Benefits from Branch Networks: Theory and Evidence from the Summary of Deposits Data

Patrycja Grzelonska[1]

University of Minnesota

November 2005

Abstract

How does the level of deposits obtained by one bank branch depend on the geography of other branches in the network? To answer this question I present a spatial model of consumer choice of depository institution. The key element in the model is the assumption that consumers can be in different places during the day and they should be able to find a nearby branch close to their homes, to where they work, shop and so on. The main implication of this model is the fact that consumers are more likely to open an account with a bank that has a lower expected distance to its branches. This in turn implies that there may be a positive relation between the level of deposits per branch and the proximity of branches. I call this effect a branch network benefit. To test this model empirically I estimate the demand for banking services using a discrete choice framework. The main component of the indirect utility function is, consistently with my model, the expected distance to a branch. I find that the expected distance to a branch is a significant factor in explaining consumers' choices of depository institutions. Decreasing the convenience of a branch network (increasing the expected distance to a branch) by merely 0.26% may lead to decrease of total deposits by 6%.

[1]Address: Department of Economics, University of Minnesota, 1035 Heller Hall, 271-19th Avenue South, Minneapolis, MN 55455; E-mail: pgrzelon@econ.umn.edu. I am especially grateful to my advisor Thomas Holmes for his guidance and support. I would also like to thank Sam Kortum, Zvi Eckstein and participants in the applied microeconomics workshop at the University of Minnesota for helpful comments.

1 Introduction

How does the level of deposits obtained by one bank branch depend on the geography of other branches in the network? In both the academic and business literature on banking, there has been much discussion about the way in which consumers value access to a network of bank branches. It's commonly assumed that the utility of the consumer from choosing a particular bank is positively related to the number of its branches in an exogenously delineated market (usually Metropolitan Statistical Area (MSA) or rural county). In this paper I present a spatial model of consumer choice of depository institution in which not only the number but also the exact locations of branches (the geography of branches) matter.

In the model each consumer is endowed with an amount of deposits that he places in the bank of his choice. The main concern for the consumer is to make sure that the distance that he will need to travel to the closest branch of his bank is as short as possible. This distance depends on both his location and the location of branches. The consumer can be in different places during the day. He can be either in a place called home, or he can be in some other places that are within 30 miles of his home (this 30 miles radius circle constitutes the consumer's "commuting" area). Given this spatial randomness each consumer has to care about the location of all the branches of a given bank. He chooses the bank with the most convenient network - the one with the lowest expected distance to a branch.

This type of behavior has implications for the relationship between the level of deposits obtained by one bank branch and the geography of other branches in the network. Branches should form a convenient, dense network for their clients: people are in different places during the day and they should be able to find a nearby branch not only close to their homes but also close to where they work, shop and so on. In this case the proximity of branches increases the number of customers who will deposit their money in each branch. I call the increase in the total deposits due to this effect a network benefit.

The goal of this paper is to find out how much consumers value convenience, and whether there are benefits of having a dense network of branches. To answer this question I estimate the demand for banking services using a discrete choice framework. The main component of the indirect utility function is, consistently with my model, the expected distance to a branch. I estimate the model with data at a rich level of geographic detail. I obtain the exact addresses of all the bank branches in the U.S. from the Summary of Deposits data (SOD) provided by the Federal Deposit Insurance Corporation (FDIC). I approximate consumers' home addresses and commuting areas using Census 2000 data. These two data sets allow me to calculate distances between consumers and branches. I find that the expected distance to a branch is a significant factor in explaining consumers' choices of depository institutions. According to my estimates, decreasing the convenience of a branch network (increasing the expected distance to a branch) by merely 0.26% may lead to the decrease of total deposits by 6%.

The fact that consumers do care about the location of all the branches of a given bank and choose banks with the lowest expected distance to a branch has important implications for issues of the antitrust regulation in the banking industry. First, it's important for the regulators to establish relevant bank markets. Commonly used exogenous delineations such as MSAs or rural counties may be too crude: two banks with networks of branches on the opposite side of the city are not really competitors. Concentration measures such as the Herfindahl-Hirschmann index that adds up squared deposits of banks in these big exogenous markets may give an incorrect picture of market power and competition intensity between banks. Second, the negative effects of the horizontal mergers between banks could be outweighed by the potential increase in the utility of consumers due to the higher convenience of networks.

The motivation for this paper is the survey-based evidence that customers do care about the location of the branches of their financial institutions. Kiser (2002) analyses the data regarding household banking behavior obtained from the 1999 Michigan Surveys of Consumers. Households that switched their depository institution at least once were asked for the primary reason for their most recent change. 51.3% of respondents answered that the change was triggered by relocation. Out of the households who didn't relocate 37.3% indicated that the change was because of more convenient location of offices and ATMs of the new institution. Households that had been with their institution at least 1 year were asked for the reasons for staying. 73.8% of respondents indicated location. Amel et al. (2002) presents the results from the data from 1998 Survey of Consumer Finances (SCF). The median distance between households and their depository institutions is 3 miles and 75% of consumers are within 10 miles.

Related literature: There are several papers that use the discrete choice framework to study the demand for banking services. However they use the information about the geography of consumers and branches in a much more limited way. Most of these studies follow the convention of bank regulators and look at the demand for banking services at the level of the exogenously chosen markets such as MSA and rural county. For example, in Dick(2002) and Knittel and Stango (2004), and Adams et al.(2005) the utility of a consumer from choosing a particular bank depends on different proxies that are supposed to account for the geography of branches, like the number of branches per square mile in the market or the number of counties (or states) in which the bank has presence. The work that is the most closely related to this is Ishii (2004). Her specification of utility depends on the distances between consumers and branches, but she only includes the distance to the nearest and second nearest branch. Most importantly however she also performs her analysis using exogenously delineated markets. Hence there might be consumers who live very close to some branches, but these branches are not in the consumers' decision set because they happen to be outside the market boundary.

The rest of this paper is organized as follows: section 2 develops the spatial model of consumer choice of depository institution and its empirical implementation, section 3 presents the data, section 4 discusses the estimation and presents the results and section 5 concludes.

2 Model of choice of depository institution

In this section I present a model of consumer behavior in which demand for banking services is spatially random: consumers are in different places throughout the day and regardless of their location they may need to obtain some banking services. This feature makes the consumer value the locations of all the branches of his bank. This in turn implies that if branches are sufficiently close the network benefits arise.

2.1 Banks

There are $j = 1, 2...J$ banks. Each bank has $B(j)$ branches branches that are distributed on a two-dimensional plane. Let B_j denote the set of branch locations for bank j. There is at most one branch operated by a particular bank at any given location. However there can be multiple branches at one location, each of them operated by a different bank.

2.2 Consumers

There are I consumers that are distributed across L discrete locations. For every consumer there is one location called "home" and he stays the fraction α of the day there (hence I will use the subscript i to refer to a location l which is a home to consumer i). The fraction $(1-\alpha)$ of the day each consumer travels. In that case he spends the fraction $(1-\alpha)\mu_{il}$ of his time in every one of the remaining locations. (for each i $\sum_{l=1,l\neq i}^{L} \mu_{il} = 1$).

