Revista de Análisis Económico, Vol. 22, Nº 2, pp. 47-74 (Diciembre 2007)

RANDOM UTILITY MODELS OF DEMAND FOR THE U.S. COMMERCIAL BANKING INDUSTRY

CESAR OROSCO*
State Street Associates

Abstract

In this paper, we specify a Random Utility Model of Demand for Deposits in the U.S. Banking Industry, assessing its particular characteristics, such as a large number of participants, a large number of markets and an unbalanced panel (many banks participate in only one market and no bank participates in all markets). We modify the standard models to incorporate the fact that deposit balances are different among consumers, in a relationship proportional to their wealth. Using a unique dataset, we estimate the model and find that characteristics other than the interest rate, such as branch density, state presence, etc. add utility to the consumer. The model is also helpful in offering a more realistic set of elasticities among the many banks present in the sample. It shows how market shares will respond depending on the market demographics and current choice set (i.e. offerings of other banks). Finally, we use the results of the model to analyze changes in welfare during the 1994-2002 period. By applying a slightly modified version of Small and Rosen's equivalent variations, we find that the consolidation process of the late 90s was welfare enhancing, particularly for the middle income consumer.

Keywords: *Random Utility Models, Discrete Choice, Financial Institutions, Banks, Bank Deposits.*

JEL Classification: *C23, C25, D4, G21, L1.*

* State Street Associates, Cambridge, MA, USA. Email: corosco@statestreet.com. The autor wants to thank the comments of an anonymous referee.

I. Introduction

The U.S. banking industry is a large and important component of the country's economy. Financial assets (which are of course, not only limited to traditional banking) account for 60% of household's assets and for 76% of their net worth.[1]

As of 2002, total assets of commercial banks exceeded $ 6.6 trillions and they held more than $ 2 trillion in deposits. The sector directly employs more than 1.7 million individuals and provides services to virtually every household and business in the country. Probably due to its strategic importance it has historically been a highly regulated industry. U.S. Bank's are under the regulatory scope of the Federal Reserve (FED), the Federal Deposit Insurance Corporation (FDIC), the Department of the Treasury, the Antitrust Division of the Justice Department, and several state agencies.

Another important characteristic is that the industry is very heterogeneous. As of 2002 there were slightly less than eight thousand commercial banks, with assets ranging from a few millions to billions of dollars.[2] To add up to this situation, markets are highly fragmented, with many banks participating in only one of just a few markets within a single state and others with presence in more than 20 states.[3] Finally, substitutes of different degrees, such as S&Ls, Thrifts and other[4] depositary institutions contribute to blur the line that defines banking services and therefore, the industry.

Slowly since the 70's and at a much faster pace in the 90's, deregulation in the industry took place. Intrastate branching deregulation began in some states before the 70's; and some interstate banking occurred as early as 1978. But the big boost in deregulation was triggered by three milestones:

i) In 1991, Congress passed the Riegle-Neal Interstate Banking and Branching Efficiency Act, allowing nationwide branching as of 1997.

ii) In 1999, The Glass-Steagal Act of 1932 and the Bank Act of 1933 were repealed, henceforth eliminating the mandatory separation of investment banking and commercial banking activities. That same year as well, the Bank Holding Company Act of 1956 was repealed, eliminating the restriction of Bank Holding Companies engaging in non-financial activities.

iii) In 2000, the Gramm-Leach-Bliley Act finally opened the doors to "universal banking", which allowed for the full integration of commercial banks and most other financial services.

The immediate –and not surprising– result of this process was a surge in mergers and acquisition both intra-industry (commercial banks merging/buying other commercial banks) and inter-industry (insurance companies, investment banks and commercial banks).

The 90's saw a yearly average of 550 mergers a year,[5] and even though most involved small banks, many mega-mergers took place (a sample of the most important ones is shown in Table 1). Literature on this phenomena abounds, touching on many aspects of the subject: Peñas and Haluk (2004) analyze the effect of mergers on bank's bonds performance; Henock (2004) finds evidence on mergers being driven as a defensive tactic over hostile takeover bids; Buch and DeLong (2004) explore the more recent trends on cross-border bank mergers; Carbo and Humphrey (2004) study the scale related costs associated with bank's mergers; Wheelock and Wilson (2004) quantify the regulatory, market, and financial characteristics that affect the probability of a bank engaging in mergers and the volume of banks it absorbs over time; Rhoades (1998) analyzes efficiency gains in mergers while in the same vein, Focarelli and Panetta (2003) find evidence that consolidation generate adverse short term changes in deposit rates, but these are reversed in the long run thanks to efficiency gains; Elfakhani, *et al.* (2003) focuses on the effects of "mega mergers"; Santomero (1999) dives into political economy by addressing the question of appropriate policy prescription in face of rapid consolidation of the financial service sector both in the U.S. and globally.

TABLE 1

SAMPLE OF LARGE MERGERS (1997-2004)

Year	Acquirer	Acquired Institution	New Name
1997	First Bank System, Inc.	U.S. Bancorp	U.S. Bancorp
1997	NationsBank Corp.	Boatmen's Bancshares, Inc.	NationsBank Corp.
1997	Washington Mutual, Inc.	Great Western Financial Corporation	Washington Mutual, Inc.
1997	First Union Corp.	Signet Banking Corp.	First Union Corp.
1998	NationsBank Corp.	Barnett Banks, Inc.	NationsBank Corp.
1998	First Union Corp.	CoreStates Financial Corp.	First Union Corp.
1998	NationsBank Corp.	BankAmerica Corp.	Bank of America Corp.
1998	Golden State Bancorp, Inc.	First Nationwide Holdings, Inc.	Golden State Bancorp, Inc.
1998	Norwest Corp.	Wells Fargo & Co.	Wells Fargo & Co.
1998	Banc One Corp.	First Chicago NBD Corp.	Bank One Corp.
1998	Travelers Group	Citicorp	Citigroup
1998	SunTrust Banks, Inc.	Crestar Financial Corp.	SunTrust Banks, Inc.
1998	Washington Mutual, Inc.	H.F. Ahmanson & Co.	Washington Mutual, Inc.
1999	Fleet Financial Corp.	BankBoston Corp.	FleetBoston Financial Corp.
1999	HSBC Holdings, plc	Republic New York Corporation	HSBC Holdings, plc
1999	Firstar Corp.	Mercantile Bancorporation, Inc.	Firstar Corp.
2000	Chase Manhattan Corp.	J.P. Morgan & Co. Inc.	J.P. Morgan Chase & Co.
2000	Wells Fargo & Co.	First Security Corp.	Wells Fargo & Co.
2001	Firstar Corp.	U.S. Bancorp	U.S. Bancorp
2001	First Union Corp.	Wachovia Corp.	Wachovia Corp.
2001	Fifth Third Bancorp	Old Kent Financial Corp.	Fifth Third Bancorp
2001	FleetBoston Financial Corp.	Summit Bancorp	FleetBoston Financial Corp.
2002	Citigroup Inc.	Golden State Bancorp, Inc.	Citigroup Inc.
2002	Washington Mutual, Inc.	Dime Bancorp, Inc.	Washington Mutual, Inc.
2003	BB&T Corp.	First Virginia Banks, Inc.	BB&T Corp.
2003	Den norske Bank ASA	Gjensidige NOR Sparebank ASA	DnB NOR Bank ASA
2004	Bank of America Corp.	FleetBoston Financial Corp.	Bank of America Corp.
2004	J.P. Morgan Chase & Co.	Bank One	JPMorgan Chase & Co.

Source: Board of Governors of the Federal Reserve System.

Changes in the industry configuration have been dramatic. The number of banks dwindled from more than 14,000 in 1975 to less than 8,000 in 2002. At the same time, the average size of banks, measured in deposits, increase from $ 364 million in 1994 to $ 837 million in 2002[6]. The ratio of small banks to large banks increased by 66%, going from 15.7 in 1994 to 26.2 in 2002.[6] The average number of banks per Metropolitan Statistical Area (MSA) has also gone from 26 to 24 in the same time-frame. This has not resulted in a reduction in the number of branches. In 1994, the total number of branches located in MSAs were 43.5 thousands (about 5.5 branches per bank), whereas in 2002, this number hovered around 55.5 thousands (about 6.5 branches per bank).

All these have definitely changed the competitive landscape of the industry and make it necessary to develop appropriate tools to assess the new market structure and its new dynamics. Issues of welfare, anti-trust regulation, M&As, Pricing and Marketing Strategies, could benefit from a model than can effectively incorporate certain particular features of the industry, such as large number of participants, differentiated products, highly segregated markets with some exclusively local competitors and many overlapping participants and heterogeneity of consumers.

So far, few attempts have been made to deal with all these characteristics at once. Many applications deal with market power and market structure in ways that only take into account some of these features; particularly interest rate effects on demand. One of the paradigms stems from the "New Empirical Industrial Organization" (NEIO) approach (Bresnahan, 1989) based on the firm profit maximization function. Examples of these are Shaffer (1989 and 1993) for an application to the U.S. and Canada, respectively; Hannan and Liang (1993) for local deposits markets in the U.S.; Suominen (1994) for the Finnish banking industry; Gruben and McComb (1996) for Mexico and Barajas *et al.* (1999) for Colombia. Bikker and Haaf (2002) measure competitive conditions and market structure using the Panzar-Rosse model, which also stems from the NEIO.

We believe that a model that can deal with the dimensionality problem of many participants, product heterogeneity and consumer variability across markets is needed. In this sense, the framework on discrete choice and random utility models of demand pioneered by McFadden (1973, 1978 and 1981) and developed and perfected by Berry (1994) and most notably Berry, Levinsohn and Pakes (1995)[7] seems to be a promising candidate to asses these issues.

Random Utility Models of Demand (henceforth RUM) have been used in several applications. BLP applied it to the U.S. Car Industry in order to analyze 20 years of national data on more than 100 car/models. Their study had a large number of products, only one market per year. Petrin (2002), focused on the case of the minivan, incorporating the extra information contained in consumer surveys to obtain more precise results. Nevo (2000) analyzed the Ready to Eat Cereal Industry; in a setting that considerable expanded the number of markets under study (over 65 markets per quarter over 20 quarters). Melnikov (2000), on the computer printer industry, introduced a more dynamic approach while Massa (2003) used a simpler version of the model (the nested logit) to analyze the Mutual Funds Industry.

