85%)	1,394,060 (16.94%)	5,131,561 (62.35%)	8,230,582
	1980	70,758 (1.48%)	745,334 (15.62%)	914,356 (19.16%)	1,150,916 (24.11%)	434,036 (9.09%)	3,315,400 (69.46%)	4,774,088
Not-For-Profit	2010	240,580 (2.92%)	1,299,791 (15.79%)	390,903 (4.75%)	182,691 (2.22%)	96,045 (1.17%)	2,210,010 (26.85%)	8,230,582
	1980	296,250 (6.21%)	806,113 (16.89%)	196,915 (4.13%)	76,715 (1.61%)	20,762 (0.43%)	1,396,755 (29.26%)	4,774,088
For-Profit	2010	207,551 (2.52%)	212,664 (2.58%)	46,910 (0.57%)	80,507 (0.98%)	341,379 (4.15%)	889,011 (10.83%)	8,230,582
	1980	27,109 (0.57%)	33,824 (0.71%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	60,933 (1.28%)	4,774,088

Table 19: Number of Activated For-profit Schools from 2005 to 2013

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013
# of for-profits	105	211	302	429	587	828	1030	1189	1309

Table 20: Summary Statistics of the Data: School and Market Characteristics

	Obs	**Mean**	**Std. Dev.**	**Min**	**Max**
Population in 2013	248	940,683.4	1,785,992	65,401	1.99*10^7
Median income in 2013	248	50,448.76	9,201.841	34,374	91,533
# **of For-profits in Markets in 2013**	248	5.2782	8.1140	1	57
# **of Community Colleges in Markets in 2013**	248	1.2510	1.9153	0	19

Table 21: Summary Statistics of the Relationship between Markets and For-Profit Schools

	Obs
# of Markets with One For-Profit	87
# of Markets with Two to Five For-Profits	101
# of Markets with More Than Five	60
# of Markets with State Regulation	41

Table 22: Parameter Estimation

θ	Estimation
α_1	33.5738***
	(0.1116)
$\gamma_{1\_Population}$	8.5388***
	(0.1539)
γ_{1_Income}	5.5967***
	(1.2937)
$\gamma_{1\_CommunityC}$	$-8.4450***$
	(0.5061)
$\gamma_{1\_Fed\_Reg}$	-0.0857**
	(0.0321)
$\gamma_{1\_State\_Reg}$	-0.0158
	(3.2263)
δ_1	-3.6357***
	(0.8264)
c	572.6663***
	(2.2385)
γ_2	$1.6397 * 10^3$
	$(6.0781 * 10^3)$
λ	0.8589**
	(0.1141)
ρ	0.3348*
	(0.1411)

