

# Demand Estimation and Consumer Welfare in the Banking Industry

Astrid A. Dick\*

November 6, 2002

## ABSTRACT

This paper estimates a structural demand model for commercial bank deposit services. Following the discrete choice literature, consumer decisions are based on prices and bank characteristics. The results, based on the US for 1993-1999, indicate that, with respect to prices, consumers respond to deposit rates, and to a lesser extent, to account fees, in choosing a depository institution. Moreover, consumers respond favorably to the branch staffing and geographic density, as well as to the bank's age, size, and geographic diversification. In light of the banks' responses to regulatory changes throughout the period, most markets experience a slight increase in welfare.

---

\*Research and Statistics, Federal Reserve Board, Washington, D.C., [astrid.dick@frb.gov](mailto:astrid.dick@frb.gov). The views expressed herein are those of the author and not necessarily those of the Federal Reserve Board. The author is grateful to Susan Athey and Nancy Rose for their insightful comments and ideas. She thanks Peter Davis, Glenn Ellison, Sara Fisher Ellison, Timothy Hannan, Paul Joskow, Whitney Newey, Scott Stern, Philip Strahan, Andrew Sweeting and Jeff Wilder as well as seminar participants at MIT and various other universities and government agencies. Any errors are the author's.

Throughout the past decade, the U.S. banking industry has experienced considerable growth and consolidation, with thousands of mergers and acquisitions, as regulatory barriers to banks' geographic expansion were fully lifted. Profit rates have remained high during the period. These facts raise a number of policy concerns. One is the impact of consolidation on competition in banking markets. This paper develops and estimates a structural model of demand for commercial bank deposit services to analyze consumer behavior, market power, and the impact of industry changes on consumer welfare.

There are several banking studies on competition related to concentration, profitability and prices, but none of them, to my knowledge, are based on structural demand frameworks, which allow for measures of consumer welfare, nor explore product differentiation. The empirical banking literature has centered until recently around the structure-conduct-performance (SCP) paradigm and the efficient structure hypothesis (ESH). Many of these studies regress profitability on concentration, which is usually taken to be exogenous, and frequently find a positive correlation between market concentration and profitability.<sup>1</sup> In particular, these papers have reported lower deposit rates, higher loan rates and higher price stickiness in more concentrated markets [e.g. Berger and Hannan (1989), Hannan (1991, 1997), Hannan and Berger (1991) and Berger (1995)]. Bresnahan (1989) analyzes the problems with these approaches, mostly related to the idea that prices, profits and concentration are all jointly endogenous, and the lack of any systematic relationships between concentration measures and economic variables of interest. Another strand of the literature has sought to analyze market conduct by estimating static Cournot models in the conjectural-variations fashion, in order to test for price-taking vs. price-setting behavior of banks<sup>2</sup> [e.g. Gelfand and Spiller (1987), Shaffer (1993) and Berg and Kim (1998)]. Bresnahan (1989) and Ti-

---

<sup>1</sup>The SCP is based on the idea that market structure (number of firms, cost structure, degree of vertical integration) determines conduct (prices, advertising), which in turn characterizes market performance (price-cost margins, efficiency, profits). The positive correlation between market concentration and profitability is taken as evidence of noncompetitive behavior in more concentrated markets. Under the ESH, firm efficiency is the driving factor in market conduct and performance, in that higher concentration and profitability is the result of more efficient firms naturally achieving larger market shares.

<sup>2</sup>The conjectural variations approach assumes that firms maximize profits over output levels based on their expectations on other firms' reactions. Given the Cournot setting, product is assumed to be homogeneous across firms.

role (1997) analyze the problems of this approach related to the irrationality of conjectured reactions in a static setting.

In an attempt to overcome the difficulties and limitations of the earlier literature, this paper develops and estimates a structural, discrete choice model of demand for banking services, taking advantage of the advances made in the industrial organization literature. This rich framework yields tools to carry out analyses of welfare, market power and consumer preferences that are not feasible under the usual approaches of the empirical banking literature. These new tools are particularly important from a policy perspective, where a fuller framework might be needed to determine proper regulation. In particular, the recent changes in the industry raise questions about the consequences of the deregulation of bank's geographic scope and the resulting consolidation on competition and consumer welfare. First of all, what characterizes consumer behavior for banking services? Second, what are the levels of competition and how have they been affected? Third, what is the net effect of all these changes for consumers?

In particular, the strategy followed here is to estimate a demand model for deposit services following the discrete choice literature,<sup>3</sup> where consumer decisions are based on prices and bank characteristics. The discrete choice approach incorporates product differentiation explicitly and helps to overcome the difficulty of estimating a large number of substitution parameters given several firms in the market. The model is estimated for the U.S. commercial banking sector over 1993-1999, using a data set that combines information from several industry sources. This work departs from the existing literature by developing a demand model derived from the consumer's utility function, and exploring product differentiation by allowing heterogeneity across banks. The model provides insight into consumer behavior and, coupled with a model of supply conduct, provides a tool for the analysis of competition and policy analysis.

Based on the estimation of logit-based models, the results indicate that, with respect to prices, consumers respond to deposit rates, and to a lesser extent, to account fees, in

---

<sup>3</sup>See, for example, the precursory papers by McFadden (1973, 1978, 1981), and also Berry (1994), Berry, Levinsohn, and Pakes (1995), and Nevo (2001).

choosing a depository institution. With respect to bank characteristics, consumer demand responds favorably to the staffing and geographic density of local branches, as well as to the age, size, and geographic diversification of banks. The paper also finds important differences across consumer and market types in the demand for banking services. For example, higher income consumers appear to value location characteristics relatively less than do lower income consumers. They also appear to care relatively more about bank size. In terms of demand across market types, urban markets appear less responsive to prices than rural markets.

Based on the estimates obtained from the logit model, I calculate demand elasticities for prices and observed bank characteristics. The median of the distribution of implied price elasticities is 6 for the interest rate on deposits. On the one hand, the account fee elasticity is below unity for most banks, which, taken in isolation, is not profit maximizing. However, this might be due to the use of checking accounts as loss leaders. The estimation of a nested logit model, on the other hand, provides a median service fee elasticity of -3.5. However, the correlation parameter of the nested logit requires the use of additional instruments. Combining the demand estimates here with pricing rules derived from a model of firm conduct enables the construction of estimates of market power, without having to rely on concentration or accounting profit measures. The implied price-cost margin (PCM) is calculated to be 25 percent for service fees, and 10 percent for the deposit rate. Furthermore, despite the dramatic changes in the industry, PCM appear to have changed little throughout the period.

Lastly, the model allows one to analyze the benefits to consumers in light of the responses of banks to the regulatory changes throughout the period. The consolidation and high levels of accounting profits have been accompanied by a change in bank characteristics and prices. This model enables one to estimate the net effect of all these changes on consumer welfare. I find that most markets experience a slight increase in welfare, with a median consumer gain of around one dollar.<sup>4</sup> Moreover, there is almost zero correlation between a market's concentration and its welfare change. As consumers are found to value several

---

<sup>4</sup>The fact that the consumer welfare increase is positive but close to zero is particularly interesting in terms of the earlier debate about interstate branching deregulation, given the concerns about the potential consumer welfare loss from greater bank consolidation, especially of those opposing deregulation.

bank attributes other than price, this exercise is at least suggestive of the bias that might arise in welfare inferences based solely on prices and concentration measures. In particular, the usual policy approach of focusing on the price effects in the case of mergers might need to acquire a broader perspective.

This paper modifies the standard discrete choice framework to fit the characteristics of a service industry such as banking. First, the analysis must confront a number of challenges, including the definition of product, quantity, and relevant geographic market, as well as the consumer decision. The relevant market is defined as geographically local, based on the empirical evidence and following the practice of the regulatory agencies. In particular, I define markets as metropolitan statistical areas (MSA) and non-MSA rural counties. Product is defined as dollar deposits, while market share is defined as the number of “average” deposit accounts, which allows for the introduction of an outside good. Second, the estimation itself has to deal with the correlation between prices and bank-specific demand shocks in order to obtain consistent estimates of the demand parameters. I define several identification strategies by using actual cost data as well as bank fixed effects. In particular, I use labor costs and other operating expenses as the main cost shifters, as well as product mix and balance sheet structure variables which add further nuance about the bank’s cost function.

The paper is organized as follows. Section 2 provides an overview of the banking industry and its regulatory framework. In section 3, the empirical framework is outlined, including the consumer decision-making process, the relevant geographic market, bank competitors, and output quantity, as well as the demand model and a model of firm conduct from which PCM can be defined. In Section 4, I describe the data and estimation, including the price instruments. Results are presented in section 5, including elasticities, PCM, and welfare effects. Lastly, section 6 concludes.

# I. The U.S. banking industry: an overview

The banking industry constitutes a major part of the U.S. economy, employing about one and a half million people and holding domestic assets of over \$5,000 billion.<sup>5</sup> Throughout the last three decades, and particularly in the nineties, the industry has undergone several changes in both its structure and regulation.<sup>6</sup> Regulatory restrictions affecting the ability of banks to diversify geographically and the range of products offered have decreased dramatically. Deregulation of unit banking and limited branch banking occurred gradually throughout 1970-1994.<sup>7</sup> In 1994, the Riegle-Neal Interstate Banking and Branching Efficiency Act was passed, permitting nationwide branching as of June 1997.<sup>8</sup>

Banks appear to have responded to these regulatory changes in a significant way. The number of commercial banks has decreased by the thousands, reaching around 8,000 in 1999 from over 11,000 a decade earlier, and most of the banks exiting have been small.<sup>9</sup> The number of mergers has averaged around 400 per year throughout the nineties.<sup>10</sup> The industry has increased to over 65,000 branches and 190,000 automated teller machines. Moreover, the distribution of bank size has changed. While the megabank<sup>11</sup> share of assets has increased to almost 30 percent, that of small banks has decreased to less than 5 percent. Based on accounting rates of return, profit margins are also high, and have been consistently at high levels throughout the nineties. As the sector increased its concentration, return on equity (the usual profit measure utilized in the industry) has remained above 15 percent.<sup>12</sup>

---

<sup>5</sup>The information described in this section has been constructed on the basis of the Federal Reserve Board Report of Condition and Income data and the Federal Deposit Insurance Corporation data on branch deposits, unless indicated otherwise.

<sup>6</sup>See Berger et al. (1995) for an extensive analysis of the U.S. banking industry throughout the period 1979-94.

<sup>7</sup>Intrastate branching as well as interstate banking through Banking Holding companies (BHC) were allowed gradually among the U.S. states. Intrastate branching deregulation began in some states even before the 1970's, while interstate banking started as early as 1978.

<sup>8</sup>As states had the option to "opt in" earlier than the June 1997 federal deadline, the Act became effective gradually among the U.S. states between 1994-1997. Only Texas and Montana opted out of interstate branching altogether.

<sup>9</sup>A small-sized bank is defined here to have assets below \$100 million.

<sup>10</sup>Based on data from the Federal Reserve Board on mergers and acquisitions.

<sup>11</sup>Megabanks are defined to have assets over \$100 billion. Despite the large number of banks in the U.S., the ten largest banks hold almost a third of national deposits.

<sup>12</sup>The profit margins are used only in a descriptive manner here. See Fisher and McGowan (1983) on the misuse of accounting rates of return.

The sample period employed here covers the years 1993-1999. Banking markets are defined as geographically local; either an MSA, in the case of urban markets, or a rural county. In the U.S. there are close to 2,600 local markets, 330 of which are MSA markets. The latter represent 83 percent of the share of U.S. dollar deposits. Looking at urban markets alone, an average market had 20 banks in 1999, which represents a decline in the number of firms for most markets since 1993. The average number of branches at the local level has gone from 131 to 140. While overall concentration in the industry has increased, as evidenced by the dramatic drop in the number of firms, there is large variation in both the levels and the changes in concentration in local markets. In 1999, the average Herfindahl-Hirschman index (HHI)<sup>13</sup> was around 1900<sup>14</sup> in urban markets,<sup>15</sup> a slight increase since 1993. According to this measure, concentration levels have increased in 47 percent of the markets.

Furthermore, bank characteristics have also changed. Table I shows market averages for some bank attributes, for 1993 and 1999. Banks have more branches per local market, and their geographic expansion has increased as banks operate in more states. Also, the average bank age has risen and the distribution of bank size has shifted to the right, as a larger proportion of banks fall in the big and medium-sized categories. In terms of prices, there is a slight increase in service fees and deposit rates.

Insert Table I about here.

---

<sup>13</sup>The Herfindahl-Hirschman index is a concentration measure constructed as the sum of the squares of the market share of deposits at the local market level. Here, following the practice of the Antitrust Division, we multiply it by a factor of 10,000, which is the index of a monopolist in a market.

<sup>14</sup>The Antitrust Division defines the threshold of a highly concentrated market at 1800. In the case of bank mergers, the Antitrust Division has used a screen of 1800/200 over the past several years. That is, in most cases they will not conduct a full investigation unless there is at least one market where: (i) the post-merger HHI is at least 1800; (ii) the merger produces a change in the HHI of at least 200.