Dick (2000), used a discrete choice logit-model to analyze the U.S. banking industry from 1993-1999 and hence constitutes a clear precursor to our study, which aims at estimating a RUM of demand for the banking industry and at generating a sensible set of elasticities that can be used in assessing and evaluating the current dynamics of the industry. We believe a RUM model is more appropriate for our purpose since it can incorporate all of the features of the industry we have discussed above. As Berry (1994) put it when discussing whether to use the logit, nested logit or random coefficients model: "The random coefficients model will be preferred when a premium is placed on estimating richer patterns of demand"

A RUM model for the U.S. Bank Industry shows promise but also poses challenges to the estimation: It needs to take into account an unprecedented number of participants (a number that falls on the thousands), a large number of markets (in our case, more than 200 MSAs per each of the 9 years of analysis) and an unbalanced panel (many banks participate in only one market and no bank participates in all markets).

The rest of this paper is organized as follows: Section II defines the theoretical model, Section III details the characteristic of the dataset used and its respective caveats as well as the estimating procedure and Section IV discusses the results of the model. Section V develops an application on welfare using the model results. Finally, we conclude and suggest a pending agenda.

II. The Theoretical Model

The formulation and notation of this model follows closely what has been applied to several other industries since the work of BLP. Here we adapt the formulation directly to the banking industry and provide only minor elaboration on certain steps taken. Readers otherwise unfamiliar with BLP and Nevo (2000) might find it useful to refer to those sources for a detailed account of our theoretical framework.

In shopping for deposit services, we assume that any given consumer can choose among a large number of firms (banks, savings associations and the like) offering such services. In particular, we assume that there are T markets, each of them with J_t firms and I_t consumers.

Suppose that consumer i, who resides in market t values deposit services offered by bank j according to the following conditional indirect utility function:[8, 9]

$$u_{ijt} = U(X_{jt}, \xi_{jt}, r_{jt}, \tau_i; \theta), \tag{1}$$

where X_{jt} is a $1xK$ vector of observable characteristics of firm j; ξ_{jt} represents the unobserved (by the econometrician) characteristics; r_{jt} is the *net deposits interest rate* (net of fees) offered by firm j; τ_i is a vector of consumer's characteristics and θ is a set of parameters.

Examples of observed characteristics of the firm would be attributes such as number of branches or size of the firm (small or large bank in terms of assets) whereas examples of unobserved characteristics would be level of courtesy of the staff, quality of management, etc.

With respect to the consumer characteristics, τ, we divide them as well into observed and unobserved. Examples of observed characteristics would be income level and employment status. Following standard terminology for this framework, we will call these observable characteristics "demographics". Unobserved characteristics could be attributes such as level of knowledge of financial products, prior experiences with financial institutions, saving habits and others.

Finally, we define D_i as a *dx1* vector of observable demographics and v_i as a *(K+1)x*1 vector of unobserved characteristics of the consumer.[10]

Now, we assume the following explicit functional form for equation (1):

$$u_{ijt} = \alpha_i r_j + X_{jt}\beta_i + \xi_{jt} + \varepsilon_{ijt}; \quad i = 1, \ldots, I_t; j = 1, \ldots, J_t; t = 1, \ldots, T \tag{2}$$

where

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i \tag{3}$$

with Π defined as a *(K+*1*) x d* matrix and Σ as *(K+*1*)x(K+*1*)* matrix of parameters[11] that govern the interaction of the demographics and the unobservable individual characteristics with the observable product characteristics[12]. Given this formulation, θ is defined as the set of parameters (α, β, Π, Σ).

Equations (2) and (3) can be rewritten as

$$\begin{gathered} u_{ijt} = \delta(x_{jt}, r_{jt}, \xi_{jt}, \alpha, \beta) + \mu_{ijt}(x_{jt}, r_{jt}, D_i, v_i\,; \Pi, \Sigma) + \varepsilon_{ijt} \\ \delta_{jt} = x_{jt}\beta + \alpha r_{jt} + \xi_{jt}, \;\; \mu_{ijt} = [r_{jt}, x_{jt}]\,(\Pi D_i + \Sigma v_i) \end{gathered} \tag{4}$$

where δ is referred to as the *mean utility*, which is common to all consumers and the sum $u_{ijt} + \varepsilon_{ijt}$ represents a zero mean heteroskedastic deviation from the mean utility and captures the effect of the random coefficients. We also assume that ε_{ijt} follows an i.i.d. extreme-value distribution.

To complete the specification of this demand system, we introduce an outside good –denoted by a zero sub-index– which captures the fact that some consumers might decide *not* to use any bank. The indirect utility from opting for this option is defined as

$$u_{i0t} = \xi_{ot} + \pi_0 D_i + \sigma_0 v_{i0} + \varepsilon_{i0t} \tag{5}$$

Since, ξ_{ot}, π_0 and σ_0 can't be identified without further assumptions or normalizing one of the inside goods, we follow the common practice of normalizing their values to zero, which is equivalent to normalizing the utility of the outside good to zero.

Following standard literature, we will assume that consumers only choose one bank for their services: the one that gives them the highest utility. In a departure from the traditional approach, we will allow the size of their deposits in such bank to vary from consumer to consumer. In particular, we will define the size of a consumer's deposit to be a fraction ω of her wealth, with ω common to all consumers.[13]

Conceptually, this can be accommodated by defining our "product unit" as one dollar of deposits held by the consumer and by allowing the purchase of several units of the product at once; with the only constraint that all her purchases are made with the same banking institution.[14] As a result, market shares must be calculated based on dollar amounts and not on other measures such as number of accounts of a particular average size.[15]

As we will see below, this feature can be easily incorporated into the model and allows a more realistic representation of market dynamics, where wealthier consumers hold larger balances in their bank accounts.

Since consumer's are defined as a vector of demographics (D_i and v_i) and product-specific shocks (ε_{i0t}, …, ε_{iJt}), the set of attributes that lead to the choice of bank j in market t is implicitly defined as

$$A_{jt}\left(x_{.t}, r_{.t}, \delta_{.t}; \Pi, \Sigma\right) = \left\{\left(D_i, v_i, \varepsilon_{i0t}, ..., \varepsilon_{iJt}\right) \middle| u_{ijt} \geq u_{ilt} \quad \forall l = 0, 1, ..., J_t\right\}, \tag{6}$$

where $x_{.t} = (x_{1t}, ..., x_{Jt})', r_{.t} = (r_{1t}, ..., r_{Jt})'$, and $\delta_{.t} = (\delta_{1t}, ..., \delta_{Jt})'$ are observed characteristics, net interest rates and mean utilities corresponding to of all firms, respectively. A_{jt} defines the set of characteristics that leads to the choice of bank j in market t.

Assuming that ties occur with zero probability, the market share of bank j is the integral over the mass of consumers' wealth portion ω in the region A_{jt}, and can be formally expressed by

$$\begin{aligned} s_{jt}(x_{.t}, r_{.t}, \delta_{.t}; \Pi, \Sigma) &= \int_{A_{jt}} dP^*(D, v, \varepsilon) \\ &= \int_{A_{jt}} dP^*(\varepsilon \mid D, v) dP^*(v \mid D) dP_D^*(D) \\ &= \int_{A_{jt}} dP_\varepsilon^*(\varepsilon) dP_v^*(v) dP_D^*(D) \end{aligned} \tag{7}$$

where $P^*(\cdot)$ denotes population's *wealth* distribution functions.[16] The second equality is an application of Bayes' rule and the last one follows from the independence assumptions previously made. Note that ω doesn't appear in equation (7), because s_{jt} is invariable to its value, as long as $\omega > 0$ and it's the same for all consumers in the same market.[17]

As Nevo (2000-Appendix) explains, when dealing with the full model ($\Pi, \Sigma \neq 0$), the predicted market shares defined by (7) and their elasticities can't be solved for analytically and must be approximated numerically. He favors the use of a *smooth simulator*, which demands fewer draws from the distributions, produces market shares that always add up to one (when including the outside share), and reduces the variance in the estimated parameters produced by the simulation by integrating the ε's analytically.[18]

In our particular case, which uses a wealth-weighted approach, Nevo's *smooth simulator* of the market shares in market t, obtained from a random sample of ns_t consumers becomes:

$$\widehat{s}_{jt} = \sum_{i=1}^{ns_t} w_{it} \frac{\exp\left(\delta_{jt} + \sum_{k=1}^{K} X_{jt}^{k}\left(\sigma_k v_i^k + \pi_{k1} D_{i1} + ... + \pi_{kd} D_{id}\right)\right)}{1 + \sum_{m=1}^{J_t} \exp\left(\delta_{mt} + \sum_{k=1}^{K} X_{mt}^{k}\left(\sigma_k v_i^k + \pi_{k1} D_{i1} + ... + \pi_{kd} D_{id}\right)\right)} \tag{8}$$

where $w_{it} = \dfrac{W_{it}}{\sum_{l=1}^{ns_t} W_{lt}}$ is the wealth-weight of consumer i in the sample of market t, W_{it} is the wealth in dollars of consumer i in market t and $(v_i^1, ... v_i^K)$ with $(D_{il}, ... D_{id})$, $i = 1, \ldots, ns_t$ are the draws from market t. The objective of incorporating consumer's wealth into the estimation of market shares is to improve their estimation. We aim at capturing the fact that wealthier consumers tend to hold, on average, larger balances and therefore, their choices have a bigger impact on the bank's market share.[19]

From our modified *smooth simulator,* the elasticity of market shares, s_{jt}, with respect to characteristic c is approximated by:

$$\eta_{jkt}^{c} \equiv \frac{\partial \widehat{s}_{jt}}{\partial x_{kt}^{c}} \frac{x_{kt}^{c}}{\widehat{s}_{jt}} = \begin{cases} \dfrac{x_{jt}^{c}}{\widehat{s}_{jt}} \displaystyle\sum_{i=1}^{ns_t} \gamma_i \widehat{s}_{ijt}\left(1 - \widehat{s}_{ijt}\right) w_{it} & , if \quad j = k \\ -\dfrac{x_{kt}^{c}}{\widehat{s}_{jt}} \displaystyle\sum_{i=1}^{ns_t} \gamma_i \widehat{s}_{ijt} \widehat{s}_{ikt} w_{it} & , if \quad j \neq k \end{cases} \tag{9}$$

where γ is either α (if c represents the net interest rate) or β (otherwise) and $\widehat{s}_{ijt} = \exp\left(\delta_{jt} + \mu_{ijt}\right) / \left[1 + \sum \exp\left(\delta_{kt} + \mu_{ikt}\right)\right]$ is the probability of consumer i choosing bank j.

