Changes in the Size Distribution of U.S. Banks: 1960–2005

Hubert P. Janicki and Edward Simpson Prescott

In 1960, there were nearly 13,000 independent banks. By 2005, the number had dropped in half, to about 6,500. In 1960, the ten largest banks held 21 percent of the banking industry's assets. By 2005, this share had grown to almost 60 percent. A great deal of these changes started during the deregulation of the 1980s and 1990s. (Figures 1 and 2 report the time paths for these two measures.)

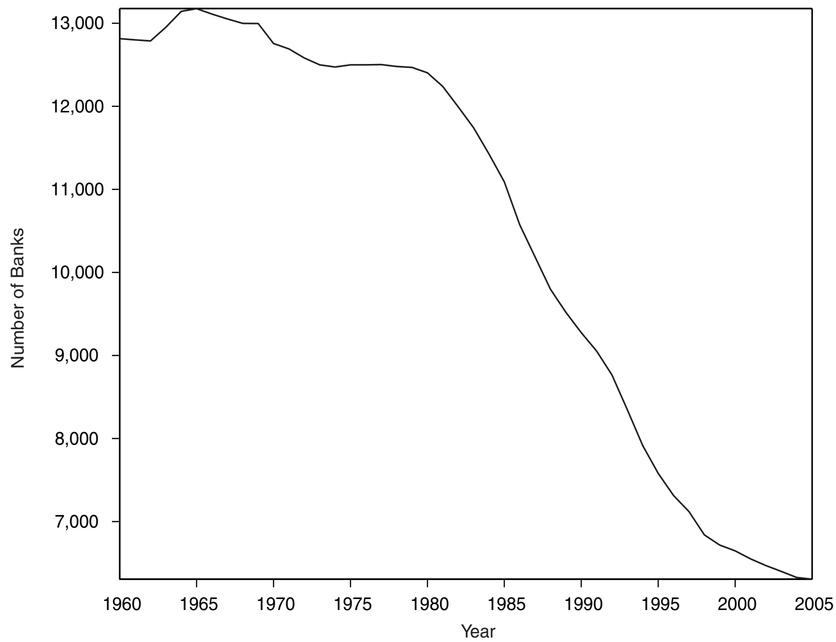
By any measure, these numbers represent a dramatic change in the bank size distribution over the 1960–2005 period. This article documents the extent of this change. It also documents the change in bank size dynamics, that is, the entry and exit of banks and the movement of banks through the size distribution. During this period, new banks formed, many more exited, either because of failure or merger, and many others changed in their size. For example, of the ten largest banks in 1960, only three were still among the top ten largest in 2005.[1]

We document these facts because they are an important step in developing a theory of bank size distribution. Although we do not provide one, such a theory would be valuable because it could be used to answer important questions such as: How costly were the pre-1980 limits on bank size? How

Hubert P. Janicki is a graduate student in economics at Arizona State University. The views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System. The authors would like to thank Huberto Ennis, Yash Mehra, Andrea Waddle, and John Weinberg for helpful comments.

[1] The ten largest banks in 1960 were Bank of America, Chase Manhattan Bank, First National City Bank of New York, Chemical Bank New York Trust Company, Morgan Guaranty Trust Company, Manufacturer's Hanover Trust Company, Bank of California, Security First National Bank, Banker's Trust Company, and First National Bank of Chicago. At the end of 2005, only three of these banks still existed: Bank of America, Citigroup (formerly National City Bank of New York), and JP Morgan Chase (an amalgamation of many of the banks listed above).