<sup>15</sup>For the case of rural markets, in 1999 there were an average of 5 banks in each market. The average Herfindahl-Hirschman index (HHI) was 4000. From 1993 to 1999, the HHI increased in 35% of the markets. Rural markets are defined as non-MSA counties, which usually means much smaller geographic areas than MSA's. As a result, they tend to show higher levels of concentration than that of MSA markets. For instance, while urban markets present no markets with only one banking firm, in 1999 there are 175 rural county markets that are monopolies. It is interesting, however, that 14 percent of these counties are in Texas and 9 percent in Montana, together representing the only states in the U.S. that opted out of the 1994 interstate branching deregulation.

## II. Empirical framework

### A. Definitions for a model of deposit services demand

#### Consumer decision:

The demand model focuses on deposit services, which include checking, savings, and time deposit accounts. While one might model demand for these products separately, deposit services cannot be disaggregated at the bank-market level.<sup>16</sup> Detailed product data exist at the bank level, but the only information collected at the branch level is for total deposits.

While this is certainly a constraint, the evidence on consumer behavior suggests that assuming that consumers cluster their purchases for deposits services within one depository institution is a reasonable approximation. Based on the Survey of Consumer Finances, consumers show preference for acquiring banking services together, in particular deposit services, and appear to cluster their purchases with their primary financial institution.<sup>17</sup> Evidence on small- and medium-sized business behavior provides similar insight.<sup>18</sup> Alternatively, one could justify this by assuming that consumers have demands for multiple banking services, and incur a fixed cost for each new firm they have to deal with. For sufficiently high fixed costs, consumers consolidate deposit services with a single bank.<sup>19</sup>

---

<sup>16</sup>One could also envision incorporating loans as part of the consumer decision in choosing a bank. In light of the empirical and anecdotal evidence, loans appear to constitute a separate area of analysis. For instance, we expect the search and purchase process for loan services to be substantially different from that of deposits, as consumers presumably are willing to “shop around” considerably more for loans (especially for those loans that are known to be more supply driven, such as commercial and industrial loans). Furthermore, the evidence on the appropriate definition of the relevant market for the case of loan services is significantly more ambiguous than it is for deposits. Also, the evidence on clustering of purchases when the set of products includes loans is not as strong, and appears to be decreasing with the expansion of markets due to technological and regulatory changes.

<sup>17</sup>See Amel and Starr-McCluer (2001). In fact, the share of households having more than one service at the primary institution rose from 57 percent in 1989 to 64 percent in 1998. See also Elliehausen and Starr-McCluer (1992).

<sup>18</sup>See Elliehausen and Wolken (1990) and Kwast et al. (1997). Another piece of evidence supportive of the approach taken here is the following. Some banks do attempt to attract customers by offering special “package deals” or by using “loss leaders” in checking accounts. Presumably, banks do this because they expect customers to buy other products, once they have been induced to buy a single service.

<sup>19</sup>Under less stringent data constraints and with the development of a more complex model of consumer behavior, we might be able to account for the entirety of consumer purchases of banking products. We might



In terms of the factors involved in the consumer decision-making process, consumers are expected to value and to be able to recognize several of a bank's attributes when searching for a deposit service provider. These include branch location and its characteristics, prices and range of services (account types, financial advising, electronic services such as ATMs and debit cards) and the bank's expertise and overall quality and effectiveness in providing services. With the advent of new technology, some consumers are likely to care about further electronic and online services.

### **Consumers:**

In terms of the customer base, there are two main types of consumers in the industry: individuals (households) and nonfinancial businesses. At the aggregate national level,<sup>20</sup> more than two thirds of checkable deposits are owned by businesses while 95 percent of time and savings deposits belong to households. Within business customers, 75 percent of time and savings deposits and 35 percent of checkable deposits are held by the nonfarm noncorporate sector, while the rest is held by the corporate sector.

A limitation of the data is that it does not allow to control directly for these different types of consumers. However, surveys of consumer finances and small- and medium-sized business finances indicate that the behavior of these two groups are similar, in that both cluster their purchases, especially of deposit services, with their primary financial institution, which is typically situated close to their own locality.

### **Relevant geographic market:**

Any analysis of consumer behavior and competition in an industry requires a definition of the relevant geographic market. If the market is defined too broadly, market power would be underestimated, and the reverse would be true if the market is defined too narrowly.

---

consider incorporating banks explicitly as multiproduct firms and allowing consumers to choose among them, given their multiproduct demands. For instance, consumers may choose to obtain one product from one bank, such as a mortgage loan, while having, say, deposit accounts with another bank, even though there might be a nonnegative fixed cost of acquiring services from another bank.

<sup>20</sup>More disaggregated data on consumer types are not available. Percentages based on the Federal Reserve Flow of Funds Accounts data for the year 1999.

Naturally, market definition has been central to regulatory agencies in dealing with antitrust cases.

The format of the available data on deposits allows for a definition of market as narrow as the MSA or rural county (and even smaller units), and as wide as the national level, given that the dollar amount of deposits is available for each bank branch in the U.S. Antitrust analysis of mergers in the banking industry has relied on the definition of market at the geographically local level, by assuming that this is representative of how most households and businesses behave when they purchase banking services. Indeed, according to the most recent evidence on consumer behavior, this appears to be appropriate, especially for deposit services. A recent article by Amel and Starr-McCluer (2001) uses data from the triennial Survey of Consumer Finances, which is sponsored by the Federal Reserve Board, and finds evidence that households continue to a substantial degree to obtain most services, especially checking accounts, at local depository institutions. In fact, they find that 90 percent of checking accounts, savings accounts and certificates of deposits (CD's) are still acquired within the local market.<sup>21</sup> Kwast et al. (1997) find that over 94 percent of small businesses use a local depository institution. Though a bit outdated for the purposes here, Elliehausen and Wolken (1990) find that 93 percent of both small and medium-sized businesses use a local commercial bank. Moreover, Rhoades (1992) and other empirical studies provide further support for the view of banking markets as geographically local.<sup>22</sup>

An important and growing debate in the industry suggests the broadening of the relevant geographic banking market to the state level. Most of the work in this area concerns loan services, however. For example, Hannan and Strahan (2000) find that, in contrast to their expectation, the state-level HHI better explains pricing of secured and unsecured loans. Radecki (1998) finds empirical evidence which is interpreted as suggestive of banking markets expansion beyond the metropolitan area and small and local market. Radecki finds that the

---

<sup>21</sup>Other deposit categories such as “time deposits greater than US\$100,000” are presumably less likely to be limited to the local geographic market.

<sup>22</sup>Rhoades uses mortgage rates in twenty U.S. cities to study the size of banking markets, and finds that these rates tend to be higher in cities where concentration is relatively high. He concludes that this significant empirical relationship provides support for the definition of local markets at the local level. Also see literature cited in the introduction for further empirical studies supportive of the definition of markets as geographically local, such as Berger and Hannan (1989).

current practice among banks in New York and other large states is to set uniform retail deposit and consumer loan rates across an entire state or large regions of a state. Using 1996 survey data, he finds that the significant correlation between retail deposit rates and concentration measures that existed at the local level in the mid-1980s has disappeared, finding instead significance at the state level. Petersen and Rajan (2002) find, based on a survey of small businesses in 1993, that the distance between firms and their lenders is growing over time.

My approach is to define the relevant banking market as geographically local, either as an MSA or a non-MSA rural county.<sup>23</sup> This is one of the finest disaggregation permitted by the data and is supported by consumer and business surveys, the bulk of the empirical literature on deposit services, and the application of antitrust policy.<sup>24</sup>

### **Commercial bank competitors:**

A broad range of institutions offer financial services. Historically, regulators classify these into depository and non-depository institutions. Depository institutions include commercial banks, thrifts, savings banks, and credit unions. Non-depository institutions include finance companies, brokerages, and mortgage lenders. All of these might be considered as potential competitors of banking firms for some product markets. The focus of this paper on the demand for deposit services and the paucity of data for depository institutions other than commercial banks limit the empirical analysis to commercial banks. As a result, any other competitors to commercial banks will be part of the outside good in the demand model.<sup>25</sup>

### **Output quantity:**

---

<sup>23</sup>In large urban areas, markets are defined as primary MSAs, as opposed to the larger consolidated MSAs.

<sup>24</sup>No market definition will ever be perfect, given the data constraints and the ever changing boundaries of markets. It is expected that wealthy households and large businesses transcend the local market limits as they are likely to purchase services from banks outside the local market.

<sup>25</sup>According to the work by Amel and Starr-McCluer (2001), based on the Survey for Consumer Finances, 98 percent of households uses a depository institution as of 1998. Thrifts, savings banks and credit unions comprise less than 6 percent in terms of U.S. deposits as of 1999. There is actually a recent paper that provides evidence for leaving these institutions outside the deposit product market. Amel and Hannan (1999) estimate a residual deposit supply in order to determine whether nonbank financial institutions should be included as participants in defining the product market relevant in antitrust analysis of proposed bank mergers, and find that only commercial banks should be included in the “antitrust market.”

While the number of accounts may be a measure of bank output, both regulators and the industry refer to output in terms of dollar volume, as this should be more representative of the activity in the market.<sup>26</sup> I therefore define market share on the basis of dollar deposit data collected at each bank branch in the U.S.<sup>27</sup>

While most consumers use a depository institution, some may purchase deposit services from institutions outside the set I consider, or forgo these services altogether. This outcome can be modeled as a decision to allocate consumers to an “outside good.” I follow the traditional approach of the discrete choice demand literature by defining the potential size of the market in terms of the local market population.

I define the relevant consumption unit<sup>28</sup> as one “average deposit account,” defined as the mean total deposits divided by the number of deposit accounts across all U.S. banks in a given year.<sup>29</sup> Taking a bank’s dollar deposit volume and dividing it by the consumption unit just defined, I obtain a bank’s quantity as the number of accounts thus defined.<sup>30</sup> Given this output quantity, market share for bank  $j$  is the number of accounts it holds, divided by the size of the (potential) market.<sup>31</sup>

---

<sup>26</sup>Suppose we define product in terms of number of accounts, for instance. Then a bank which has a relatively large number of student accounts with one dollar each, would have a large market share, yet little presence in the market in terms of volume.

<sup>27</sup>The volume should capture the average of annual flows, including accounts that open and close throughout the year, as consumers presumably enter and exit the market continuously. However unlikely, though, one could envision a situation where a bank has no new accounts throughout a year, in which case one would find less consumer responsiveness to bank attributes, as these customers are locked-in and confront switching costs. Similarly, one would find more responsiveness in the case of all new accounts.

<sup>28</sup>This raises two issues: (i) consumers have heterogeneous demands for deposit services, and (ii) the magnitude of an individual consumer’s demand might depend on the price. None of these are dealt with explicitly here. By defining a consumption unit, one is effectively averaging over consumers’ heterogeneous demand. This is as reasonable or unreasonable here as it is elsewhere in the literature, even though the problem might be more acute in this setting.

<sup>29</sup>More precisely, I calculate  $k = \text{deposits}/\text{number of accounts}$  for each bank and then take the average of  $k$  across all U.S. banks.

<sup>30</sup>The natural definition of market share would be to divide each bank’s dollar deposits by the total deposits in the market. However, this presents a practical problem, as the inclusion of an outside good is required here. Therefore, I convert dollar volumes into number of accounts sold. I do this by dividing each bank’s dollar deposits by  $M * k$ , where  $M$  represents the number of people in the market and  $k$  the average dollar deposit account, instead of dividing dollar deposits by the total dollar deposits in the market. Also note that I use here the stock as opposed to the flow of product. These two have been assumed to be proportional in the empirical banking literature. There are no data available on flows.

<sup>31</sup>In some markets, it will be the case that the shares actually add up to a number larger than one, whenever the market size, given by the local population, is not broad enough. In this case, one can either:

## B. Demand model

My model of deposit services demand is designed to reflect as closely as possible the nature of consumer decision making in choosing a depository institution, given the constraints in the available data. As pointed out earlier, the data do not allow me to distinguish between different types of consumers, such as households and businesses. Additionally, the only data available at the local level, where I define the relevant geographic market, are the total dollar amount of deposits. As a result, it is not possible to distinguish between checking and time and savings accounts. Detailed banking data are only available at the consolidated bank level, which aggregates across all local markets served by the bank. The format of the data drives a number of modelling choices, but seems unlikely to significantly distort the interpretation of the results.

### B.1. Model

Demand is derived and estimated following a discrete choice approach. This methodology allows for the exploration of the degree of product differentiation in the industry explicitly, and it solves the dimensionality problem present when many firms are in the market, as in the case of the banking industry.<sup>32</sup>

Following Lancaster (1971) and the work of McFadden (1973, 1978, 1981), under the discrete choice approach the utility an individual derives from consuming a given product is a function of two sets of factors: individual and product characteristics. Let consumer

---

(i) drop from the sample those markets that have this feature; (ii) consider multiplying the market size by a factor  $\lambda > 1$ , as has been sometimes done in the literature. I follow both approaches in Section 5.

<sup>32</sup>There are several approaches to demand estimation. For instance, in the constant elasticity case, with  $N$  products, there are  $N^2$  parameters to estimate in a traditional demand model. In the banking industry this is particularly problematic because there are usually many banking firms to be considered in a single market, even under restrictive market size definitions. Another possibility is the multi-stage budgeting approach following Hausman et al. (1994), who construct a multi-level demand system for differentiated products. This approach requires an a priori grouping of products (brands) into exhaustive and mutually exclusive sets. Following Lancaster (1971), under the discrete choice approach consumers are endowed with preferences over product characteristics, which is key as the parameter space is largely reduced. This is convenient in the present setting as there are many fewer bank characteristics than banks. Examples of this framework are the multinomial and nested logit (see Anderson et al. (1992)) and the random coefficients model.

specific characteristics be denoted by  $\zeta$ , and overall product characteristics by  $C$ , where  $C = (x, \xi, p)$ , with  $x$  representing observed product characteristics,  $\xi$  unobserved product characteristics, and  $p$  prices. It is assumed that firms and consumers observe all relevant characteristics, unlike the researcher, who may observe some characteristics but not all. Thus, one may write utility for consumer  $i$  for product  $j$  as  $U(\zeta_i, p_j, x_j, \xi_j; \theta_D)$  where  $\theta_D$  represents the demand parameters to be estimated.