Each consumer opens an account in one of the banks. Regardless of his location he may need to obtain banking services, in which case he goes to the closest branch of his bank. He derives utility v_{ij} from theses services. However there is a cost γ per unit of distance traveled if the consumer is not in the same location as the closest branch of his bank. Hence consumer i's net utility from having an account in bank j is given by:

$$U_{ij} = v_{ij} - \gamma(\alpha D(i,j) + (1-\alpha)\sum_{l=1,l\neq i}^{L} \mu_{il} D(l,j),$$

where $D(l,j) = \min_{l' \in B_j}\{DIST_{l,l'}\}$ is the distance between a particular location l and the closest branch of bank j. Each consumer deposits his money in the bank that gives him the highest net utility. These deposits are assigned to the branch that is the closest to his home.

2.3 Network benefit

In this model the consumer is in different places throughout the day. This means the consumer cares about the whole network of branches of his bank. If consumer stayed at one location all the time (so $\alpha = 1$), he would only care about the distance from there to the closest branch. Since he may also be in other places, he values the locations of all the branches, he cares about the expected distance that he needs to travel.

Banks that maximize the total level of deposits have to maximize the number of customers for whom its network is convenient. The proximity of one branch to another generates a network effect, and it can expand the number of customers who will deposit their money in each branch.

Definition 1 *Network benefit is the increase in the number of consumers, who will deposit their money in one branch due to the proximity of another branch.*

General results that relate the network benefits to the distribution of consumers and location of banks and branches are complex.

2.4 Simple example

To highlight the main forces in the model above, let's look at a simplified version of the model with only one bank and two branches. Consumers are distributed uniformly on a line (here, unlike in the model above, the location space for consumers is continuous). Every consumer derives the same utility v from branch services. There is a cost γ per unit of distance traveled if the consumer is not in the same location as the closest branch of his bank. Consumer will open an account in this bank if his net utility is at least 0[2].

First consider the case in which every consumer stays at home the whole time, $\alpha = 1$ (Figure 1). As long as the two branches are at least $2\frac{v}{\gamma}$ apart, each branch attracts consumers from $\frac{v}{\gamma}$ distance. If the branches are located closer than $2\frac{v}{\gamma}$, the markets start to overlap, one branch cannibalizes the consumers from the other one.

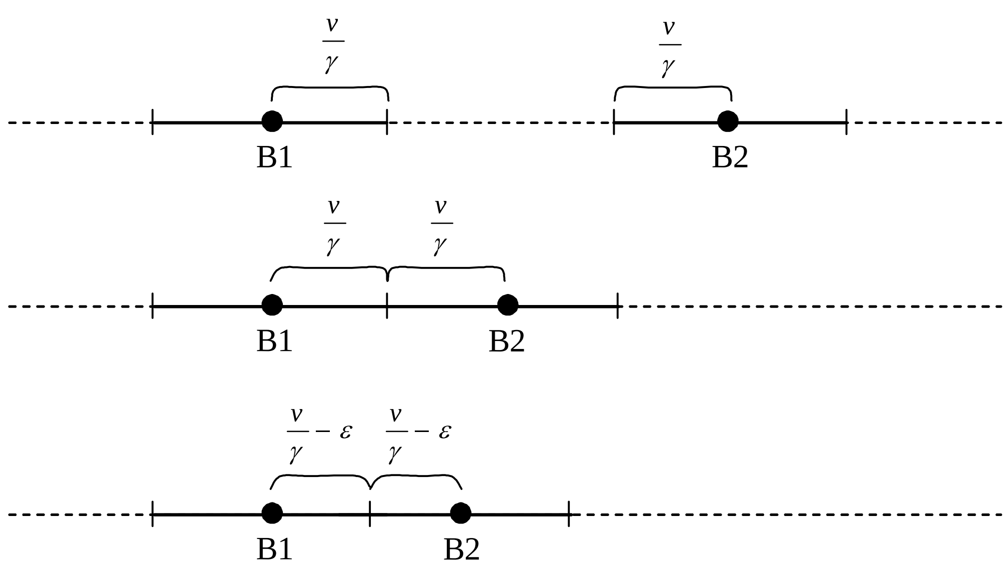


Figure1 : Bank market with $\alpha = 1$

[2]Suppose that the networks of competitors are such that each consumer can get this 0 net utility from signing up with them.

Now consider the case in which consumers is in different place throughout a day (Figure 2) Every consumer can be in three different locations: the fraction α of his time he is at home, the fraction $(1-\alpha)$ of his time he travels. In the latter case he spends half of his time in a place that is a distance λ (assume $\lambda < \frac{v}{\gamma}$) to the right of his home and another half of his time in a place that is a distance λ to the left of his home. Now the number of consumers that each branch can attract depends on the distance between both branches In particular, consider a consumer located $\frac{v}{\gamma} + (1-\alpha)\lambda$ away from the branch B1, on the right side. When the branches are far from each other, say $4\frac{v}{\gamma}$, it's always closer for this consumer to go to the branch B1. The expected distance that this consumer has to travel to get to the branches of this bank is:

$$ED = \alpha(\tfrac{v}{\gamma} + (1-\alpha)\lambda) + \tfrac{(1-\alpha)}{2}(\tfrac{v}{\gamma} + (1-\alpha)\lambda - \lambda) + \tfrac{(1-\alpha)}{2}(\tfrac{v}{\gamma} + (1-\alpha)\lambda + \lambda) = (\tfrac{v}{\gamma} + (1-\alpha)\lambda).$$

Since it is bigger than $\frac{v}{\gamma}$ this consumer will not sign up for an account with this bank.

However if the branches are located $2\frac{v}{\gamma} + 2(1-\alpha)\lambda$, this consumer goes to the branch B1 when he is at home and when he travels λ to the left, but he goes to the branch B2 when he travels to the right. The expected distance that this consumer has to travel to get to the branches of this bank is::

$$ED = -\alpha(\tfrac{v}{\gamma} + (1-\alpha)\lambda) - \tfrac{(1-\alpha)}{2}(\tfrac{v}{\gamma} + (1-\alpha)\lambda - \lambda) - \tfrac{(1-\alpha)}{2}(\tfrac{v}{\gamma} + (1-\alpha)\lambda - \lambda) = -\tfrac{v}{\gamma}.$$

Since it is equal to $\frac{v}{\gamma}$ this consumer will open an account with this bank. The same logic also holds for anybody located between $\frac{v}{\gamma}$and $\frac{v}{\gamma} + (1-\alpha)\lambda$ away from the branch B1 on the right side, and from the branch B2 on the left side. This additional demand from outside the $\frac{v}{\gamma}$ range is what I call a network benefit.