Figure 1 Total Number of Independent Banks



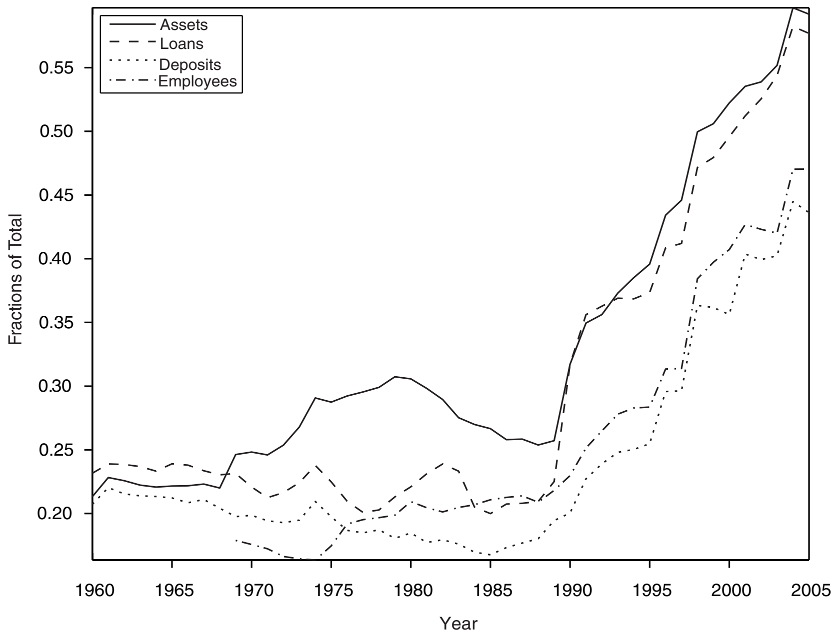
Notes: All banks and bank holding companies that are under a higher level holding company are treated as a single independent bank. Section 3 gives a more precise definition of an independent bank.

will the bank size distribution continue to evolve? And will there be more concentration? If so, should policy do anything about it?

Our analysis of the size distribution emphasizes fitting the data to the lognormal and Pareto distributions. These distributions are utilized because they are commonly used to describe skewed distributions and frequently have been used to describe firm size distribution. As we will demonstrate, the lognormal poorly fits the upper right tail of the size distribution. The Pareto distribution fits better with this part of the distribution, but the quality of the fit is much better before deregulation.

We examine bank size dynamics along several dimensions. First, we determine whether the data satisfies Gibrat's Law, that is, whether growth is independent of firm size. We find that Gibrat's Law is a good description of the data during the 1960s and 1970s, before deregulation, but not a good description afterward. After the 1970s, large banks grew faster than small banks, though more so in the 1980s and 1990s than they did during the

Figure 2 Market Shares of Ten Largest Banks



Notes: The definition of a bank is given in Section 3. Data on number of employees starts in 1969.

2000–2005 period. Second, we document that entry into banking was remarkably stable over the entire period. Entry is cyclical but averages about 1.5 percent of total operating banks. Finally, we calculate transition matrices, that is, the probability a bank will move from one size category to another, over each of the decades. Following Adelman (1958) and Simon and Bonini (1958), we use these transition probabilities and the entry data from 2000–2005 to forecast continued changes in the size distribution. The forecast predicts a continued decline in the total number of banks, but at a much slower rate than in the 1980s and 1990s, followed by a leveling off in the decline. It also predicts that there will still be a large number of small banks as well as a sizable number of mid-size banks. If the present trends continue, the transition in banking that began in the 1980s is slowing down and coming to an end.

1. LITERATURE

In many industries, the distribution of firm size is highly skewed to the right, that is, there are many small firms and a few large ones. One distribution that has this characteristic and is frequently used to describe firm size distribution is the lognormal distribution. A random variable is lognormally distributed if the logarithm of the random variable is normally distributed.[2]

Early studies of firm size distribution, namely Gibrat (1931), found that the lognormal distribution fit the empirical data fairly well. Gibrat (1931) also found evidence that firm growth was independent of firm size. This latter finding, often called Gibrat's Law or the Law of Proportionate Effect, was important because a statistical process that satisfies it would generate a lognormal distribution in the long run.

Later studies have found mixed support for Gibrat's findings. In particular, studies in the 1980s found that the proportional rate of growth of a firm conditional on survival decreases in size. Sutton (1997) is a good survey of these results.

Another category of distributions used to fit the size distribution is based on the power law. This category takes the form

$$f(x) = cx^{-\alpha},$$

where $x > 0$ and $c > 0$. In economics, the Pareto distribution is a power law distribution often used to describe highly skewed data.[3] It is similar to the lognormal, but with a thicker right tail.