Standard errors are in parenthesis*** p<0.01, ** p<0.05, * p<0.1

Figure 10: Map for Active For-Profit Schools in 2004

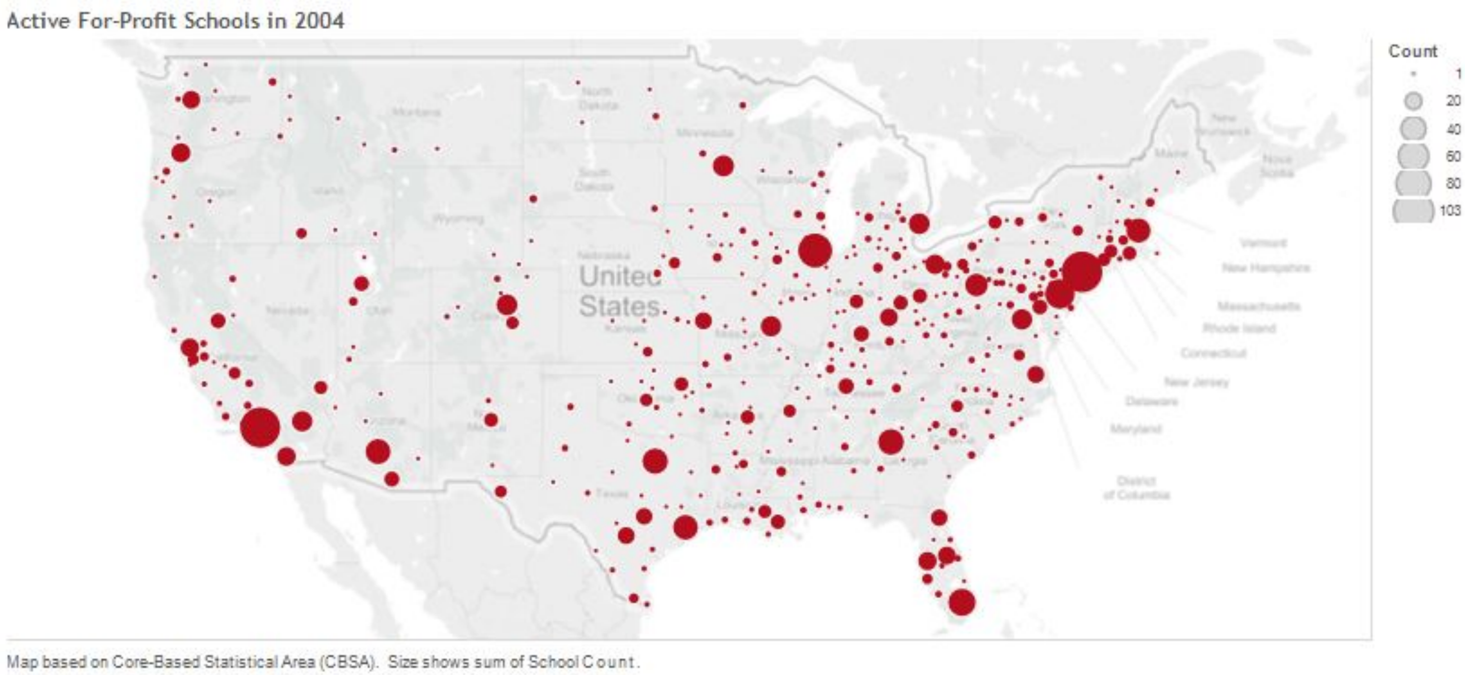


Figure 11: Map for Active For-Profit Schools in 2013

Figure 12: Map for Number of School Entry between 2005 and 2013

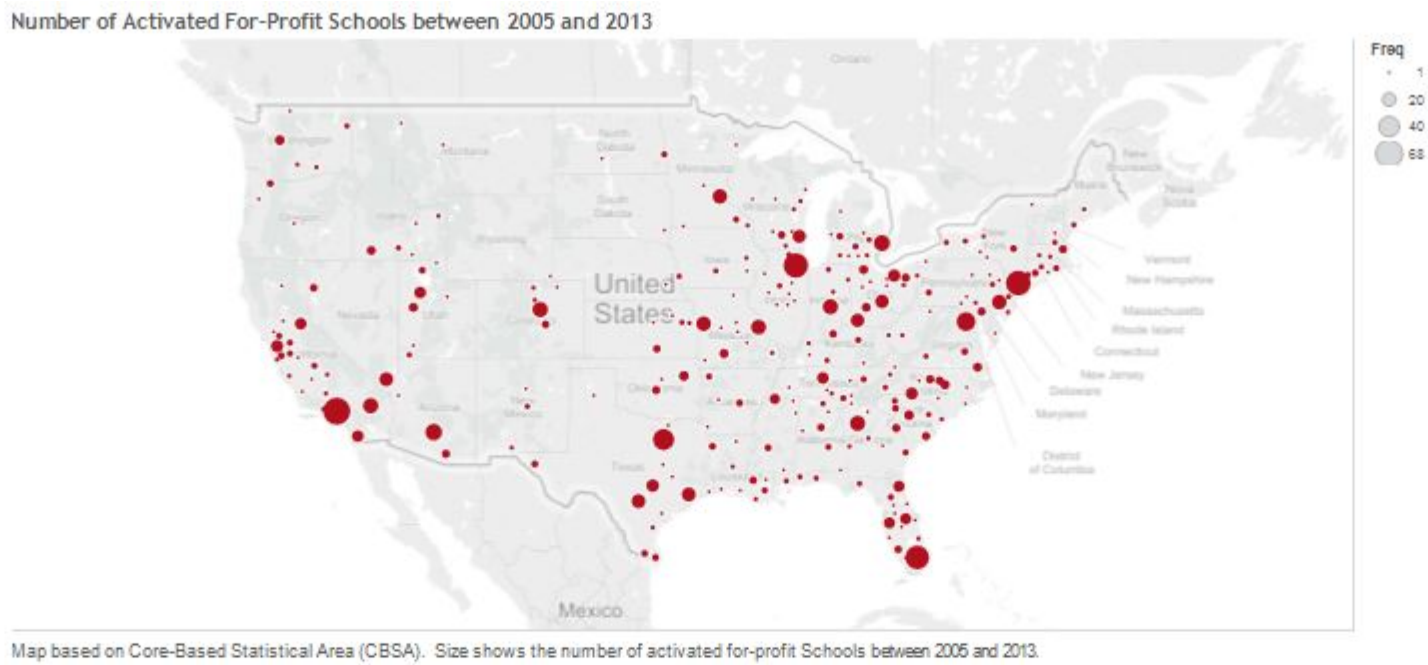


Table 23: Results for Counterfactual: State Regulation

Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
For-Profits Entry (Data)	105	211	302	429	587	828	1030	1189
For-Profits Entry (Simulation)	91	188	253	379	586	752	941	1211
For-Profits Entry (State Regulation Counterfactual)	91	188	253	379	586	752	941	1211

Table 24: Results for Counterfactual: Federal Regulation

Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
For-Profits Entry (Data)	105	211	302	429	587	828	1030	1189
For-Profits Entry (Simulation)	91	188	253	379	586	752	941	1211
For-Profits Entry (Fed Regulation Counterfactual)	91	188	252	374	583	747	941	1211

Table 25: Results for Counterfactual: President Obama Proposal for Community Colleges

Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
For-Profits Entry (Data)	105	211	302	429	587	828	1030	1189
For-Profits Entry (Simulation)	91	188	253	379	586	752	941	1211
For-Profits Entry (Community Colleges Counterfactual)	27	65	122	235	376	424	503	605