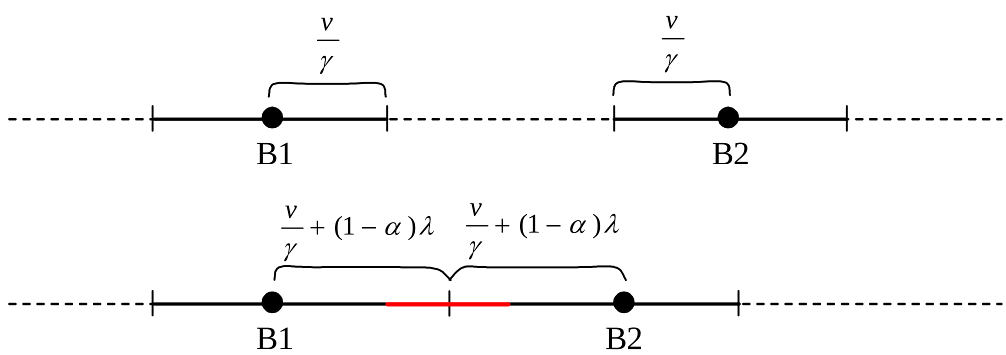


Figure 2 : Bank market with $\alpha < 1$

Figure 3 shows how the number of consumers per branch changes with distance in this example (for derivation see Appendix 1).

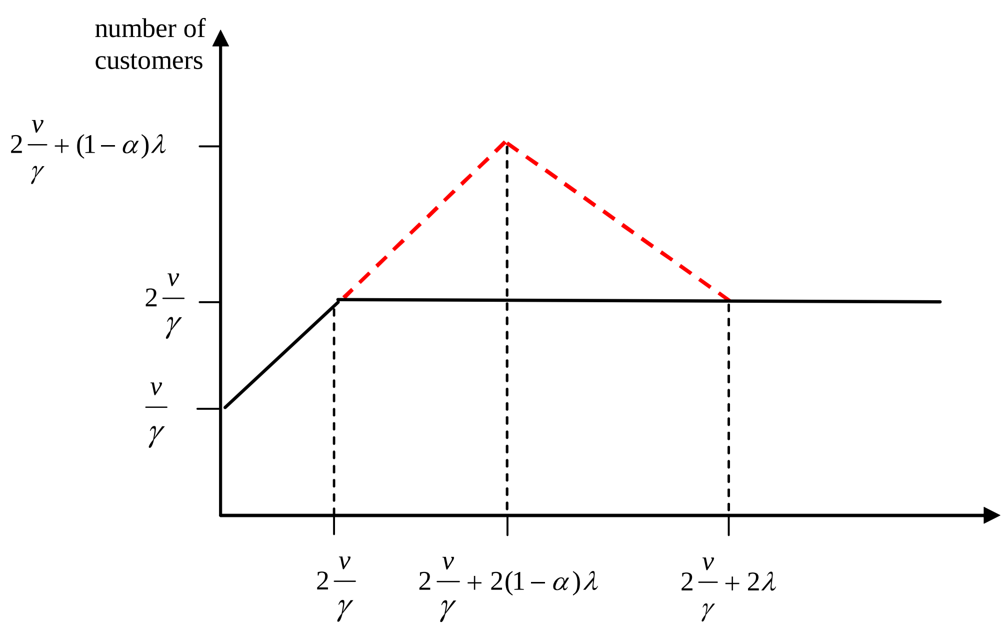


Figure 3 : Number of consumers per branch as a function of distance between two branches

The solid line corresponds to the $\alpha = 1$ case. The bank needs only to worry about the cannibalization effect. In the $\alpha < 1$ case the dashed line replaces the corresponding solid segment underneath. Now the bank can generate branch network benefits if the branches sufficiently close.

Proposition 1 *When the consumer stays at home the whole time, $\alpha = 1$, the customer neighborhoods of two branches overlap (cannibalization effect arises) whenever they are closer than $2\frac{v}{\gamma}$.*

When the consumer is in different places throughout the day, $\alpha < 1$,

a) the cannibalization effect arises whenever two branches are closer than $2\frac{v}{\gamma}$,

b) the network effect arises if the branches are located between $2\frac{v}{\gamma}$ and $2\frac{v}{\gamma} + 2\lambda$.

c) the network effect is the highest at the distance $2\frac{v}{\gamma} + 2(1-\alpha)\lambda$ and is equal to $(1-\alpha)\lambda$.

In the two-dimensional space with uniformly distributed consumers the same logic holds. If the branches are far from each other there are no network benefits. Consumers patronizing one particular branch are located in a circle of $\frac{v}{\gamma}$ radius around it (Figure 4).

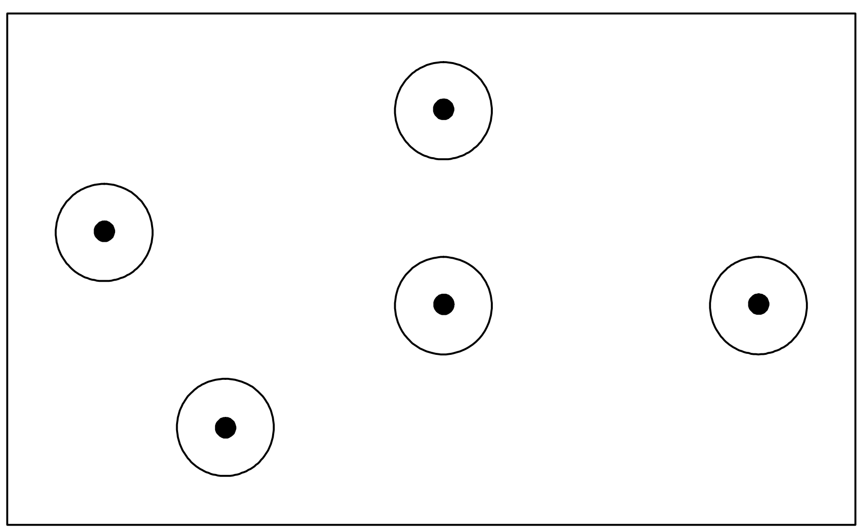


Figure 4 : Branch network without network benefits

However if the branches are located sufficiently close, network benefits arise. The circles representing customer neighborhoods expand (Figure 5).

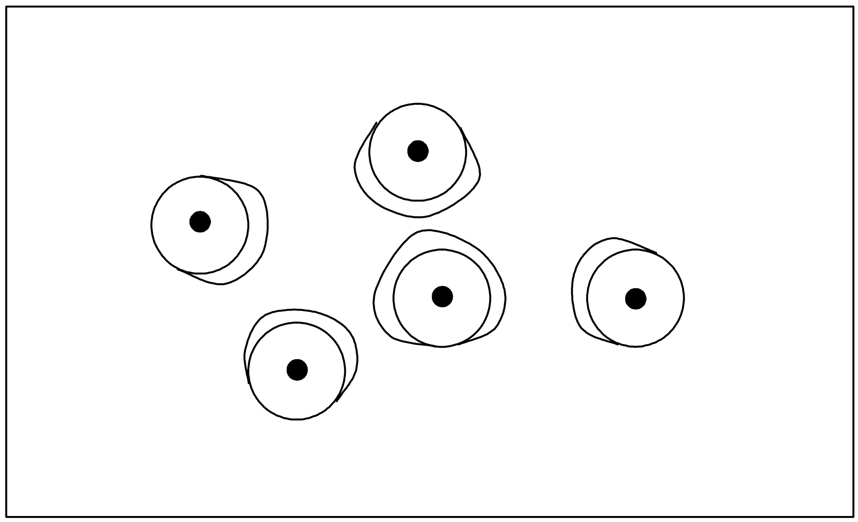


Figure 5 : Branch network with network benefits

2.5 Empirical Implementation

In my model a key source of the network benefits is the consumer's value of convenience represented by the expected distance to travel to a branch. In order to test whether these benefits indeed arise I measure to what extent the desire for convenience affects consumers' demand for banking services.