Power law distributions have been used in the sciences to fit data in a wide variety of applications. Newman (2005) surveys various applications of the power law, including studies of word frequency, magnitude of earthquakes, diameter of moon craters, intensity of solar flares, and population of cities. One property observed in many applications is that the size of the r-th largest observation is inversely proportional to its rank.[4] It is observed so frequently that it is called Zipf's Law. Like the lognormal distribution, Zipf's Law can be generated with appealing assumptions on the dynamics. For example, Simon and Bonini (1958) study entry dynamics by assuming a constant probability of entry and show that the distribution follows a power law in the upper tail. See Gabaix (1999) for a detailed study of Zipf's Law.

[2] The probability density function of a lognormal distribution is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{x}e^{-(log(x)-\mu)^2/(2\sigma^2)},$$

where $x > 0$, μ and σ are the mean and standard deviation, respectively, of the natural log of x.

[3] The probability density function of a Pareto distribution is

$$f(x) = kx_{min}^k x^{-(k+1)}$$

where $x > x_{min}$, $x_{min} > 0$ and $k > 0$. The Pareto distribution is usually expressed as the probability that random variable X is greater than x, that is, $Prob(X \geq x) = x_{min}^k x^{-k}$, which is also a power law.

[4] Formally, $x_r = cr^{-\alpha}$, where observation x_r has rank r, c is a constant, and α is close to 1. We solve for r, normalize the equation by dividing by N (where the Nth ranked observation is x_{min}) and simplify to obtain the counter cumulative function: $r/N = 1 - F(x) = (x_{min}/x)^{1/\alpha}$. This is the Pareto distribution with coefficient $k = 1/\alpha$.

Whether Zipf's Law fits U.S. firm size data is a matter of some debate. Using 1997 U.S. census data, Axtell (2001) finds that it fits the firm size distribution. Using a different data set, however, Rossi-Hansberg and Wright (2006) find that it does not fit so well. They also find that establishment growth and exit rates decline with size.

The banking literature has long been interested in the size distribution of banks, partly because of the large degree to which laws and regulations limited bank size. Recent studies include Berger, Kashyap, and Scalise (1995), Ennis (2001), and Jones and Critchfield (2005).

While the banking literature has long noted that bank size distribution is skewed, it has not typically tried to fit this using the previously mentioned category of distributions. This absence has made it difficult to compare bank size distribution with that of other industries.

There is part of this literature, however, that tests for Gibrat's Law in the banking industry but for smaller samples than those used in this study. Alhadeff and Alhadeff (1964) analyze growth in assets of the 200 largest U.S. banks between 1930–1960 and find the largest banks grew more slowly than the banking system itself. They find, however, that the top banks that survived throughout the sample period grew faster than the system as a whole and attribute this to mergers among the largest banks. Rhoades and Yeats (1974) analyze growth among U.S. banks by deposits for 1960–1971 and find that the largest banks grew more slowly than the whole banking system. In a study of the 100 largest international banks from 1969 to 1977 by assets, Tschoegl (1983) finds that growth rates of banks are roughly independent of size, but that growth rates exhibit positive serial correlation. Saunders and Walter (1994) use international data for the 200 largest banks for the 1982–1987 period and reject Gibrat's Law, finding that the smaller banks grow faster than larger banks in terms of assets. More recently, Goddard, McKillop, and Wilson (2002) find evidence that during the 1990s in the United States, large credit unions grew faster than their smaller counterparts.

As we discussed, the firm growth results are important because they ultimately determine the distribution of firm sizes. One interesting strand in the literature calculates a Markov transition matrix of movement between size categories and then calculates stationary size distributions. Early examples include Hart and Prais (1956), Simon and Bonini (1958), and the study of the steel industry by Adelman (1958). The only study that we are aware of that applies this technique to banking is Robertson (2001).

More recently, the literature has attempted to generate firm size dynamics, and ultimately a size distribution, from models of maximizing behavior of firms. Sutton (1997) surveys several such models. One paper in this literature is the learning model of Jovanovic (1982), in which new firms receive productivity shocks, learn about them over time, and then decide whether to continue or exit. Another prominent example is Hopenhayn (1992).