Figure 13: Counterfactual Experiments: Number of For-Profit Entry

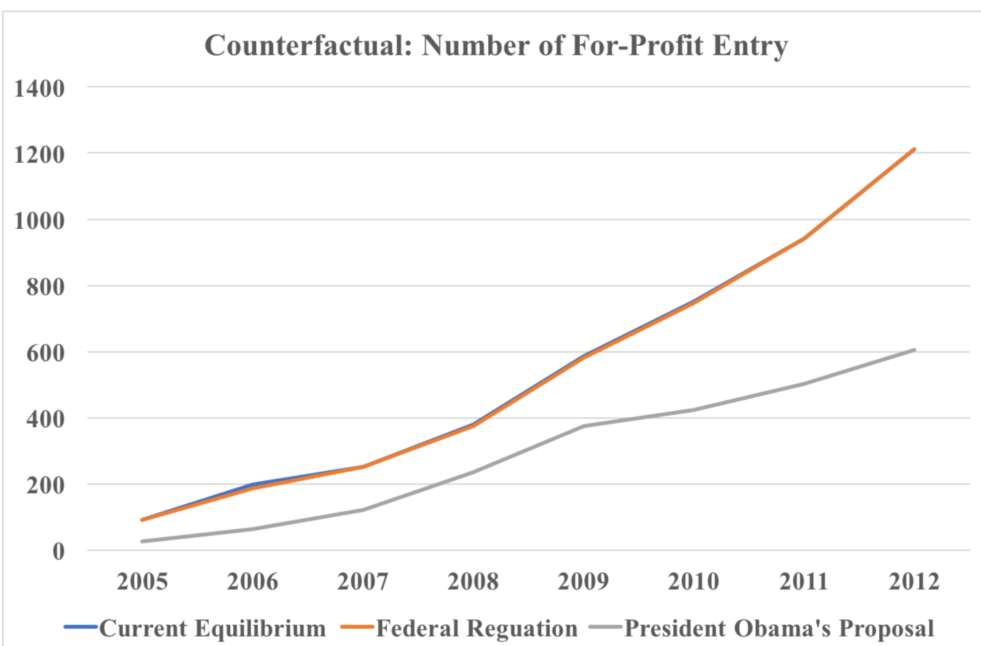
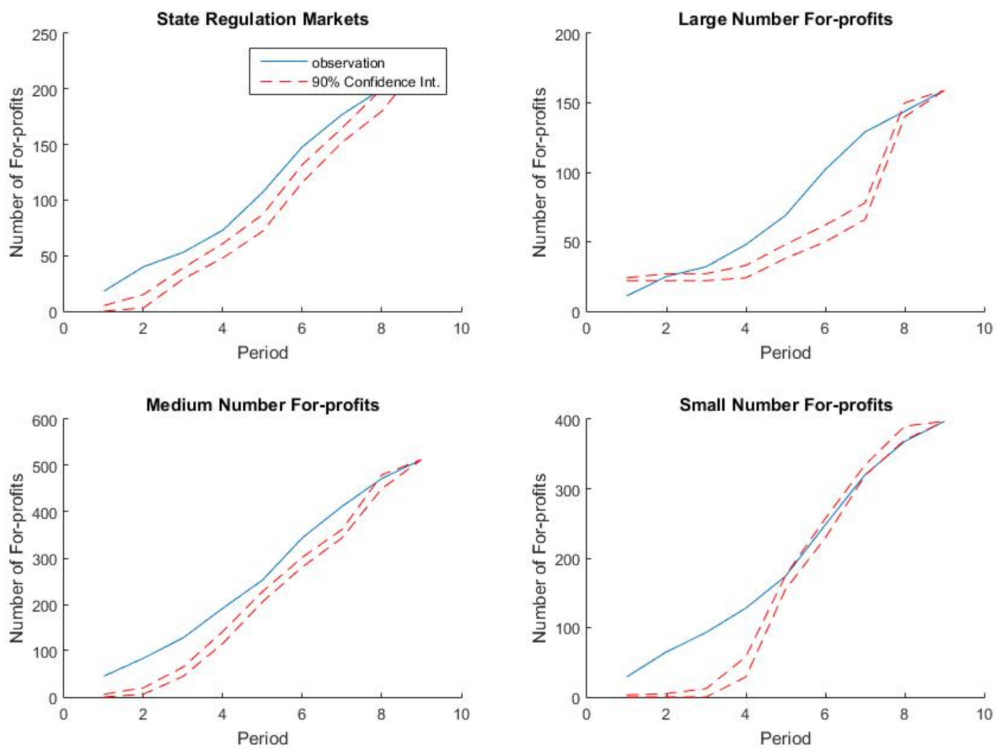


Figure 14: Number of Activations per Period by Different Kinds of Markets



References

[1] **Agarwal, Sumit and Robert Hauswald,** "Distance and Private Information in Lending," *The Review of Financial Studies,* 2010, 23(7), 2757-2788.

[2] **Amel, Dean F. and Kennenth P. Brevoort,** "The Perceived Size of Small Business Banking Markets," *Journal of Competition Law and Economics,* 2005, 1(4), 771-784.

[3] **Andrews, D.W. and G. Soares,** "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection" *Econometrica,* 2010, 78(1), 119-157.

[4] **Banerjee, Abhijit V. and Esther Duflo,** "Giving Credit Where It Is Due," *Journal of Economic Perspectives,* 2010, 24(3), 61-80.