The demand for banking services estimation follows the discrete choice literature. The indirect utility of consumer i from branch j is given by:

$$V_{ij} = \beta d_j - \gamma ED_{ij}(\alpha, \mu_{i1}, ...\mu_{il}) + \eta_{ij},$$

where d_j is a bank dummy and η_{ij} is a consumer-branch specific unobservable (hence I split the ν_{ij} into βd_j and η_{ij}). In the model, the consumer is choosing a bank j in which he opens his account. Given his choice of the bank, his deposits are assigned to the branch j_b that is closest to his home.

As I pointed out in the introduction, this specification differs from the previous work in the way the geography enters the indirect utility. First, I'm not imposing any exogenous delineation of bank markets. Second, the utility of a consumer from choosing a particular bank doesn't depend on geography-proxies like the number of branches per square mile in the market, but instead it depends on the exact distances between the consumer and the branches of that bank.

I account for different characteristics of banking services at different bankes only through bank fixed effects. The banking services should be fairly homogenous across different branches of the same bank. Banks usually offer the same variety of basic products, the interior design of branches is standardized across locations. This fixed effect captures the total utility not only from observable characteristics, but also from unobservable attributes such as brand image or advertising.

I assume that $\mu_i's$ are functions of distance and population. A consumer living in location i is more likely visit a place that is closer and that has a larger population. In particular:

$$\mu_{il} = \exp(-\tau DIST_{il}) POP_{il}.$$

Hence:

$$ED_{ij} = (\alpha DIST_{ij_b} + \frac{(1-\alpha)}{\sum_{l=1,l\neq 1}^{L} \exp(-\tau DIST_{\hat{i}i}) POP_l} \sum_{l=1,l\neq 1}^{L} \exp(-\tau DIST_{il}) POP_l DIST_{l\hat{j_b}}).$$

I assume that η_{ij} has an i.i.d. type I extreme value distribution. Thus the probability that the consumer living in location i will choose bank j is given by:

$$P_{ij} = \frac{\exp(\beta d_j - \gamma ED_{ij})}{\sum_{j=1}^{J} \exp(\beta d_j - \gamma ED_{ij})}.$$

As a quantity variable, I use the total amount of deposits in each branch, not the number of accounts. Hence I assume then that each consumer is endowed with an amount of deposits that he places in the branch of his choice, and that this amount is proportional to the consumer's income.

This assumption is in line with the SCF that indicates that higher income households hold higher account balances. Hence the deposits of consumer i in the bank j are given by:

$$D_{ij} = \lambda INC_i P_{ij}.$$

The type I extreme value functional assumption about the error terms implies that an individual consumer's demand exhibits Independence from Irrelevant Alternative (IIA) property. However it's not the case for the aggregate branch demand or aggregate bank demand. The ratio of demand for branch j_b and j_b' (or bank j and j') is given by:

$$\frac{\sum_{i=1}^{I} \lambda INC_i P_{ij}}{\sum_{i=1}^{I} \lambda INC_i P_{ij'}} = \frac{\sum_{i=1}^{I} \lambda INC_i \frac{\exp(\beta d_j - \gamma ED_{ij})}{\sum_{j=1}^{J} \exp(\beta d_j - \gamma ED_{ij})}}{\sum_{i=1}^{I} \lambda INC_i \frac{\exp(\beta d_{j'} - \gamma ED_{ij})}{\sum_{j=1}^{J} \exp(\beta d_j - \gamma ED_{ij})}}.$$

If the geography didn't enter the utility, the consumers characteristics would be only included in the error term that is integrated out from the P_{ij}'s. In that case the $\sum_{j=1}^{J} \exp(\beta d_j - \gamma ED_{ij})$ terms in the numerator and denominator would cancel and this ratio would only depend on the characteristics of branches j_b and j_b' (or bank j and j'). With geography however this is not the case and this ratio depends on the characteristics of all of other branches (banks). Here the geography leads to complex substitution patterns among the branches (banks).

The utility from banking services doesn't depend on their prices. This is in contrast with the previous literature on estimating demand for banking services [3]. Not including any prices is equivalent to conjecturing that banks engage mainly in an non-price competition; they try to win consumers through accessibility embedded in the ED_{ij} and quality of services embedded in the dummy variables. From the consumers' point of view they should care relatively more about such "durable" characteristics, because they face high switching costs once they choose their banks.

[3] For example Dick (2002) uses interest rate paid on deposits and service charges on checking accounts. Ishii (2004) uses interest rate paid on deposits and surcharrge fees (fees charged by other bank for using their ATMs). Knittel and Stango (2004) use service charges on checking accounts and surcharrge fees.

There is also an issue of data availability. In reality banks offer very complicated pricing schedules[4]. Many offer very low or zero deposit rates and no fees for most types of transactions, others offer significantly positive interest rates but also positive transaction fees. The trade-off between these two types of prices always seem to be in this direction: the higher the interest rate the higher the fees. Since the actual data is not available most authors use some proxies constructed using balance sheet information. For example the deposit interest rate is calculated as the ratio of interest expenses on deposits to deposits. These are bank-level proxies and can not be separately identified from the bank fixed effects. Thus, in my approach they are subsummed in the fixed effects estimates.

Finally the unbiased estimation of the above demand specification requires that the included bank characteristics are exogenous. In this case the number and location of branches should not be driven by the level of the deposits. It seems reasonable to think of banks choosing the location of new branches so that they can attract new customers as opposed to responding to their existing customers' needs. In line with the model it's beneficial to have the branches close together (network effect), but not so that the customer neighborhoods would overlap (cannibalization effect). There might be however some unobserved factors that may lead banks with high levels of deposits to open more branches. If this is a case one might be concerned about the endogenity problem.

3 Data

To estimate the model I use two main sources of data: Summary of Deposits (SOD) data provided by the Federal Deposit Corporation (FDIC) and Census 2000.

The SOD data contains information about all branches of commercial and savings banks as

[4]Stavins (1999).

well as savings institutions in the U.S.[5]. For each branch it provides the exact address and the amount of deposits. This data is collected annually (as of June 30 of each year) and is available in electronic form (starting from 1994) on the FDIC website (http://www2.fdic.gov/sod/).

The deposits information is for all types of accounts (checking, time and saving) and for all types of customers (households and businesses). According to SCF (Amel et al.(2002)) households tend to cluster their demand for depository services at one institution. The behavior of small businesses if found to be similar to that of individuals (Kwast et a. (1997): they chose close institutions and obtain most of their services in one institution. The behavior of big firms could be potentially different.

I use the address information to find the distances between customers and branches. For that purpose I need the geographical coordinates of the addresses. This data is not provided. Instead I match the addresses with the coordinates using Geolytics software (25% are geocoded with only ZIP code precision). Figure 6 shows all the branches of all banks in Iowa.