2. LEGAL AND REGULATORY LIMITS ON BANK SIZE

Describing bank size distribution is particularly important because of the many legal and regulatory limits on bank size that existed through the 1970s and were removed during the 1980s and 1990s. As we will see, the removal of these barriers coincide with dramatic changes in size dynamics and the size distribution.

In 1960, banks could not branch across state lines and some states even forbade branching within a state.[5] The 1966 Douglas Amendment to the 1956 Bank Holding Company Act allowed interstate banking only with expressed authorization by participating states. However, no state allowed interstate banking at the time and the amendment was not even exercised until 1978 when Maine allowed out-of-state bank holding companies (BHCs) to operate within the state.[6] Over the next 12 years, many of the intrastate and interstate restrictions were removed by the states.[7]

The remaining interstate banking restrictions were removed by the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994. The act permitted bank holding companies to acquire banks in any state and, beginning in 1997, allowed interstate bank mergers. (See Kane [1996] for a summary of the act.) More recently, the Gramm-Leach-Bliley Act of 1999 allowed banks to engage in nonbanking financial activities such as insurance.

3. THE DATA

Many banks operate under a bank holding company structure. A bank holding company is a legal entity that i) directly or indirectly owns at least 25 percent of the bank's stock, ii) controls the election of a majority of a bank's directors, or iii) is deemed to exert controlling influence of bank policy by the Federal Reserve (Spong 2000). Many bank holding companies have multiple banks and even other holding companies under their control. Historically, this legal organization was used to avoid some of the restrictions on branching (Mengle 1990). In many cases, a bank holding company would operate many activities jointly. For this reason, we follow Berger, Kashyap, and Scalise (1995) and treat all banks and bank holding companies under a higher level holding

[5] The 1927 McFadden Act forbade interstate branching by federally chartered banks. Later, the Federal Reserve extended the ruling to include all state-chartered banks that are regulated by the Federal Reserve.

[6] There were some means around these restrictions. For example, the 1956 Bank Holding Company Act did not limit the location of *nonbank* subsidiaries of bank holding companies, so some banks had a cross-state network of *nonbranch offices* that would specialize in activities like lending. Also, some exemptions were allowed for the acquisition of insolvent banks from government deposit insurance funds (Kane 1996).

[7] Jayaratne and Strahan (1997) list when states removed restrictions to interstate banking and intrastate branching.

company as a single independent banking enterprise. For convenience, we will typically refer to each of these entities as a bank.

Data on banks are taken from the Reports on Condition and Income (the "Call Report") collected by federal bank regulators. We use fourth quarter data on all commercial banks in the United States. We look specifically at commercial banks and exclude savings banks, savings and loan associations, credit unions, investment banks, mutual funds, and credit card banks. Individual commercial banks that belong to a holding company are then grouped according to a unique bank holding company regulatory number, and their assets, deposits, loans, and employees are summed and replaced by one entry in our data set. Our data set, therefore, tracks bank holding companies and independent commercial banks not affiliated with a holding company. It does not distinguish between a merger and a failure. Both events are treated as an exit.

We use four measures of bank size. The first is commercial bank assets. For this variable, we have data from 1960–2005. Prior to 1969, we only have domestic holdings, but after 1969, we have foreign and domestic holdings. The next two size measures are domestic holdings of deposits and loans. For both of these variables, we have data from 1960–2005. The final size measure is the number of domestic employees.[8] For this last measure, we only have data for 1969–2005. Assets and loans are adjusted to include off-balance-sheet items starting in 1990. (See the Appendix for details.) All the variables are adjusted for the total aggregate size of that variable in each year. In particular, firm size data are converted into market share numbers for that year and then multiplied by the total quantity in the banking industry of that variable in 2004. The market share adjustment facilitates comparison across years, while the scaling by 2004 aggregate quantities gives a sense of the size in terms of recent quantities.

