Explaining the Geographic Scope of Banks: A Role for Depositor Heterogeneity?

Nathan H. Miller*
University of California - Berkeley

January 2008

Abstract

The commercial banking industry is characterized by banks that differ greatly in their geographic scope. I argue that this diversity may arise due to depositor heterogeneity. I develop a theoretical model in which depositors prefer banks of greater scope but differ in their willingness-to-pay, and show that the unique subgame perfect equilibrium features a scope distribution. Structural estimation of the model firmly supports the notion that depositor heterogeneity exists and is substantial, and empirical predictions specific to the model hold in the data. The theoretical model extends naturally to any industry in which consumers value firm scope (or firm size).

Keywords: firm size, geographic scope, commercial banks, depositor heterogeneity
JEL classification: G2, L1, L2

*I thank Allen Berger, Joseph Farrell, Richard Gilbert, Ashley Langer, Juan Lleras, John Sutton, Sofia Villas-Boas, Catherine Wolfram, and seminar participants at the University of California, Berkeley for valuable comments. Please address correspondence to: Department of Economics, University of California-Berkeley, 512 Evans Hall #3800, Berkeley, CA 94720-3880.

1 Introduction

The commercial banking industry is characterized by banks that differ greatly in their scope. For example, contrast the Bank of America, which operates in 207 metropolitan markets, to the more than 3,000 commercial banks that operate within a single metropolitan market.[1] I present a new explanation for this diversity based on depositor heterogeneity. The structure of the argument is as follows. I first show theoretically that bank scope diversity arises naturally when depositors prefer banks of greater scope but differ in their willingness-to-pay. I then estimate the model using the simulated generalized method of moments. The results suggest that depositor heterogeneity exists and is substantial. Finally, I test empirical predictions distinct to the theoretical model; the results suggest that depositor heterogeneity is an important determinant of the observed bank scope distribution. The explanation for the observed scope diversity developed here is compatible with other theories, in the sense that the results do not rule out other determinants of bank scope.[2]

The theoretical model features banks that compete for depositors within a single geographic market (the "inside" market) and invest their deposits in a competitive asset market. Each bank first determines its scope, i.e., the number of outside markets in which it operates. Then, conditional on the realized scope choices, each bank sets a deposit interest rate. The depositors travel to each outside market with some exogenous probability and, in the event of travel, incur a cost unless their bank operates in the outside market. I let the probability of travel have a distribution among depositors, which introduces differences in willingness-to-pay for scope, and assume that each depositor selects the bank that maximizes expected utility. The unique subgame perfect equilibrium is characterized by a distribution of bank scopes, the support of which is limited only by the number of outside markets to which depositors may travel. Furthermore, banks of greater scope set lower deposit rates (i.e., higher prices) and maintain higher deposit shares within the inside market. As I discuss in detail below, this conforms to a second well-documented stylized fact in the industry.

In the empirical implementation, I first estimate a structural model of depositor choice using the simulated generalized method of moments (e.g., Berry, Levinsohn and Pakes 1995, Nevo 2001) and test the first-order assumption of the theoretical model, namely that depositors prefer banks of greater scope but differ in their willingness-to-pay. Estimation exploits nearly 40,000 bank-market-year observations over the period 2001-2006, as well as

[1]Data are from the 2006 Summary of Deposits.

[2]The most prominent explanation for scope diversity invokes the notion that small banks have comparative advantages evaluating "opaque" small business loans and comparative disadvantages evaluating other loans (e.g., Stein 2002; Cole, Goldberg and White 2004; Berger et al 2005). I discuss this explanation below.

individual-level data from 2000 Consumer Population Survey (CPS) March Supplement. The data suggest that the mean depositor annually values a unit increase in bank scope at 42.90 cents. The number may be substantial, as it implies that the mean depositor values the scope of Bank of America at 88.41 dollars more than that of a single-market bank. The data also suggest that scope valuations increase in depositor income, and a statistical test firmly rejects the null hypothesis that no depositor heterogeneity exists. The estimation procedure itself may be of interest to banking economists because, to my knowledge, it is the first demand system that flexibly incorporates heterogenous depositor preferences for observables.

I then evaluate three empirical predictions of the theoretical model that are not easily generated under (realistic) alternative assumptions. The first empirical prediction is that a bank should be less likely to enter an outside market if its original market already features banks of greater scope. Within the context of the theoretical model, such entry may be disadvantageous because it lessens scope differentiation and intensifies deposit rate competition. To test the prediction, I examine the entry decisions of 3,658 single-market banks over the sample period. I develop a reduced-form model of entry and estimate the model using standard logit regression techniques. The results are strongly consistent with the empirical prediction. For example, the results imply that a hypothetical one standard deviation increase in the number of two-market banks (i.e., banks with branches in exactly two markets) in the original market reduces the probability that a single-market bank enters an outside market by an average of 76.54 percent.

Any bank that enters an outside market must decide which outside market to enter, and the second empirical prediction deals with this choice. The theoretical model suggests that a bank should be less likely to enter a specific outside market (conditional on entry to some outside market) if another bank of similar scope already exists in both the original market and the specific outside market. To test the prediction, I examine the 351 single-market banks that enter a second market over the sample period. I again develop a reduced-form model of entry and estimate the model using conditional logit regression techniques, controlling directly for market synergies and the profitability of the outside markets. The results support the empirical prediction. For example, the results imply that single-market banks are at least 35 percent less likely to enter a specific outside market if a two-market bank already operates in the original and the specific outside market.

The third prediction is that the number of banks within a given geographic market should increase with that market's nearness to other markets. The prediction follows the theoretical result that the support of the bank scope distribution (and thus the number of banks) within a market is determined by the number of relevant outside markets. Under the intuition that the probability of travel decreases in the distance between markets, one would expect markets that are near others to provide greater opportunities for scope differentiation and support more banks. I proxy the nearness of a geographic market with the mean distance (in miles) between the market and all other markets. The prediction holds in the data – markets that are near other markets support more banks, even controlling for market income, population, and land area. However, the relationship is driven solely by the differences in the number of single-market banks; there is no apparent relationship between market nearness and the number of multimarket banks. It is therefore difficult to conclude that the finding provides substantive support for the theoretical model.

The theoretical model extends naturally to any industry in which consumers value firm scope (or firm size) but differ in their willingness-to-pay. Two prominent examples include the commercial airline industry and the wireless telephone industry. In the commercial airline industry, consumers that accumulate frequent flier models may prefer airlines that fly to many destinations (e.g., Delta Air Lines) over airlines that fly to fewer destinations (e.g., Alaska Airlines). In the wireless telephone industry, consumers that travel may prefer providers that offer national coverage (e.g., Verizon Wireless) over the many providers that offer only local or regional coverage (e.g., Cellcom in northeast Wisconsin). The theoretical model can therefore be interpreted as extending the well-developed literature on optimal firm size. Of course, whereas the bulk of this literature invokes various supply-side frictions to explain firm size (for example, consider the transaction-cost models of Williamson [1965, 1979, 1985] and Klein, Crawford and Alchian [1978] and the control-rights models of Grossman and Hart [1986], Hart and Moore [1990] and Hart [1995]), the theoretical model developed here invokes demand-side factors and is thus closer to the vertical differentiation models of Shaked and Sutton (1982, 1983).

The paper proceeds as follows. Section 2 develops two stylized facts about the commercial banking industry and discusses the relevant literature. Section 3 presents the theoretical model and derives empirical predictions. Section 4 describes the data, estimates the structural model of depositor choice, and tests the empirical predictions. Section 5 concludes.

2 Two Stylized Facts

The theoretical model conforms to two stylized facts of the commercial banking industry: 1) a distribution of bank scope exists and 2) banks of greater scope offer lower deposit rates yet capture greater deposit market shares within markets. I develop these facts here and discuss the relevant literature. I defer detailed discussion of the data for expositional convenience.

Panel A of Table 1 shows the total number of commercial banks with branches in at least one metropolitan Core Based Statistical Area (CBSA) and Panel B shows the mean number of commercial banks per CBSA.[3] Each panel is tabulated by bank scope over the period 2001-2006. Most banks operate within a single CBSA (e.g., nearly 80 percent in 2006) and the vast majority operate in fewer than six CBSAs (e.g., 97 percent in 2006). The diversity of bank scope is pronounced within individual CBSAs – nearly 75 percent of the CBSA-year combinations over the sample period include at least one bank in each scope category, and the average CBSA-year combination features 11.21, 5.20, 2.22, and 3.80 banks that operate in 1, 2-5, 5-20 and more than 20 CBSAs, respectively. The distribution is relatively stable through the sample period. Although banks of greater scope are more common in 2006 than 2001, the changes appear to be of only second-order magnitude. It may therefore be reasonable to conjecture that bank scope diversity is real and not simply part of an out-of-equilibrium transition.[4]

[Table 1 about here.]

The extant literature explains the coexistence of large and small banks as a product of comparative advantages in loan underwriting technologies. In particular, larger banks may exploit economies to scale in processing easily quantifiable information (i.e., "hard information"). By contrast, smaller banks may better handle qualitative information that is difficult to transmit and/or verify across layers of bureaucracy (i.e., "soft information"), because they can properly align loan officer research incentives (Stein 2002) and/or mitigate agency problems (Berger and Udell 2002). Some empirical evidence supports this hypothesis. Smaller banks typically allocate a far greater proportion of their assets to small business loans, which may be more difficult to evaluate via hard information (e.g., Berger, Kashyap, and Scalise 1995). More directly, the quantitative financial statements of loan applicants are less predictive of subsequent underwriting decisions at smaller banks (Cole, Goldberg and White 2004), and smaller banks also tend to have closer, more personal, more exclusive, and

[3]The Office of Management and Budget defines a metropolitan CBSA to be a geographic areas that contains at least one urban area of 50,000 or more inhabitants. A CBSA also includes surrounding counties that meet specific commuting requirements.

[4]Deregulation and technological advances increased the efficient scope of banking activity during the 1980s and 1990s. The result was a protracted period of consolidation: for example, Berger (2003) reports that the number of banks, inclusive of rural banks, decreased from 14,392 to 8,016 over the period 1984-2001. Berger, Kashyap and Scalise (1995) and Berger (2003) provide excellent reviews of the relevant banking literature.

longer relationships with their borrowers (Berger et al. 2005).[5]

The extent to which comparative advantages in loan underwriting technologies fully explain the bank scope distribution is not clear. From a theoretical standpoint, the framework predicts a bimodal scope distribution characterized by distinctly large and small banks that specialize in the analysis of hard and soft information, respectively. The observed distribution, by contrast, suggests a prominent role for "medium-sized" banks. Further, recent empirical evidence suggests that only 20 percent of small business loans issued by banks with assets under $1 billion are evaluated with soft information underwriting technologies (Berger and Black 2007). The market for soft information loans, by itself, may not be sufficient to support the operation of small banks.

I now turn to the second stylized fact. Table 2 shows mean deposit interest rates and market shares, tabulated by bank scope. The means are based on 45,785 bank-CBSA-year observations over the period 2001-2006. I calculate the deposit interest rates as interest expenses over total deposits, and the market shares as a proportion of all commercial bank deposits in the CBSA-year. As shown, deposit rates decrease with scope and market shares increase in scope. The differences are dramatic: for example, the average single-market bank offers a deposit rate that is 58 percent higher than the average bank with branches in more than twenty CBSAs, yet it captures less than one-fifth the market share.