In this paper I focus on one particular state, Iowa. In the future I will extend this analysis to the whole U.S.. Table 1 and 2 show the summary statistics for banks in Iowa. Most of them are fairly small in terms of number of branches. The mean number of branches is merely 3.17. Almost 40% of all banks (167 out of 433) have only one branch. Banks with two branches constitute 25% and with 3 to 5 branches another 27%. The top three banks are: U.S.Bank with 97, Wells Fargo with 72, and Commercial Federal Bank with 38 branches. The mean amount of deposits is 38.2 millions dollars.

The information about consumers comes from Census 2000. Ideally I would like to have information about individuals, where they live, where they work, where they shop, and so on. Lacking that type of data, I use Census information. I approximate an individual consumer by

[5]More information on what type of branches are included in the analysis is provided in the Appendix 2.

a bunch of people that live in a small area called a block group, hence a block group becomes a synonym for a consumer in my model[6]. The Census provides home addresses for such consumers in the form of geographical coordinates of each block group. I assume that consumers travel to other block groups that are within a distance of 30 miles from home. To compute distances between consumers (block groups) and bank branches I use the Haversine formula from Sinott (1984) for calculating the distance between two point on a sphere.

The block groups and counties in Iowa are shown in Figure 7. Counties are shaded differently, while block groups are the smallest shapes. Tables 3 and 4 provide the summary statistics. The total number of block groups is 2627. The mean area is 0.006 square miles. The mean population is 1,113.94 but the variation is pretty big, ranging from 3 all the way to 7,582. The mean per capita income is 19,115 dollars, the range again is substantial from 3,225 up to 68,079 dollars. On average there are 32.56 banks present in the 30 miles radius circle around each block group.

To get an idea whether there is any evidence of network benefits I do the following exercise. For each bank I measure the proximity of its branches. To do so, for each branch of a given bank, I calculate its distances to the remaining branches. Next I choose the minimum of these distances. This is a distance of this branch to the whole network of the bank. Finally I take the mean of all the minimal distances. Next I compute the average deposits per branch for each bank. and I plot them against the proximity measure. The result is shown in Figure 8. There is a slight negative relation between the two variables, which is suggestive that keeping the branches not too far from each other may in fact increase deposits per branch. This is of course a crude experiment. Like to the simple example from section 2.4, it abstracts from differences in the population and per capita income across the state and the number of competitors in the neighborhood of each bank.

[6]Block group is literally a group of blocks, it's a subdivision of a tract that in turn is a subdivision of a county.

4 Estimation

For a given vector of parameters $\theta = (\beta, \gamma, \alpha, , \tau, \lambda)$ I can calculate the predicted deposits per branch using:

$$D_j = \sum_i \lambda INC_i P_{ij}.$$

Let ε_{j_b} be the difference between the actual and predicted deposits at branch j_b:

$$\varepsilon_{j_b} = \ln(D_{j_b}^{data}) - \ln(D_{j_b}(\theta)).$$

I assume the discrepancy is normally distributed measurement error. The actual data is likely to contain measurement error. Banks use different methods of allocating total deposits in all the accounts across branches. They may assign them to the office in closest proximity to the account holder's address, the office where the deposit account is most active or the office of origination of the account. The deposits data includes also the deposits of businesses, which I don't model explicitly. Hence:

$$\varepsilon_{j_b} \tilde{} N(0, \sigma^2).$$

I look for the parameters that maximize the loglikelihood function given by:

$$LnL = -\frac{\# j_b}{2} \ln(2\Pi\sigma^2) - \frac{1}{2} \frac{\sum_{j_b} (\varepsilon_{j_b})^2}{\sigma^2}.$$

Estimates: The parameter estimates are shown in table 5. The estimate of γ, which measures the importance of the expected distance in utility, is 0.37. The probability of being home, α, is 0.7. The distance decay parameter τ that enters the probabilities of traveling to other block groups is 0.5. The estimate of the variance of measurement error, σ^2, is 1.5. The correlation between actual deposits per branch and the deposits predicted by the model is 0.25.

Given α and τ the expected distance that consumer travels to a branch is 3.35 miles.

The elasticity of demand with respect to the expected distance to a branch is given by:

$$EDE_{ij} = \frac{\partial D_{ij}}{\partial ED_{ij}} \frac{ED_{ij}}{D_{ij}} = \lambda INC_i \frac{\partial P_{ij}}{\partial ED_{ij}} \frac{ED_{ij}}{\lambda INC_i P_{ij}} = -\gamma(1 - P_{ij})ED_{ij}.$$

With the estimate of γ of 0.37 the average elasticity is -5.84.

Table 6 shows how, given these estimates, the probability of choosing a given bank changes with the expected distance to a branch. For the analysis I assume that there are just two competitors, banks A and B, with the same quality of services (the same bank fixed effect). Consider the first row where the expected distance to a branch of bank A is set to 0, and the expected distance to a branch of bank B is varied. As the expected distance to a branch of competitor B increases from 0 to 3 miles the probability of choosing bank A increases from 50 to 75%. When the expected distance goes up by another 3 miles this probability is 90% and when it reaches 10 miles this probability is close to 1. The following rows show what happens when the expected distance to branch A is increased. Since the competitors differ only by the expected distance to a branch this table is symmetric.

Branch network benefits: To asses how the deposits per branch are affected by the proximity of other branches I do the following experiment. I pick one bank. For this bank I force the consumers to always visit the branch that is closest to their home, regardless weather they are at home or they travel. In that way I decrease the convenience of this bank, each customer has a longer expected distance to travel. Using my estimates I compute the predicted deposits for the branches of that bank. I compare them with the predicted deposits in the original model. I repeat this procedure for all the banks in my sample. This experiment can be viewed from the perspective of each consumer as a decrease in the density of branch networks to such an extent that they are always closer to their home branch than to any other branch (graphically it's like switching from the network depicted in Figure 5 to the network depicted in Figure 4).

The results from this exercise can bee seen in Table 7. Depending on the size of the network the increase in the expected distance ranges from 0.25% for banks with only two branches up to 3.85% for the biggest bank in Iowa in terms of branches, U.S.Bank. The corresponding decrease in the total deposits ranges from 3.67% to 15.57%. However this relation is not monotonic. The 15.57% decrease in the deposits occurs for the bank which has only 9 branches and for which the expected distance decreased only by 2.79%. On average the expected distance increases by 0.26% and the total deposits of each bank decrease by 6%. Hence the benefits from having convenient branch networks are significant.

Alternative specifications: To see to what extents including the detailed geography of consumers and branches improves the estimation of the demand for banking services over the previous work I also estimate some alternative specifications. The results are shown in Table 8.

In the first one I set the parameter γ equal to 0 to see the explanatory power of the expected distance compared to the model with the fixed effect only. The estimate of the measurement error σ^2 is higher, 2.04, as compared to 1.50 in the model with the geography of branches.

In the next specification I set the parameter α equal to 1. This is a model in which the geography matters but only in the limited way. Consumers care only about the distance from home to the closest branch of each bank. The estimate of γ is 0.35, hence is pretty close to the original model. This is not surprising given that the estimate of α is as high as 0.7. The estimate of the measurement error σ^2 is slightly higher than in the original specification and equal to 1.52. To check whether the model with richer geography is significantly different from the model with α constrained to be 1, I do the likelihood ratio test. The hypothesis that the restriction is true is rejected even with a 0.995 confidence level.