Consumers are interested in purchasing deposit services from a bank.<sup>33</sup> Assume that  $t = 1, \dots, T$  markets are observed, each with  $i = 1, \dots, I_t$  consumers and  $j = 1, \dots, J_t$  firms. Let the utility function take on a linear form such that the conditional indirect utility of consumer  $i$  from choosing bank  $j$ 's services in market  $t$  is

$$u_{ijt} \equiv \delta_{jt} + \epsilon_{ijt} \equiv p_{jt}^d \alpha^d - p_{jt}^s \alpha^s + x_{jt} \beta + \xi_j + \epsilon_{ijt}, \quad (1)$$

where  $p_{jt}^d$  and  $p_{jt}^s$  represent interest rate paid by banks on deposits and service charges on checking accounts, respectively,  $x_{jt}$  is a  $K$ -dimensional row vector of observed product characteristics for product  $j$  in market  $t$ ,  $\xi_j$  represents unobserved product characteristics (depicted as a mean across consumers), and  $\epsilon_{ijt}$  is a mean zero random disturbance. The  $K+2$  dimensional vector  $\theta_D = (\alpha^d, \alpha^s, \beta)$  represents the taste parameters.

Each consumer chooses the bank that maximizes her utility, so that consumer  $i$  chooses bank  $j$  whenever  $U(\zeta_i, p_j, x_j, \xi_j; \theta_D) \geq U(\zeta_i, p_r, x_r, \xi_r; \theta_D)$ , for  $r = 0, 1, \dots, J$ , where  $r = 0$  represents the outside alternative,<sup>34</sup> such that a consumer that chooses it remains outside

---

<sup>33</sup>I assume that consumers choose the proportion of assets they allocate to deposit services prior to choosing a bank. As discussed earlier, I also assume that consumers want to obtain checking, time and savings deposit accounts from the bank they choose. However, one could just as well assume that consumers want only one of these deposit services. This may be the case as long as consumers care about the same set of prices and observed characteristics when choosing a deposit service, whether they choose only a checking account or a time deposit or both. One might ask why a consumer that is only obtaining a checking account, say, might care about the interest rate on time deposits. As discussed earlier, the evidence that consumers tend to cluster their purchases of financial services in their local institution is particularly strong for the case of deposit services. In this respect, it can be argued therefore that consumers internalize the prices of other products into their decision because they expect to purchase them in the future from this same bank, given a positive fixed cost of switching to another bank.

<sup>34</sup>We introduce an outside good for two reasons: 1) without such good, an overall, proportional price increase would leave market shares unchanged, which is unreasonable because consumers might choose not to purchase services at all, or they might simply choose another type of depository institution, such as a

of this market and allocates expenditure to other things.<sup>35</sup> The distribution of consumer characteristics  $\zeta$  is assumed to be known.<sup>36</sup> The consumer's choice rule defines implicitly the set of  $\zeta_i$  that results in purchase of product  $j$ . Let

$$A_j = \{\zeta : U(\zeta, p_j, x_j, \xi_j; \theta_D) \geq U(\zeta, p_r, x_r, \xi_r; \theta_D) \text{ for } r = 0, 1, \dots, J\} \quad (2)$$

be the set of values of  $\zeta$  that induce choice of good  $j$ . Aggregate demand is thus obtained by integrating out the choice function over the distribution of  $\zeta$  in the population. It is then easy to see that the market share of good  $j$  is the probability that  $\zeta_i$  falls into region  $A_j$ , for all  $i$ . Given a density  $f(\zeta)$  for  $\zeta$ , market share is (assuming ties occur with zero probability):

$$s_j(p, x, \xi; \theta_D) = \int_{\zeta \in A_j} f(\zeta) d\zeta \quad (3)$$

With a market of size  $M$ , demand for bank  $j$  is then given by  $M s_j(p, x, \xi; \theta_D)$ .

A simplifying yet restrictive assumption regarding consumer heterogeneity frequently made in solving the integral in (3) is that the random error  $\epsilon_{ij}$  is i.i.d. extreme value with the distribution function  $\exp(-\exp(-\epsilon))$ , and that it enters utility only through an additive-separable form, as in (1). In this manner, the convenient closed form solution of the multinomial logit model is obtained for the market share of product  $j$ :

$$s_j(\delta) = \frac{\exp(\delta_j)}{\sum_{r=0}^{J+1} \exp(\delta_k)} \quad r = 0, 1, \dots, J. \quad (4)$$

While this result is appealing for its computational simplicity, it imposes restrictive substitution patterns, as own- and cross-price elasticities depend only on market shares.<sup>37</sup>

---

thrift; 2) in order to obtain a simple specification in the logit model, by assuming that the mean utility of the outside good is zero.

<sup>35</sup>Note that we have dropped the market subscript  $t$ .

<sup>36</sup>The distribution of consumer characteristics does not have to be assumed parametrically since it may also be empirically estimated from demographic data.

<sup>37</sup>One implication from the utility model, which assumes the additive separability between the two terms in (1), such that the  $\delta$  term depends on product characteristics only, and the disturbance term solely on consumer characteristics, is that substitution patterns only depend on the  $\delta_j$ 's. Each  $\delta$  vector is associated with a unique vector of market shares, so that cross-price elasticities between any two products depend only on their market shares, and observable characteristics of the products do not play a role in the price elasticity.

The nested logit model reduces this problem by allowing the distribution of consumer characteristics to depend on the unknown parameter  $\sigma$ , to be estimated. As a result, market shares and the implied mean utilities vary with  $\sigma$ . In particular, the nested logit model allows consumer tastes to be correlated across products  $j$ . This requires an a priori grouping of products into  $G+1$  exhaustive and mutually exclusive sets (including the outside good). For product  $j \in G_g$ , consumer  $i$ 's utility is given by:

$$u_{ijt} \equiv \delta_{jt} + \varsigma_{ig} + (1 - \sigma)\epsilon_{ijt} \quad (5)$$

As  $\sigma$  approaches one, the correlation of consumer preferences across products in group  $g$  approaches one as well. While this is certainly a more flexible model than the basic logit, it comes at the cost of an increase in the parameters to be estimated and so the number of required instruments.

As outlined in Cardell (1991) and Berry (1994), 5 can be interpreted as a random coefficients model<sup>38</sup> involving random coefficients  $\varsigma_{ig}$  only on group-specific dummy variables. Allowing  $d_{jg}$  to be a dummy variable equal to one if  $j$  is in  $G_g$  and zero otherwise, 5 becomes:

$$u_{ijt} \equiv \delta_{jt} + \sum_g [d_{jg}\varsigma_{ig}] + (1 - \sigma)\epsilon_{ijt} \quad (7)$$

In this simple manner, the nested logit model introduces significant flexibility by allowing consumer preferences to be correlated within product categories. See Berry (1994) for a detailed outline of this model.

---

<sup>38</sup>The random coefficients model allows for even more complicated interactions of consumer heterogeneity and observed product characteristics. If the coefficient on observed product characteristic  $k$  is allowed to vary per consumer,  $\beta_{ik}$ , where  $\beta_{ik} = \bar{\beta}_k + \sigma_k \nu_{ik}$ , one obtains the random coefficients model where each individual assigns a different utility level to each observable product characteristic (or some of them). The utility function takes on the following form:

$$U(\zeta_i, p_j, x_j, \xi_j; \theta_D) \equiv p_j^d \alpha^d - p_j^s \alpha^s + x_j \bar{\beta} + \xi_j + \sum_k \sigma_k x_{jk} \nu_{ik} + \epsilon_{ij} \quad (6)$$

where  $\nu$  is i.i.d. across individuals and characteristics. With the introduction of interactions between consumer and observed product characteristics, market shares are no longer driven by mean utility levels alone. In terms of substitution patterns, this means that consumers will substitute towards similar products. See Berry, Levinsohn and Pakes (1995) for more details.



## C. Supply

In the current setting, a firm is an FDIC-insured commercial bank.<sup>39</sup> Suppose there are  $J$  banks in the market, which are assumed to be price setters.<sup>40</sup> Firms maximize profits over two prices: interest rate paid on deposits  $p^d$ , and service charge on checking accounts  $p^s$ . Assume firms obtain the loan interest rate  $p^l$  elsewhere using the funds they generate on the deposit side. Let the profit for a bank  $j$  be:

$$\pi_j(p, z, \xi, \omega_j, \theta) = (p_j^l + p_j^s - p_j^d - mc_j)Ms_j(p, x, \xi; \theta_D) - F_j \quad (8)$$

where  $q_j = Ms_j(p, x, \xi; \theta_D)$  is the observed output quantity of deposit services of the bank,  $mc_j$  represents marginal costs,  $\gamma$  is a vector of unknown cost side parameters and  $\theta = (\theta_D, \gamma)$ . Note that I specify the cost function as  $C_j(q_j, mc_j, \gamma) = mc_j * q_j(p, x, \xi; \theta_D) + F_j$  where  $F_j$  represents the fixed cost of production.<sup>41</sup>

Assuming the existence of a pure strategy interior equilibrium and strictly positive prices at equilibrium, a bank maximizes profits given its attributes and the prices and attributes of competing banks. The price vector satisfies the following first order conditions for each bank  $j = 1, \dots, J$ :

$$p_j^l - p_j^d + p_j^s - mc_j = \frac{s_j(p, x, \xi; \theta_D)}{\left| \frac{\partial s_j(p, x, \xi; \theta_D)}{\partial p_j^s} \right|} \quad (9)$$

$$p_j^l - p_j^d + p_j^s - mc_j = \frac{s_j(p, x, \xi; \theta_D)}{\left| \frac{\partial s_j(p, x, \xi; \theta_D)}{\partial p_j^d} \right|}. \quad (10)$$

---

<sup>39</sup>A commercial bank is a business that accepts deposits of money subject to withdrawal on demand or at the end of a specified period and employs that money primarily to grant credit, among other activities such as buying and selling negotiable instruments. A commercial bank is therefore a multiproduct firm. As discussed earlier, I concentrate here on deposit services.

<sup>40</sup>The number of banks may vary by market, as indicated earlier. I just drop the market subscript  $t$  here.

<sup>41</sup>While not critical to the main exercise carried out here, note that I have assumed constant marginal costs, which is a standard assumption. This assumption should be reasonable within the relevant range. Also, the form of the cost function implicitly assumes economies of scale. The existence of economies of scales in banking has been documented in several studies, such as that by Berger and Mester (1997) and the many references therein related to this literature.

The PCM are thus defined as  $(p_j - mc_j)/p_j$ . Estimates of these can be obtained by using the estimated demand parameters.

### III. Data and estimation

#### A. Data

The data come from several sources. The data on bank characteristics derived from balance sheet and income statement are taken from the Report on Condition and Income (“Call Reports”) from the Federal Reserve Board. The data on branch deposits used in the construction of local market shares, as well as the number of branches, are obtained from the Federal Deposit Insurance Corporation (FDIC).<sup>42</sup> Demographic data at the MSA and county level are taken from both the U.S. Census and the Bureau of Economic Analysis. The sample covers the period 1993-1999.<sup>43</sup> All U.S. local markets are included in the sample, with about 300 MSA markets and 2000 rural markets on average per year.<sup>44</sup>

An observation is defined as a bank-market-year combination in the estimation exercises. The bank attributes are chosen based on the available data and on the belief that they are important and recognizable to the consumer. Summary statistics are provided in Appendix A, which also contains a description of the variables. Because most of the available data measure observed bank characteristics at the bank rather than bank-market level, most attributes for banks that operate in more than one market offer no market variation within a given year. For instance, price for bank  $j$  is the same across all markets within a year. This may be less restrictive than it first appears. Radecki (1998), for example, finds that the current practice among banks in New York and other large states is to set uniform retail

---

<sup>42</sup>Furthermore, I use a data set from the Federal Reserve Board on Banking Holding Companies (BHC) to construct an indicator variable for whether the bank belongs to a BHC.

<sup>43</sup>The data are taken from the second quarter reports of each year. I choose this quarter because one of the variables of interest is only reported then.

<sup>44</sup>For 1999, for instance, the largest MSA in the sample is Los Angeles-Long Beach, CA, with more than 9 million people, and the smallest is Enid, OK with approximately 57,000. The average MSA population is around 660,000. For the case of rural markets, the largest is Ulster, NY with about 167,000, and the smallest is Arthur, NE with 400 people. The average population in rural markets is around 24,000 people.

deposit and consumer loan rates across an entire state or large regions of a state. Moreover, 70 percent of banks operate in only one local market, and as a result their headquarter's data exactly fit local market data.