[Table 2 about here.]

One common explanation for the negative relationship between deposit rates and bank scope is that larger banks have superior access to wholesale funds and substitute away from deposits (e.g., Kiser 2004; Hannan and Prager 2004; Park and Pennacchi 2004). However, the wholesale funds hypothesis predicts that deposit shares decrease with scope, and is therefore incomplete at best. Potentially closer is the observation of Bassett and Brady (2003) that the deposit growth of small banks exceeded that of large banks over the period 1990-2001. Indeed, the combination of sticky deposit supply – due to switching costs and/or other factors – and appropriate initial conditions could generate the patterns shown in Table 2. Although there is some empirical evidence supporting the existence of deposit supply stickiness (e.g., Sharpe 1997, Kiser 2002a, 2002b), to my knowledge no work has attempted to quantify its importance for the banking industry structurally.

[5] Similarly, Liberti and Mian (2006) show that the underwriting decisions of loan officers may weight soft information more heavily than those of bank managers.

3 Theoretical Model

3.1 The equilibrium concept

The theoretical model is based on a three stage non-cooperative game. The timing of the game is as follows: In the first stage, banks decide whether to enter a single geographic market, which I label the inside market. Entry decisions become common knowledge at the end of the first stage. In the second stage, banks that enter the inside market choose whether to establish a presence in each of M outside markets. Finally, banks set deposit rates given the second stage actions and compete for deposits within the inside market.

Strategies consist of actions to be taken in each of the three stages. Strategies are therefore of the form: do not enter the inside market; or enter the inside market, establish a presence in $0, 1, \ldots,$ or M outside markets (conditional on the first stage actions), and set a deposit interest rate (conditional on the first and second stage actions).

Banks that enter the inside market receive payoffs (profits) based on a depositor choice model introduced below. Banks that do not enter the inside market receive payoffs of zero.

The solution concept is that of subgame perfect equilibrium (e.g., Selton 1975). An n-tuple of strategies forms a subgame perfect equilibrium if it forms a Nash equilibrium in every stage-game. To solve the game, I first analyze deposit rate competition in the third stage, and then turn to the second and first stages.

3.2 Deposit rate competition

Suppose that $j = 1, 2, \ldots, J$ banks enter the inside market. Each bank is characterized by the number of outside markets in which it is present (m_j), and chooses a deposit rate r_j^D to maximize profits:

$$r_j^D = \arg\max(r^L - r_j^D) N s_j(r_\cdot^D, m_\cdot), \tag{1}$$

where r^L is the fixed interest rate obtained from investments in a competitive lending market, N is the size of the inside market, s_j is the share of deposits obtained from the inside market, and $r_\cdot^D$ and $m_\cdot$ are vectors of the deposit rates and bank scopes, respectively. Without loss of generality, I normalize the size of the inside market to one.

Deposit shares are determined by a continuum of depositors that differ only in the probability with which they travel to outside markets. This probability of travel, which I denote as α, is distributed according to some cumulative distribution function $F(\alpha)$ with support between zero and one. Depositors split the probability of travel evenly across the outside markets, so that the probability of travel to each outside market is simply α/M.[6] Depositors that travel to an outside market from which their bank is absent pay a cost γ to participate in a competitive ATM market. Thus, the expected utility that depositor i receives from bank j takes the form

$$\mathrm{E}[u(\alpha, j)] = (1-\alpha) r_j^D + \alpha \frac{m_j}{M} r_j^D + \alpha \left(1 - \frac{m_j}{M}\right) (r_j^D - \gamma), \tag{2}$$

where the term $\frac{m_j}{M}$ is the conditional probability of travel to an outside market in which bank j has a presence, given that travel occurs. Combining terms yields the tractable expression

$$\mathrm{E}[u(\alpha, j)] = r_j^D + \gamma\alpha \left(\frac{m_j}{M} - 1\right). \tag{3}$$

The utility representation has the interpretation that depositors that will not travel (i.e., $\alpha = 0$) consider only the deposit interest rate. Importantly, all depositors at least weakly prefer banks of greater scope for a given deposit interest rate. The set-up therefore fits within the class of vertical differentiation models first analyzed by Shaked and Sutton (1982, 1983) and Tirole (1988, Section 2.1).[7]

Depositors select the bank that provides the greatest expected utility. This implicitly defines the set of travel probabilities that corresponds to the selection of bank j, and integrating over the set yields an expression for the market share of bank j:

$$s_j = \int_{[\alpha| \; u_{ij} \geq \; u_{ik} \forall \; k=1,\ldots, \; J]} \partial F(\alpha). \tag{4}$$

Ranking banks in increasing order of scope, such that $m_1 < m_2 < \cdots < m_J$, there exists a travel probability α_j such that depositors characterized by α_j are indifferent between bank j and bank $j-1$ at the relevant deposit rates and sizes, i.e., $\mathrm{E}[u(\alpha_j, j)] = \mathrm{E}[u(\alpha_j, j-1)]$. For banks $j > 1$, these indifference travel probabilities have the expression:

$$\alpha_j = \frac{r_{j-1}^D - r_j^D}{\gamma \left(\frac{m_j - m_{j-1}}{M}\right)}. \tag{5}$$

[6]Alternatively, it is possible model depositors that differ in their market-specific travel probabilities. For example, one New Yorker may be more likely to visit Chicago than Boston while another may be more likely to visit Boston than Chicago. The results presented here extend naturally.

[7]Vertical differentiation models typically analyze firms that choose their "quality" and then compete in prices for consumers that differ in their willingness-to-pay. In these models, the support of consumer preferences tends to bound the number of profitable firms (e.g., Shaked and Sutton 1983, 1984; Motta 1993) and prices, market shares, and profits generally increase in quality (e.g., Shaked and Sutton 1982; Choi and Shin 1992; Donnenfeld and Weber 1992; Wauthy 1996; Lehmann-Grube 1997).

Depositors with the travel probability $\alpha < \alpha_j$ strictly prefer bank $j-1$ to bank j, whereas depositors with $\alpha > \alpha_j$ strictly prefer bank j to bank $j-1$. The indifference travel probability for the bank of least scope has the special form $\alpha_1 = \frac{r_1^D}{\gamma(1-m_1/M)}$, which reflects the possibility that not all depositors select a bank.

Further progress requires an evaluation of the market shares, and I let α have uniform density in order to facilitate an analytic solution. The uniform density simplifies the market share equations: $s_j = \alpha_{j+1} - \alpha_j$ for $0 \leq j < J$ and $s_J = 1 - \alpha_J$.[8] One can then differentiate the profit function to obtain the first order conditions that characterize any stage-game equilibrium:

$$
\begin{aligned}
r_1^D &= \begin{cases} \frac{1}{2}(r^L + r_2^D) & \text{if } r_1^D \geq 0 \\ \frac{1}{2}(r^L + \left(\frac{M-m_1}{M-m_2}\right) r_2^D) & \text{if } r_1^D < 0 \end{cases} \\
r_j^D &= \frac{1}{2}\left(r^L + \left(\frac{m_{j+1}-m_j}{m_{j+1}-m_{j-1}}\right) r_{j-1}^D + \left(\frac{m_j - m_{j-1}}{m_{j+1}-m_{j-1}}\right) r_{j+1}^D\right), \quad 1 < j < J \\
r_J^D &= \frac{1}{2}\left(r^L + r_{J-1}^D - \gamma\left(\frac{m_J - m_{J-1}}{M}\right)\right)
\end{aligned} \tag{6}
$$

A number of results follow. The first is that any two banks of identical scope necessary trigger undifferentiated Bertrand competition in the deposit interest rate:

Lemma 1. *Let two banks of equal scope exist ($m_j = m_k$). In any stage game Nash equilibrium (if it exists), these banks set their deposit rates to the competitive loan interest rate ($r_j^D = r_k^D = r^L$). Any bank l with $m_l \leq m_j = m_k$ does not earn positive profits.*

It follows that banks can earn positive profits only through scope differentiation. The result places a natural upward bound on the number of banks that can earn positive profits. The bound is determined by the extent to which differentiation is feasible:

Corollary 1. *In any stage game Nash equilibrium (if it exists), no more than $M+1$ banks earn positive profits.*

The logic is intuitive. If more than $M+1$ banks exist then at least two banks must be of equal scope and Lemma 1 applies.

[8] The use of the uniform distribution in models of vertically differentiated firms is not typically critical to the results (e.g., Gabszewicz et al. 1981).

The next results develop the equilibrium relationships between scope, deposit rates, and market shares. It is apparent from the expected utility specified in Equation 3 that depositors at least weakly prefer banks of greater scope at any given deposit interest rate:

Lemma 2. *Let bank j have greater scope than bank k ($m_j > m_k$). If bank j sets a deposit rate at least as large as bank k ($r_j^D \geq r_k^D$) then bank k has zero market share and does not earn positive profits.*

Lemma 2 helps deliver the key result that deposit interest rates decrease in scope and market shares increase in scope.

Proposition 1. *Any stage game Nash equilibrium (if it exists) features deposit rates that decrease strictly in scope, i.e., $r_j^D > r_{j+1}^D$ for any $j < J$. Further, if $J = M+1$ and $m_j = j-1$ then any stage game Nash equilibrium features market shares and profits that increase strictly in scope, i.e., $s_j < s_{j+1}$, and $\pi_j < \pi_{j+1}$ for any $j < J$.*

Proposition 1 implies that any stage game Nash equilibria (and thus any subgame perfect equilibrium) may be consistent with the second stylized fact developed in Section 2. The proposition may also help reconcile the empirical finding of Berger and Mester (2003) that bank mergers enhance profit productivity but not cost productivity, controlling for market power. Under Proposition 1, increases in scope may increase profit even with no changes in production technology.

The final stage game result is that a unique stage game Nash equilibrium does indeed exist, provided the number of banks does not exceed the upward bound for profitability and the banks are suitably differentiated in scope.

Lemma 3. *Suppose that $J \leq M+1$ and $m_j \neq m_k$ for all $j \neq k$. Then a unique stage game Nash equilibrium exists in which all J banks have positive profits.*

3.3 Entry and competition in bank scope

I now analyze competition in the second and first stages, in turn.

In the second stage, banks that enter the inside market choose their scope, given the number of inside market entrants (J) and the number of outside markets (M). The characterization of the subgame perfect strategies in the two stage subgame depends on the relative number of entrants and outside markets. In particular, I develop separate results for the cases in which $J \leq M+1$ and $J > M+1$:

Lemma 4. *If $J \leq M+1$ then the two stage subgame has a unique class of subgame perfect equilibria in which each bank differs in scope.*

Lemma 5. *If $J > M+1$ then the two stage subgame has a unique class of subgame perfect equilibria in which there exists at least one bank with each possible scope $1, 2, \ldots, M$, and $m_j = m_k$ for some j and k.*

The results follow naturally from the third stage deposit rate competition. So long as the number of entrants is no greater than the upward bound established in Corollary 1, any bank that is not differentiated in scope has a profitable deviation in the second stage. On the other hand, if the number of entrants exceeds the upward bound then at least one bank is undifferentiated and earns no profits in the third stage, yet has no profitable deviation in the second stage.