Note how the inclusion of wealth-weighted market shares affects the result of the formula: the response of wealthier consumers weighs more on the change in market shares given an exogenous change. Of course, lower income consumers weight less individually, but are more numerous as a group. In terms of policy, this highlights even more the importance of taking into consideration the income distribution of a market when analyzing its dynamics.

III. The Dataset and Estimation Procedure

In order to estimate our model, we need the following variables: market shares, deposit interest rates, fees, and other observable product characteristics at the market level. In addition, we need demographic information from consumers. Finally, suitable instruments are also incorporated.

Market definition and bank's characteristics

In order to define market shares, a definition of market is needed. Several partitions are available here and data from the FDIC Summary of Deposits (SOD) allows us to calculate market shares in partitions that go from the national level to zip code areas. We choose to work with the most common approached for analysis of demand deposits, which is to define markets as eminently local. Studies by Starr-McCluer (2001), Kwast *et al*. (1997), Wolken (1990) and Rhoades (1992) present evidence that Metropolitan Statistical Areas (MSAs) and counties are better suited to encompass the main aspects of banks' behavior.[20]

Even though the SOD gives us market share information for all FDIC insured institutions (not only commercial banks), as well as number of branches at the local level, the remaining characteristics must be obtained from the FED Reports on Condition and Income (Call Reports), which contain information only at the bank level. For these reasons, we will assume that the characteristics of a particular bank in any market it participates, is fairly homogeneous and hence, it's a good approximation to use the aggregate information on Call Reports. This assumption seems reasonable for several characteristics, which are perceived by consumers at the national level; such as size and age of the bank. It's also true by definition on others such as number of states where the bank is present, the location of its headquarters or if it belongs to a holding company. It is also true for small banks that operate in only one MSA. But there is not any a-priori reason why it should hold true for its "pricing structure" (i.e. the interest rate paid on deposits and fees charged on a particular market). Information on interest rate and fees at the local level is not available at the extent we are interested.[21]

We opted for estimating interest rate on deposits and fees from balance sheet and income statement information contained in the Call Reports. This is certainly an important assumption, since it's clear that data calculated this way is subject to large measurement error and also because we are assuming that the same rates and fees applies to every single market covered by a bank.[22]

Demographic data

Demographics are obtained from the 2000 Census information. We focus our analysis of demographics on the household level, and assume they explain most of the demand drivers for traditional bank deposits. A more sophisticated analysis should find a way to incorporate firms as a second class of consumers.[23]

We used information on household income, sex, race, education level, number of children, employment and marital status and time used on commuting. Unfortunately, from the publicly available census data it's not possible to identify rural counties nor MSAs with population lower than 100,000. For these reasons, we excluded from our dataset these unidentifiable markets. Also, since we used each characteristic as a deviation from it's national mean, we assume that even though certain variables certainly changed over time, it's distribution remained constant, allowing us to use our demeaned variables not only in the Census year (1999-2000), but also in the rest of the years in our sample.

Finally, we need information of consumer's wealth. Unfortunately, this data is not directly available from the Census. To overcome this problem, we assume that current income is a good proxy of a consumer's wealth.[24]

Outside share

Because we are using a wealth-weighted approach to estimate market shares, the more common approach of defining market size as the result of computing number of potential consumers, times one standard-sized purchase per consumer is not applicable. Moreover, when this approach is applied to bank deposits, it can yield market sizes smaller than actual size in many markets (particularly in big financial hubs, such as New York).

Our take on market size was to estimate it indirectly through an implied outside share for banks. From the FDIC data, we know total deposits held in commercial banks and its close competitors (S&Ls, Thrifts, etc.). Also, from the Survey of Consumer Finances, we get the percentage of households that don't have a bank account for every year of our sample. Putting these two pieces together, we estimate how much of the total market is captured by the banks present in our sample.[25]

Instruments

In order to estimate the model, instruments are needed to correct the correlation of the interest rate with the error term. We looked for instruments that can capture a bank's cost structure (*cost shifters*) as well as ones that can capture the within markets variability.[26] Such instruments include the loan rate, equity levels, average size of a deposit account, overhead and fixed assets, bad loans provisions, liquidity ratios, level of industrial loans, number of accounts larger than $ 100K, percentage of transaction accounts with respect to total accounts and interest income. We also used instruments of the type suggested by BLP, which help us with the within market variability. BLP consider as valid instruments for a firm, the sum of characteristics of other firms competing in the same market. Since our dataset is unbalanced, we used the average of the competitors' characteristics instead of the simple sum. Lastly, we used time dummies for each year in the sample.[27]

Final dataset

Our final dataset consists of 17,770 observations encompassing an average of 191 MSAs for each year between 1994 and 2002. It covers a very dynamic period for the industry, offering much valuable variability in the characteristics that helps in obtaining better estimates. Table 2 shows a summary of the dataset structure.

TABLE 2

DATASET STATISTICS

Variable	Median	Mean	Max	Min	Stdev
MKTDEP: Total Deposits in local market ($ Millions)	2,517.24	3,565.50	22,029.54	417.01	3,074.64
NSAVINST: Number of FDIC Sav. Instit. in the local market	18.00	20.01	63.00	5.00	9.06
NBANKS: Number of FDIC Banks in the local market	15.00	16.60	54.00	4.00	7.70
NSAMPLE: Nº of banks in the market present in dataset	11.00	11.50	25.00	4.00	3.70
NTOTBRSI: # Branches of all Sav. Inst. in local market	70.00	94.02	419.00	11.00	68.90
NTOTBRBK: # Branches of all Banks in local market	61.00	81.20	377.00	10.00	60.39
BKTOTDEP: Total Depostis of the Bank (NatI Level - Millions)	414.37	15,491.41	383,627.00	9.70	50,125.06
TOTASST: Total Assets of the Bank (National Level - Millions)	500.13	23,354.69	584,270.33	11.13	76,525.38
TOTLOAN: Total Loans of the Bank (National level - Millions)	315.93	14,886.47	379,714.00	1.96	47,474.16
TOTDEP: Total Deposits of the Bank in Local Mkt (Millions)	130.32	264.96	7,677.13	5.20	426.17
BKTOTBR: Number of Branches of Bank (National Level)	13.00	248.92	4,288.00	1.00	640.69
NUMEMPLY: Number of Employees of Bank (National Level)	204.67	6,239.36	138,000.00	4.67	18,638.03
NBRANCH: Nº of branches of bank in the local market	4.00	6.63	82.00	1.00	7.69
OUTSHARE: Outside share (%)	24.92	25.38	49.97	9.10	9.99
SHARE: Market share of bank in local market	4.77	8.09	54.92	1.00	8.17
FULSHARE: Market share adjusted by outside option	4.25	7.21	49.92	0.87	7.29
POP: Total population in MSA	243.82	345.30	1,918.01	101.54	275.86
SQMLAREA: Total area in MSA (Sq. Miles)	1,569.59	2,1 29.47	39,719.10	62.43	3,094.84
HOUSUNIT: Number of Housing Units in MSA	105.39	143.78	786.30	40.61	114.07
LANDAREA: Total Land Area in MSA (Sq. Miles)	1,455.67	1,998.38	39,368.64	46.69	3,046.30
XAGE: Age of the Bank (in Decades)	7.85	7.40	21.85	0.10	4.49
XBIG: Equals 0 if TOTDEP < 1 Billion, 1 otherwise	0.00	0.43	1.00	0.00	0.50
XBRDNSAR: # of Branches/Landarea (SQ Miles)	3.01	5.89	406.94	0.04	13.94
XBRDNSBK: % Dev. from mean # branches of banks in Mark	1.00	1.27	8.63	0.05	0.97
XBRDNSSI: % Dev. from mean # branches of Sav. Inst in Mkt	1.03	1.34	8.77	0.05	1.04
XDEPFEE: Fees on Transaction Accounts	0.01	0.01	0.03	0.00	0.00
XDEPNETR: XDEPRATE - XDEPFEE	0.03	0.03	0.12	0.00	0.01
XDEPRATE: Interest Rate on Deposits	0.03	0.03	0.13	0.00	0.01
XEMPLOY: (Avg. Nº of Employees/1000)/Nº of Branches	0.16	0.20	5.22	0.02	0.17
XHHLOCHQ: Equals 1 if High Holder is based in loc mkt state(s)	1.00	0.72	1.00	0.00	0.45
XHHPRES: # of States where the High Holder has a Sav. Inst.	1.00	4.70	33.00	1.00	5.86
XHOLDING: Equals 1 if Bank belongs to a Financial Holding	1.00	0.90	1.00	0.00	0.30
XLOCHQT: Equals 1 if Bank is based in local mkt state(s)	1.00	0.87	1.00	0.00	0.33
XPRES: # of States where the Bank has a presence	1.00	2.62	25.00	1.00	4.19
XSMALL: Equals 0 if TOTDEP < 100MM, 1 otherwise	0.00	0.21	1.00	0.00	0.41
Z100KACC: Dep > 100 K/Tot Dep	30.24	30.67	94.50	1.32	11.28
ZAVGACC: Tot Dep/N. Accts of Bank (national Ievel) ($ Thsnds	9.08	10.46	216.78	1.09	7.38
ZEQUIP: Furnit & Equipment/Assets (%)	0.46	0.49	6.02	−0.25	0.22
ZEQUITY: Equity/Total Assets (%)	8.38	8.94	28.04	1.05	2.39
ZINDLOAN: Loans to lndiv (IncI. Mortgage)/Gross Loans (%)	42.76	43.65	100.00	0.22	16.46
ZINFRAS: Fixed Premises/Assets (%)	1.56	1.80	12.93	0.01	1.03
ZINTDEP: Interest Bearing Dep/Total Deposits (%)	83.43	82.59	99.98	32.71	7.32
ZINTINC: Interest Income/Total Operating Income (%)	86.67	84.99	140.34	11.23	8.60
ZLABOR: Wages/Assets (%)	1.57	1.64	10.71	0.05	0.58
ZLIQUID: Cash, Securities & Fed Funds/Assets (%)	22.81	24.60	90.41	0.13	11.63
ZLOANR: Interest Rate on Loans	0.09	0.09	0.35	0.00	0.02
ZOTHREXP: Other Expenses/Assets (%)	1.20	1.36	43.89	0.03	0.80
ZOVERHD: (Wages + Furnit & Equipment + Other)/Asset (%)	3.30	3.50	51.46	0.24	1.23
ZPROVIS: Provisions/Total Gross Loans (%)	1.40	1.54	15.93	0.06	0.71
ZTRANSAC: Transaction Accounts/Total Deposits (%)	23.97	25.02	73.17	0.00	10.40