Two prices are observed: service charges on checking accounts and the interest rate paid on deposits. Following common practice, these are imputed from deposit revenues, in the case of checking accounts, and from deposit expenses, in the case of time and savings deposits, adjusted by the corresponding stock.<sup>45</sup>

A central part of this work is to illustrate the extent to which consumers view banks as heterogeneous. Several observed bank characteristics are included. First, the number of local branches per square mile and the number of employees per branch should be important bank attributes to consumers. The number of employees at the branch should affect consumers' decisions given that it is correlated with wait time. It might also capture the value of human interaction to consumers who find technological access to their bank more intimidating, and the types of services specific to bank branches.

I control for bank size through a set of dichotomous variables: large, medium, and small (the latter omitted in the regressions).<sup>46</sup> These capture factors that affect the service clients may receive as a consequence of larger size, such as a larger infrastructure and better facilities and know-how, as well as a lower probability of failure that is sometimes ascribed to larger banks. I also introduce the number of states in which the bank has presence, which should measure the value attached to network size and geographic diversification. The age of the bank, measured by the number of years since the bank first began operations, may reflect the value attached to the perceived degree of experience and expertise of a bank, about which a consumer has some level of cognizance.

---

<sup>45</sup>Data on actual interest rates is sometimes collected by surveys, but based on a small sample of around 300 banks.

<sup>46</sup>The size categories are defined, somewhat broadly, as follows: medium-sized banks have assets between 100 million and 300 million, while large banks have assets above 300 million. The definition is based on the FFIEC form that banks are supposed to report to the regulatory authority.

## B. Estimation

Let  $S_j$  represent the observed market share for firm  $j$ , while  $s_j$  is the market share predicted by the model, as introduced earlier. The following equations hold exactly at the true values of  $\delta$ :

$$S_j = s_j(\delta) \quad j = 1, \dots, J \quad (11)$$

Assuming the distribution of consumer unobservables is known so that market shares depend only on mean utility levels, the vector-valued equation  $S = \vec{s}(\delta)$  can be inverted to produce

$$\delta = (\vec{s})^{-1}(S) \quad (12)$$

One can obtain the above expression either analytically, such as in the logit case, or numerically, such as by contraction mapping.<sup>47</sup> Having done this, the observed market shares  $S$  uniquely determine the means of consumer utility for each good.

In particular, if one normalizes to zero the mean utility of the outside good  $\delta_0$ , as well as take logs, one obtains the following expression for  $\delta_j$ , given the predicted market shares of the logit model presented in (4):

$$\ln(S_j) - \ln(S_o) = \delta_j \equiv x_j\beta - p_j\alpha + \xi_j. \quad (13)$$

Thus,  $\delta_j$  is uniquely identified from a simple algebraic calculation using observed market shares. Given the simple linear model derived above, one can estimate  $\alpha$  and  $\beta$  in (13) with ordinary least squares, by regressing  $\delta_j$  on  $(x_j, p_j)$ , as well as deal with the potential endogeneity of price using standard linear instrumental variables techniques.

Under the nested logit model, it is necessary to determine an a priori grouping of banks that are believed to be correlated in terms of the consumer preferences over them. I divide here banks into those that are geographically diversified, by operating in more than one state, and those that have branches in exactly one state. This grouping is reasonable in light of the almost bimodal distribution of banks' geographic presence, since banks that operate in no

---

<sup>47</sup>See Pakes (1986).

more than one state tend to be in only a single local market, while those that have presence in more than one state tend to operate in many local markets within each state. Note that while nationwide branching deregulation began in 1994, even as early as 1993 (the beginning of the sample), most states had agreements with neighboring states that allowed for banks to cross state borders. Given such grouping, under the nested logit model expression (13) becomes:

$$\ln(s_j) - \ln(s_o) = x_j\beta - \alpha p_j + \sigma \ln(\bar{s}_{j/g}) + \xi_j \quad (14)$$

where  $\ln(\bar{s}_{j/g})$  represents the market share of product  $j$ , which belongs to group  $g$ , as a fraction of the total group share. This term is clearly endogenous and as a result, instruments are necessary to obtain a consistent estimate of  $\sigma$ . As indicated in Berry (1991), the characteristics of other firms in the group might be used for such purpose.

### C. Instruments

Under a unique equilibrium for all possible parameter and data values, the first order conditions (9) and (10) imply for each bank  $j$  the “reduced-form” price  $p_j^{RF}(z, \xi, \omega, \theta)$ , where price is either service fees or interest rate on deposits. What is important here is that prices are a function of the unobserved product characteristics  $\xi$ . Though the researcher might not know the values of  $\xi$ , it is reasonable to assume that banks do know them, in particular when they set prices, as well as consumers when they make decisions. Prices, as a result, are likely to be correlated with them.

Unobserved bank characteristics are variables such as the bank’s service quality, reputation for customer service, prestige, expertise, and even reputation aspects related to its soundness as a financial institution and ability to counter systemic financial distress. The correlation of prices and the unobserved characteristics will require treatment of endogeneity for consistent estimation; that is, instrumental variables (IV) estimation will have to be used.

Assume that the demand unobservable is mean independent of both observed bank characteristics and cost shifters. Letting both observed bank characteristics  $x$  and cost shifters  $w$  enter the matrix of instruments  $z$ , one has, formally:

$$E[\xi_j|z] = 0. \tag{15}$$

Price and quantity are determined jointly, while I take as given and do not model the determination of bank characteristics and costs.

The main set of instruments will be derived from actual cost data.

### *Cost shifters*

In the set of cost shifters, I include direct cost variables related to operating costs, as well as credit risk, product mix and balance sheet structure variables. The direct cost variables<sup>48</sup> included here represent all three types of operating costs: labor, expenses on premises and fixed assets, and “other” operating expenses. There are plausible situations in which these instruments might fail the independence assumption. Suppose that the cost variables actually include a hidden quality component. For instance, salaries might be a proxy for service quality, since more skilled employees will tend to be more expensive. Then salaries would be correlated with unobserved quality, and the independence assumption would be violated. To explore this possibility, I use an alternative measure of labor costs by constructing for each bank a measure of “potential labor costs,” based on the wage rate resulting from the weighted average over all markets in which the bank operates (where the weight is the bank’s deposit share in each market out of its total deposits). The other two direct cost variables are expected to have no significant quality component as they are costs such as electricity

---

<sup>48</sup>The banking industry is a service sector, and the production of its products is intensive in labor. However, it is an information-based industry, and as a result it also relies heavily on information technology equipment and software. In fact, with the evolution of technology, banks continue to introduce less labor-intensive systems in their provision of services, such as ATMs and “supermarket” branches. Other bank inputs are fixed assets such as branches and other materials involved in the provision of these services.

and other utilities, building maintenance, legal fees, insurance, amortization of assets, and ordinary repairs.<sup>49</sup>

Credit risk, another cost shifter, is represented by loan provisions,<sup>50</sup> which are reserves constructed based on the bank management's evaluation of its loan portfolio risk, and also following regulatory guidelines. Loan provisions represent the percentage of loans that the bank expects to become unrecoverable. While credit risk might be thought of as being mainly a cost related to loans, banks might cross-subsidize loans by shifting some of this cost towards other products.<sup>51</sup>

As additional cost shifters, I include product mix variable ratios and indicators for diverse balance sheet items, related to liquidity, loans, and net worth, as well as indicators for whether the bank operates in any rural markets or belongs to a banking holding company. In particular, I incorporate the following balance sheet items, adjusted by assets: cash, federal funds and securities ("liquidity"), real estate and loans to individuals, and net worth. One would expect these indicators to introduce further nuance about the bank's cost function. Banks hold very diverse product portfolios, and these might have different effects on costs. For instance, banks that belong to a banking holding company may systematically have different marginal costs from those that operate independently, as well as those banks

---

<sup>49</sup>Expenses on premises include items such as depreciation and amortization charges against assets; ordinary repairs to premises and maintenance; insurance expense related to the use of premises; equipment and furniture; property taxes; utilities. "Other" expenses include travel and expense allowances; insurance; Federal Deposit Insurance, Financing Corporation (FICO), and Comptroller of the Currency assessment expense; legal fees; net losses (gains) from the sale or other disposal of assets; amortization of intangible assets; data processing and advertising.

<sup>50</sup>There is an alternative measure of credit risk, which is ex-post risk represented by charge-off losses. This measure, however, while it guarantees exogeneity, is less correlated with price because it captures realized deviations from expected risk, and should therefore not be taken into account in the bank's calculation of risk.

<sup>51</sup>We expect risk to be a choice variable of the bank prior to price determination. It might be possible for future credit risk to increase due to an increase in price given the adverse selection mechanism that could operate under higher interest rates. However, as rational agents, banks should be aware of this potential problem when pricing, and therefore it is very unlikely such occurrence would take place. See Stiglitz and Weiss (1981) where banks respond to this problem by rationing credit.

One situation in which credit risk would not be an appropriate instrument is if it stands for a particular type of specialization that consumers find valuable. It is unlikely that credit risk is of significance to potential depositors, given that deposits are insured in the U.S. up to \$100,000 by the FDIC for member banks (the sample contains only FDIC insured banks). A more subtle story where the exogeneity assumption would be violated is for the case of banks engaging in mass marketing of specific consumer risk types. Even though consumers might not be aware of the bank's risk portfolio, prices could still be correlated with the unobserved demand component. This occurrence, however, is not expected to be prevalent.

that have presence in rural markets. Banks might also differ in their management talent which is responsible for determining how much liquidity to carry in the bank's portfolio, and how large the bank's capitalization should be, and these factors might cause banks to confront different marginal costs of production. One would expect these to be bank choice variables, econometrically exogenous to the pricing decision, and therefore uncorrelated with unobserved valuation.

Suppose however that the specialization of a bank in certain types of activities affects demand. In this case, the assumption of independence would be violated. However, most of the variables included in the IV set are absolutely unrelated to product specialization, such as cash, liquidity holdings, and bank capitalization. Furthermore, it is not very likely that consumers seeking a deposit services provider care and/or are aware about the bank's loan portfolio structure. The rest of the variables mentioned above should have little meaning from the perspective of a potential depositor.

#### *BLP instruments*

The common instrumental variables approach in the discrete choice literature, based on the work by Berry, Levinsohn and Pakes (1995) (henceforth BLP), has been to use the characteristics of other products in the market as instruments for price. I refer to these as BLP instruments. This relies on the assumption that product characteristics other than price are exogenous, and therefore orthogonal to unobserved demand. Given the location of products on the characteristics space, price will be correlated with the characteristics of other products, as can be seen in equations (9) and (10). The argument, derived from the oligopoly pricing model presented earlier, is that products that have close substitutes will have lower markups while other products located further away from rival ones will have higher prices



relative to cost.<sup>52</sup>

### *Instruments for the correlation parameter*

As mentioned in section B.1, the nested logit model, while adding flexibility to the demand model, requires additional instruments for the identification of the additional parameter  $\sigma$ . In lack of a better instrument, I use here the characteristics of other banks in the nest, as usually done in the literature and discussed in the section of BLP instruments. In particular, I use the branch density, branch personnel, bank age, and size of other banks in the group as instruments for the within-group share of a given bank.

## **D. Discussion of firm choice**

While the decision of the consumer has already been outlined in detail, as well as the necessary instruments given the unobservable and observable variables, it might be worth discussing in more depth the bank choice in terms of some of the bank characteristics that affect consumer decisions and therefore enter the specification here.

While clearly all bank characteristics are eventually endogenous and are a bank's choice variables, it may be useful to analyze the process through which they change, and whether that could potentially affect the validity of the specifications used in the current analysis. While the need to instrument price has already been analyzed, there are some bank characteristics that enter the model specification that could make the analysis less valid.

---

<sup>52</sup>See Bresnahan (1987) for further details on the development of this concept. Under certain properties of exchangeability of  $\xi$ , the first order basis function associated with characteristic  $z_{jk}$ , other than itself, is

$$\sum_{r \neq j, r \in J} z_{rk}. \quad (16)$$

Exchangeability of the demand and pricing function (or partial exchangeability of the latter) is required so that the function of  $\xi$  does not depend on the ordering of rival products but just on their characteristics. For each demand characteristic taken to be exogenous we compute the term in the equation above to use as an instrument. For instance, if the characteristic is branch density of bank  $j$ , the instrument vector for bank  $j$  includes the branch density of bank  $j$  as well as the sum of the branch density across rival banks in the market under consideration. These two terms times the number of exogenous characteristics gives the number of moment conditions (see BLP and the references therein for more details).

In particular, the bank characteristics related to size such as branch density, and whether the bank is in the large or medium sized category, have been assumed to play a part in the consumer decision and to be exogenous in the specification.

In terms of the branch density, the underlying assumption for including it is that branch density is not driven by the bank's local market shares, but rather, is the bank's choice. For instance, the bank sets a growth target in local markets or decides to be a high-quality bank that offers a number of branches to its consumers, regardless of its local market shares. This appears to be a reasonable assumption, both in light of the anecdotal evidence and economic intuition.

The evidence suggests that banks open branches to attract new customers and therefore increase their market share, as opposed to responding to their existing client base needs. Branches can affect a bank's demand by attracting new customers that might have otherwise chosen another bank. That is, new consumers choose a bank because it has opened up a new branch, as opposed to banks putting a branch in an area where they have already acquired a certain number of customers. There is plenty of anecdotal evidence about how banks hope to woo customers using their branches, usually with stylish merchandising and customer service.<sup>53</sup> In fact, branches are to banks a form of advertising itself, as banks become more visible to consumers through them.<sup>54</sup> Other pieces of evidence on banks' expansion mechanisms suggest that some banks use the "hub and spoke" expansion model, opening branches in other retail stores, because they cost less and provide more exposure (from Memphis Business Journal, April 2002).