To complete the analysis, in the first stage some number of banks choose whether to enter the inside market given the number of outside markets (M) and with full knowledge of competition in the subsequent stages. In order to eliminate equilibria in which banks enter and then fail to earn positive profits in the third stage, I introduce an arbitrarily small entry cost $\epsilon > 0$. The main result of the theoretical model follows immediately.

Proposition 2. *For any sufficiently small $\epsilon > 0$ and a number of potential entrants $N > M$, there exists a unique subgame perfect equilibrium in which exactly $M+1$ banks enter the inside market. Each bank enters a different number of outside markets and earns positive profits.*

Proposition 2 establishes that the unique subgame perfect equilibrium is consistent the first stylized fact presented in Section 2, namely that a diversity of bank scopes exists. This diversity exists despite the fact that every depositor at least weakly prefers banks of greater scope for a given interest rate. The presence of depositor heterogeneity allows larger banks to reduce their deposit rates and still attract depositors with high scope valuations. Smaller banks, meanwhile, maintain a (relatively) unprofitable niche in which they attract depositors with low scope valuations.[9]

[9]The result that some smaller firms may prefer their niche position has empirical parallels. For example, Bajari, Fox and Ryan (2006) note that small cellular carriers often do not agree to charge roamers low per-minute rates.

3.4 Testable empirical predictions

To the extent that differences in depositor willingness-to-pay exist and are important, the theoretical model has a number of empirical predictions. The first two predictions correspond to the stylized facts developed above:

> **Prediction 1.** A diversity of bank scopes should exist within markets.

> **Prediction 2.** Banks of greater scope should have lower deposit interest rates and higher market shares.

Section 2 shows that these predictions hold in the banking data over the period 2001-2006. Although a number of alternative theories can together generate these predictions, the theoretic model presented here may retain its appeal because it provides a single intuitive explanation. The next predictions further differentiate the theoretical model:

> **Prediction 3.** A bank with branches in market m should be less likely to enter an outside market if banks of greater scope already operate in market m.

> **Prediction 4.** Conditional on entry to some outside market, a bank with branches in market m should be less likely to enter market n if another bank of similar scope already exists with branches in markets m and n.

> **Prediction 5.** The number of banks within a market should increase with the market's nearness to other markets.

The third and fourth predictions follow the theoretical result that scope differentiation improves profitability (e.g., Lemmas 1 and 4). The fifth prediction follows the intuition that depositors are more likely travel to nearby markets than distant ones. Markets that are near others have more outside markets that are relevant, and thus offer superior opportunities for scope differentiation (e.g., Corollary 1).

4 Data and Empirical Implementation

4.1 Data

The bulk of the data used in this study comes from the Summary of Deposits and the December and June Call Reports. The Summary of Deposits tracks the location of all commercial bank branches and deposits and is maintained by the Federal Deposit Insurance Corporation (FDIC). The Call Reports contain the balance sheets and income statements of commercial banks and are maintained by the Federal Financial Institutions Examination Council (FFIEC). I compile data from these sources over the period 2001-2006. The data yield 26,905 observations on the bank-year level and 45,785 observations on the bank-CBSA-year level; each observation is in June of its respective year.

Table 3 presents summary statistics at the bank-year and bank-CBSA-year level. The main quantities of interest are the deposit interest rates, the market shares, and the bank scopes. I calculate the deposit rates as interest expenses (incurred over the previous year) over deposits (averaged over the previous year), the market shares as deposits over the sum of all commercial bank, thrift, and credit union deposits, and the scopes as the numbers of CBSAs in which the banks have branches. To be clear, the deposit interest rates and the scopes vary on the bank-year level, and the market shares vary on the bank-CBSA-year level.[10] The mean bank-year observation has a deposit rate of 0.019, an average market share of 0.025, and scope of 1.70. The means of the bank-year-CBSA observations more heavily weight banks that operate in many CBSAs. The lower deposit rate mean (0.017) and higher market share mean (0.047) are consistent with the stylized fact that banks of greater scope offer lower deposit rates and have higher market shares.

Turning to the remaining commercial bank variables, the mean bank-year observation has 1.235 billion dollars in gross total assets. The assets are funded, in part, by an average of 0.76 billion dollars in deposits and invested in an average of 0.72 billion dollars worth of loans. Finally, the mean bank-year observation has 2.94 branches in each of its markets, employs 23.41 people per branch, and charged off roughly five million dollars of loans over the previous year. The mean bank-CBSA-year observation has much higher gross total assets, loans, deposits, and branch density because the mean more heavily weights the larger banks.

[Table 3 about here.]

The statistical tests of Predictions 3 and 4 require some aggregation to the CBSA-year level, and Table 4 presents summary statistics for the 2,160 CBSA-year observations in the data over the period 2001-2006. As shown, the mean CBSA-year contains 9.49 billion dollars in deposits, spread among 21.02 banks and 152.11 bank branches. The mean CBSA

[10]The lack of CBSA-specific deposit rates may not hinder empirical analysis. More than 80 percent of the bank-year observations have branches in a single CBSA, and larger banks tend to set uniform deposit rates across CBSAs (e.g., Radecki 1998; Heitfield 1999). Although the regulatory reports do not report the location of ATMs, Dick (forthcoming) reports a correlation coefficient between the number of branches and the number of ATMs of 0.80, based on data obtained from a large ATM network.

population, median household income (in dollars), and land area are 645.47, 48.03, and 22.11 thousand, respectively. Implicit in the data construction is the notion that metropolitan CBSAs approximate well the relevant geographic markets for depository services. Recent research provides some support: Amel and Starr-McCluer (2001) show that the median household travels only three miles to its depository institution and that roughly 90 percent of checking and savings accounts are held by local depository institutions. Kwast, Starr-McCluer and Wolken (1997) report similar numbers for small businesses. Finally, Heitfield (1999) and Heitfield and Prager (2004) show that average deposit rates vary across cities, and interpret the results as consistent with local market competition.

[Table 4 about here.]

4.2 The estimation of depositor willingness-to-pay

In this section, I develop and estimate a structural model of depositor choice and test the first-order assumption of the theoretical model, namely that depositors prefer banks of greater scope but differ in their willingness-to-pay.[11] To improve the empirical relevancy of the theoretical model, I augment the expected utility specification of Equation 3 such that it more flexibly accommodates bank differentiation and depositor heterogeneity:

$$\mathrm{E}[u_{ijmt}] = \beta_i^D r_{jt}^D + \beta^M \left(\frac{m_{jt}}{M} - 1\right) + x'_{jmt}\beta^X + \xi_{jmt} + \epsilon_{ijmt}, \tag{7}$$

where i, j, m, and t index depositors, banks, markets, and years, respectively. The vector x_{jmt} includes observable bank characteristics, the scalar ξ_{jmt} is the mean depositor valuation of bank j, and the scalar ϵ_{ijmt} allows for depositor-specific taste shocks. Within the context of the theoretical model, the parameter β_i^D is the inverse travel probability (i.e., $\beta_i^D = 1/\alpha$) and the parameter β^M is the annualized cost of participating in the competitive ATM market (i.e., $\beta^M = \gamma$.) I combine these parameters to calculate willingness-to-pay for a unit increase in scope, given by $\beta^M/(M\beta_i^D)$.[12] The final parameter vector, β^X, captures the contribution of bank observables to expected utility.

The depositor-specific parameter β_i^D allows for heterogeneity in the willingness-to-pay for bank scope. I model the parameter as having the multivariate normal distribution

[11] The structural model requires its own set of assumptions, and I evaluate throughout the extent to which estimation is robust to different specifications and identification strategies.

[12] I convert willingness-to-pay into dollar terms by multiplying by 3,800, the median dollar amount held in transaction accounts (Bucks, Kennickell and Moore 2006).

conditional on depositor income:

$$\beta_i^D = \beta^D + \pi y_i + \sigma \nu_i, \qquad \nu_i \sim \mathrm{N}(0,1), \tag{8}$$

where y_i is income, ν_i captures unobserved demographics, the parameter β^D is the mean depositor interest rate valuation, the parameter π allows this mean valuation to vary across incomes, and σ is a scaling vector. This formulation permits a simple statistical test for the presence of heterogeneity in willingness-to-pay: under the null hypothesis of no heterogeneity, the demographic parameters π and σ are jointly zero.

As in the theoretical model, I permit depositors to select the outside option, i.e., elect not to choose a commercial bank. I let the outside option be thrifts and credit unions, and use thrift and credit union balance sheets to determine the outside option market shares directly.[13] The use of alternative outside option shares, calculated as population (times a constant of proportionality) less commercial bank deposits, returns similar parameter estimates. I denote the outside option as $j = 0$. Since the mean valuations of the outside option are not separably identifiable, I let the expected utility obtainable from the outside option be $\mathrm{E}[u_{i0mt}] = \epsilon_{i0mt}$.

I again maintain the assumption that depositors select the bank that provides the greatest expected utility. This defines the set of depositor attributes $\{y_{im},\, \nu_{im},\, \epsilon_{i.mt}\}$ that correspond to the selection of bank j in market m and year t. Integrating over the set yields an expression for the deposit market shares:

$$s_{jmt} = \int_{[(y_i,\, \nu_i,\, \epsilon_{ijmt})|\ u_{ijmt} > u_{ikmt} \forall\ k=0,\ldots,\ J]} \partial F(\epsilon) \partial F(\nu) \partial F(y), \tag{9}$$

where $F(\cdot)$ denotes the relevant population distribution functions. Specific distributional assumptions on $F(\cdot)$ enable evaluation of the integral via analytic and/or numerical methods. Throughout, I let ϵ_{ijmt} be distributed *iid* with the extreme value type I density. The distributional assumption on ϵ is itself not restrictive, as mixed logit models can approximate any random utility model, to any degree of accuracy, given appropriate choices of regressors and random parameter distributions (McFadden and Train 2000).[14]

[13]The use of thrifts and credit unions as the outside good is somewhat compelling. Amel and Starr-McCluer (2001) show that commercial banks, thrifts and credit unions together account for 98 percent of all checking and savings accounts. Further, lumping the individual thrift and credit union institutions together may not unduly restrict the substitution patterns of interest: Adams, Brevoort and Kiser (2007) estimate a median cross interest rate elasticity between commercial banks and thrifts and only -0.002.

[14]Somewhat more troublesome, given the empirical evidence regarding depositor switching costs (e.g., Sharpe 1997; Kiser 2002a, 2002b), is the assumption that the taste shocks are independent over time.

As a prelude to the mixed logit estimation, I first impose the restriction that depositors have homogenous tastes for observables (i.e., I impose $\pi = \sigma = 0$). The restriction makes integration over the demographics y and ν unnecessary. The resulting logit demand system follows from analytical integration over the depositor-specific taste shocks:

$$\log(s_{jmt}) - \log(s_{0mt}) = \beta^D r^D_{jt} + \beta^M \left(\frac{m_{jt}}{M} - 1\right) + x'_{jmt}\beta^X + \xi_{jmt}. \tag{10}$$

The equation can be estimated by treating the mean valuation (ξ_{jmt}) as an unobserved error term. The logit demand system is problematic because it specifies unrealistic elasticities (e.g., Berry 1994; Berry, Levinsohn and Pakes 1995; Nevo 2001) and does not permit the test of depositor heterogeneity that is of interest here. Nonetheless, its estimation is much less computationally burdensome than that of the mixed logit model, and may also provide substantial intuition regarding the validity of the identification strategy.[15]

Table 5 shows the logit results. The Column 1 results are computed with OLS. The coefficients have the expected signs: depositors appear to prefer to prefer banks with higher deposit rates, greater scopes and branch densities, and more employees per branch. The deposit rate coefficient of 12.78 corresponds to a median deposit rate elasticity of only 0.18. One might expect these numbers to understate the true depositor responsiveness to deposit interest rate changes. If higher quality banks (i.e., those with high mean valuations) tend to offer lower deposit rates, then the assumed orthogonality between the deposit rates and mean valuations fails and the estimated coefficient should be too small.[16]

[Table 5 about here.]