IV. Results

As a starting point, we estimated the basic logit-model, which doesn't allow for consumer heterogeneity. As explanatory variables, we try to use those that help better describe the different dimensions from which consumers derive utility. We also used eight dummy variables all but one year of the sample, to (imperfectly) account for different regimes.[28]

As we explained in the previous section, the interest rate and fees variables seemed to contain some noise in their calculation. We obtain better results when used a "net rate" variable, which subtracts the calculated fees rate from the calculated offered deposit rate. This variable seems to summarize appropriately the information consumer's take into account when choosing a banking institution, and helps to average out the measurement error.[29] The results are shown in Table 3. The fit of 41% is reasonably good given the heterogeneity of the data (among other features: unleveled panel, many small banks and many highly specialized banks).

TABLE 3

LOGIT ESTIMATION

	Coefficient	Standard Deviation
Intercept	–3.4243	(0.051)**
Net Deposit Rate	12.3447	(1.001)**
Relative Branch Density	0.5592	(0.007)**
Size Dummy	0.0174	(0.002)**
Nº States w/presence	0.2168	(0.023)**
Local Hdqters Dummy	0.1875	(0.023)**
Holding Dummy	0.4329	(0.018)**
Employment Expses	0.6518	(0.040)**
Age of Bank	0.0015	(0.002)
R-Squared:	0.41	
Nº Observations	17,770	
Nº of Market/Years	1,720	

** Denotes 99% Significance.

Besides the standard caveats of unrealistic substitution patterns,[30] the model captures the "right" interactions (sign) of our key variables. Consumers seem to value positively the net return in their deposit balances and the results are statistically significant despite the fact that poor quality of the measured variable is affecting the results.

Besides the net interest rate, most of the other variables are statistically significant and seem to explain a large part of consumers' behavior.[31, 32] One of the key variables is relative branch density, which, at 99% significance level, highlights he fact that time and/or convenience is highly valued by consumers.

Another interesting result is that consumers seem to be more inclined to conduct their business with banks perceived as local (i.e. the high holders headquarters are located in the state where consumers reside). One could speculate that many behavioral and financial issues are in place here: banks headquartered in the state have more (and visible) ways to stimulate the local economy, via jobs, community work, etc., which can result in a more favorable perception of the consumer towards the institution.

The employment expenses variable, which functions as a proxy for quality of service, has the expected sign and is significant. Finally, size and numbers of states where the bank is present also are perceived as positive qualities for a bank. We didn't have any a priori expectation for the sign of these variables, since different explanations could be attributable to either sign.

4.1 The full model

As noted before, the main objective of this paper is to allow for heterogeneity of consumers and capture more realistic substitution patterns. Using census information, we analyzed potential interaction with all the variables and demographic characteristics of head of households such as income, commuting time, sex, age, race, employment status, marital status, education level and parenting status. Several specifications were tested and they suggested that most of the possible interactions have little explanatory power beyond the one provided by the logit model.[33] Nevertheless, interacting two of these demographic characteristics with three of our chosen bank's characteristics proved to be very relevant and seemed to capture most of the heterogeneity of consumers: income level and commuting time interact –in some cases strongly– with net interest rate, bank size and branch density; a result that we believe carries strong economic sense. Table 4 summarizes the results.

TABLE 4

FULL MODEL - INCOME WEIGHTED ESTIMATION

	Mean Coeff	Random Shock	Income	Commuting Time
Intercept	−3.1805	0.2938	−1.8078	−3.3163
	*(0.15)***	*(6.79)*	*(0.48)***	*(2.81)*
Net Deposit Rate	14.8839	−0.6224	0.8124	
	*(3.29)***	*(0.36)*	*(21.56)*	
Relative Branch Density	0.7103	−2.8826		4.114
	*(0.02)***	*(0.36)***		*(2.08)**
Size Dummy	0.0134	−0.262	1.1511	
	*(0.003)***	*(17.91)*	*(0.33)***	
# States Presence	0.144			
	*(0.03)***			
Local Hdqters Dummy	0.2054			
	*(0.04)***			
Holding Dummy	0.2345			
	*(0.05)***			
Employment Expses	0.6974			
	*(0.06)***			
Age of Bank	0.0001			
	(0.00)			
GMM Function:	0.25			
Deg. of Freedom	3			
Nº Observations	17,770			
Nº of Market/Years	1,720			

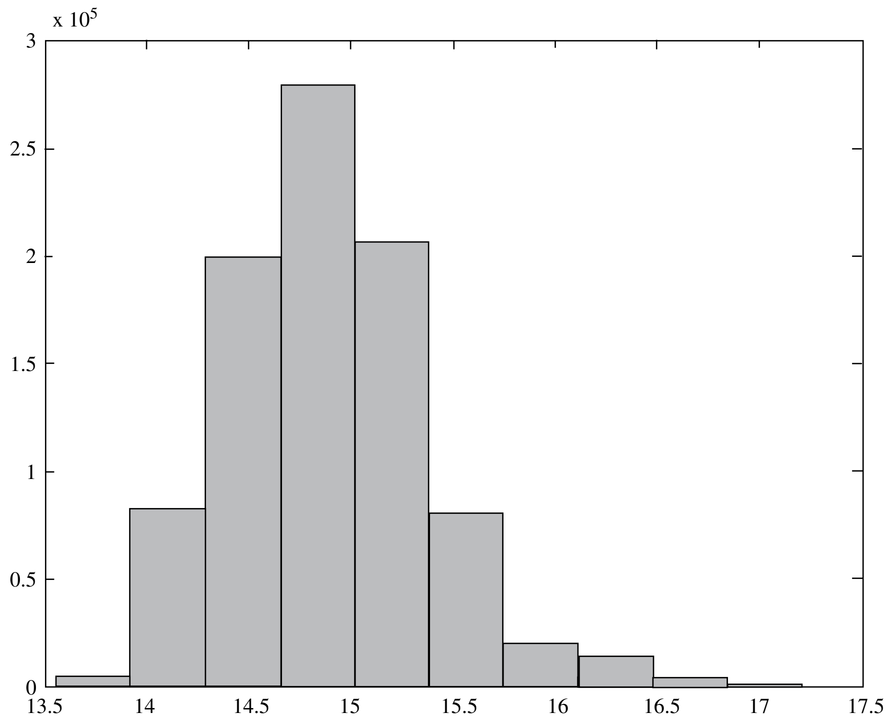
* Denotes 95% significance. ** Denotes 99% significance.

Firstly, the lack of statistical significance in most of the random shock coefficients suggests that most of the heterogeneity is explained by the observed demographics. The fact that several other specifications that included more interactions carried little explanatory power backs up this observation.[34]

The interactions with demographics shown in Table 4 are quite intuitive and three of them are statistically significant. In particular, they suggest that wealthier households care more about returns on their deposits. This can be justified by noticing that they carry larger balances and have more options to shop around when it comes to interest rates, like savings accounts, CDs, etc. Also, wealthier households are usually more financially savvy and try to maximize the returns on their investments. It must be noted, though, that the random coefficient here is not significant, but as stated before, there are measurement problems with the interest rate variable. The distribution of the random coefficient on interest rate is shown in Figure 1. The range of variability goes from 13.56 (least wealthy households) to 17.21 (wealthiest households). Even at the lower levels, it exhibits the correct (positive) sign.

FIGURE 1

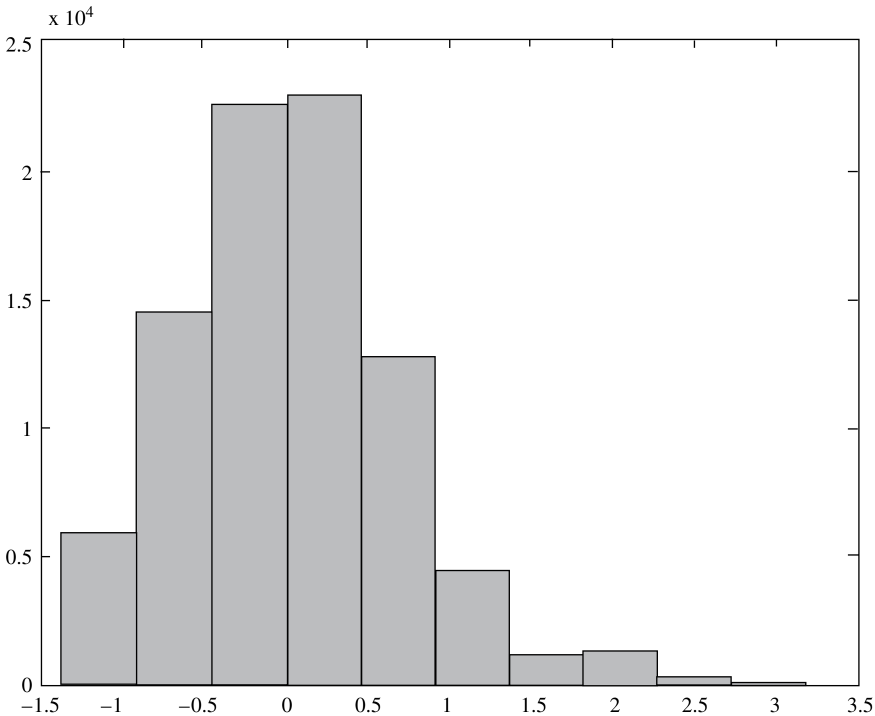
DISTRIBUTION OF NET INTEREST RATE COEFFICIENT



The results also suggest that larger banks have an edge with wealthier households, which can be justified by the fact that larger banks usually offer a more complete and sophisticated set of financial instruments that complement deposits aimed at satisfying the more complex needs of these consumers. In this sense, the income/size relationship can be uncovering the unobservable (for the econometrician, given the dataset) existence of "boutique" services available to wealthier customers. It is also worth noting that the distribution of the coefficient –shown in Figure 2 has some of the less wealthy consumers actually preferring smaller banks to larger ones. A feasible explanation would be that this kind of consumers carry little balances and don't care for sophisticated services as long as they get good and inexpensive service. In many cases, a small, local bank might be able to better fit the bill than a larger, multi-state mega bank.