While the above might be fairly obvious, it may be important to highlight it here, since under a different scenario about bank branch expansion, the branch component could bias the results. If the growth process of bank size is one where banks open branches in order to gain market share, and as a result become "larger" banks, it is then meaningful and appropriate to include branch density in the specification without the need of instrumenting

---

<sup>53</sup>As part of this strategy, many banks use soothing colored-interiors, and have even tried installing coffee shops and "investment bars" within their branches (in "Bank branches take a page from retail's book," San Francisco Business Times, Sept. 2001).

<sup>54</sup>For instance, many banks put clocks in their branches to get consumers to notice them.

for it, as banks will be perceived as higher utility banks once they are big and have many branches. Presumably, these banks become large because they provide higher quality from the perception of the consumer, with the number of branches playing a role in this perception of service quality.

In terms of the dichotomous variables for size, that is, whether the bank is large or medium in terms of assets, the underlying assumption for including them is that while a bank may be large as a whole, the market share it has in market  $t$  has no feedback effect, within the relevant time range, on which size category the bank belongs to. This assumption is plausible for many reasons.<sup>55</sup> First, the variable is actually not perfectly correlated with assets, since it is defined using the Federal Reserve Board's definition on what form the bank has to report, which while correlated with size, is not updated frequently. In fact, there is also almost no variation in terms of which size category a bank belongs to across the years in the sample, as only 17 percent of banks ever change category in the sample. Moreover, while there is variation in terms of the market shares that a given bank has across the local markets it serves, size is defined at the bank level (and in terms of assets as opposed to deposits). Again, the reason for including this variable is not capturing size itself, but rather, because it is expected to capture certain features associated with larger banks, such as service infrastructure, diversity of products, and expertise. In fact, not controlling for these variables probably worsens the endogeneity of prices and adds omitted variables bias.

---

<sup>55</sup>Given the nature of the size variable, it is worth to estimate the model without it in order to determine whether it affects results. I will, in fact, drop it from the analysis to do so, and as will be seen later, it does not appear to change the results, and therefore I choose to keep both the big and medium variables as part of the final specification.

## IV. Results

Table II<sup>56</sup> presents results for all banking markets, while Table III shows a similar set of results for urban banking markets alone, to allow for the possibility that the latter might be significantly different from rural markets. All columns correspond to the logit model, except for column (vi) which is based on the nested logit model. Columns (i) and (ii) show the OLS results, while the rest of the columns show the IV estimation results, where the price variables, service fees and the deposit rate, are instrumented for. All standard errors are adjusted for heteroskedasticity and for the fact that there may be correlation between observations of the same bank. All regressions include state and year effects.

### A. Logit specification

The logit specification reduces the model to the estimation of  $\ln(s_j) - \ln(s_o)$  on prices and product characteristics. Columns (i), (iii) and (iv) show results for the basic specification based on prices and observed bank characteristics. Both the OLS and the IV model show that consumers respond to prices, as the coefficients are highly significant and of the expected sign. Service fees enter utility negatively while the interest rate received by consumers on their deposits enters positively.<sup>57</sup> All the observed bank characteristics are highly significant

---

<sup>56</sup>As discussed earlier, some markets have their market shares add up to a number larger than one. The results shown here drop these markets from the sample, given the lack of prior knowledge about what the  $\lambda$  factor should be in determining potential market size. However, I have elsewhere estimated the specifications including those markets, by multiplying the market size by a factor  $\lambda > 1$ , as has been the practice in the literature whenever necessary. Indeed, results do not appear to be sensitive in any significant manner to the exclusion of these markets from the estimation, nor to the specific  $\lambda$  factor used to increase the potential market size. There are only 11 MSA markets that have to be removed under this methodology accounting for all years (in 1999, for instance, only two MSA markets are removed: San Francisco, CA and Wilmington-Neward, DE-MD) and several rural county markets. Also note that results are not affected when known money centers such as New York, Chicago, San Francisco, Delaware, etc, are removed from the estimation.

<sup>57</sup>Prices may be computed in terms of end-of-period stock of deposits or in terms of the average stock. While I use the end-of-period stock here, results are robust to using the alternative measure (results not reported). Note also that adjusting the deposit rate for the risk free rate does not affect results.

Also, I introduce the interest rate on loans in the specifications to explore whether consumers respond to the general level of loan rates charged by the bank when choosing a deposit service provider (results not shown). Given that switching costs of dealing with a new bank should be nontrivial, consumers may take into account other product attributes when purchasing deposit services, foreseeing that they will likely purchase other services from the same bank. The results suggest that is a factor in the data. The coefficient on loan price is negative and very precisely estimated, while the rest of the results are not affected.

and appear to affect utility positively. This is also what one would expect, given that consumers should show non-satiation for these characteristics throughout the possible ranges. Specifically, consumers appear to value the staffing of branches, the number of branches per square mile in their locality,<sup>58</sup> the number of years the bank has been in operation (which might proxy for experience)<sup>59</sup>, as well as larger bank size and the geographic diversification of the bank as measured by the number of states in which the bank operates. Furthermore, in order to make sure that the size variables are not affecting the results, as mentioned earlier, I drop these variables from the specification to find that, indeed, they do not change the results. This can be appreciated in Appendix B, where results for various specifications are reported without the size variables.

The OLS fit is rather low, indicating that around 70 percent of the variation in mean utility is due to unobserved factors. As discussed earlier, one would not expect to be able to account for all the bank attributes consumers value. Given these unobservable bank characteristics, which should be known to market participants and therefore taken into account in the price setting process, I instrument for prices, since some of these unobserved components might be correlated with price. If prices are higher when unobserved quality is higher, for instance, one might not observe market share responses to higher prices. Column (iii) presents the first IV specification, based on cost shifters such as labor costs,<sup>60</sup> operating expenses on premises and fixed assets as well as other expenses, credit risk, and some product mix and balance sheet structure indicators which should contain further information about the bank's cost function.<sup>61</sup> The introduction of instrumental variables generates considerable increase

---

<sup>58</sup>As mentioned earlier, all of the explanatory variables, with the exception of branch density, are defined at the bank level and therefore do not vary across markets within a given year and bank. One might be concerned that the market specific deviation from mean demand characteristics embedded in the error term may be correlated with the market specific deviation contained in branch density. To account for this possibility, I reestimate the IV model by using the bank's average branch density in each year as opposed to the bank's market-specific branch density. I find that the results are not affected (results not reported).

<sup>59</sup>Note that capping age at 25, which is usually taken to be the "mature" age of a bank, provides similar results.

<sup>60</sup>I use the weighted average of the wage rate across a bank's markets to obtain a measure of labor costs. One could alternatively use the actual salaries paid by the bank. As discussed earlier, these might contain a quality component, as more skilled workers, which in turn might provide better service to consumers, should be more expensive to the bank. Results are actually not affected by the use of this latter measure, but market wages are preferred as a potentially cleaner measure of costs.

<sup>61</sup>I also obtain estimates by removing credit risk from the instrument set, in order to explore whether any of the results are being driven by it, and find that this is not the case. While the rest of the operating cost

in the absolute magnitude of the price coefficients and leaves the rest of the coefficients more or less unchanged, which should be the case under the assumption of exogeneity of observed bank characteristics other than price.<sup>62</sup>

The first stage regressions show a reasonable fit, with the exogenous variables explaining over 40 percent of the variation in prices (results shown in Appendix C). Almost all of the price instruments have highly significant coefficients. Costs of operation appear to be positively correlated with service fees, and negatively correlated with the rate paid by banks on deposits, as one would expect.

Column (iv) shows results from incorporating BLP-type instruments based on the observed bank characteristics,<sup>63</sup> as described earlier, to the first set of instrumental variables. The BLP instruments are based on the following characteristics of the rival banks in the local market: staffing and geographic density of branches, bank age, number of states of bank presence, and size. These instruments have no power for identification of prices by themselves (results not reported), and when added to the earlier instruments set, results are similar to those obtained before.

### *Demographics*

Columns (ii) and (v) show the results from adding income and population variables, in order to control for demographic differences across local markets. The results are insightful, and reinforce the idea of differentiation in the industry. In particular, I incorporate income per capita and population density, and introduce several interactions of these variables with bank characteristics. In the case of population, I only interact it with branch density. The idea is to explore whether consumers in more densely populated areas value relatively more

---

variables should be true cost shifters, “other” expenses does include advertising, which could be correlated with unobserved valuation. While it is presumably a small component of this cost item, which includes hundreds of other types of costs, I nevertheless drop the variable from the instrument set and the results continue to hold, even though the fit of the first-stage regressions decreases slightly (results not reported).

<sup>62</sup>Note that even though I have assumed the exogeneity of observed bank characteristics, the identification does not rely on it.

<sup>63</sup>While there is almost no variation within a given year for the characteristics of a bank (except for branch density in a local market), one can still obtain variation as the set of banks in each market varies, as well as from the time variation in the sample.

the number of branches in the local market, given that one might expect branches to get more crowded, as well as have longer lines and therefore higher waiting time for service.

Upon introducing demographic variables, the fit of the model increases considerably. Results are similar to those obtained earlier for both the OLS and the IV specifications, with the exception that now there are demographic variables interacted with observed bank characteristics, which appear to describe interesting features about consumer behavior. Richer consumers<sup>64</sup> appear to care less about the number of branches per square mile, the number of employees in each branch, and the age of the bank. Moreover, results suggest that these consumers value relatively more the geographic scope of the bank, as evidenced by the number of states in which the bank operates. This is what one would expect, given that wealthier consumers (both high-income individuals and large businesses) are thought to rely less on branches to carry out financial transactions with their bank. It also provides evidence confirming the presumption of the regulatory agencies about the scope of the relevant bank market being larger for higher income clients. Moreover, wealthier consumers appear to value more heavily larger-sized banks, also a reasonable result as one might expect bigger banks to be able to offer features that higher income clients may find particularly useful, such as wider product diversification and service expertise, in which larger banks might have a competitive advantage. Results also suggest that more populated areas are more sensitive to the local branch density, as the marginal effect of an additional branch in a densely populated market should be larger due to the higher crowding of branches.

## B. Nested logit

The nested logit model is more flexible than the logit, and while it does not allow for the complexity in demand patterns that the random coefficients provides, as pointed out by Berry (1994), it does allow for nontrivial interactions between product and consumer characteristics

---

<sup>64</sup>One might alternatively interpret this coefficient as a market specific effect. That is, consumers in wealthier markets behave differently than consumers in lower income markets. Clearly, if consumers in a market were homogeneous in terms of income, both interpretations would be identical.

and may be preferred when a large penalty is placed on computational complexity.<sup>65</sup> Column (vi) of Table II reports results from estimating the IV nested logit model. The basic results are qualitatively similar to those obtained earlier, with the exception of the size variable “big” which is no longer significant. Note that the variable on geographic diversification has now been dropped as an explanatory variable, since this is the criterion used in the nesting strategy. The correlation parameter  $\sigma$  appears to be close to one, indicating that consumer preferences across geographically diversified banks (as well as within the group of “local” banks), are highly correlated, suggesting that the nests are appropriate.<sup>66</sup>

The estimation of the  $\sigma$  parameter requires additional instruments, as discussed earlier. In particular, I use the BLP-type instruments, that is, the characteristics of other banks in the group, as an instrument for the within-market share of a given bank. A caveat is that these instruments might not be all that appropriate for the identification of this additional parameter. For instance, the BLP instruments did not have the power necessary to identify prices by themselves, as discussed earlier.

### C. Urban markets

Table III presents results for urban (MSA) markets only. Results are similar to those obtained earlier, though the fit of the model is higher.<sup>67</sup> The price coefficients are significant and of the expected sign. Once demographic characteristics are introduced, results are also similar, in that wealthier consumer value branch density and age relatively less than lower income consumers, and care relatively more about bank size. The nested logit results for urban

---

<sup>65</sup>I have tried estimating the random coefficients model, but the results are extremely sensitive to starting values. Furthermore, as convergence is sometimes not reached, the model with random coefficients might be too demanding of the data.

<sup>66</sup>Note that the specification with demographics is not estimated given the high correlation between demographic variables and the number of geographically diversified banks in a given banking market – which is the nesting criterion. In particular, market income and population show a strong and positive correlation with the proportion of banks that operate in more than one state, and as a result, the within-group market shares of all banks tend to be smaller.

<sup>67</sup>For instance, in the OLS specification (results not reported), almost 50 percent of the variation in market share appears to be explained by prices and observed bank characteristics alone.



markets are also similar to the previous findings, though the bank age or experience effect is no longer estimated precisely.

Insert Table II about here.

Insert Table III about here.

## D. Further specifications: Bank fixed effects

An alternative in trying to control for unobserved quality is the use of bank-specific dummy variables. Given data constraints, one might expect that many characteristics relevant to the consumer's decision will not be accounted for explicitly in the empirical model, and as a result one might want to consider the introduction of firm-level effects. Following Nevo (2001), and introducing a time subscript  $s$ , suppose one lets the error term be instead  $\xi_j + \Delta\xi_{jts} + \epsilon_{ijts}$ , which introduces firm dummy variables in order to control for the vertical component  $\xi_j$ . A bank dummy variable would capture the unobserved product characteristics that do not vary across market-year combinations within a given bank. Note that even when bank fixed effects are introduced, the time deviation from mean demand characteristics, which are correlated with price, would still be left in the error term. As a result, one still needs to instrument for this leftover component.