I employ two sets of instruments to help mitigate this potential endogeneity problem. The first set relies on the notion that loan-side conditions may affect deposit pricing (e.g., Kiser 2004) but have no direct effect on depositor valuations. In particular, I proxy the demand for funds and loan cost:

$$z_{1,jt} = \frac{\text{LOANS}_{jt}}{\text{GTA}_{jt}}, \quad \text{and} \quad z_{2,jt} = \frac{\text{CHG}_{jt}}{\text{LOANS}_{jt}}. \tag{11}$$

Banks with higher loans-to-assets ratios may be more likely seek additional deposits with

[15]The mixed logit estimation requires demographic information from the Consumer Population Survey (CPS). The information is available for a subset of the CBSAs in the full sample. For consistency, I use the subset to estimate the logit model. The results are robust to the use of the full sample.

[16]Indeed, Dick (2002) and Knittel and Stango (2007) estimate median interest rate elasticities for the commercial banking industry of 1.70 and 1.20, respectively. Elasticities below one in magnitude are often considered inconsistent with profit maximization.

which to fund loans, and banks with higher charge-offs may have greater lending costs and less incentive to attract additional deposits. Turning to the second set of instruments, under the maintained assumption that the non-deposit rate characteristics are exogenous, competitor non-price characteristics provide natural instruments. Similarly to Berry, Levinsohn and Pakes (1995), I calculate the sum of competitor characteristics and then average this sum across each bank's markets:

$$z_{3,jt} = \frac{1}{m_{jt}} \sum_{l \in M_{jt}} \left(\sum_{k \neq j,\, k \in J_{lt}} w_{klt} \right), \tag{12}$$

where w_{ijt} is a vector of bank observables that includes $\left(\frac{m_{jt}}{M} - 1\right)$ and x_{jmt}, M_{jt} is the set of markets in which bank j has branches, and J_{mt} is the set of banks with branches in market m. In Table 5, I refer to the two sets of instruments as "Loans" and "CC," respectively.[17]

I show the baseline 2SLS logit results in Column 2. As expected, the estimated deposit rate coefficient of 105.06 is much larger, and the implied median deposit rate elasticity is a more plausible 1.47. A statistical comparison to the OLS results, ala Hausman (1978), easily rejects the null that the deposit rate is exogenous to depositor mean valuations. The bank scope coefficient of 4.35 remains statistically different than zero and is consistent with depositors that prefer banks of greater scope. The coefficients imply an annual willingness-to-pay for a unit increase in bank scope of 43.70 cents. The number may be substantial. For example, it implies that depositors value the scope of Bank of America (branches in 207 CBSAs in 2006) at roughly 89.97 dollars more than that of a single-market bank. By way of comparison, the 2006 data suggest that the median account of 3,800 dollars would earn 33.86 fewer dollars at Bank of America than at the average single-market bank.[18]

The consistency of the logit estimation depends on 1) the validity of the instruments and 2) the exogeneity of the non-deposit rate characteristics. I examine these assumptions in the remaining columns. To start, Column 3 estimates the model using only the loan-side instruments. The results are quite similar to the baseline. If one is willing to assume that the loan-side instruments are valid, then a comparison to the baseline results yields a statistical

[17]The model does not provide intuition regarding the form of the relationship (if any) between the deposit interest rate and the instruments. I use third-order polynomials in each instrument to flexibility estimate the first stage.

[18]The results are robust to a number of alternative specifications. For example, the inclusion of market or year fixed effects (which limits identification derived from outside good market shares), the inclusion of the deposit fee rate as an additional endogenous regressor (i.e., fee income over deposits), and the use of the alternative outside good (calculated as population times a constant of proportionality less commercial bank deposits), do not substantially alter the results.

test for the validity of the competitor characteristic instruments (e.g., Hausman 1978; Ruud 2000). Intuitively, the Column 3 estimates are consistent given the exogeneity of the loan-side instruments. Under the null hypothesis that the competitor characteristics are also valid, the baseline coefficients should be similar. The data do not reject the null (p-value= 0.983). Column 4 estimates the model using only the competitor characteristic instruments. The results are again similar to the baseline, and the data fail to reject the null that the loan-side instruments are valid, given the validity of the competitor characteristic characteristics (p-value= 0.764). Together, the results provide some evidence that the instruments may be valid.

Columns 5 and 6 help evaluate the exogeneity of bank scope. One might expect the baseline estimated scope coefficient to overstate the true effect. If higher quality banks (i.e., banks for which depositors have higher mean valuations) tend to enter more markets, then the assumed orthogonality between scope and the mean valuations fails and the estimated coefficient should be too large. To help address this concern, I decompose the mean valuations into bank fixed effects and market-year specific valuations:

$$\xi_{jmt} = \xi_j + \Delta\xi_{jmt}, \tag{13}$$

and estimate the bank fixed effects directly.[19] The scope coefficient is then identified from changes in scope, i.e. off of bank entry. For illustration, consider a bank j that operates in CBSA m during period t and in both CBSAs m and n during period $t+1$. Two comparisons identify the scope coefficient. The first comparison is that of bank j's market share in CBSA m across periods. An increase in market share would suggest that scope provides value to depositors. The second comparison is that of bank j's market share in CBSA m during period t and bank j's market share in CBSA n during period $t+1$. A market share that is higher in the latter instance would again suggest that depositors value scope. The second comparison is confounding empirically, however, because bank entrants tend to have small market shares initially (e.g., Berger and Dick 2007), due to switching costs and/or other factors.

Column 5 presents the results of the bank fixed effects specification. The bank scope

[19]The procedure yields consistent estimates provided that a bank's quality is fixed across markets. In the implementation, I estimate separate effects for the 1,212 banks that have branches in at least two CBSAs at some point during the sample, and estimate a shared effect for the remaining banks. The restriction greatly eases the computational burden, and the procedure should still mitigate bias due to scope endogeneity. The parameters of interest are not well-identified empirically when a separate fixed effect is estimated for each single-market bank because the effects overtax the data (there are six observations per single-market bank).

coefficient of 2.31 is more than 45 percent smaller than the baseline coefficient and is no longer statistically different than zero.[20] However, it is not clear whether the reduction is due to the correction of endogeneity bias or due to the confounding market shares of recent entrants. To address the matter, I add lagged market share to the specification, which controls for the out-of-equilibrium effects associated with entry. Column 6 presents the results. The bank scope coefficient of 4.39 is larger, close to the baseline coefficient, and statistically significant. Together, the Column 5 and 6 results suggest that bank scope endogeneity may be unimportant in the baseline specification. More generally, the results shown in Columns 3 through 6 provide empirical support for the identification strategy.

I now return to the mixed logit case, in which depositors are permitted to have heterogenous tastes for observables. Estimation requires numerical integration over demographic characteristics (as in Equation 9). I let income have the lognormal distribution within CBSAs and estimate the income distributions using individual-level data from the 2000 CPS March Supplement. I then draw 200 quasi-random incomes for each CBSA from the appropriate estimated distribution. I proxy the unobserved demographics with 200 quasi-random draws from a standard normal distribution. To ease interpretation, I normalize income to have mean zero and variance one across all CBSAs, following Nakamura (2006).[21]

With the numerical integration in hand, I am able to obtain the vector of mean depositor valuations, ξ, for any set of proposed parameters $\theta = (\beta^D, \beta^M, \beta^X, \pi, \sigma)$ via a contraction mapping algorithm (e.g., Berry 1994; Berry, Levinsohn and Pakes 1995, Nevo 2001). The parameters are then identified by the usual assumption, namely that $\mathrm{E}[Z'\xi(\theta^*)] = 0$, where Z includes the instruments and θ^* includes the true population parameters. The simulated generalized method of moments (SGMM) estimator takes the form:

$$\widehat{\theta}_{SGMM} = \arg\min_{\theta\in\Theta} \xi(\theta)' Z \Phi^{-1} Z' \xi(\theta), \tag{14}$$

where Φ^{-1} is a consistent estimate of $\mathrm{E}[Z'\xi(\theta^*)\xi(\theta^*)'Z]$. I compute the standard errors with the usual formulas (e.g., Hansen 1982; Newey and McFadden 1994).[22]

[20] The data reject the null that the bank fixed effects are jointly zero (p-value= 0.00).

[21] CPS data are available for 259 CBSAs. Estimation uses 38,993 of 45,785 bank-CBSA-year observations. I use Halton numbers to capture income and unobserved demographics. Train (1999) and Bhat (2001) find that the simulation variance caused by 100 Halton quasi-random numbers is smaller than the simulation variance caused by 1,000 random draws. In estimating the CBSA-specific income distributions, I consider the household income of individuals over 18 years in age. I trim the bottom 1.5 percent of the sample to eliminate negative and zero incomes, which are incompatible with the lognormal assumption.

[22] I follow the procedure outlined in Nevo (2000) to perform the contraction mapping and estimation. I apply a clustering correction that allows for the consistent estimation of the standard errors in the presence of arbitrary correlation patterns between observations from the same bank.

Table 6 presents the results of the mixed logit regression. The mean depositor valuations (β's), shown in the first column, are similar to those of the baseline logit results. The mean valuations for the deposit interest rate and bank scope are 115.92 and 4.70, respectively, and the implied mean willingness-to-pay for a unit increase in bank scope is 42.92 cents. Again, this number may be substantial, as it implies that the mean depositor values the scope of Bank of America (branches in 207 CBSAs in 2006) at roughly 88.41 dollars more than that of a single-market bank. The results also suggest that depositors may prefer banks with greater branch densities and more employees per branch, though the first coefficient is smaller in magnitude (vis-a-vis the logit results) and the second is no longer statistically significant.

[Table 6 about here.]

The next columns show estimates of depositor heterogeneity around the mean deposit rate valuation. The estimated deposit rate parameter standard deviation (σ) is small and not statistically significant. By contrast, the coefficient on the interaction with depositor income (π) is large and statistically different than zero. The coefficient suggests that a one standard deviation increase in income lowers the deposit rate valuation by 65.58, so that higher income depositors are less price-sensitive. A joint statistical test rejects the null of no heterogeneity (i.e., $\sigma = \pi = 0$) at any conventional level. Figure 1 graphs the estimated willingness-to-pay distribution.[23] The skew of the distribution reflects the result that the lognormally distributed depositor incomes matter more to deposit rate valuations than the normally distributed unobserved demographics. The 25th, 50th, and 75th percentiles of the empirical distribution are 31.30, 35.22, and 44.56 cents, respectively, so that a depositor at the 75th percentile has a willingness-to-pay that is 42.34 percent greater than a depositor at the 25th percentile. Overall, the results are strongly consistent with the notion that depositor heterogeneity exists and is substantial.