FIGURE 2

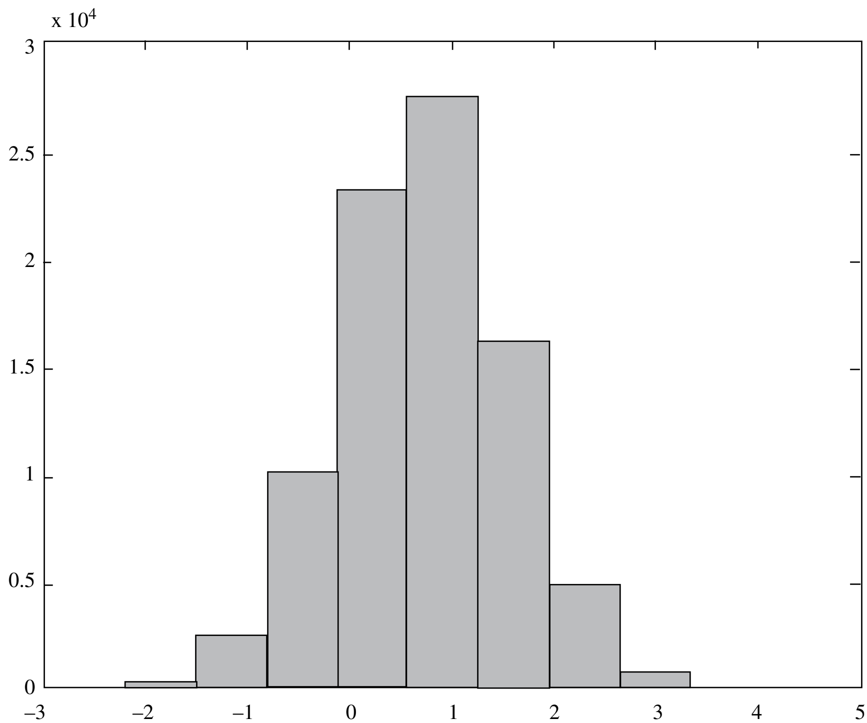
DISTRIBUTION OF SIZE COEFFICIENT



Perhaps the most intuitive result of all is the positive relationship between commuting time and branch density. If we see commuting time as a variable inversely related to time available for income-producing tasks (i.e. paid jobs), errands and leisure, we could easily see that the longer the time committed to commuting the more valuable a higher branch density becomes to the consumer (our dataset starts in 1994 and ends in 2002, and hence it probably fails to capture the explosion in online banking, which might weaken this relationship in the future). Worth noting is the fact that the only random shock that exhibits statistical significance is the one on the branch density variable, suggesting that some unobservable factors might be important in capturing consumer's heterogeneity and its interaction with the branch density variable. For instance, retirees and people who work at home do not spend any time commuting to work, making this variable inappropriate as a proxy of the value of their time. It is also possible that this is one of the factors that cause the branch density coefficient distribution shown in Figure 3 to have a small fraction of consumers (less than 20%) with the "wrong" sign (negative). We believe this is a statistical anomaly, since branch density should be valued positively in all cases.[35] Nevertheless, as we will see below, this anomaly doesn't get fully transmitted to the estimated elasticities.

FIGURE 3

DISTRIBUTION OF COMMUTING TIME PARAMETER



4.2 Sensitivity analysis

We now move to analyze own and cross elasticities of our sample. Our model uncovers the possibility that "interest rate elasticity" takes the back seat (or at most shares the driving seat) to other meaningful interactions such as branch density elasticity; therefore, we focus on these two variables. It would also be interesting to analyze the elasticity of size, since our previous results suggest that size doesn't *always* matter. Unfortunately, our model can't offer much further insights, since dummies are not continuous variables.

The distribution of the own-net interest rate elasticity is shown in Figure 4. The median value of the own interest rate elasticity is 0.36. If we divide the sample between large banks and smaller banks (defined by our size dummy variable), we find that larger banks exhibit a smaller value: 0.32 for large banks vs. 0.40 for small banks. This agrees –under the restrictive assumptions that drive the calculation of Lerner Indexes–[37] with the notion that larger banks have more market power. Examples of banks that fall in the bottom 1% percentile of the net interest rate elasticity distribution are Bank of America, Wells Fargo and PNC Bank. It is also worth mentioning that during the second half of the 90's, the mean value of the net interest rate elasticity was around 0.40 with un upward trend; but in the last year of our sample, 2002, this trend was reversed and the elasticity had gone down to 0.25. This fact is interesting since it matches a typical trend on industry concentration: at first, when deregulation is enacted, competition increases and market powers falls; but later, after a process of mergers and acquisitions and consolidation of the new industry structure, the survivors are stronger, larger and with more market power.

The own elasticity of branch density has a median of 0.47, with a standard deviation of 0.95, suggesting that the decision on expanding the numbers of branches can be better made by analyzing the particular characteristics of the market. The distribution is shown in Figure 5 (only a few banks, less than 5% of the sample, exhibit the "wrong" sign: we believe this is just an abnormality due to poor fit for those particular banks. In fact, 80% of the banks with the "wrong" sign are "small" banks, whose fit is harder to achieve). Examples of markets where an aggressive branch expansion policy had a lot of potential to be very successful in capturing market share in the year 2002 are Naples, FL and Daytona Beach, FL. In these two markets, Bank of America and Wachovia (which competed in several markets in 2002) had branch densities larger than 3.6, which is in the top 5% of the distribution. For the same reason, in Daytona Beach, their smaller competitors, such as Coquina Bank and Cypress Bank, are so far behind in number of branches, that any marginal opening of new ones seems worthless.[38] Their branch elasticity is in the bottom 5% of the distribution for the whole dataset, with a value practically equal to zero.

In order to provide a picture of cross elasticities, we chose the two of the largest banks and found out how close competitors they are in some representative markets.

The results, shown in Table 5, indicate the way Bank of America competed with Wachovia in both interest rate and branch density during the year 2002, the last in our sample. For instance, *ceteris paribus*, this information suggests that a 90% increase in Bank of America's net interest rate in the Daytona Beach-FL market has the same effect over Wachovia a 1% expansion in Bank of America's number of branches. The same relationship becomes 12% net rate to 1% branches in the Tallahassee-FL market. In other words, when Bank of America seeks to compete against Wachovia, branch expansion is less efficient in Tallahassee than in Daytona Beach.

Daytona Beach has a high branch density, and Bank of America and Wachovia are the top two banks in number of branches in the locality, with 30 and 29, respectively. They are competing very closely for the commanding lead in market share and branch density. An expansion in the number of branches of Bank of America can consolidate its leadership position in terms of branches, tilting consumer's choice towards

FIGURE 4

NET INTEREST RATE OWN-ELASTICITY DISTRIBUTION

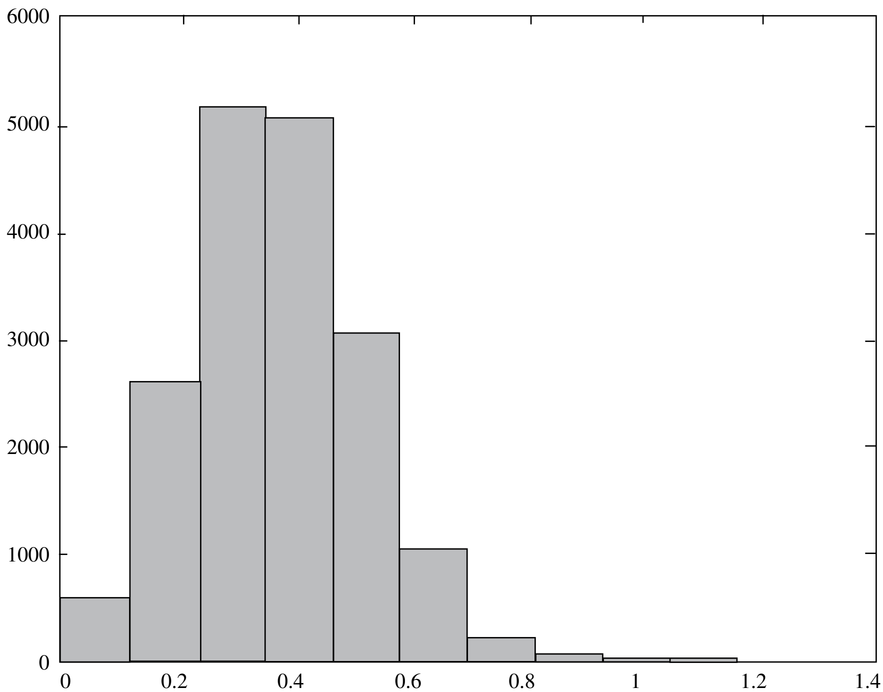
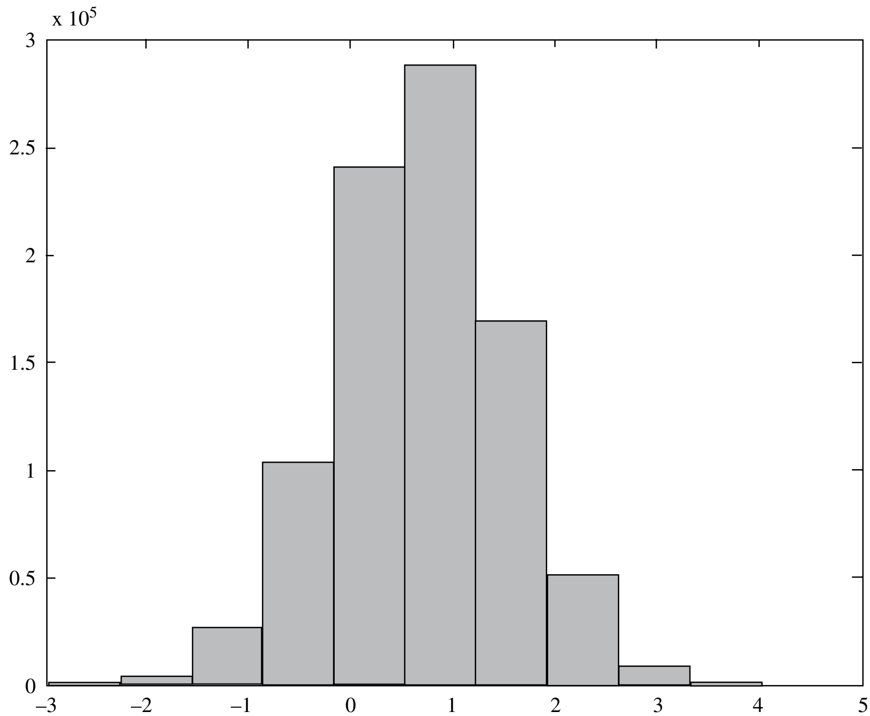