In the third specification, instead of the exact account of the locations of branches with respect to the consumer, his utility depends merely on the number of branches of a given bank within 30 miles from him:

$$V_{ij} = \beta d_j - \kappa(\text{no of branches within 30 miles}) + \eta_{ij}.$$

The estimate of γ is -0.04, hence the number of branches seems to have almost no explanatory power. The estimate of the measurement error σ^2 is higher, 1.80, as compared to 1.50 in the model with the geography of branches.

In the last specification the geography is even more crude. The consumer's utility depends now on the number of branches of a given bank in the county in which consumer lives:

$$V_{ij} = \beta d_j - \kappa(\text{no of branches in the county}) + \eta_{ij}.$$

This model is the closest in spirit to how the regulatory agencies perceive the competition between banks: all the banks in such exogenously delineated areas are equal competitors regardless of the specific geography of the branches. The estimate of the measurement error σ^2 now is 3.82, much bigger than 1.50 in the original model and 1.80 in the model with the number of branches within 30 miles from the consumer.

5 Conclusion

In this paper I presented a model of consumer choice of depository institution. The key element of this model is the assumption that consumers are in different places throughout a day. This element is necessary to understand the benefit of a network. Ideally a consumer would like to open an account with an institution that has a dense network of branches, so that the expected distance they need to travel to get to its branches is minimal. This type of behavior has implications for the level of deposits one branch in the network can obtain depending on the location of all other branches in the network: the higher density of branches should lead to higher deposits per branch. I call this effect a network benefit.

To empirically test this model I estimated the demand for banking services using a discrete choice framework. The main component of the indirect utility function is, consistently with my model, the expected distance to a branch. I found out that the expected distance to a branch is a significant factor in explaining consumers' choices of depository institutions. According to my estimates decreasing the convenience of a branch network (increasing the expected distance to a branch) by merely 0.26% may lead to decrease of total deposits by 6%. Hence the benefits from having convenient branch networks are significant.

Most of the previous studies that estimate the demand for banking services use the information about geography of consumers and branches in a limited way. My results indicate that accounting for the exact locations of branches and not imposing any exogenous market delineations improves the prediction of the model for deposits by branch. Hence banking markets don't necessarily correspond to these commonly used exogenous delineations. A sensible definition of geographic markets is necessary for regulatory agencies that are in charge of constructing meaningful measures of market power and competition intensity between banks.

References

[1] Adams, Robert M., Kenneth P. Brevoort, and Elisabeth K.Kiser (2005), Who Competes with Whom? The Case of Depository Institutions, Federal Reserve Board, Finance and Economics Discussion Series working paper.

[2] Aguirregabiria, Victor, and Gustavo Vicentini (2004), Dynamic Spatial Competition Between Multi-Store Retailers, Boston University, working paper.

[3] Amel, Dean F., and Martha Starr-McCluer (2002), Market Definition in Banking: Recent Evidence, *The Antitrust Bulletin*, 47(1), 63-89.

[4] Berry, Steven T. (1994), Estimating Discrete Choice Models of Product Differentiation, *Rand Journal of Economics*, vol.25, 242-262.

[5] Berry, Steven T., James Levinsohn and Ariel Pakes (1995), Automobile Prices in Market Equilibrium, *Econometrica*, vol.60(4), 889-917.

[6] Dick, Astrid A. (2002), Demand Estimation and Consumer Welfare in the Banking Industry, Federal Reserve Bank of New York working paper.

[7] Holmes, Thomas J. (2004), The Location of Sales Offices and the Attraction of Cities, *Journal of Political Economy*, vol.133(3), 551-581.

[8] Holmes, Thomas J. (2005), The Diffusion of Wal-Mart and Economies of Density, University of Minnesota, working paper.

[9] Hotelling, Harold (1929). Stability in Competition, *Economic Journal*, vol.39, 41-57.

[10] Ishii, Joy (2004), Interconnection and Compatibility in Network Industries: ATM Networks in the Banking Industry, Harvard University, working paper.

[11] Kiser, Elisabeth K. (2002), Household Switching Behavior at Depository Institutions: Evidence from Survey Data, *The Antitrust Bulletin*, 47(4), 619-640.

[12] Knittel,Christopher R., and Victor Stango (2004), Incompatibility, Product Attributes and Consumer Welfare: Evidence from ATMs, University of California, Davis, working paper.

[13] Kwast, Myron L., Martha Starr_McCluer and John D. Volken (1997), Market Definition and the Analysis of Antitrust in Banking, *The Antitrust Bulletin*, 42(4), 973-995.

[14] McFadden, Daniel (1974), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka, ed., *Frontiers in Econometrics*, 105-142, New York: Academic Press

[15] Salop, Steven (1979), Monopolistic Competition with Outside Good, *Bell Journal of Economics*, vol.10, 141-156.

[16] Sinnot, Roger W. (1984), Virtues of the Haversine, *Sky and Telescope*, 68(2), 159.

[17] Stavins, Joanna (1999), Checking Accounts: What Do Banks Offer and What Do Consumers Value, *New England Economic Review* March/April, 3-13.

6 Appendix 1

Number of customers per branch in the example from section 2.4.

When $\alpha = 1$ the marginal consumer comes from the distance $x = \frac{v}{\gamma}$. If the distance between the two branches x is less or equal to $2\frac{v}{\gamma}$, then the number of customers per branch is $\frac{v}{\gamma} + \frac{x}{2}$. If the distance between the two branches x is more than $2\frac{v}{\gamma}$, then the number of customers per branch is $2\frac{v}{\gamma}$.

When $\alpha < 1$ and the distance between the two branches x is less or equal to $2\frac{v}{\gamma}$, then the number of customers per branch is $\frac{v}{\gamma} + \frac{x}{2}$, Now suppose that x starts to increase. If the branches are still located close enough that there are some consumers who will visit both branches, then for the marginal consumer, who is located x' far from the left branch, it must be that:

$$\tfrac{1-\alpha}{2}(v - \gamma(x' - \lambda)) + \alpha(v - \gamma x') + \tfrac{1-\alpha}{2}(v - \gamma(x - x' - \lambda)) = 0.$$

Hence:

$$x' = \tfrac{1}{\alpha}\tfrac{v}{\gamma} + \tfrac{1-\alpha}{\alpha}\lambda - \tfrac{1-\alpha}{\alpha}\tfrac{x}{2}.$$

The number of customers is maximized when $x = 2x'$, hence:

$$x' = \tfrac{v}{\gamma} + (1 - \alpha)\lambda.$$

So for $x \leq 2\frac{v}{\gamma} + 2(1 - \alpha)\lambda$, the number of customers per branch is given by $\frac{v}{\gamma} + \frac{x}{2}$.

If $2\frac{v}{\gamma} + 2(1 - \alpha)\lambda < x \leq 2\frac{v}{\gamma} + 2\lambda$, the number of customers per branch is given by $\frac{v}{\gamma} + (\frac{1}{\alpha}\frac{v}{\gamma} + \frac{1-\alpha}{\alpha}\lambda - \frac{1-\alpha}{\alpha}\frac{x}{2})$.