Columns (iv) and (v) in Appendix B show the results from introducing bank fixed effects in the logit model. The qualitative results are similar to those obtained earlier, with the exception of bank age which is no longer estimated precisely. The coefficients on prices are statistically significant and of the expected sign. However, their magnitudes are considerably smaller than those estimated in earlier specifications. When I instrument for prices, given that one might be concerned about the time deviation from mean demand characteristics still in the error term, as discussed earlier, the price coefficient increases for the deposit rate, but the coefficient on service charges remains smaller than unity.

There are several possible explanations for these results. One of them has to do with the within-bank variation relative to the across bank variation in the sample, since the fixed

effects regressions are identified from within-bank variation over time. This kind of variation is small in the sample, as evidenced by the analysis of variance. For instance, of the total variation accounted for by both across time and across bank variation in the sample, the time variation explains less than 5 percent. When a test is carried out for whether the bank fixed effects are jointly significant, the null hypothesis of all fixed-effects equal to zero cannot be rejected at any reasonable level of confidence.

A further explanation might be related to the short-run consumer response to prices. Under the existence of switching costs, which one would expect may be significant in banking,<sup>68</sup> consumers may not respond immediately to price changes. The fixed effects regression might be capturing this kind of response, given that they are identified through the within-bank variation.<sup>69</sup> On the other hand, the earlier specifications, which are weighted toward between-bank variation, may be indicative of consumers' behavior over a lengthier period of time. Results that indicate low consumer sensitivity to changes in service charges, especially in the short-run, fit well with the anecdotal evidence on the reluctance of consumers to switch deposit service provider.

## E. Elasticities and PCM

Table IV presents the distribution of price elasticities based upon the estimates obtained earlier for the logit and nested logit model, which facilitate the interpretation of the coefficient magnitudes. Under the logit model,<sup>70</sup> the median elasticity of service fees is -0.6, while the median deposit rate elasticity is 5.9. The elasticity of service fees, which is below unity, is clearly not profit maximizing: given this elasticity value, banks should respond by raising

---

<sup>68</sup>It is not difficult to argue that consumers find it costly to switch to another bank, given the time and inconveniences involved in such process. Furthermore, the development of a relationship between bank and customer, valuable to both parties involved, is known to be one of the most unique features of commercial banking.

<sup>69</sup>A more damaging interpretation might be related to the specification form used here.

<sup>70</sup>Under the logit model as defined in section 3.2, the price elasticities of the market shares defined by equation (4) are

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\alpha p_{jt}(1 - s_{jt}) & \text{if } j = k \\ \alpha p_{kt} s_{kt} & \text{otherwise.} \end{cases} \quad (17)$$

fees until they reach an elastic portion of the demand curve. However, there are plausible explanations for why banks might choose not to raise service fees. First, there is abundant anecdotal evidence about banks using low or zero-fee checking accounts as a loss leader, that is, as a way to attract and lock in consumers that will later on proceed to purchase other services offered by the bank. Second, banks may consider raising service fees as too conspicuous to both consumers as well as regulators. Yet another possibility is that the relevant price elasticity is simply that of the bundle of all deposit products.

Under the nested logit model,<sup>71</sup> the improvement in the distribution of elasticities is remarkable for the case of service fees, with less than 10 percent of the distribution below unity. The median elasticity is -3.5 for service fees, indicating that a one percent increase in fees would lead to a 3.5 percent decrease in market share. In terms of the deposit interest rate, the elasticities have almost doubled, with a median of 10.9. This suggests that a one percent increase in the interest rate paid on deposits would lead to an 11 percent increase in market share. The nested logit model, however, does require the use of additional instruments to estimate the correlation parameter.<sup>72</sup>

Insert Table IV about here.

Table V shows the mean elasticities for prices and bank attributes, for both the overall sample and urban markets only. A striking difference between urban and rural markets is that the former show a much lower responsiveness to both service fees and deposit rates, both under the logit and nested logit models. A possible reason for this is that urban consumers might face higher switching costs. The below-unity elasticities for urban markets might result from reasons already mentioned, such as loss leader checking accounts. The rest of the elasticities, for the bank attributes, are interpreted in a similar manner to the price elasticities: for instance, a one percent increase in the number of local branches of a bank would lead to an increase of less than one percent in that bank's market share.

---

<sup>71</sup>See Berry (1994) for a derivation of elasticities based on the nested logit model.

<sup>72</sup>While elasticity results vary across the logit and nested logit models, it is worth mentioning that qualitative results on consumer preferences and the results on consumer welfare change, shown later in the paper, are similar under both models.

Insert Table V about here.

Given the estimated demand parameters, and assuming the model of supply conduct introduced earlier, PCM may be calculated. Table VI shows the average PCM for both the logit and nested logit model, for account fees and the interest rate on deposits. While it is not meaningful to obtain a PCM for service fees under the logit model, the implied PCM for the interest rate on deposits is 17 percent. The nested logit model, by providing a distribution of elasticities that are mostly above one, allows for the computation of a PCM for service fees. The average PCM for account fees is 25 percent, while for the deposit interest rate is almost 10 percent. Furthermore, despite the dramatic changes in the industry, PCM appear to have changed little throughout the period. Comparing 1993 to 1999, there is an increase in the PCM of service fees from 24 percent to 27 percent, while the PCM on deposits stays around 10 percent.

Insert Table VI about here.

## F. Welfare effects and other implications

The finding that product differentiation is part of deposit demand, as consumers appear to value several bank attributes other than price, can have important consequences in terms of consumer welfare considerations. As seen earlier in Table I, while service fees have increased slightly since 1993, other bank attributes such as the number of branches (even controlling for population growth), geographic expansion, size and age, have also increased in magnitude. Under this setting, welfare effects might still be positive, even if prices increase.<sup>73</sup>

To explore this possibility, I carry out a simple calculation of welfare change in the sample period. Following Small and Rosen (1981), in the context of the discrete choice model, welfare effects of changes in the choice set between periods  $s$  and  $s - 1$  in some

---

<sup>73</sup>One could alternatively adjust the deposit rate for the risk free rate, but that would require adjusting for the maturity of the deposit portfolio of each bank. A back-of-the-envelope calculation indicates that results are not affected much by this adjustment, even though the return over the risk free rate has decreased from 1993 to 1999.

market  $t$  are measured as the expected equivalent variation ( $EV$ ) of the changes. The latter is defined as the amount of money that would make consumers indifferent, in expectation, between facing the two choice sets. Then, one has that

$$EV = S_s(p_t, x_t; \theta_D) - S_{s-1}(p_t, x_t; \theta_D) \quad (18)$$

where  $S(p, x; \theta_D) = \ln[\sum_j^J \exp(\delta_j(p_j, x_j; \theta_D))]$ .

Taking the average over the expected equivalent variations across markets both in 1993 and in 1999, I find that there is a consumer gain of \$1.60 throughout the period. Table VII shows some percentiles for the welfare change in local markets throughout 1993-1999. Looking at the distribution of expected variations across markets, I find that 93 percent of local markets have experienced increases in benefits. For instance, out of a total of 301 urban markets, 289 show positive variations. Under the nested logit model, results on consumer welfare are similar, though slightly lower in terms of the increase in welfare. According to the distribution of expected variations from this model, 78 percent of local markets have experienced increases in benefits. Furthermore, there is almost zero correlation between the welfare change in a market and its HHI. This type of exercise is at least suggestive of the kind of bias that might arise in welfare inferences based solely on prices and concentration measures.

Insert Table VII about here.

While markets are assumed here to be geographically local, the results might still provide insight into the relevant market definition, which is a central issue to regulatory agencies in terms of designing regulation, investigating potential antitrust cases and carrying out merger analysis. In particular, local market bank variables, such as branch density, appear to be highly significant for the consumer decision. Presumably, this is a sign that banking markets continue to be geographically local, at least for the bulk of consumers. In fact, the generally good fit of the model, which is defined at the level of the MSA and rural county, might be a fair indication that the relevant geographic market is demarcated appropriately.

The demand estimates are also useful in analyzing the supply side, in terms of the implied bank incentives for horizontal differentiation.<sup>74</sup> For example, given consumer tastes for branches, over time banks might find it profitable to increase their branch networks either through direct investment or through mergers that would enhance this attribute. Also, they might want to increase the staffing of their branches. Furthermore, banks may strategically decide to merge with neighboring state banks, as consumers value geographic diversification. This is an interesting implication, since the antitrust authorities have tended to be more permissive when it comes to allowing mergers between banks whose markets do not overlap. In fact, once geographic restrictions were lifted in the nineties, banks appear to have responded to these incentives. Not only there has been a dramatic merger wave in the industry following deregulation, but banks have now more branches in each local market, cover more local markets, and have presence in a larger number of states.

## V. Concluding remarks

This paper estimates the demand for deposit services in the U.S. commercial banking industry. The results provide insight on consumer behavior in choosing a deposit institution, as consumers are found to respond not only to prices but to several bank attributes as well. In particular, consumers are sensitive to both account fees and deposit interest rates. Consumer demand responds positively to the staffing and geographic density of local branches, as well as to the age, size and geographic diversification of banks. There are also important differences across consumer and market types in the demand for banking services. Higher income consumers appear to value location characteristics relatively less than do lower income consumers, and bank size relatively more. In terms of demand across market types, urban markets appear less responsive to prices than rural markets.

Understanding the form of demand and consumer behavior in banking has several immediate uses. The use of a structural model of demand, which incorporates product dif-

---

<sup>74</sup>While observed bank attributes are assumed exogenous in the econometric specification, over time banks are expected to change the characteristics of their services.

ferentiation, provides utility derived elasticities, as well as a framework for the analysis of policy. In particular, I find that the median of the distribution of implied price elasticities is -0.6 for the case of service fees and 5.9 for the interest rate on deposits. The finding of a service fee elasticity below unity for most banks is interpreted as resulting from the use of checking accounts as loss leaders, among other reasons. The introduction of the nested logit model improves the distribution of price elasticities by shifting it to the right, with a median service fee elasticity of -3.5 and a median deposit rate elasticity of 10.9. The implied PCM for service fees and deposit rates show an average of 25 and 10 percent, respectively, and change little from 1993 to 1999. Furthermore, the estimates of consumer preferences across bank characteristics can be used to analyze the effects of recent banking industry consolidation on consumers, which has tended both to reduce the number of banks and increase the number of local branches in many markets. Throughout the sample period, 1993-1999, most markets appear to have experienced a slight welfare gain, with a median consumer gain of \$1.16. I find that this kind of exercise is at least suggestive of the type of bias that might arise in welfare inferences based solely on prices and concentration measures. It is also suggestive of the kinds of data that are required to analyze the benefits to consumers due to regulatory change and for other related empirical research in the industry. In terms of the recent deregulation of banks' geographic diversification, the results indicate that the current regulatory framework might be appropriate, at least on the deposit side (since the scope of this work leaves out the loan side). However, within the current regime, it might be useful to consumers to require banks to publish detailed account fee data to provide greater transparency of bank fees.

As possible extensions, the analysis provides a framework for counter-factual exercises related to merger effects. In order to do this, however, the supply side has to be carefully analyzed and be jointly estimated with demand. In the banking industry, this would be particularly interesting in light of the extensive applied research that has been carried out in banking cost function estimation. While there is usually very little prior knowledge on the technical characteristics of the production side, the banking industry enjoys a wealth of results and methods from the empirical literature on cost functions.<sup>75</sup>

---

<sup>75</sup>See Berger and Mester (1997) and the references therein.

Another possible extension is to use this framework to study the prevalence of consumer switching costs in the industry, which one would expect to be high, at least for some services. If in the future data become available that track individual consumers' balances, one could potentially answer questions related to this type of costs. One way to do this would be to use the time variation of the data to determine how much of a change in the mean utility value it takes for the consumer to switch banks. The reason why this might be promising in this setting is related to the specific nature of banking services. Unlike other industries where the consumer's purchase is a one point transaction, banks hold consumer balances that vary with transactions as well as the consumer's willingness to continue the relationship with her current bank. Lastly, this approach may also be used in the study of other banking services, such as loans.

## References

- [1] Adams, R.M., Roller, L. and R.C. Sickles (2002). "Market Power in Input and Output Markets: An Empirical Application to Banking." Federal Reserve Board, manuscript.
- [2] Amel, D.F. and T.H. Hannan (1999). "Establishing Banking Market Definitions through Estimations of Residual Deposit Supply Equations." *Journal of Banking and Finance*, 23:1667-1690.
- [3] Amel, D.F. and M. Starr-McCluer (2001). "Market Definition in Banking: Recent Evidence." *Board of Governors of the Federal Reserve System, Division of Research and Statistics*, Washington, D.C..
- [4] Anderson, S.P. and A. de Palma (1992). "Multiproduct Firms: A Nested Logit Approach." *Journal of Industrial Economics*, 3:261-276.
- [5] Anderson, S., A. de Palma and J. Thisse (1992). *Discrete Choice Theory of Product Differentiation*. Cambridge: MIT Press.