[Figure 1 about here.]

4.3 Tests of the Empirical Predictions

4.3.1 The decision to enter an outside market

The theoretical model predicts that a bank should be less likely to enter an outside market if its original markets already feature banks of greater scope. Intuitively, the presence of

[23] Roughly five percent of depositors have deposit rate valuations near or below zero. The corresponding willingness-to-pay for these depositor is quite large (or quite small) and is not shown.

larger banks creates an environment in which entry would lessen scope differentiation and intensify deposit rate competition. I examine the empirical prediction using data on the entry decisions of single-market banks. Within this specific context, the empirical prediction can be precisely formulated. Suppose a single-market bank operates in market m. Then the single-market bank should be less likely to enter any second market if more two-market banks already operate in market m.

To test the empirical prediction, I construct a regression sample of single-market banks over the period 2001-2005. The 14,469 bank-year observations in the sample include 3,658 single-market banks that, combined, operate in 191 CBSAs.[24] Of these single-market banks, 351 enter a second market during the sample period. Banks that enter a second market drop out of the sample after entry. The estimation procedure is standard logit model. Let $j = 1, 2, \ldots, J$ single-market banks, each based in an original market m_j, determine whether to enter a second market in period $t+1$. I represent the latent utility of entry and the observed entry choice as

$$v^*_{j,t+1} = \text{NBANKS}'_{mt}\lambda_1 + w'_{jt}\lambda_2 + \gamma_{mt} + \eta_{jt} \quad \text{and} \quad v_{j,t+1} = 1\{v^*_{j,t+1} > 0\}, \tag{15}$$

respectively. The vector NBANKS_{mt} captures the number of banks in market m during period t that have branches in exactly one, exactly two, three through ten, and more than ten CBSAs. I refer to these four variables as NBK1_{mt}, NBK2_{mt}, NBK3_{mt} and NBK4_{mt}, respectively. The primary variable of interest is NBK2_{mt}; the theoretical model implies that its coefficient should be negative. The vector w_{jt} includes controls at the bank-year level and the vector γ_{mt} includes market and year fixed effects. The parameter vectors λ_1 and λ_2 can be consistently estimated with standard logit regression under the assumption that the error term η_{jt} has the extreme value type I density.[25]

Before turning to the results, some discussion of the specification may be fruitful. First, market fixed effects solve the basic identification problem that markets may differ in their ability to support banks of greater scope (potentially for unobservable reasons). Such market

[24] The remaining CBSAs never have a single-market bank enter a second market. I exclude the corresponding 3,744 bank-year observations from the regression sample. The use of market fixed effects makes the inclusion/exclusion of these observations irrelevant for estimation because the fixed effect parameters perfectly predict observed entry.

[25] Two econometric points may be of interest. First, the estimation of market and year fixed effects within the limited dependent variable framework does not introduce incidental parameters bias because the ratio of observations to parameters converges to infinity with the number of banks per market-year. Overall, I estimate 201 parameters using 14,469 observations. Second, I cluster the standard errors at the market level to account for arbitrary correlation patterns among bank-year observations in the same market.

heterogeneity, if unaccounted for in the regression specification, would bias the NBK2_{mt} coefficient upwards and against the empirical prediction. The upward bias occurs because markets that better support banks of greater scope are likely to have more banks of greater scope; further, single-market banks in these markets may find entry into an outside market more profitable. The inclusion of market fixed effects controls directly for this confounding influence. Second, the bank-year control variables include the deposit and branch market shares of the single-market branch in its original market. One might expect banks that have more substantial market shares in their original markets to be more likely to enter a second market.

The first column of Table 7 presents the baseline results. The NBK2_{mt} coefficient is negative and statistically different than zero, consistent with the theoretical model. The coefficient is also substantial in magnitude. For example, a hypothetical one standard deviation increase in NBK2_{mt} reduces the probability that a single-market bank enters a second market by an average of 76.54 percent.[26] Interestingly, the entry choices of single-market banks appear uncorrelated with the number of banks in the original market that have branches in *more* than two markets (the corresponding regression coefficients are small and not statistically different than zero). One might infer that differentiation in geographic scope has quickly diminishing returns. Turning quickly to the control variables, single-market banks that have larger deposit and branch market shares in their original markets are more likely to enter a second market, though only the branch market share coefficient is statistically significant (results not shown). Overall, the regression results strongly support the prediction that banks should be less likely to enter an outside market if their original markets already feature banks of greater scope. The results also underscore the empirical importance of this prediction.

[Table 7 about here.]

At the risk of digression, Columns 2 through 4 show the results when the market fixed effects and/or the bank control variables are excluded from the specification. As discussed above, one might expect the estimated NBK2_{jt} coefficient to be less negative (or even posi-

[26]The mean number of two-market banks in the regression sample is 8.14. The standard deviation is 6.39. To measure the average percent change in entry probabilities, I first calculate the probability of entry for each bank-year observation. This probability, call it p_{jt1}, has the simple logit expression $p_{jt1} = \exp(x'_{jt}\widehat{\lambda})/\left(1+\exp(x'_{jt}\widehat{\lambda})\right)$, where the vectors x_{jt} and $\widehat{\lambda}$ include the regressors and estimated coefficients, respectively. I then calculate the probability of entry for each observation given a hypothetical increase in the number of two-markets banks present in the original market. This probability, call it p_{jt2}, has an expression analogous to that of p_{jt1}. The average percent change is then simply $\frac{1}{JT}\sum_{j=1}^{J}\sum_{t=1}^{T}\left(p_{jt2}-p_{jt1}\right)/p_{jt1}$.

tive) when market fixed effects are excluded, due to the confounding influence of unobserved market heterogeneity. This is precisely what happens. Column 2 omits the market fixed effects; the specification is otherwise identical to that of Column 1. The resulting NBK2_{jt} coefficient is nearly zero. The coefficients on NBK1_{jt}, NBK3_{jt}, and NBK4_{jt} also change in the predicted directions. Finally, Column 3 omits only the bank control variables and Column 4 omits both the market fixed effects and the bank control variables. In each case, the exclusion of the bank controls has little effect on the estimated regression coefficients. One might conclude that bank heterogeneity has little effect on the main results.

4.3.2 The choice of which outside market to enter

Any bank that enters an outside market must decide which outside market to enter. The second empirical prediction deals with this choice between outside markets. In particular, the theoretical model predicts that a bank with branches in market m should be less likely to enter market n (conditional on entry somewhere) if another bank of similar scope exists with branches in both market m and market n. I again turn to the entry decisions of single-market banks to provide a specific context in which to test the empirical prediction.

To motivate the empirical strategy, consider the following simplified setting: a single-market bank (bank j) is based in an original market (m_j) and enters one of two outside markets (n_1 or n_2). Suppose that a two-market bank already exists with branches in m_j and n_1. Then the theoretical model suggests that bank j is more likely to enter n_2 because it offers greater scope differentiation. One could test the hypothesis by regressing the observed entry decision on an indicator variable, call it 2CBSA_{jn}, that equals 1 if the two-market bank exists in markets m_j and n, and 0 otherwise. Provided that the two-market bank is randomly assigned between n_1 and n_2, the regression coefficient provides a consistent test of the empirical prediction. Of course, this condition is unlikely to hold in practice. The two-market bank may operate in m_j and n_1 because n_1 is more profitable than n_2, and/or because greater synergies exist between m_j and n_1 than between m_j and n_2. Both possibilities threaten to bias the regression coefficient upwards, i.e., against the empirical prediction, and I control directly for these potentially confounding factors.

The actual empirical model generalizes the simplified setting to allow for many banks and many outside markets. Let $j = 1, 2, \ldots, J$ single-market banks, based in the original markets m_j, determine which of $n = 1, 2, \ldots, N$ outside markets to enter. Each single-market banks enters the outside market that provides the greatest value:

$$v_{jn} = \lambda_1 2\text{CBSA}_{jn} + \lambda_2 \text{SYNERGY}_{jn} + \text{PROFITS}_n' \lambda_3 + \eta_{jn}, \tag{16}$$

where the scalar 2CBSA_{jn} equals one if a two-market bank already exists with branches in markets m_j and n, and zero otherwise, the scalar SYNERGY_{jn} represents the synergies between the markets m_j and n, and the vector PROFITS_n captures the profitability of market n. The parameters λ_1, λ_2, and λ_3 can be consistently estimated with conditional logit regression (e.g., Greene 2003, Section 19.7) under the assumption that the bank-specific error η_{jn} has the extreme value type I density. The theoretical model implies that $\lambda_1 < 0$.

I take the empirical model to the 351 instances in which a single-market bank entered an additional CBSA over the period 2001-2005. The regression observations are combinations of single-market banks and outside markets: the 351 single-market banks and 359 outside CBSAs form 126,009 regression observations. The dependent variable equals one if the single-market bank entered the outside market, and zero otherwise. The independent variable of interest, 2CBSA_{jn} is directly observable. I proxy the synergies between the home and outside markets using the proportion of all banks in the original market that also have branches in the respective outside market. To proxy the profitability of the outside markets, I include the number of outside market banks with branches in exactly one, exactly two, three through ten, and more than ten CBSAs. I also include second-order polynomials in median income, population, and land area. I lag the synergy and profitability controls to mitigate any potential endogeneity concerns.

The first column of Table 8 presents the baseline results. The 2CBSA_{jn} coefficient is negative and statistically different than zero, consistent with the empirical prediction of the theoretical model. The coefficient is sizable in magnitude and suggests that a single-market bank is on average 35 percent less likely to enter market n if a two-market bank already exists in m_j and n.[27] This may actually understate the true effect: to the extent that the control variables imperfectly measure synergies and profits, the true 2CBSA_{jn} parameter is likely more negative than the estimated coefficient. Overall, the regression result strongly

[27] To be clear, I calculate the probability that bank j enters each outside market n, alternately toggling 2CBSA_{jn} to be 0 and 1, and holding the remaining regressors constant. The percent change in probability that bank j enters the outside market n due to the presence of a two-market bank in markets m_j and n is

$$\frac{Pr(n|\, j, 2\text{CBSA}_{jn} = 1) - Pr(n|\, j, 2\text{CBSA}_{jn} = 0)}{Pr(n|\, j, 2\text{CBSA}_{jn} = 0)}.$$

The probabilities have the familiar logit closed form solutions. I report the average percent change among the regression observations.

supports the empirical prediction of the theoretical model.

[Table 8 about here.]

Again at the risk of digression, Columns 2 through 4 show the results when the synergy control and/or the profit controls are excluded from the specification. As discussed above, one might expect the estimated $2CBSA_{jn}$ coefficient to be less negative (or even positive) due to omitted variables bias in each of these regressions. That is precisely what happens. Column 2 omits the controls that proxy the outside market profitability. The resulting $2CBSA_{jn}$ coefficient remains negative but is smaller in magnitude (-0.241) and not statistically different than zero. Column 3 omits the synergy controls, and Column 4 omits both sets of controls. The resulting $2CBSA_{jn}$ coefficient is positive and large in both cases. These alternative regressions may demonstrate the importance of controlling for market synergies and other factors when testing or estimating scope effects.