If $x > 2\frac{v}{\gamma} + 2\lambda$, the number of customers per branch is given by $2\frac{v}{\gamma}$.

7 Appendix 2

A branch is any location, or facility, where deposit accounts are opened, deposits are accepted, checks paid, and loans granted. Examples of branches are: brick and mortar locations, detached drive-in facilities, seasonal offices, offices on military bases or government installations, paying/receiving stations or units, and Internet and Phone Banking locations where a customer can open accounts, make deposits and borrow money. Not considered as branches are: Automated Teller Machines (ATM), Consumer Credit Offices, Contractual Offices, Customer Bank Communication Terminals (CBCT), Electronic Fund Transfer Units (EFTU), and Loan Production Offices.

Types of branches that I use in the estimation are categorized in the SOD data as BRSETYP 11, 12 and 23. These are:

BRSETYP=11 - Full Service - Brick and Mortar Office - Accept deposits, make loans, open/close accounts, loan officer on site, normal hours, full-time staff; may have safe deposit facilities on site. The site may be owned by the institution or may be leased by the institution.

BRSETYP=12 - Full Service - Retail Office - Accept deposits, make loans, open/close accounts, loan officer on site, normal hours, full-time staff, located in a retail facility such as a supermarket or department store; may have safe deposit facilities on site.

BRSETYP=23 - Limited Service - Drive-Through/Facility Office - Accept deposits and payments; however, may not offer other services. This may be branches own facility, located within a retail establishment or a detached drive-through branch.

Table 1: Summary statistics for banks in Iowa

	mean	std dev.	min	max
Number of branches	3.17	6.41	1	97
Mean deposits (millions)	38.2	58.4	4	1099

Source: SOD 2004

Table 2: Banks in Iowa by number of branches, mean deposits per branch and mean total deposits (millions)

no of banks	no of branches	mean deposits	total deposits
167	1	45.1	45.1
107	2	40.3	80.6
63	3	26.8	80.4
37	4	27.0	108
16	5	29.1	145.5
10	6	27.7	166.2
9	7	27.3	191.1
4	8	59.3	474.4
1	9	32.7	294.3
7	10	40.3	403
2	11	25.9	284.9
1	12	73.9	886.8
2	14	41.1	575.4
1	15	51.5	772.5
1	16	28.0	448
1	19	27.2	516.8
1	21	13.4	281.4
1	38	26.2	995.6
1	72	74.7	5378.4
1	97	41.5	4025.5
total=433	total=1372		

Source: SOD 2004

Table 3: Summary statistics for block groups in Iowa

	mean	std dev.	min	max
Population	1113.94	601.77	3	7582
Per capita income	19115.10	5962.45	3225	68079
Area (square miles)	6.02	9.60	0.01	59.46
Number of banks with a branch within 30 miles	32.56	11.28	6	58

Source: Census 2000

Table 4: Distribution of block groups by population (thousands) and number of banks with a branch within 30 miles in Iowa

	0-0.5	0.5-1	1-1.5	1.5-2	2-2.5	2.5-5	>5
6-12	0	46	10	2	0	0	0
12-18	0	79	39	3	0	0	0
18-24	5	301	158	26	4	3	0
24-30	6	341	227	44	14	9	0
30-36	2	233	126	27	7	4	0
36-42	0	114	90	27	9	8	0
42-48	5	121	99	26	21	20	5
48-54	15	111	140	55	19	14	5
>54	30	3	2	1	0	1	0

Source: Census 2000 and SOD 2004

Table 5: Estimation results - parameters

Parameters	Estimates (std errors)
γ - expected distance to a branch	0.37 (0.03)
α - probability home	0.7 (0.48)
τ - distance decay	0.5 (0.28)
σ^2 - measurement error	1.50 (0.05)

Table 6: Probability of choosing bank A depending on the expected distance to its branches (first column) and the expected distance to the branches of competitor B (first row)

	0	1	2	3	4	5	6	7	8	9	10	15	20	30
0	0.5	0.59	0.67	0.75	0.81	0.86	0.9	0.93	0.95	0.96	0.97	1	1	1
1	0.41	0.5	0.59	0.67	0.75	0.81	0.86	0.9	0.93	0.95	0.96	0.99	1	1
2	0.33	0.41	0.5	0.59	0.67	0.75	0.81	0.86	0.9	0.93	0.95	0.99	1	1
3	0.25	0.33	0.41	0.5	0.59	0.67	0.75	0.81	0.86	0.9	0.93	0.99	1	1
4	0.19	0.25	0.33	0.41	0.5	0.59	0.67	0.75	0.81	0.86	0.9	0.98	1	1
5	0.14	0.19	0.25	0.33	0.41	0.5	0.59	0.67	0.75	0.81	0.86	0.97	1	1
6	0.1	0.14	0.19	0.25	0.33	0.41	0.5	0.59	0.67	0.75	0.81	0.96	0.99	1
7	0.07	0.1	0.14	0.19	0.25	0.33	0.41	0.5	0.59	0.67	0.75	0.95	0.99	1
8	0.05	0.07	0.1	0.14	0.19	0.25	0.33	0.41	0.5	0.59	0.67	0.93	0.99	1
9	0.04	0.05	0.07	0.1	0.14	0.19	0.25	0.33	0.41	0.5	0.59	0.9	0.98	1
10	0.03	0.04	0.05	0.07	0.1	0.14	0.19	0.25	0.33	0.41	0.5	0.86	0.97	1
15	0	0.01	0.01	0.01	0.02	0.03	0.04	0.05	0.07	0.1	0.14	0.5	0.86	1
20	0	0	0	0	0	0	0.01	0.01	0.01	0.02	0.03	0.14	0.5	0.97
30	0	0	0	0	0	0	0	0	0	0	0	0	0.03	0.5

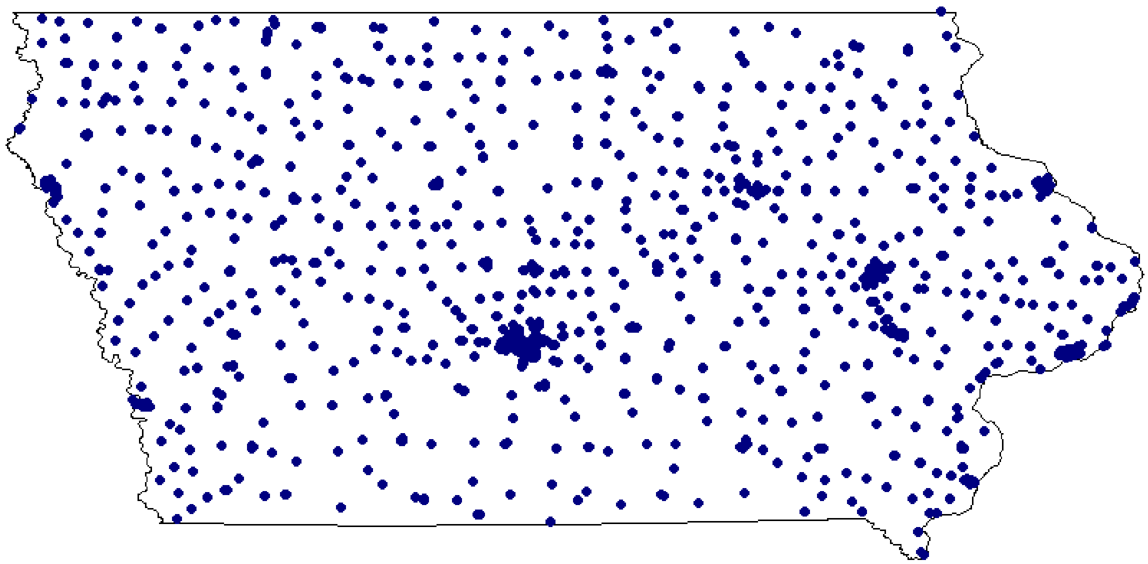
Table 7: Benefits from branch networks - percentage change in the total deposits due to percentage change in the expected distance