- [6] Berg, S.A. and M. Kim (1998). "Banks as Multioutput Oligopolies: An Empirical Evaluation of the Retail and Corporate Banking Markets." *Journal of Money, Credit and Banking*, 30:135-153.
- [7] Berger, A. (1995). "The Profit-Structure Relationship in Banking - Tests of Market-Power and Efficient-Structure Hypotheses." *Journal of Money, Credit, and Banking*, 27:404-31.
- [8] Berger, A. and T.H. Hannan (1989). "The Price-Concentration Relationship in Banking." *Review of Economics and Statistics*, 71:291-299
- [9] Berger, A.N, A.K. Kashyap and J.M. Scalise (1995). "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been." *Brooking Papers on Economic Activity*, 2:55-218.
- [10] Berger, A.N. and L.J. Mester (1997). "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?" *Journal of Banking and Finance*, 21:895-947.
- [11] Berry, S.T. (1994). "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, 25:242-262.
- [12] Berry, S., J. Levinsohn and A. Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, 63:841-890.
- [13] Bresnahan, T.F. (1987). "Competition and Collusion in the American Automobile Industry: The 1955 Price War (in Oligopoly Pricing)." *Journal of Industrial Economics*, 35:457-482.
- [14] Bresnahan, T.F. (1989). "Empirical Studies of Industries With Market Power." *Handbook of Industrial Organization*, R. Schmalensee and R.D. Willig., eds.. Amsterdam:Elsevier Science Publishers B.V. Vol. II, Chapter 17.
- [15] Cardell, N.S. (1991) "Variance Components Structures for the Extreme Value and Logistic Distributions." Washington State University, mimeo.

- [16] Elliehausen, G.E. and J.D. Wolken (1990). "Banking Markets and the Use of Financial Services by Small and Medium-Sized Businesses." *Federal Reserve Board of Governors Staff Studies*, 160.
- [17] Elliehausen, G.E. and J.D. Wolken (1992). "Banking Markets and the Use of Financial Services of Households." *Federal Reserve Bulletin*, 76:801-817.
- [18] Fisher, F.M. and J.J. McGowen (1983). "On the Misuse of Accounting Rates of Return to Infer Monopoly Profits." *American Economic Review*, 73:82-97.
- [19] Gelfand, M.D. and P.T. Spiller (1987). "Entry Barriers and Multiproduct Oligopolies." *International Journal of Industrial Organization*, 5:101-113.
- [20] Hannan, T.H. (1991). "Bank Commercial Loan Markets and the Role of Market Structure: Evidence from Surveys of Commercial Lending." *Journal of Banking and Finance*, 15:133-149.
- [21] Hannan, T.H. (1997). "Market Share Inequality, the Number of Competitors, and the HHI: An Examination of Bank Pricing." *Review of Industrial Organization*, 12:23-35.
- [22] Hannan, T.H. and A.N. Berger (1991). "The Rigidity of Prices: Evidence from the Banking Industry." *American Economic Review*, 81:938-945.
- [23] Hannan, T.H. and J.N. Liang (1993). "Inferring Market Power from Time-Series Data: The Case of the Banking Firm." *International Journal of Industrial Organization*, 11:205-218.
- [24] Hannan, T.H. and P.E. Strahan (2000). "Are Bank Loan Markets Really Local?" Federal Reserve Board of Governors, unpublished manuscript.
- [25] Hausman, J., G. Leonard and J.D. Zona (1994). "Competitive Analysis with Differentiated Products." *Annales D'Economie et de Statistique*, 34:159-180.
- [26] Kwast, M., M. Starr-McCluer and J. Wolken (1997). "Market Definition and the Analysis of Antitrust in Banking." *The Antitrust Bulletin*, XLII No. 4:973-995.

- [27] Lancaster, K.J. (1971). *Consumer Demand: A New Approach*. New York: Columbia University Press.
- [28] McFadden, D. (1973). "Conditional Logit Analysis of Qualitative Choice Behavior." *Frontiers of Econometrics*. Ed. by P. Zarembka. New York: Academic Press.
- [29] McFadden, D. (1978). "Modeling the Choice of Residential Location." *Spatial Interaction Theory and Planning Models*. Ed. by A. Karlqvist, et al. Amsterdam: North-Holland.
- [30] McFadden, D. (1981). "Econometric Models of Probabilistic Choice." *Structural Analysis of Discrete Data*. Ed. by C. Manski and D. McFadden. Cambridge: MIT Press.
- [31] Nevo, A. (2001). "Measuring Market Power in the Ready-To-Eat Cereal Industry." *Econometrica*, 69:307-342.
- [32] Pakes, A. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*, 54:755:784.
- [33] Petersen, M.A. and R.G. Rajan (2002). "Does Distance Still Matter? The Information Revolution in Small Business Lending." *Journal of Finance*, forthcoming.
- [34] Radecki, L.J. (1998). "The Expanding Geographic Reach of Retail Banking Markets." *Federal Reserve Bank of New York Economic Policy Review*, Vol. 4, No 2.
- [35] Rhoades, S.A. (1992). "Evidence on the Size of Banking Markets from Mortgage Rates in Twenty Cities." *Federal Reserve Board Staff Studies*, 162.
- [36] Shaffer, S. (1993). "A Test of Competition in Canadian Banking." *Journal of Money, Credit, and Banking*, 25:49-61.
- [37] Small, K.A. and H.S. Rosen (1981). "Applied Welfare Economics with Discrete Choice Models." *Econometrica*, 49:105:130.
- [38] Stiglitz, J. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information." *American Economic Review*, 71:393-410.

[39] Tirole, J. (1997). The Theory of Industrial Organization. Cambridge: MIT Press.

**Table I**  
**Market Average for Bank Prices and Characteristics: 1993 vs 1999**

Each entry is calculated as a simple average across all banking markets. See Appendix A for a description of the variables. Source: Federal Reserve Report on Condition and Income; U.S. Census; Bureau of Economic Analysis.

Variable	1993	1999
<i>Service fees</i>	0.538%	0.565%
<i>Deposit interest rate</i>	3.06%	3.40%
<i>Number of employees per branch</i>	17	16
<i>Branch density</i>	0.0037	0.0039
<i>Age of bank</i>	72	80
<i>Number of states of bank's operations</i>	1.0	2.3
<i>Big (1 = yes)</i>	25.7%	38.2%
<i>Medium (1 = yes)</i>	16.5%	20.8%

**Table II**  
**Estimation Results: All Banking Markets**

The logit model consists of the estimation of the log difference of bank  $j$ 's market share  $s_j$  and the outside good's market share  $s_o$  ( $\ln(s_j) - \ln(s_o)$ ) on prices and product characteristics; the nested logit model is similar but also includes, as an explanatory variable, the market share of bank  $j$ , which belongs to group  $g$ , as a fraction of the total group share. State and year effects are included in all specifications. Estimated standard errors, robust and corrected for within bank dependence, are in parentheses. See text and Appendix A for description of variables and instrument sets.

Explanatory Variable	OLS		LOGIT		NESTED	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
<i>Service fees</i>	-91.038 (5.397)**	-75.602 (4.402)**	-139.675 (18.729)**	-120.638 (18.537)**	-167.229 (13.232)**	-64.834 (6.359)**
<i>Deposit rate</i>	31.367 (1.745)**	15.256 (1.391)**	198.222 (9.736)**	204.980 (9.632)**	67.641 (7.089)**	29.402 (3.245)**
<i>Employees per branch</i>	0.001 (0.000)**	0.012 (0.004)**	0.001 (0.000)*	0.001 (0.000)*	0.011 (0.004)**	0.001 (0.000)**
<i>Branch density</i>	42.366 (5.371)**	121.935 (14.313)**	44.789 (5.823)**	44.675 (5.812)**	121.775 (15.228)**	11.699 (1.399)**
<i>Bank's age</i>	0.011 (0.001)**	0.012 (0.001)**	0.011 (0.001)**	0.011 (0.001)**	0.010 (0.001)**	0.001 (0.000)**
<i>Number of states</i>	0.037 (0.017)*	-0.071 (0.021)**	0.114 (0.024)**	0.112 (0.024)**	-0.024 (0.025)	
<i>Big</i>	0.639 (0.047)**	-1.951 (0.129)**	1.016 (0.064)**	1.010 (0.064)**	-1.462 (0.151)**	-0.025 (0.024)
<i>Medium</i>	0.329 (0.032)**	-1.017 (0.108)**	0.948 (0.054)**	0.974 (0.054)**	-0.707 (0.117)**	0.115 (0.020)**
<i>Employees * Income</i>		-0.003 (0.001)**			-0.003 (0.001)*	
<i>Branch density * Income</i>		-19.458 (2.561)**			-19.551 (2.749)**	
<i>Age * Income</i>		-0.004 (0.001)**			-0.003 (0.001)**	
<i>No. of states * Income</i>		0.037 (0.010)**			0.035 (0.010)**	
<i>Big * Income</i>		1.287 (0.062)**			1.141 (0.071)**	
<i>Medium * Income</i>		0.710 (0.050)**			0.647 (0.052)**	
<i>Income</i>		-3.706 (0.134)**			-3.526 (0.141)**	
<i>Income squared</i>		0.398 (0.028)**			0.375 (0.029)**	
<i>Population * Branch density</i>		0.347 (0.149)*			0.496 (0.161)**	
<i>Population density</i>		-82.324 (3.482)**			-77.423 (3.718)**	
$\ln(\bar{s}_j/g)$						0.927 (0.006)**
Adj. R-Sq. / Overident. test	0.29	0.51	Accept	Accept	Accept	Accept
First-stage R-sq. (Service fees)			0.40	0.41	0.41	0.40
First-stage R-sq. (Deposit rate)			0.43	0.43	0.43	0.42
Instruments			Cost	Cost+BLP	Cost	Cost
Observations	103380	102949	103360	103360	102929	103438

\*significant at 5%; \*\* significant at 1%.

**Table III**  
**Estimation Results: Urban Banking Markets Only**

The logit model consists of the estimation of the log difference of bank  $j$ 's market share  $s_j$  and the outside good's market share  $s_o$  ( $\ln(s_j) - \ln(s_o)$ ) on prices and product characteristics; the nested logit model is similar but also includes, as an explanatory variable, the market share of bank  $j$ , which belongs to group  $g$ , as a fraction of the total group share. State and year effects are included in all specifications. Estimated standard errors, robust and corrected for within bank dependence, are in parentheses. See text and Appendix A for description of variables and instrument sets.

Explanatory Variable	LOGIT				NESTED	
	OLS		IV	IV	IV	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
<i>Service fees</i>	-27.442 (5.199)**	-27.238 (4.883)**	-64.069 (13.106)**	-65.891 (13.244)**	-71.388 (11.257)**	-28.415 (5.653)**
<i>Deposit rate</i>	8.201 (1.906)**	1.622 (1.698)	49.740 (7.123)**	55.520 (7.126)**	15.310 (6.366)*	10.370 (3.275)**
<i>Employees per branch</i>	0.002 (0.000)**	0.006 (0.002)*	0.002 (0.000)**	0.002 (0.000)**	0.005 (0.002)*	0.001 (0.000)**
<i>Branch density</i>	39.043 (6.697)**	75.439 (13.573)**	40.507 (7.055)**	40.650 (7.094)**	75.431 (14.038)**	16.815 (2.787)**
<i>Bank's age</i>	0.004 (0.001)**	0.010 (0.002)**	0.004 (0.001)**	0.004 (0.001)**	0.008 (0.002)**	0.001 (0.000)
<i>Number of states</i>	0.056 (0.012)**	-0.011 (0.032)	0.080 (0.015)**	0.082 (0.015)**	0.011 (0.033)	
<i>Big</i>	1.826 (0.064)**	0.909 (0.197)**	1.879 (0.067)**	1.886 (0.068)**	1.042 (0.202)**	0.547 (0.031)**
<i>Medium</i>	0.912 (0.036)**	0.527 (0.158)**	1.035 (0.046)**	1.054 (0.046)**	0.568 (0.162)**	0.331 (0.026)**
<i>Employees * Income</i>		-0.001 (0.001)			-0.001 (0.001)	
<i>Branch density * Income</i>		-10.416 (2.743)**			-10.349 (2.825)**	
<i>Age * Income</i>		-0.003 (0.001)**			-0.003 (0.001)**	
<i>No. of states * Income</i>		0.021 (0.013)			0.019 (0.013)	
<i>Big * Income</i>		0.335 (0.074)**			0.290 (0.076)**	
<i>Medium * Income</i>		0.166 (0.060)**			0.161 (0.061)**	
<i>Income</i>		-3.215 (0.195)**			-3.145 (0.197)**	
<i>Income squared</i>		0.483 (0.033)**			0.468 (0.034)**	
<i>Population * Branch density</i>		0.924 (0.184)**			0.994 (0.191)**	
<i>Population density</i>		-55.581 (2.893)**			-54.487 (2.985)**	
$\ln(\bar{s}_j/g)$						0.683 (0.013)**
Adj. R-Sq. / Overident. test	0.47	0.56	Accept	Accept	Accept	Accept
First-stage R-sq. (Service fees)			0.37	0.37	0.38	0.36
First-stage R-sq. (Deposit rate)			0.41	0.41	0.41	0.41
Instruments			Cost	Cost+BLP	Cost	Cost
Observations	41498	41498	41481	41481	41481	41481

\*significant at 5%; \*\* significant at 1%.

**Table IV**  
**Price Elasticity Percentiles**

The entries correspond to the indicated percentiles of the distribution of price elasticities across banks, based on the estimation results shown in Table II, columns (iii) and (vi).