4.3.3 Market nearness and the number of banks

Finally, the theoretical model generates the empirical prediction that the number of banks within a given market should increase with its "nearness" to other markets. To examine the prediction empirically, I proxy nearness inversely with the mean distance in miles between a CBSA and all other CBSAs.[28] The average CBSA has a mean distance of 1060 miles, and the sample standard deviation is 349 miles.

The empirical prediction holds in the data. Figure 2 plots the univariate relationship between CBSA nearness and the number of banks for all 2,160 CBSA-year observations over the period 2001-2006. The three CBSAs with the greatest mean distance (Honolulu, Anchorage, and Fairbanks) also have very few banks (e.g., 8, 5, and 5 banks, respectively, in 2006). The CBSA with the most banks (Chicago-Nashville-Jolie, with 231 banks in 2006) has one of the shortest mean distances. Among all CBSAs, the correlation coefficient between the number of banks and mean distance is -0.11. Further, an OLS regression of the number of commercial banks per CBSA-year observation on mean distance and second-order control polynomials in CBSA median household income, population, and land area yields a mean distance coefficient of -0.014 that is statistically different than zero at standard levels (standard error = 0.003).[29] The regression coefficient is also substantial in magnitude: a one

[28]Formally, the proxy is $\frac{1}{M}\sum_{n\neq m}\text{DIST}_{mn}$, where DIST_{mn} is the miles between CBSA m and CBSA n.

[29]I cluster the standard errors at the CBSA level to account for heteroscedasticity and arbitrary correlation patterns between observations from the same CBSA.

standard deviation increase in mean distance corresponds to a 24 percent reduction in the number of banks when evaluated at the mean.

[Figure 2 about here.]

More detailed investigation calls into question the extent to which the relationship between CBSA nearness and the number of banks supports the theoretical model, however. The empirical prediction is generated by the theoretical model because markets near many others may offer greater opportunities for scope differentiation. One would naturally expect the relationship between nearness and the number of banks to be driven by differences in the number of multimarket banks rather than the number of single-market banks because, by definition, single-market banks do not exploit opportunities for scope differentiation. The opposite holds in the data – the relationship between CBSA nearness and the number of banks appears to be driven primarily by differences in the number of single-market banks.

Figure 3 shows a scatterplot of CBSA nearness and the number of single-market banks for the 2,160 CBSA-year observations over the period 2001-2006. It is clear that, on average, CBSAs that are near others support more single-market banks. By way of contrast, Figure 4 shows the scatterplot of CBSA nearness and the number of multimarket banks. No relationship is apparent. Further, an OLS regression of the number of multimarket banks on mean distance and second-order control polynomials in median household income, population, and land area yields a mean distance coefficient of -0.00003 that is small in magnitude and not statistically different than zero (standard error = 0.0007). Thus, while the empirical prediction of the theoretical model holds in the data – CBSAs near many others do support more banks – it is not apparent that the prediction holds because CBSAs in close proximity to many others offer greater opportunities for scope differentiation. It is therefore difficult to conclude that the relationship between CBSA nearness and the number of banks provides substantive support for the theoretical model.

[Figure 3 about here.]

[Figure 4 about here.]

5 Conclusion

I argue that the observed distribution of commercial bank scopes may arise, in part, due to depositor heterogeneity. From a theoretical standpoint, bank scope diversity arises naturally when depositors prefer banks of greater scope but differ in their willingness-to-pay. Structural estimation, based on a model of depositor choice, suggests that depositors value scope and that substantial heterogeneity in willingness-to-pay exists. Finally, reduced-form tests support three distinct empirical predictions of the theoretical model. Overall, the theoretical and empirical results build the case that depositor heterogeneity is an important determinant of the bank scope distribution. The results are compatible with other theories, in the sense that the results do not rule out other determinants of bank scope. Thus, the relative importance of depositor heterogeneity remains to be explored and could be the subject of further research. The estimation of structural models may help fill this gap, and recent advances in estimation techniques may help make such endeavors feasible (e.g., Weintraub, Benkard and Van Roy 2007). The work of Ishii (2004) on ATM network supply may represent a first step towards a suitable modeling framework.

References

[1] Adams, Robert M., Kenneth P. Brevoort, and Elizabeth K. Kiser. 2007. Who competes with whom: the case of the depository institutions. *Journal of Industrial Economics*, 55: 141-167.

[2] Amel, Dean F. and Martha Starr-McCluer. 2001. Market definition in banking: Recent evidence. *Journal of Banking and Finance*, 23: 1667-1690.

[3] Bajari, Patrick, Jeremy T. Fox, and Stephen Ryan. 2006. Evaluating wireless carrier consolidation using semiparametric demand estimation. Mimeo.

[4] Berger, Allen N. 2003. The economic effects of technological progress: evidence from the banking industry. *Journal of Money, Credit, and Banking*, 35: 141-176.

[5] Berger, Allen N. and Lamont K. Black. 2007. Bank size and small business finance: tests of the current paradigm. Mimeo.

[6] Berger, Allen N. and Astrid Dick. 2007. Entry into banking markets and the early-mover advantage. *Journal of Money, Credit, and Banking*, 39: 775-807.

[7] Berger, Allen N., Anil K. Kashyap, and Joseph M. Scalise. 1995. The transformation of the U.S. banking industry: what a long, strange trip it's been. *Brookings Papers on Economic Activity*, 55-201.

[8] Berger, Allen N. and Loretta J. Mester. 2003. Explaining the dramatic changes in performance of US banks: technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation*, 12: 57-95.

[9] Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein. 2005. Does function follow organizational form: evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76: 237-269.

[10] Berger, Allen N. and Gregory F. Udell. 2002. Small business credit availability and relationship lending: the importance of bank organisational structure. *Economic Journal*, 112: 32-53.

[11] Berry, Steven T. 1994. Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25: 242-262.

[12] Berry, Steven T., James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica*, 63: 841-890.

[13] Bhat, Chandra R. 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research, Part B*, 35: 677-693.

[14] Brady, Thomas F. and William F. Bassett. 2002. What drives the persistent competitiveness of small banks? FEDS Working Paper No. 2002-28.

[15] Bucks, Brian K., Arthur B. Kennickell, and Kevin B. Moore. 2006. Recent Changes in U.S. Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances. *Federal Reserve Bulletin*, 92: A1-A38.

[16] Choi, Chong Ju, and Hyun Song Shin. 1992. A comment on a model of vertical product differentiation. *Journal of Industrial Economics*, 40: 229-231.

[17] Cole, Rebel A., Lawrence G. Goldberg and Lawrence J. White. 2004. Cookie cutter vs. character: the micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis*, 39: 227-251.

[18] Dick, Astrid A. 2002. Demand estimation and consumer welfare in the banking industry. *Finance and Economics Discussion Series*, Federal Reserve Board, 2002-58.

[19] Dick, Astrid A. Forthcoming. Market size, service quality, and competition in banking. *Journal of Money, Credit and Banking.*

[20] Donnenfeld, Shabtai and Shlomo Weber. 1992. Vertical product differentiation with entry. *International Journal of Industrial Organization*, 10: 449-472.

[21] Gabszewicz, Jaskold J., Avner Shaked, John Sutton, and J. F. Thisse. 1981. Price competition among differentiated products: a detailed study of a Nash equilibrium. ICERD Discussion Paper No. 37, London School of Economics.

[22] Greene, William H. (2003). *Econometric Analysis*, 5th ed. Prentice-Hall, Upper Saddle River, NJ.

[23] Grossman, Sanford J. and Oliver D. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94: 691-719.

[24] Hart, Oliver D. 1995. *Firms, Contracts, and Financial Structure*. Oxford University Press, Oxford.

[25] Hart, Oliver D. and John Moore. 1990. Property rights and the nature of the firm. *Journal of Political Economy*, 98: 1119-1158.

[26] Hannan, Timothy H. and Robin A. Prager. 2004. Multimarket bank pricing: an empirical investigation of deposit interest rates. *Journal of Economics and Business*, 58: 256-272.

[27] Hannan, Timothy H. and Robin A. Prager. 2005. Retail deposit fees and multimarket banking. *Journal of Banking and Finance*, 30: 2561-2578.

[28] Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica*, 50: 1029-1054.

[29] Hausman, Jerry. 1978. Specification tests in econometrics. *Econometrica*, 46: 1251-1271.

[30] Heitfield, Eric A. 1999. What do interest rate data say about the geography of retail banking markets? *Antitrust Bulletin* 44: 333-347.

[31] Heitfield, Eric A. and Robin A. Prager. 2004. The geographic scope of retail deposit markets. *Journal of Financial Services Research*, 25: 37-55.

[32] Ishii, Joy. 2004. Interconnection pricing and compatibility in network industries: ATM networks in the banking industry. Working paper, Stanford University Graduate School of Business.

[33] Kiser, Elizabeth K. 2002a. Predicting household switching behavior and switching costs at depository institutions. *Review of Industrial Organization*, 20: 349-365.

[34] Kiser, Elizabeth K. 2002b. Household switching behavior at depository institutions: evidence from survey data. Working Paper, Federal Reserve Board.

[35] Kiser, Elizabeth K. 2004. Modeling the whole firm: the effects of multiple inputs and financial intermediation on bank deposit rates. Working Paper, Federal Reserve Board.

[36] Klein, Benjamin, Robert C. Crawford, and, Armen A. Alchian. 1978. Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, 21: 297-326.

[37] Knittel, Christopher R. and Victor Stango. 2007. Incompatibility, product attributes and consumer welfare: evidence from ATMs. Mimeo.

[38] Kwast, Myron L., Martha Starr-McCluer, and John D. Wolken. 1997. Market definition and the analysis of antitrust in banking. *The Antritrust Bulletin*, 42: 973-995.

[39] Lehmann-Grube, Ulrich. 1997. Strategic choice of quality when quality is costly: The persistence of the high-quality advantage. *RAND Journal of Economics*, 28: 372-384.

[40] Liberti, Jose and Atif Mian. 2006. Estimating the effect of heirarchies on information use. Mimeo.

[41] McFadden, Daniel and Kenneth Train. 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15: 447-470.

[42] Motta, Massimo. 1993. Endogenous quality choice: price vs. quantity competition. *Journal of Industrial Economics*, 41: 113-131.

[43] Nakamura, Emi. 2007. Accounting for incomplete pass-through. Mimeo.

[44] Nevo, Aviv. 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy*, 9: 513-548.

[45] Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69: 307-342.

[46] Newey, Whitney K. and Daniel L. McFadden. 1994. Estimation in large samples. *The Handbook of Econometrics*, Vol. 4, ed. by Daniel L. McFadden and Robert F. Engle. Amsterdam: North-Holland.

[47] Park, Kwangwoo and George Pennacchi. 2004. Harming depositors and helping borrowers: the disparate impact of bank consolidation. *Proceedings of the 40th Annual Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago.

[48] Radecki, Lawrence J. 1998. The expanding geographic reach of retail banking markets. *Federal Reserve Bank of New York Economic Policy Review*, 15-33.

[49] Ruud, Paul A. 2000. *An Introduction to Classical Econometric Theory.* Oxford University Press, New York, NY.

[50] Selton, Reinhard. 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4: 25-55.

[51] Sharpe, Stephen A. 1997. The effect of consumer switching costs on prices: a theory and its application to the bank deposit market. *Review of Industrial Organization*, 12: 79-94.