FIGURE 5

BRANCH DENSITY OWN ELASTICITY DISTRIBUTION



the new clear leader. On the other hand, in Tallahassee, Bank of America has only 7 branches and Wachovia only 6. Also, there are 4 other banks with more branches, with 17 branches being offered by the largest bank in the market, Capital City Bank. In this situation, Bank of America and Wachovia remain in the mid pack group, and a little deviation in branches doesn't do much to help either firm to steer away from the pack and take the leadership positions.

These values can also be very useful when analyzing expansion strategies. In a way, they can help quantify by how much the sum is larger than the parts in a successful merger or acquisition. If two banks merge, the resulting branch density in markets where they used to compete will be higher, giving an extra boost to the new entity's market share.

TABLE 5

CROSS ELASTICITIES BY MSA MARKET IN THE YEAR 2002

Elasticity of…. … with respect to ….	Interest Rate		Branch Density		Equiv. Ratio	
	Bank of America	Wachovia	Bank of America	Wachovia	Bank of America	Wachovia
	Wachovia	Bank of America	Wachovia	Bank of America	Wachovia	Bank of America
Asheville, NC	−0.02	−0.03	−0.26	−0.53	11.95	20.13
Athens, GA	−0.03	−0.01	−0.03	−0.03	1.06	6.65
Augusta-Aiken, GA-SC	−0.02	−0.02	−0.27	−0.29	11.44	16.49
Charleston-North Charlest, SC	−0.04	−0.02	−0.97	−0.92	23.12	43.91
Charlottesville, VA	−0.05	−0.02	−0.58	−0.38	12.72	19.76
Columbia, SC	−0.04	−0.03	−0.72	−1.04	19.15	39.48
Danville, VA	−0.02	0.00	0.00	0.00	−0.17	−1.28
Daytona Beach, FL	−0.06	−0.02	−2.11	−2.17	34.42	90.03
Fayetteville, NC	−0.02	−0.01	−0.07	−0.09	2.70	7.91
Fort Myers-Cape Coral, FL	−0.06	−0.02	−1.80	−1.71	31.50	111.04
Gainesville, FL	−0.04	−0.02	−0.48	−0.54	11.65	23.49
Goldsboro, NC	−0.01	−0.02	−0.01	−0.06	1.49	3.80
Greenville, NC	−0.02	−0.02	−0.10	−0.24	5.00	12.65
Greenville-Spartanburg-Anderson, SC	−0.03	−0.02	−0.73	−1.13	26.54	63.55
Hickory-Morganton-Lenoir, NC	−0.01	−0.01	−0.03	−0.04	5.02	3.58
Jacksonville, NC	−0.04	−0.01	−0.03	−0.05	0.76	3.86
Johnson City-Kingsport-Br	−0.01	−0.01	−0.02	−0.03	1.99	5.80
Lakeland-Winter Haven, FL	−0.05	−0.02	−1.36	−0.79	26.88	49.37
Lynchburg, VA	−0.01	−0.02	−0.05	−0.10	6.90	5.80
Macon, GA	−0.03	−0.01	−0.16	−0.19	5.91	32.98
Myrtle Beach, SC	−0.02	−0.01	−0.34	−0.24	14.42	26.59
Naples, FL	−0.07	−0.02	−2.35	−1.96	34.78	93.20
Ocala, FL	−0.03	−0.01	−0.13	−0.13	4.26	21.50
Pensacola, FL	−0.02	−0.01	−0.12	−0.13	5.96	20.51
Punta Gorda, FL	−0.05	−0.01	−0.40	−0.18	8.86	22.49
Raleigh-Durham-Chapel Hil	−0.01	−0.02	−0.30	−0.45	20.55	24.68
Roanoke, VA	−0.02	−0.02	−0.06	−0.11	3.12	5.52
Sarasota-Bradenton, FL	−0.04	−0.01	−0.82	−0.60	21.86	105.16
Savannah, GA	−0.05	−0.02	−0.89	−0.88	18.19	39.87
Sumter, SC	−0.02	−0.02	−0.06	−0.11	3.56	5.99
Tallahassee, FL	−0.04	−0.01	−0.16	−0.12	3.93	11.60
Wilmington, NC	−0.02	−0.01	−0.13	−0.18	5.71	12.03

V. Effects of Banking Deregulation on Welfare

Since our model encompasses the most intense period of deregulation and a large part of the resulting intra-industry moves during the late 90s and early 2000s, we have a favorable framework to analyze changes in household's welfare. An important advantage of discrete choice models that incorporate characteristics other than price (or in our case, interest rate), is that they allow for a better understanding of changes in consumer's utility in times of change of the industry's landscape. For instance, during our sample period the number of banks greatly diminished, but at the same time, the number of branches increased. Interest rates moved, and larger banks took over smaller ones. By analyzing this change in *bank's characteristics* through our model, we can provide insights into the net results on consumers' utility.

Small and Rosen (1981) developed a basic framework for analyzing welfare changes in discrete choice models. Their basic equation from estimating equivalent variations (EV) must be slightly modified and reinterpreted given the particular features of our model. Mainly, we can't express EV as dollar amounts, since our utility function defined in equation (2) doesn't allow for the derivation of marginal utility of income, but for marginal utility of *net interest rate*. In this sense, we can use Small and Rosen's framework to estimate if there was an increase or decrease in welfare between time t and t' and quantify such change as the *equivalent variation in net interest* rate paid to deposits *(EVNIR)* that a consumer should have received in order to remain indifferent between the two choice sets. In other words, from equation (2), we can derive:

$$EVNIR = \frac{1}{\alpha_{it}} \left\{ E\left[U_i \left(X_{jt'}; \xi_j; r_{jt'}; \tau_{it}; \theta \right) \right] - E\left[U_i \left(X_{jt}; \xi_j; r_{jt}; \tau_{it}; \theta \right) \right] \right\} \tag{10}$$

as the *EVNIR* of consumer i for a particular market between starting time t and final time t'; and where

$$E\left[U_i \left(X_{t'}; \xi_j; r_{t'}; \tau_{it}; \theta \right) \right] = \sum_{j=1}^{J_t} s_{ijt} u_i \left(X_{jt'}; \xi_j; r_{jt'}; \tau_{it}; \theta \right) \tag{11}$$

and s_{ijt} is defined as for equation (9).

With equation (10) at hand, we proceeded to estimate the national median *EVNIR* between the years of our sample, with focus on 1994 and 2002, the first and final years. By national median, we mean the median of all consumers in all markets in our sample that are registered in both years of comparison.[39] We also divided the analysis among the lower 33%, middle 33% and top 33% of consumers with respect to their household income, and found the median *EVNIR* of each group to see how these welfare changes were distributed. Table 6 presents the corresponding *EVNIR* between 1994 and 2002 and suggests that, on average, welfare has improved for every income group. On the aggregate level, the increase in welfare is equivalent to a 0.47% increase in the net rate paid to household deposits held in 1994.

TABLE 6

EVNIR 1994-2002

Period	All	Low 33	Mid 33	Top 33
94-02	0.47%	0.54%	0.65%	0.27%

Table 6 also suggests that MSA population in the middle income group was the group most benefited, with an *EVNIR* of 0.65% whereas the top 33 income group was the least benefited. In a way, it makes sense that the wealthiest are heavy users of *financial services* that expand well beyond basic bank services (e.g. mutual funds and other investments, financial planning, etc.), and have had enough leverage to have their basic *banking* needs properly catered since the beginning by banks which consider them "premium" clients. Conversely, low-income households are less intense users of banking services and hold lower deposit balances; but have some potential large welfare gains if and when they become more incorporated into the banking network. In other words, banking services are of the outmost importance for America's middle class, and this group is the most prone to either gain or suffer with changes in the banking industry.