[5] **Bailey, Badway and Gumport,** "For-Profit Higher Education and Community Colleges," *National Center for Postsecondary Improvement,* 2001.

[6] **Barr, Michael S.,** "Credit Where it Counts: The Community Reinvestment Act and its Critics," *New York University Law Review,* 2005, 80(2), 513-652.

[7] **Barros, Pedro Pita,** "Multimarket Competition in Banking, with an Example from the Portuguese Market," *International Journal of Industrial Organization,* 1999, 17(3), 335-352.

[8] **Barwick, P. J.** "What Happens when Walmart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica,* 2008, 76(6), 1263-1316.

[9] **Berger, Allen N. and Gregory F. Udell,** "Relationship Lending and Lines of Credit in Small Firm Finance," *The Journal of Business,* 1995, 68(3), 351-381.

[10] **Berger, Anthony Saunders, Joseph M. Scalise, and Gregory F. Udell,** "The Effects of Bank Mergers and Acquisitions on Small Business Lending," *Journal of Financial Economics,* 1998, 50(2), 187-229.

[11] **Berger, Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein,** "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks," *Journal of Financial Economics,* 2005, 76(2), 237-269.

[12] **Berry. Steve T.,** "Estimation of a Model of Entry in the Airline Industry," *Econometrica,* 1992, Vol. 60, 889-917.

[13] **Berry, Steve T.,** "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics,* 1994, 25(2), 242-262.

[14] **Berry, Steve, James Levinsohm, and Ariel Pakes** "Automobile Prices in Equilibrium," *Econometrica,* 1995, 63(4), 841-890.

[15] **Boot, Arnoud W. A. and Anjan V. Thakor,** "Can Relationship Banking Survive Competition?," *The Journal of Finance,* 2000, 55(2), 679-713.

[16] **Brevoort, Kenneth P., John A. Holmes, and John D. Wolken,** "Distance Still Matters: The Information Revolution in Small Business Lending and the Persistent Role of Location, 1993-2003" 2010.

[17] **Burgess, Robin and Rohini Pande,** "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment," *The American Economic Review,* 2005, 95(3), 780-795.

[18] **Cellini and Riegg,** "Financial Aid and For-Profit Colleges: Does Aid Encourage Entry?" *Journal of Policy Analysis and Management,* 2010, Vol. 29, 526-552.

[19] **Cellini and Goldin C.,** "Does Federal Student Aid Raise Tuition? New Evidence on For-Profit Colleges"*NBER Working Paper 17827.*

[20] **Cetorelli, Nicola and Philip E. Strahan,** "Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets," *The Journal of Financial* 2006, 61(1), 437-461.

[21] **Chiappori, Pierre-André, David Perez-Castrillo, and Thierry Verdier,** "Spatial Competition in the Banking System: Localization, Cross Subsidies and the Regulation of Deposit Rates," *European Economic Review,* May 1995, 39(5), 889-918.

[22] **Davis, Peter,** "Spatial Competition in Retail Markets: Movie Theaters," *Working Paper,* MIT Sloan.

[23] **Deming, Goldin and Lats, L.** "The For-Profit Postsecondary School Sector: Numble Critters or Agile Predators?" *Journal of Economics Perspectives*, 2012, Vol. 26, 139-164.

[24] **Downes and Greenstein**"Understanding the Supply Decision of Nonprofits: Modeling the Location of Private School" RAND Jounral of Economics, 1996, Vol. 27, 365-390.

[25] **Drexler, Alejandro and Antoinette Schoar,** "Do Relationships Matter? Evidence from Loan Officer Turnover," 2012.

[26] **Fisman, Raymond, Daniel Paravisini, and Vikrant Vig,** "Cultural Proximity and Loan Outcomes," 2012.

[27] **Garmaise, Mark J. and Tobias J. Moskowitz,** "Bank Mergers and Crime: The Real and Social Effects of Credit Market Competition," *The Journal of Financial* 2006, 61(1), 495-538.