no of branches	change in distance	change in deposits
1	0	0
2	0.25	3.67
3	0.52	6.3
4	0.55	7.58
5	0.57	7.35
6	0.93	8.47
7	1.1	8.75
8	2.3	10.06
9	2.79	15.57
10	1	6.7
11	1.14	7.85
12	2.48	14.09
14	1.75	6.88
15	1.92	8.9
16	0.8	4.39
19	1.7	8.04
21	1.98	9.02
38	2.18	9
72	3.1	8.13
97	3.85	9.91
total=1372		

Table 8: Alternative specifications

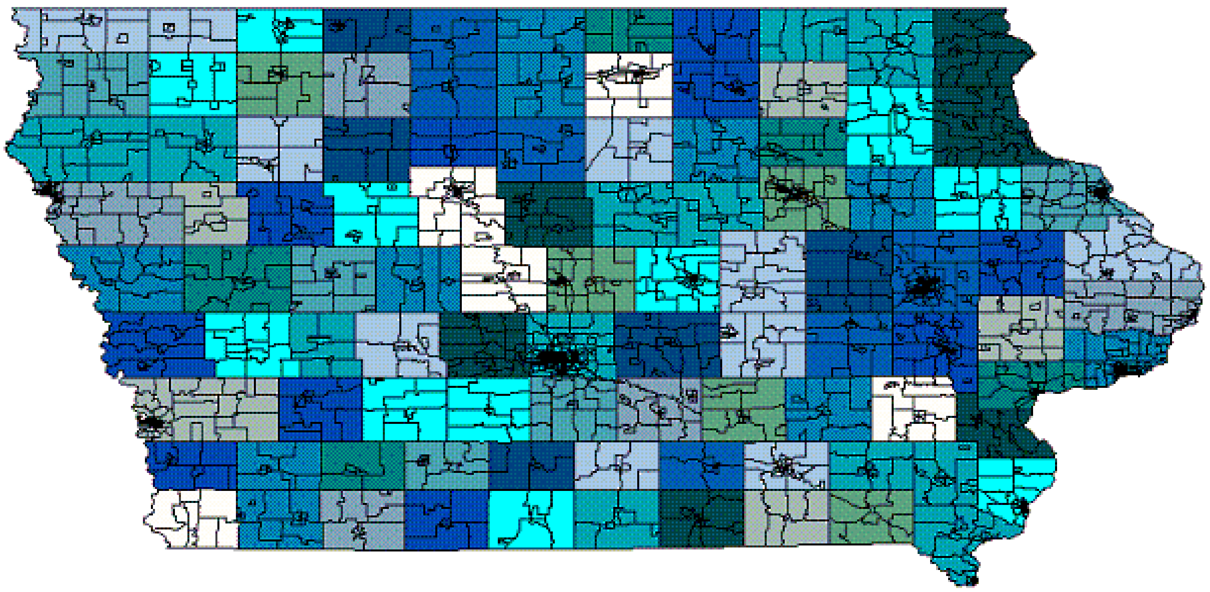
	$\gamma = 0$	$\alpha = 1$		# br. - 30 mi.	# br. - county
γ		0.35 (0.03)	κ	-0.04 (0.005)	0.2 (0.07)
σ^2	2.04 (0.07)	1.52 (0.05)	σ^2	1.80 (0.06)	3.82 (0.14)

Figure 6: Bank branches in Iowa



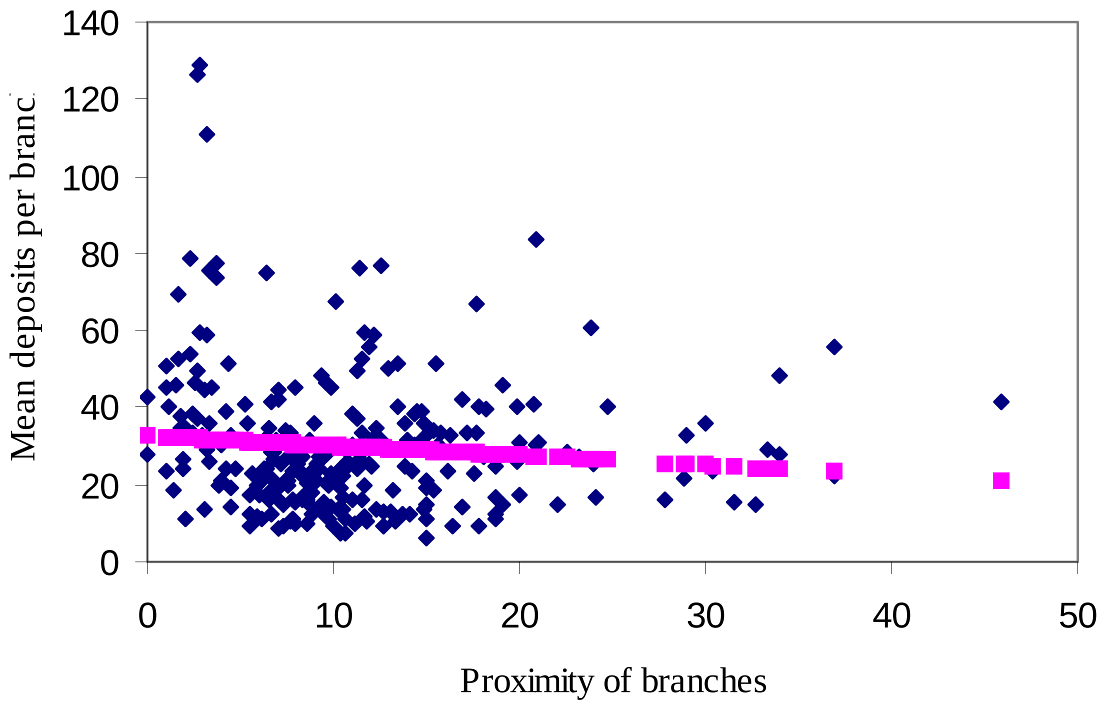
Source: SOD 2004

Figure 7: Block groups and counties in Iowa



Source: Census 2000

Figure 8: The proximity of branches in the network (miles) vs mean deposits per branch (millions)



Source: Census 2000 and SOD 2004