FIGURE 6

EVNIR WELFARE INDEX 1994-2002

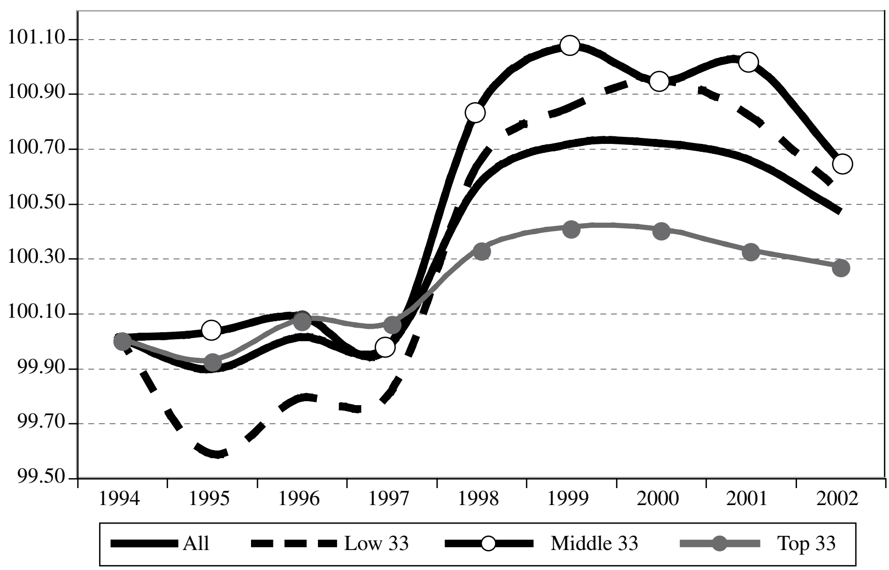


TABLE 7

EVNIR BY MSA MARKET, 1994-2002

EVNIR < -2%	-2% <= *EVNIR* < 0%	0% <= *EVNIR* < 2%	*EVNIR* > 2%	
Albany, GA	Erie, PA	Waco, TX	Greeley, CO	Canton-Massillon, OH
State College, PA	Elkhart-Goshen, IN	Fresno, CA	Lancaster, PA	Toledo, OH
Dover, DE	Eugene-Springfield, OR	Utica-Rome, NY	Green Bay, WI	Montgomery, AL
Yuma, AZ	Lafayette, LA	Lake Charles, LA	Sheboygan, WI	Flint, MI
Gainesville, FL	Pensacola, FL	Birmingham, AL	Glens Falls, NY	Terre Haute, IN
Flagstaff, AZ-UT	Corpus Christi, TX	Cedar Rapids, IA	Sharon, PA	Hamilton-Middletown, OH
Williamsport, PA	Las Cruces, NM	Monroe, LA	Rochester, MN	Fort Walton Beach, FL
Louisville, KY-IN	Syracuse, NY	Lima, OH	Wausau, WI	Billings, MT
Amarillo, TX	Tuscaloosa, AL	Tyler, TX	Goldsboro, NC	Mobile, AL
Reading, PA	Altoona, PA	Muncie, IN	Jacksonville, NC	Johnstown, PA
Albuquerque, NM	Lafayette, IN	Naples, FL	Longview-Marshall, TX	Ocala, FL
Medford-Ashland, OR	Fargo-Moorhead, ND-MN	Athens, GA	Sumter, SC	Decatur, AL
El Paso, TX	Bremerton, WA	Roanoke, VA	Galveston-Texas City, TX	Greenville, NC
Lubbock, TX	Tulsa, OK	St. Joseph, MO	Portland, ME	Fort Collins-Loveland, CO
Iowa City, IA	Merced, CA	Asheville, NC	Abilene, TX	Joplin, MO
Duluth-Superior, MN-WI	Anniston, AL	Chico-Paradise, CA	Fort Myers-Cape Coral, FL	Buffalo-Niagara Falls, NY
Jamestown, NY	Eau Claire, WI	Brazoria, TX	Jackson, TN	Lakeland-Winter Haven, FL
Jackson, MS	Laredo, TX	South Bend, IN	Springfield, MO	Columbus, GA-AL
Tucson, AZ	Savannah, GA	Gary, IN	Columbia, MO	Lynchburg, VA
Lexington, KY	Ann Arbor, MI	Reno, NV	St. Cloud, MN	Daytona Beach, FL
Wichita Falls, TX	Yakima, WA	Dothan, AL	Boulder-Longmont, CO	Macon, GA
Boise City, ID	Mansfield, OH	Waterloo-Cedar Falls, IA	Kenosha, WI	Alexandria, LA
Tallahassee, FL	Sioux City, IA-NE	Santa Fe, NM	Hattiesburg, MS	Bloomington, IN
Lincoln, NE	Decatur, IL	Yuba City, CA	Lansing-East Lansing, MI	Punta Gorda, FL
Springfield, IL	La Crosse, WI-MN	Provo-Orem, UT	Columbia, SC	Houma, LA
Madison, WI	York, PA	Rockford, IL	Anchorage, AK	Kokomo, IN
	Myrtle Beach, SC	Yolo, CA	Benton Harbor, MI	Charlottesville, VA
	Odessa-Midland, TX	Salem, OR	Chattanooga, TN-GA	Baton Rouge, LA
	Panama City, FL	Colorado Springs, CO	Dayton-Springfield, OH	Augusta-Aiken, GA-SC
	Kankakee, IL	Redding, CA	Rochester, NY	Danville, VA
	Modesto, CA	Wilmington, NC	Florence, AL	Rocky Mount, NC
	Janesville-Beloit, WI	Knoxville, TN		Fayetteville, NC
	Champaign-Urbana, IL	Hagerstown, MD		
	Pueblo, CO	Racine, WI		
	Spokane, WA	Fort Smith, AR-OK		
	Stamford-Norwalk, CT	Fort Wayne, IN		
	Huntsville, AL			

Figure 6 shows an *EVNIR* index built with base 1994 = 100 for all income classes and gives an idea on how welfare evolved due to changed in bank's characteristics from 1994-2002. It provides further insight into the evolution of the impact of deregulation on welfare.[40] It confirms our initial finding that the middle income group was the most benefited, but also shows that deregulation benefits ripened in the late 90s and stalled (and maybe declined) towards the last two or three years of our sample. It raises the question if benefits of deregulation were fully realized after a massive wave of mergers and acquisitions and further industry consolidation might not be welfare enhancing anymore.

Welfare analysis by MSA regions

We have already found some evidence that, overall, welfare increased during the deregulation period of 1994-2002. Our model also offers some insights into the geographical distribution of welfare changes. With this purpose, we calculated the *EVNIR* between 1994 and 2002 in all markets for which we have data for both years. The numbers are reported in Table 7 and are sorted by the size of the *EVNIR* gain. The results indicate that welfare gains were not uniformly distributed across markets and about 39% of the analyzed markets even experienced a welfare loss. A simple inspection doesn't reveal a clear geographical pattern of these losses, so it's likely that the interaction of many factors is responsible for the results. In any case, these results could be used to identify particular markets where a more thorough look at the effects of deregulation is granted. For instance, one could focus on the characteristics of markets such as Albany-GA, State College-PA and Dover-DE (the three worst welfare losers) and compare them to the characteristics of Danville-VA, Rocky Mount-NC and Fayetteville-NC (the three largest welfare winners).

VI. Conclusions and Pending Agenda

The deregulation process experimented by the banking industry has generated a fundamental shift in the industry's structure. In the last 30 years, the number of banks in the U.S. has reduced by 40%, and in the 90s mergers have averaged 550 per year. These changes and the already important part that financial institutions play in the economy makes a model able to capture its market dynamics very relevant.

The industry has some distinctive characteristics, such as a large number of participants (a number that falls on the thousands), a large number of markets (in our case, more than 200 MSAs per each of the 9 years of analysis) and an unbalanced panel (many banks participate in only one market and no bank participates in all markets). These characteristics make discrete choice models an appealing alternative, which is not without challenges. One small modification we incorporated into the estimation of a BLP-type model was to allow consumer's to "purchase" deposits of different sizes, directly proportional to their wealth.

The logit version of the model offers a first glimpse of the many dimensions that are relevant when choosing a bank for deposits. Besides interest rate and fees, consumer's also value –and some times highly– aspects such as branch density, size of the bank, national network (states where the institution is present), local presence (where headquarters are located) and quality of service (measured as the ratio of wages to deposits). But it's the full model that captures the heterogeneity of consumer's taste, with respect of two key variables: income and commuting time. Specifically, wealthier consumer's value more high interest rates as well as larger banks. They also put a premium on their time, obtaining more utility from a denser branch network than less wealthy consumers do.

The full model is also helpful in offering a more realistic set of elasticities among the many banks present in the sample. It shows how market shares will respond depending on the market demographics and current choice set (i.e. offerings of other banks). For instance, when analyzing Bank of America's branch expansion in capturing Wachovia's market share, we found that such policy is more efficient in Daytona Beach than in Tallahassee.

Between 1994 and 2002, on average, consumer's experienced a welfare gain equivalent to a 0.47% increase in the rate paid to their deposit balances. Middle-income groups were the most benefited of all, with lower income groups coming next. We also found that most welfare gains seemed to have occurred during the late 90s, and after that, gains stagnated or even decreased slightly.

Also, welfare changes were not homogeneous since some markets gained more than others. Moreover, about 39% of the markets in our sample experienced a welfare loss.

Discrete choice theory and RUM models in particular are well suited to analyzing the banking industry. We believe ours has proven useful in such task, but there's still plenty of room for improvement. For instance, allowing for different fractions of income to be diverted into savings would make for a more realistic model. Incorporating a well specified cost function could allow for better assessment of market power; adding firms as holders of banks deposits into the "demographics" could increase the explaining power of the model. Finally, a joint estimation with the model for deposits with a RUM model for the loans market at a national level (since that seems to be the relevant "local market" for loans) could provide a much more complete picture of the industry.

Notes

1 Flow of Funds Accounts, Balance Sheet of Households and Non Profit Organization, 1995-2002.
2 As of 2002, Bank of America, the largest U.S. commercial bank in terms of deposits, held $ 562 billions in assets and $ 375 billions in deposits.
3 As of 2002, Wells Fargo had a presence in 24 of the 50 states of the Union.
4 Modeste (1997) even discusses the substitutability of Savings Deposits and Mutual Funds.
5 Data from the Chicago FED on mergers and acquisitions.
6 In this paper, we will define a small bank as holding less than $ 100 million in deposits, and a large bank as holding more than $ 1 billion.
7 Throughout this paper, we will refer to their paper as BLP.
8 The name is due to the fact that the consumer gets the utility *conditional* on choosing firm *j*.

[9] For the sake of simplicity, we will henceforth avoid sub indexes "t" which indicates the market where the consumer lives and the firm operates on a given year (i.e. "t" is a time-geography index) except in the cases when the market we are referring to is not obvious.

[10] We set the size of v at Kx1 because it will give us greater flexibility to analyze the possible interactions of the firm's observable characteristics (including the net deposit rate, r).

[11] Later, in order to simplify these interactions, we will assume that Σ is diagonal.

[12] We assume that each firm offers one and only one type of product (deposit services). Therefore, it equates to the same to talk about either firm or product characteristics.

[13] A more precise approach would be to assume a distribution of w and make it endogenous to the model. This poses a series of estimation challenges, given the already highly non-linear structure in the model.

[14] This assumption basically states that consumers work only with one banking institution, which seems reasonable and is backed up by evidence such as the Survey of Consumer Finances and Starr and McCluer (2001).