[28] **Gilje, Erik,** "Does Local Access to Finance Matter?," January 2012.

[29] **Greenstone, Michael, Alexandre Mas, and Hoai-Luu Q. Nguyen,** "Do Credit Market Shocks affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and 'Normal' Economic Times," 2014.

[30] **Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini,** "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation," *The Journal of Financial* 2010, 65(3), 796-828.

[31] **Holmes, T. J.** "The Diffusion of Wal-Mart and Economics of Density," *Econometricaw,* 2011, 79, 253-302.

[32] **Houde J., Newberry P. and Seim K.** "Economies of Density in E-Commerce: a Study of Amazon's Fulfillment Center Network," *Working Paper,* 2017.

[33] **Jayaratne, Jith and Philip E. Strahan,** "The Finance-Growth Nexus: Evidence from Bank Branch Deregulation," *The Quarterly Journal of Economics* 1996, 111(3), 639-670.

[34] **Kerr, William R. and Ramana Nanda,** "Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship," *Journal of Financial Economics,* October 2009, 94 (1), 124-149.

[35] **Mcfadden, D.** "A method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica,* 1989, 57, 992-1026.

[36] **Munoz, Ana Patricia and Kristin F. Butcher,** "Using Credit Reporting Agency Data to Assess the Link between the Community Reinvestment Act and Consumer Credit Outcomes," 2013.

[37] **Nguyen, Hoai-Luu Q,** "Are Credit Markets Still Local? Evidence from Bank Branch Closings," *American Economic Journal: Applied Economics,* 2009, 11(1): 1-32.

[38] **Pakes, Ariel, and David Pollard,** "Simulation and the Asymptotics of Optimization Estimators," *Econometrica,* 1989, 57, 1027-1057.

[39] **Pakes, Ariel, Jack Porter, Katherine Ho, and Joy Ishii** "Moment Inequalities and Their Application," *Working Paper,* Harvard University.

[40] **Park, Kwangwoo and George Pennacchi,** "Harming Depositors and Helping Borrowers: The Disparate Impact of Bank Consolidation," *Review of Financial Studies,* April 2009, 22 (1), 1-40.

[41] **Peek, Joe and Eric S. Rosengren,** "Bank Consolidation and Small Business Lending : It's Not Just Bank Size that Matters," *Journal of Banking & Finance,* August 1998, 22 (6-8), 799-819.

[42] **Peek, Joe and Eric S. Rosengren,** "Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States," *The American Economic Review,* 2000, 90 (1), 30-45.

[43] **Petersen, Mitchell A. and Raghuram G Rajan,** "The Benefits of Lending Relationships: Evidence from Small Business Data," *The Journal of Finance,* 1994, 49 (1), 3-37.

[44] **Peek, and Raghuram G. Rajan** "The Effect of Credit Market Competition on Lending Relationships," *The Quarterly Journal of Economics,* 1995, 110 (2), 407-443.

[45] **Sapienza, Paula,** "The Effects of Banking Mergers on Loan Contracts," *The Journal of Finance,* 2002, 57 (1), 329-367.

[46] **Schmidt-Dengler, P.** "Empirical Anaysis of Dynamic Models with Multiple Agents," *Ph.D. Dissertation,* 2005.

[47] **Schmidt-Dengler, P.** "The Timing of New Technology Adoption: The Case of MRI," *Society for Economic Dynamics,* 2006, No. 3.

[48] **Strahan, Philip E. and James P. Weston,** "Small Business Lending and Bank Consolidation: Is There Cause for Concern?," *Current Issues in Economics and Finance,* 1996, 2 (3).

[49] **Strahan, and Philip E.,** "Small Business Lending and the Changing Structure of the Banking Industry," *Journal of Banking & Finance,* August 1998, 22 (6-8), 821-845.

[50] **Takahashi Y.** "Estimating a War of Attrition: The Case of the U.S. Movie Theater Industry," *American Economic Review,* 2015, Vol. 105, 2204-41.