[52] Shaked, Avner and John Sutton. 1982. Relaxing price competition through product differentiation. *Review of Economic Studies*, 49: 3-13.

[53] Shaked, Avner and John Sutton. 1983. Natural oligopolies. *Econometrica*, 51: 1469-1484.

[54] Shaked, Avner and John Sutton. 1984. Natural oligoloies and international trade. *Monopolistic Competition and International Trade*, ed. by H. Kierzkowski. Oxford: Clarendon Press.

[55] Stein, Jeremy. 2002. Information production and capital allocation: Decentralized versus heirarchical firms. *Journal of Finance*, 67: 1891-1921.

[56] Tirole, Jean. 1988. *The Theory of Industrial Organization*, MIT Press, Cambridge.

[57] Train, Kenneth. 1999. Halton sequences for mixed logit. Working paper, University of California, Berkeley.

[58] Wauthy, Xavier Y. 1996. Quality choice in models of vertical differentiation. *Journal of Industrial Economics*, 44: 345-353.

[59] Weintraub, Gabriel Y., C. Lanier Benkard, and Benjamin Van Roy. 2007. Markov perfect industry dynamics with many firms. Mimeo.

[60] Williamson, Oliver E. 1975. *Markets and Heirarchies: Analysis and Antitrust Implications*. Collier Macmillan, New York.

[61] Williamson, Oliver E. 1979. Transaction-cost economics: The governance of contractual relations. *Journal of Law and Economics*, 22: 233-61.

[62] Williamson, Oliver E. 1985. *The Economics Institutions of Capitalism*. Free Press, New York.

A Proofs

Proof of Lemma 1. The first order conditions for bank j and bank k simplify to $r_j^D = (1/2) * (r^L + r_k^D)$ and $r_k^D = (1/2) * (r^L + r_j^D)$, respectively, when $m_j = m_k$ for any j. Solving the conditions yields $r_j^D = r_k^D = r^L$. It follows that $\pi_j = \pi_k = 0$. Finally, by Equation 3, depositors strictly prefer banks j and k to any bank l characterized by $m_l < m_j = m_k$ at the deposit rate r^L, so the market share and profit of bank l are zero. □

Proof of Lemma 2. By Equation 3, expected utility is increasing in deposit rate and bank scope. Deposits prefer bank j to bank k, and bank k has zero market share. □

Proof of Proposition 1. I start with the claim that deposits rate decrease in scope, i.e., $r_j^D > r_{j+1}^D$ for any $j < J$. The first order conditions for profit maximization imply that $r_j^D > (1/2) * (\frac{m_{j+1}-m_j}{m_{j+1}-m_{j-1}} r_{j-1}^D + \frac{m_j - m_{j-1}}{m_{j+1}-m_{j-1}} r_{j+1}^D)$ for $1 < j < J$. Equilibrium must therefore be characterized either by deposit rates that decrease strictly in scope, i.e. $r_1^D > r_2^D > \cdots > R_J^D$ or by deposit rates that increase strictly in scope, i.e. $r_1^D < r_2^D < \cdots < R_J^D$. But the latter cannot characterize equilibrium. For example, if $r_{J-1}^D < r_J^D$ then bank $J-1$ has zero market share by Lemma 2 and a sufficient deposit rate increase restores positive market share; such an action is infeasible only when $r_J^D = r^L$, in which case bank J earns zero profits and prefers to lower its deposit rate. Thus, it must be true that $r_1^D > r_2^D > \cdots > R_J^D$ in any equilibrium.

I now turn to the claim that market shares and profits increase in scope, i.e. $s_j < s_{j+1}$ and $\pi_j < \pi_{j+1}$ for any $j < J = M + 1$ with $m_j = j - 1$. The proof is by contradiction. Suppose that $s_j > s_{j+1}$. Then, by the definition of market share, $\alpha_{j+1} - \alpha_j > \alpha_j - \alpha_{j-1}$, and noting that $m_k - m_{k-1} = 1$ for any k,

$$2r_j^D - r_{j-1}^D - r_{j+1}^D > 2r_{j+1}^D - r_j^D - r_{j+2}^D,$$

from Equation 5. Rearranging terms yields

$$r_j^D - r_{j+1}^D > \frac{1}{2}(r_{j-1}^D + r_{j+1}^D) - \frac{1}{2}(r_j^D + r_{j+2}^D),$$

and, substituting the first order conditions specified in Equation 6 into the left hand side it must be that

$$\frac{1}{2}(r^L + \frac{1}{2}(r_{j-1}^D + r_{j+1}^D)) - \frac{1}{2}(r^L + \frac{1}{2}(r_j^D + r_{j+2}^D)) > \frac{1}{2}(r_{j-1}^D + r_{j+1}^D) - \frac{1}{2}(r_j^D + r_{j+2}^D).$$

Rearranging and canceling yields the simple condition $r_j + r_{j+2} > r_{j-1} + r_{j+1}$, which contradicts the finding presented above that $r_{j-1} > r_j$ and $r_{j+1} > r_{j+2}$. The claim that profits increase in scope is trivially true from Equation 1, given that deposit rates decrease in scope and market shares increase in scope. □

Proof of Lemma 3. Since in equilibrium $r_1^D > r_2^D > \cdots > R_J^D$ (Lemma 3) and $r_j^D > (1/2) * (\frac{m_{j+1}-m_j}{m_{j+1}-m_{j-1}} r_{j-1}^D + \frac{m_j - m_{j-1}}{m_{j+1}-m_{j-1}} r_{j+1}^D)$ for $1 < j < J$, it must be that $\alpha_{j+1} > \alpha_j$ by Equation 5, and the market share of bank j is positive. Turning to banks 1 and J, the equilibrium characteristics $r_1^D > r_2^D$ and $r_{J-1}^D > r_J^D$ imply that $\alpha_2 > 0$ and $\alpha_J < 1$, so banks 1 and J have positive market shares. The condition $r_1^D > r_2^D$ necessarily implies $r^L > r_1^D$, by Equation 6. Therefore, $r^L > r_j^D$ and $s_j > 0$ for all j, and there exists at least one stage game Nash equilibrium in which all banks have positive profits. Finally, rearranging the first order conditions such that $Ax = b$, where A is a matrix of coefficients, x is a vector of deposit rates, and b is a vector of solutions, it is apparent by inspection that A is nonsingular for any J. Applying Cramer's Rule, the first order conditions generate a unique stage game Nash equilibrium. □

Proof of Lemma 4. If $J = M + 1$ then the two stage subgame has a unique subgame perfect equilibrium in which each bank differs in scope. Suppose that each bank enters a different number of outside markets. By Lemma 3, the third stage features a unique Nash equilibrium in which each bank earns positive profits. If a bank deviates in the second stage then it has a scope that is equal to that of exactly one other bank and, by Lemma 1, earns zero profits in the third stage. The strategy profile in which each bank differs in scope is therefore subgame perfect. Next, suppose that some banks choose to be of equal scope. These banks earn zero profits by Lemma 1, but have a profitable deviation available in the second stage. The strategy profile in which each bank differs in scope is therefore the unique subgame perfect equilibrium.

If instead $J < M + 1$ then the two stage subgame has a unique class of subgame equilibria in which each bank differs in scope. Suppose that each bank enters a different number of outside markets. By Lemma 3, the third stage features a unique Nash equilibrium in which each bank earns positive profits. The class of strategy profiles in which each bank differs in scope is therefore subgame perfect. Uniqueness follows the reasoning of the $J = M + 1$ case. □

Proof of Lemma 5. Because $J > M + 1$, there must exist at least one bank j such that $m_j = m_k$ for some bank k. By Lemma 1, banks j and k, as well as any banks of lesser scope, earn zero profits in the third stage Nash equilibrium. If the scope space is covered, i.e., there is at least one bank with each possible scope, then bank j has no profitable deviation in the second stage. The class of equilibria is unique: if the scope space is not covered then bank j has a profitable deviation in which is selects the unoccupied scope in the second stage. □

Proof of Proposition 2. Suppose that $M + 1$ banks enter the inside market. Then, by Proposition 1 and Lemma 2, these banks subsequently enter different numbers of outside markets and earn positive profits in the third stage. The outcome is subgame perfect because no bank has a profitable deviation in any subgame. Next, suppose that $J > M + 1$ banks enter the inside market. By Lemma 4, at least one bank must earn zero profits in third stage; and, given the entry cost ϵ, this bank prefers not to enter the inside market. Lastly, if $J < M + 1$ banks enter the inside market then at least one non-entrant prefers to enter the inside market. The subgame perfect equilibrium is therefore unique. □

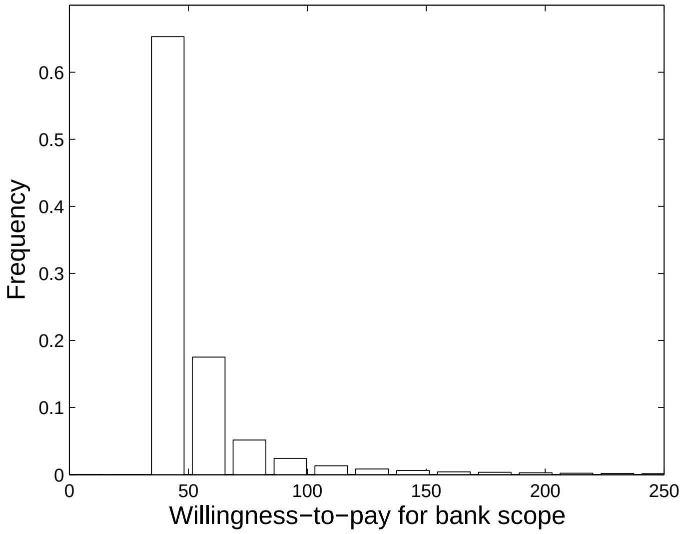

Figure 1: The frequency distribution of the willingness-to-pay for a unit increase in bank scope (based on Table 6). Willingness-to-pay is measured in cents per year.

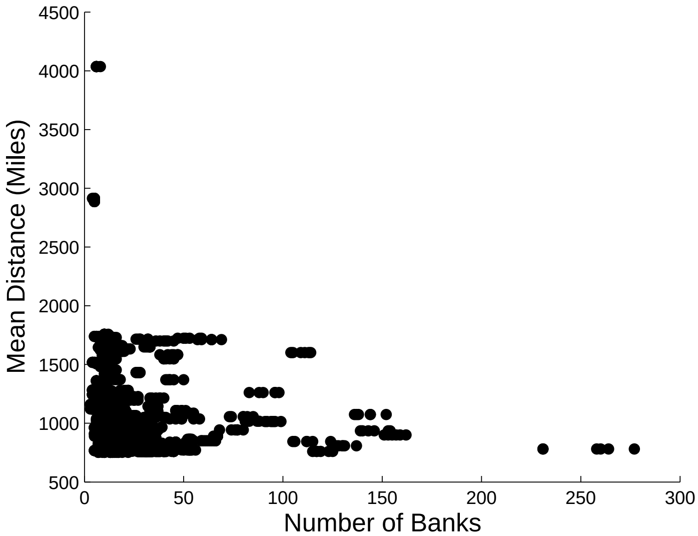

Figure 2: Market nearness and the number of banks. Each point represents a single CBSA-year observation. The vertical axis is the average distance in miles between the CBSA and all other CBSAs. The horizontal axis is the number of banks in the CBSA-year.