[15] The choice of number of accounts, which has been used by other studies, carries certain limitations. In particular, the correspondence between population size and number of accounts in certain markets become difficult to reconcile (mainly due to the differences in income within MSAs while trying to impose a uniform account size in dollar terms for the whole sample) and often results in a negative outside share. The common approach has been to adjust upwards the potential market size of problematic MSAs by an arbitrary factor.

[16] Traditionally, $P^*(\cdot)$ denotes population distribution functions (e.g. distribution of consumers). In our case, in order to allow for consumers to purchase dollar amounts proportional to their wealth, we need to integrate not over consumer's in the relevant market, but map the market's *wealth* to its demographics and integrate over the set defined by $A(.)$ to obtain the corresponding dollar-weighted market share.

[17] In other words, consumers know ex-ante that they will save a fraction W of their wealth but still must choose a bank for such balances (or another savings alternative) based on which provides them with their preferred mix of characteristics. Total market size becomes the fraction ω of the aggregate market's wealth. Market shares are the percentage of that money that goes into each bank. As long as $\omega > 0$ and is the same for every consumer, market shares are invariant to its value.

[18] The remaining induced variance is caused by the simulation of the D's and the v's.

[19] Note that we are estimating market share for dollar amounts and not from another definition such as number of accounts. In the latter, a small account has the same effect as a large account when calculating market shares, making the use of consumer's wealth in the previous equation incorrect.

[20] The issue is less clear when it comes to analyzing the loan side. It seems that firms are more able to shop for loans without being restricted to their own locality.

[21] Surveys are available, but they only cover a very small sub-sample of banks.

[22] Some evidence supports this assumption, though Radecki (1998) shows that banks tend to maintain a homogeneous pricing structure across their relevant markets. Also, a large number of banks in our sample participate in only one market, making the assumption irrelevant for those specific cases.

[23] Adding the firm's dimension to demographic posses several additional problems, mainly because it's less evident which local market is relevant to a firm. Fortunately, consumers account for 95% percent of time and savings deposits. Firms hold the majority (about 2/3s) of checking accounts.

[24] This is a strong assumption, since it might underestimate the weight of certain groups of consumer's such as retirees, or temporarily unemployed consumers. Nevertheless, these groups are usually in a dis-saving stage, perhaps with the exception of the very wealthy.

[25] Another option is to use the Flow of Funds Accounts and define market size as the total of Financial Assets held by households. Nevertheless, this number encompasses many assets that are not necessarily close substitutes of deposits (such as stocks).

[26] It's difficult to find instruments that meet both requirements. Most of the cost shifters are obtained from balance sheet information (FED Call Reports) and hence are at the bank level.

[27] Nevo (2000) strongly recommends the use of "brand specific" dummies to capture the characteristics that do not vary by market. Unfortunately, in our dataset, most of our instruments do not vary by market, and we would have to omit them due to its co-linearity with the dummies. This would leave us without enough instruments to identify the model (due to the extra instruments needed for the interest rate variable and the random coefficients).

[28] As explained in the introduction, during our sample period, 1994-2002, many important regulatory changes took place. These are discrete events that can't be easily quantified.

[29] From income statements and balance sheets, measurement might be off in both interest rates and fees, but, it's probably more accurate when ranking banks from "more profitable" accounts to "less profitable" accounts from the consumer point of view.

[30] This issue is widely discussed in the literature. See, for instance, Nevo (2000) and BLP.

[31] Anyone convinced that "pricing" summarizes the entire information set consumer's need to make a decision could argue that if interest rates were measured more accurately, this might not have been the case. Nevertheless, we believe that the dynamics of the market and the characteristic of the service are better explained (statistically and economically) by the interactions shown in our model.

[32] For brevity, we don't show the statistics for our year - dummy variables, but most of them are significant at the 99% level.

[33] All these specifications were estimated using our modified "income weighted" approach as well as the traditional one, which assumes only equal sized purchases by every consumer.

[34] Nevertheless, it can also be the case that what is rejected is the extreme value distribution of the error term.

[35] One could argue that some consumers might shy away from banks with a large branch proliferation in benefit of small, more "intimate" banks. Even though this might be a possible fringe case, we believe that the size-dummy should better capture this phenomenon.

[36] Even though not exclusive for the banking industry case, it is an interesting feature of this market how banks compete in variables other than "price" (which in this case is analogous to interest rate), reshaping their own characteristics to present themselves as more attractive to consumers. Also, it gives banks more flexibility, since it poses less potential complications to modify characteristics selectively and tailored to each particular market than to carry over a substantially different interest rate policy for every market (which also, has the potential to be arbitraged by specialized institutions).

[37] Freixas and Rochet (1999), p. 58.

[38] Bank of America and Wachovia were the leaders in that market and had 30 and 29 branches, respectively. Coquina and Cypress, two small local banks, had less than 3 and 2 branches, respectively.

[39] Since our panel is unbalanced also in MSAs for different years, a minor percentage of them (about 10%) can't be compared and must be discarded to "square" the panel in its MSA dimension.

[40] To construct the index, we performed a year by year analysis fixing 1994 as our base with a value of 100. We increased the index of the corresponding year by the formula 100x(1+*EVNIR*).

References

AMEL, D. F. and T. HANNAN (1999). "Establishing Banking Market Definitions Through Estimation of Residual Deposit Supply Equations". *Journal of Banking and Finance*, 23 (11), pp. 1667-1690.

BAIN, J. (1951). "Relation of the Profit Rate to Industry Concentration: American Manufacturing, 1936-1940". *Quarterly Journal of Economics*, 65 (3), pp. 293-324.

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

BERRY, S.; J. LEVINSOHN and A. PAKES (1995). "Automobile Prices in Market Equilibrium". *Econometrica*, 63, pp. 841-890.

BRESNAHAN, T. (1989). "Empirical Studies of Industries with Market Power". R. Schmalensee and R.D. Willig, eds. *Handbook of Industrial Organization*, Elsevier Publishers B.V. Vol II, Chapter 17.

BUCH, C. M. and G. DELONG (2004). "Cross-Border Bank Mergers: What Lures the Rare Animal?". *Journal of Banking and Finance*, 28 (9), pp. 2077-2102.

DICK, A. (2001). "Demand Estimation and Market Power in the Banking Industry". *Mimeo, MIT.*

ELFAKHANI, S.; R. GHANTOUS, I. BAALBAKI (2003). "Mega-mergers in the US Banking Industry". *Applied Financial Economics*, 13 (8), pp. 609-22.

EVANOFF, D. and E. ÖRS (2002). "Local Market Consolidation and Bank Productive Efficiency". *Federal Reserve Bank of Chicago*, Working Paper 2002-25.

FREIXAS X, and H. CH. ROCHET (1999). *The Microeconomics of Banking*, The MIT Press.

GRUBEN, W. C. and R. P. McCOMB (1996). "Privatization, Competition and Supracompetition in the Mexican Commercial Banking System". *Federal Reserve Bank of Dallas*.

HANNAN, T. and J. N. LIANG (1993). "Inferring Market Power from Time-Series Data: The Case of the Banking Firm". *International Journal of Industrial Organization*, 11 (2), pp. 205-218.

MASSA, M. (2003). "How Do Family Strategies Affect Fund Performance? When Performance-Maximization is Not the Only Game in Town". *Journal of Financial Economics*, 67 (2), pp. 249-304.

McFADDEN, D. (1973). "Conditional Logit Analysis of Qualitative Choice Behavior". *Frontiers of Econometrics*. Ed. by P. Zarembka. New York: Academic Press.

McFADDEN, D. (1981). "Econometric Models of Probabilistic Choice". C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data*. Cambridge: MIT Press.

McFADDEN, D. (1978). "Modeling the Choice of Residential Location". In: Karlqvist, *et al*., eds., *Spatial Interaction Theory and Planning Models.* Amsterdan: North-Holland.

MELNIKOV, O. (2001). "Demand for Differentiated Durable Products: The Case of the U.S. Computer Printer Market". *Mimeo,* Yale University.

MODESTE, N. and M. MUHAMMAD (1997). "The Substitutability of Saving Deposits and Mutual Equity Funds: The Empirical Evidence". *International Review of Economics and Business*, 44 (4), pp. 785-92.

NEVO, A. (2000). "A Practitioner's Guide to Estimating Random-Coefficients Logit Models of Demand". *Journal of Economics and Management Strategy*, 9 (4), pp. 513-548.

NEVO, A. (2001). "Measuring Market Power in the Ready-To-Eat Cereal Industry". *Econometrica*, 69 (2), pp. 755-784.

PETERSEN, M. A. and R. G. RAJAN (2001). "Does Distance Still Matter? The Information Revolution in Small Business Lending". *Mimeo*, March.

PETRIN, A. (2002). "Quantifying the Benefits of New Products: The Case of the Minivan". *Journal of Political Economy*, 110 (4), pp. 705-729.

RADECKI, L.J. (1998). "The Expanding Geographic Reach of Retail Banking Markets". *Federal Reserve Bank of New York Economics Policy Review*, 4 (2), pp. 15-34.

RHOADES, S. A. (1996). "Bank Mergers and Industry-Wide Structure, 1980-1994". *Staff Study 169,* Board of Governors of the Federal Reserve System.

RHOADES, S. A. (1998). "The Efficiency Effects of Bank Mergers: An Overview of Case Studies of Nine Mergers". *Journal of Banking and Finance*, 22 (3), pp. 273-91.

SANTOMERO, A. M. (1999). "Bank Mergers: What's a Policymaker to Do?". *Journal of Banking and Finance*, 23 (2-4), pp. 637-43.

SMALL, K. A. and H. A. ROSEN (1981). "Applied Welfare Economics with Discrete Choice Models". *Econometrica,* 49 (1), pp. 105-130.

STANGO, V. and CH. KNITTEL (2003). "Compatibility and Pricing with Indirect Network Effects: Evidence from ATMs". *Working Papers Series, Chicago FED*, WP 2003-33.

SUOMINEN, M. (1994). "Measuring Competition in Banking: A Two-Product Model". *Scandinavian Journal of Economics*, 96 (1), pp. 95-110.

WHEELOCK, D. C. and P. W. WILSON (2004). "Consolidation in US Banking: Which Banks Engage in Mergers?". *Review of Financial Economics*, 13 (1-2), Special Issue, pp. 7-39.