Price		10%	25%	Median	75%	90%
<i>Service fees</i>	Logit	-1.223	-0.936	-0.639	-0.417	-0.263
	Nested	-7.143	-5.188	-3.469	-2.140	-1.151
<i>Deposit rate</i>	Logit	3.995	4.969	5.941	6.886	7.704
	Nested	5.140	8.233	10.904	13.299	15.163



**Table V**  
**Elasticity Means**

The entries represent the mean elasticity across banks for the indicated bank characteristic, based on the estimation results shown in Table II and Table III, columns (iii) and (vi), for all markets and urban markets, respectively.

Variable	All markets		Urban markets	
	Logit	Nested	Logit	Nested
<i>Service fees</i>	-0.719 (-0.88,-0.55)	-3.975 (-4.69,-3.26)	-0.367 (-0.49,-0.24)	-0.492 (-0.67,-0.31)
<i>Deposit rate</i>	5.859 (5.27,6.45)	10.498 (8.02,12.97)	1.491 (1.04,1.93)	0.946 (0.29,1.60)
<i>Employees per branch</i>	0.016 (0.00,0.03)	0.106 (0.04,0.17)	0.034 (0.02,0.05)	0.060 (0.03,0.09)
<i>Branch density</i>	0.150 (0.11,0.19)	0.438 (0.34,0.54)	0.167 (0.11,0.22)	0.195 (0.13,0.26)
<i>Bank's age</i>	0.701 (0.61,0.79)	0.578 (0.17,0.98)	0.235 (0.16,0.31)	0.100 (-0.02,0.22)
<i>Number of states</i>	0.153 (0.09,0.21)		0.108 (0.07,0.15)	

95% confidence interval in parenthesis.

**Table VI**  
**Price Cost Margins**

Variable	All markets	
	Logit	Nested logit
Service fees	—	25.2% (21.3-30.7)
Deposit rate	17.1% (15.5-19.0)	9.5% (7.7-12.5)

95% confidence intervals in parenthesis.

**Table VII**  
**Local Market Consumer Welfare Change Percentiles (1993-1999)**

The entries correspond to the indicated percentiles of the distribution of consumer welfare change from 1993 to 1999 across banking markets, based on the equivalent variation calculation:  $EV = S_s(p_t, x_t; \theta_D) - S_{s-1}(p_t, x_t; \theta_D)$  where  $S(p, x; \theta_D) = \ln[\sum_j^J \exp(\delta_j(p_j, x_j; \theta_D))]$ . Estimates are based on Tables II and III, columns (iii) and (vi).

		10%	25%	Median	75%	90%
<i>All markets</i>	Logit	\$0.13	\$0.60	\$1.13	\$1.70	\$2.25
	Nested	\$-0.14	\$0.00	\$0.15	\$0.33	\$0.53
<i>Urban markets</i>	Logit	\$0.24	\$0.40	\$0.64	\$0.89	\$1.18
	Nested	\$-0.11	\$0.00	\$0.12	\$0.29	\$0.41

## Appendix A: Summary Statistics

Variable	Mean	St. Dev.	Min	Max
<b>ALL MARKETS</b>				
<i>Market share</i>	8.54%	10.57%	0.00%	99.98%
<i>Service charge</i>	0.56%	0.33%	0.00%	4.95%
<i>Deposit interest rate</i>	3.24%	0.78%	0.02%	7.35%
<i>Number of employees per branch</i>	19	40	1	4187
<i>Branch density</i>	0.0038	0.0092	0.0000	0.4069
<i>Age of bank</i>	71	40	0	215
<i>Number of states of bank's operations</i>	1.5	2.0	1.0	17.0
<i>Big (1 = yes)</i>	33.30%			
<i>Medium (1 = yes)</i>	20.34%			
<i>Labor costs (weighted average of wage rate)</i>	24463	5551	559	60022
<i>Expenses on premises and fixed assets / Assets</i>	0.46%	0.23%	0.00%	5.72%
<i>Other expenses / Assets</i>	1.26%	0.72%	0.00%	41.76%
<i>Credit risk / Assets</i>	0.16%	0.36%	-21.51%	21.78%
<i>Bank operates in at least one rural area (1 = yes)</i>	74.30%			
<i>Most business loans are small (1 = yes)</i>	32.06%			
<i>Most agricultural loans are small (1 = yes)</i>	43.11%			
<i>Banking holding company indicator</i>	83.82%			
<i>Cash / Assets</i>	5.30%	3.42%	0.05%	64.96%
<i>Federal funds + securities / Assets</i>	30.94%	14.19%	0.00%	98.43%
<i>Real estate loans / Assets</i>	33.01%	13.46%	0.00%	92.58%
<i>Loans to individuals / Assets</i>	9.89%	7.15%	0.00%	96.67%
<i>Net worth / Assets</i>	9.46%	3.31%	-6.25%	83.06%
<i>Income per capita</i>	21278	5106	3495	56643
<i>Population per square mile</i>	293	686	0.1	11855
Number of observations (bank-market-year)	103380			
<b>URBAN MARKETS</b>				
<i>Market share</i>	2.59%	4.56%	0.00%	60.74%
<i>Service charge</i>	0.59%	0.38%	0.00%	4.95%
<i>Deposit interest rate</i>	3.08%	0.81%	0.02%	7.03%
<i>Number of employees per branch</i>	22	62	1	4187
<i>Branch density</i>	0.0045	0.0121	0.0000	0.4069
<i>Age of bank</i>	61	44	0	215
<i>Number of states of bank's operations</i>	1.4	1.9	1.0	17.0
<i>Big (1 = yes)</i>	36.21%			
<i>Medium (1 = yes)</i>	23.27%			
<i>Labor costs (weighted average of wage rate)</i>	27879	5130	1732	60022
<i>Expenses on premises and fixed assets / Assets</i>	0.52%	0.26%	0.00%	5.72%
<i>Other expenses / Assets</i>	1.37%	0.81%	0.00%	41.76%
<i>Credit risk / Assets</i>	0.19%	0.40%	-5.28%	21.78%
<i>Bank operates in at least one rural area (1 = yes)</i>	35.99%			
<i>Most business loans are small (1 = yes)</i>	20.55%			
<i>Most agricultural loans are small (1 = yes)</i>	35.18%			
<i>Banking holding company indicator</i>	80.37%			
<i>Cash / Assets</i>	5.70%	3.64%	0.06%	62.44%
<i>Federal funds + securities / Assets</i>	29.84%	13.82%	0.00%	98.43%
<i>Real estate loans / Assets</i>	35.22%	13.78%	0.00%	92.58%
<i>Loans to individuals / Assets</i>	9.21%	7.75%	0.00%	96.06%
<i>Net worth / Assets</i>	9.27%	3.64%	-6.25%	83.06%
<i>Income per capita</i>	25007	4854	11538	56643
<i>Population per square mile</i>	652	974	12	11855
Number of observations (bank-market-year)	41498			

Source: Federal Reserve Report on Condition and Income; U.S. Census; Bureau of Economic Analysis.

## Description of Variables

Variable	Description
<i>Share</i>	(Local market dollar deposits / "Average" U.S. deposit account) / Market population
<i>Service fee</i>	Service charge on deposit accounts / Deposits
<i>Deposit interest rate</i>	Interest expense on deposits (includes interest on time, savings and NOW accounts) / Deposits
<i>Loan interest rate</i>	Interest income on loans / Loans
<i>Employees per branch</i>	Number of bank employees / Number of branches
<i>Branch density</i>	Number of branches in local market / Square miles of local market
<i>Bank's age</i>	Years since beginning of bank's operations
<i>Big (1=yes)</i>	Bank with assets over US\$300M
<i>Medium (1=yes)</i>	Bank with assets of US\$100M-300M
<b>Instruments</b>	
<i>Labor costs</i>	Weighted average of annual wage in local markets where bank operates
<i>Expenses on premises and fixed assets / Assets</i>	Expenses including: utilities, repairs, taxes, insurance, equipment, etc.
<i>Other expenses / Assets</i>	Other expenses including: legal fees, amortization, advertising, etc.
<i>Credit risk / Assets</i>	Provisions for loans / Assets
<i>Bank operates in at least one rural market (1=yes)</i>	Whether bank operates in at least one non-MSA rural county
<i>Most business loans are small</i>	If all or substantially all of the dollar volume of loans secured by nonfarm nonresidential properties and commercial and industrial loans have amounts of US\$100,000 or less
<i>Most agricultural loans are small (1=yes)</i>	If all or substantially all of the dollar volume of loans secured by nonfarm farmland and loans to finance agricultural production and other loans to farmers have amounts of US\$100,000 or less
<i>Banking holding company indicator (1=yes)</i>	Whether bank is owned by banking holding company
<i>Cash / Assets</i>	
<i>Federal funds + securities / Assets</i>	
<i>Real estate loans / Assets</i>	
<i>Loans to individuals / Assets</i>	
<i>Net worth / Assets</i>	

Appendix B  
Further Specifications

Variable	Logit			Fixed effects logit	
	IV - drop size variables			OLS	IV
	(i)	(ii)	(iii)	(iv)	(v)
<i>Service fees</i>	-57.821 (19.647)**	-128.169 (14.182)**	-0.771 (0.309)*	-0.225 (0.012)**	-0.757 (0.302)*
<i>Deposit rate</i>	224.404 (9.995)**	75.994 (7.588)**	68.256 (4.588)**	5.519 (0.826)**	65.010 (4.451)**
<i>Employees per branch</i>	0.002 (0.001)**	0.016 (0.005)**	0.002 (0.000)**	0.002 (0.000)**	0.002 (0.000)**
<i>Branch density</i>	49.121 (6.474)**	136.069 (17.402)**	19.756 (0.873)**	19.673 (0.870)**	19.734 (0.872)**
<i>Bank's age</i>	0.013 (0.001)**	0.007 (0.001)**	0.003 (0.001)*	0.003 (0.001)*	0.002 (0.001)
<i>Number of states</i>	0.129 (0.023)**	-0.058 (0.028)*	0.016 (0.005)**	-0.007 (0.005)	0.019 (0.005)**
<i>Big</i>				0.384 (0.045)**	0.336 (0.046)**
<i>Medium</i>				0.215 (0.025)**	0.347 (0.028)**
<i>Number of employees*Income</i>		-0.004 (0.001)**			
<i>Branch density*Income</i>		-22.059 (3.084)**			
<i>Age*Income</i>		0.000 (0.001)			
<i>No. of states*Income</i>		0.063 (0.012)**			
<i>Big*Income</i>					
<i>Medium*Income</i>					
<i>Income</i>		-3.582 (0.138)**			
<i>Income squared</i>		0.481 (0.028)**			
<i>Population*Branch density</i>		0.834 (0.221)**			
<i>Population density</i>		-80.370 (4.285)**			
Adj. R-Sq. / Overident. test	Accept	Accept	Accept	0.10	Accept
First st. R-Sq. (Serv.fees)	0.40	0.41	0.26		0.26
First st. R-Sq. (Deposit rate)	0.40	0.40	0.31		0.32
Instruments	Cost	Cost	Cost		Cost
Observations	103360	102929	103360	103380	103360

NOTE.— State and year effects included. Estimated standard errors, robust and corrected for within bank dependence, are in parentheses. \*significant at 5%; \*\* significant at 1%.

Appendix C  
First-Stage Results

Variable	Table II, column (iii)	
	Service fees	Deposit rate
<i>Number of employees per branch</i>	0.000 (0.000)*	0.000 (0.000)**
<i>Branch density</i>	0.008 (0.002)**	-0.012 (0.004)**
<i>Bank's age</i>	0.000 (0.000)	0.000 (0.000)*
<i>Number of states</i>	0.000 (0.000)**	0.000 (0.000)**
<i>Big</i>	0.000 (0.000)**	-0.002 (0.000)**
<i>Medium</i>	0.000 (0.000)	-0.004 (0.000)**
<i>(Weighted) average wage</i>	0.000 (0.000)**	0.000 (0.000)**
<i>Expenses on premises and fixed assets</i>	0.310 (0.043)**	-0.351 (0.047)**
<i>Other expenses</i>	0.057 (0.011)**	-0.050 (0.012)**
<i>Credit risk</i>	-0.045 (0.008)**	0.125 (0.013)**
<i>Bank operates in at least one rural area (1 = yes)</i>	0.000 (0.000)*	0.001 (0.000)**
<i>Most business loans are small (1 = yes)</i>	0.000 (0.000)**	0.000 (0.000)**
<i>Most agricultural loans are small (1 = yes)</i>	0.000 (0.000)	0.000 (0.000)
<i>Banking holding company indicator</i>	0.000 (0.000)*	-0.001 (0.000)**
<i>Cash / Assets</i>	0.011 (0.001)**	-0.041 (0.002)**
<i>Federal funds + securities / Assets</i>	0.000 (0.000)	-0.006 (0.001)**
<i>Real estate loans / Assets</i>	-0.001 (0.000)*	0.002 (0.001)**
<i>Loans to individuals / Assets</i>	0.005 (0.001)**	-0.004 (0.001)**
<i>Net worth / Assets</i>	-0.008 (0.001)**	-0.023 (0.002)**
Adj. R-Sq.	0.40	0.43
Observations	103360	103360

NOTE.— State and year effects included. Estimated standard errors, robust and corrected for within bank dependence, are in parentheses. \*significant at 5%; \*\* significant at 1%.