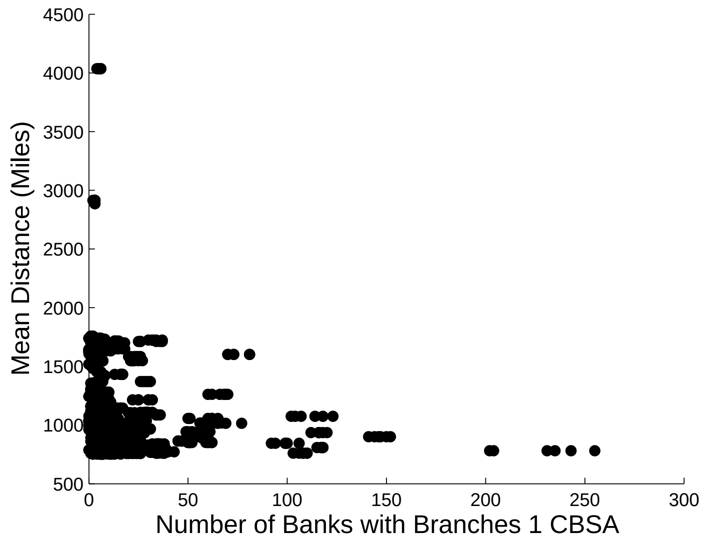

Figure 3: Market nearness and the number of banks with branches in only one CBSA. Each point represents a single CBSA-year observation. The vertical axis is the average distance in miles between the CBSA and all other CBSAs. The horizontal axis is the number of banks in the CBSA-year that have branches only one CBSA.

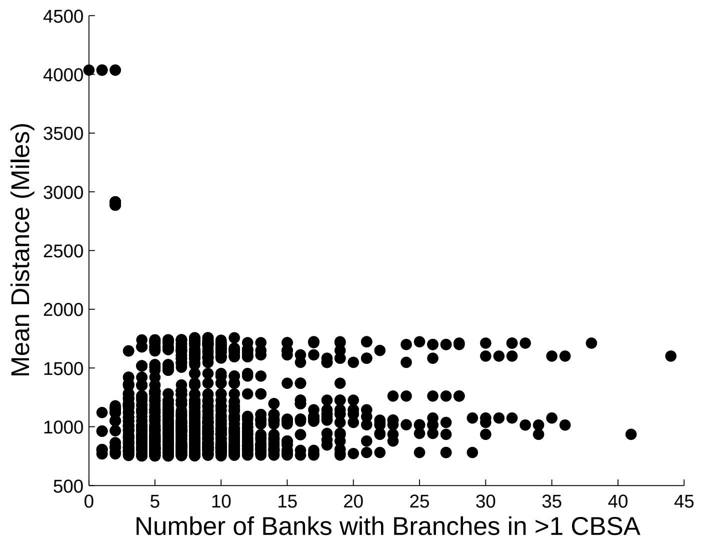

Figure 4: Market nearness and the number of banks with branches in more than one CBSA. Each point represents a single CBSA-year observation. The vertical axis is the average distance in miles between the CBSA and all other CBSAs. The horizontal axis is the number of banks in the CBSA-year that have branches in more than one CBSA.

Table 1: The Bank Scope Distribution, 2001-2006

Panel A: The total number of banks

# of CBSAs	2001	2002	2003	2004	2005	2006
1	3,806	3,694	3,645	3,558	3,510	3,480
2-5	700	725	758	753	781	831
6-20	74	77	82	89	91	95
≥21	23	26	27	28	24	28

Panel B: The mean number of banks per CBSA

# of CBSAs	2001	2002	2003	2004	2005	2006
1	11.29	11.09	10.98	10.68	10.64	10.71
2-5	4.94	5.07	5.19	5.21	5.45	5.83
6-20	2.41	2.46	2.62	2.72	2.88	2.89
≥21	3.49	3.69	3.82	3.96	3.90	4.22

The data include all commercial banks with branches in at least one CBSA, over the period 2001-2006.

Table 2: Deposit Rates and Market Shares

# of CBSAs	Deposit Rate	Market Share
1	0.019	0.020
2-5	0.018	0.039
6-20	0.014	0.075
≥21	0.012	0.114

The table shows mean deposit rates and market shares by bank scope. The data include all commercial banks with branches in at least one CBSA over the period 2001-2006, for a total of 45,785 bank-CBSA-year observations.

Table 3: Commercial Bank Summary Statistics

Units of Observation:		Bank-Year		Bank-CBSA-Year	
Variable	Description	Mean	St. Dev.	Mean	St. Dev.
Deposit pricing and market share					
r^D_{jt}	Deposit rate	0.019	(0.011)	0.017	(0.011)
s_{jmt}	Market share	0.025	(0.053)	0.047	(0.078)
Bank scope					
m_{jt}	# of CBSAs	1.702	(5.167)	17.389	(39.024)
$\left(\frac{m_{jt}}{M} - 1\right)$	Normalized scope	-0.998	(0.014)	-0.954	(0.109)
Other bank variables					
GTA_{jt}	Gross total assets	1.235	(17.017)	38.269	(124.512)
LOANS_{jt}	Loans	0.721	(9.217)	22.211	(68.094)
DEPS_{jt}	Deposits	0.762	(9.366)	24.197	(78.223)
BRDEN_{jmt}	Branch density	3.935	(8.190)	7.176	(18.933)
NEMP_{jt}	# of employees	23.412	(32.002)	24.404	(29.377)
CHG_{jt}	Charge-offs	0.005	(0.107)	0.152	(0.542)

Summary statistics for 26,905 bank-year observations and 45,785 bank-CBSA-year observations over the period 2001-2006. Gross total assets, deposits, loans, and charge-offs are in billions of 2000 dollars.

Table 4: CBSA Summary Statistics

Units of Observation:		CBSA-Year	
Variable	Description	Mean	St. Dev.
Number of commercial banks, branches, and deposits			
BANK_{mt}	Total banks	21.202	(24.156)
BRANCH_{mt}	Total branches	152.114	(296.166)
DEP_{mt}	Total deposits	9.491	(33.771)
Other CBSA variables			
POP_m	Population	645.468	(1487.606)
INC_m	Median HH income	48.033	(7.692)
MIL_m	Land area	22.111	(22.545)

Summary statistics for 2,160 CBSA-year observations over the period 2001-2006. Total deposits are in billions, and population, median household income, and land area are in thousands.

Table 5: Logit Regression Results

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Deposit interest rate						
r_{jt}^D	12.780***	105.060***	115.269***	115.846***	98.576***	100.821***
	(1.436)	(6.983)	(13.591)	(7.721)	(5.704)	(5.603)
Bank scope						
$\left(\frac{m_{jt}}{M} - 1\right)$	2.455***	4.347***	4.557***	4.568***	2.313	4.387**
	(0.333)	(0.572)	(0.632)	(0.605)	(2.091)	(2.134)
Control variables						
BRDEN_{jmt}	1.434***	1.463***	1.466***	1.466***	1.377***	1.015***
	(0.048)	(0.045)	(0.045)	(0.044)	(0.060)	(0.053)
NEMP_{jt}	0.062*	0.073*	0.074*	0.074*	-0.002	-0.040
	(0.033)	(0.041)	(0.043)	(0.043)	(0.040)	(0.038)
$\text{lag}(s_{jmt})$						8.824***
						(0.693)
Imputed willingness-to-pay for scope						
WTP	2.033	0.437	0.418	0.418	0.248	0.464
R^2	0.479	0.286	0.240	0.238	0.437	0.443
1st stage F-test	·	226.96***	71.73***	318.56***	214.67***	287.16***
Instruments	·	Loans+CC	Loans	CC	Loans+CC	Loans+CC
Fixed effects	·	·	·	·	Bank	Bank

Results from logit estimation. The data include 38,993 bank-CBSA-year observations. The dependent variable is $\log(s_{jmt}) - \log(s_{0mt})$, where s_{jmt} is the market share of bank j and s_{0mt} is the outside good market share. All regressions include a constant. Standard errors are clustered at the bank level and shown in parenthesis. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Table 6: Mixed Logit Regression Results

Variable	Means (β's)	St. Dev (σ)	Income (π)
Deposit interest rate			
r^D_{jt}	115.923***	2.263	-65.580***
	(8.178)	(41.228)	(18.997)
Bank scope			
$\left(\frac{m_{jt}}{M} - 1\right)$	4.701***		
	(0.627)		
Control variables			
BRDEN_{jmt}	1.070***		
	(0.044)		
NEMP_{jt}	0.070		
	(0.045)		
GMM objective		14,268.73	
% of r^D_{jt} coefficients<0		5.46	

Results from mixed logit estimation. The data include 38,993 bank-CBSA-year observations. The regression also includes an intercept. The instruments are loan-side measures and competitor characteristics. Standard errors are clustered at the bank level and shown in parenthesis. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Table 7: The Decision to Enter an Outside Market

Variables	(1)	(2)	(3)	(4)
The number of banks in the original market, by bank scope				
NBK1_{mt}	0.032	-0.018***	0.029	-0.021***
	(0.041)	(0.006)	(0.039)	(0.007)
NBK2_{mt}	-0.232***	0.001	-0.236***	-0.003
	(0.074)	(0.023)	(0.075)	0.025)
NBK3_{mt}	0.034	0.027*	0.031	0.021
	(0.069)	(0.015)	(0.068)	(0.015)
NBK4_{mt}	-0.069	0.061***	-0.072	0.027
	(0.080)	(0.022)	(0.080)	(0.022)
Market fixed effects	yes	no	yes	no
Bank controls	yes	yes	no	no
Pseudo R^2	0.131	0.069	0.121	0.048

Results from logit regressions. The data include 14,469 bank-year observations from 3,658 single-market banks over the period 2001-2005. The dependent variable equals one if the single-market bank enters a second market in the subsequent year, and zero otherwise. The variables NBK1_{mt}, NBK2_{mt}, NBK3_{mt} and NBK4_{mt} are the number of banks in the original market with branches in exactly one, exactly two, three through ten, and more than ten CBSAs, respectively. The bank controls include the single-market bank deposit and branch market shares within the original market. The regressions also include year fixed effects. Standard errors are clustered at the market level and shown in parenthesis. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Table 8: The Choice of which Outside Market to Enter

Variables	(1)	(2)	(3)	(4)
$2CBSA_{jn}$	-0.442**	-0.241	4.460***	4.625***
	(0.226)	(0.218)	(0.139)	(0.132)
Synergy Control	yes	yes	no	no
Profit Controls	yes	no	yes	no
Pseudo R^2	0.463	0.450	0.151	0.126

Results from conditional logit regressions. The observations are combinations of the 351 single-market banks that enter an outside market over the sample period and the 359 outside CBSAs. The dependent variable equals one if the single-market bank entered the outside market, and zero otherwise. The variable $2CBSA_{jn}$ equals one if a two-market bank already exists with branches in the home and outside market, and zero otherwise. The synergy control is the proportion of banks in the original market that also have branches in the outside market. The profit controls include the number of outside market banks with branches in exactly one, exactly two, three through ten, and more than ten CBSAs, as well as second-order polynomials in the outside market median income, population, and land area. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.