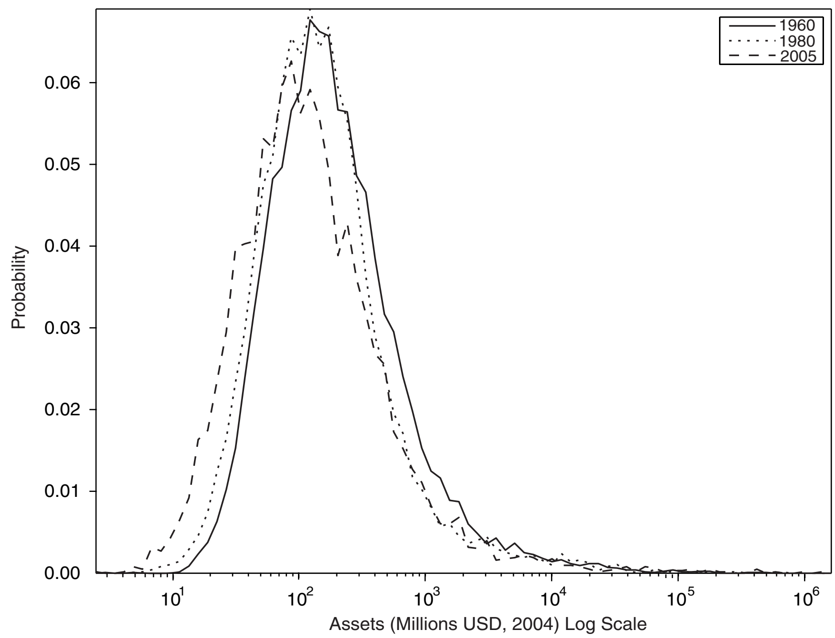
4. THE SIZE DISTRIBUTION

The size distribution of banks has always been skewed, but it has become more so since the 1960s. Figure 3 shows the distribution of assets for 1960, 1980, and 2005. Each year is normalized by the total assets in that year relative to 2004 so that the distributions are comparable over time. The distribution is plotted on a *log* scale. Because the size distribution is so skewed to the right, the log scale—or something similar—is needed to fit all the banks on the graph.

Figure 3 demonstrates that there are a large number of small banks and a few large banks. As is evident, there is a shift in the distribution to the left

[8] The Call Report counts employees in terms of full-time equivalents.

Figure 3 Change in Bank Size Distribution Over Time



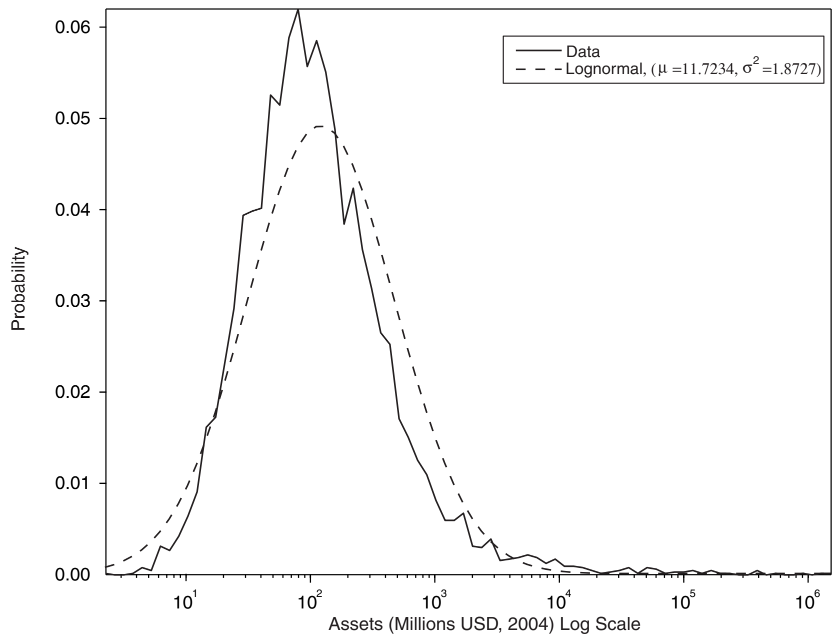
Notes: Each line is a probability distribution of bank size as measured by assets for a given year.

over time. Since we have scaled assets in each year to be the same scale, this change means that a higher fraction of the assets are being held by the small number of large banks, as indicated earlier in Figure 2.

Visually, the graphs suggest that the distribution might be accurately represented by a lognormal distribution. Figure 4 reports the actual distribution and an estimated lognormal distribution for assets in 2005, where the lognormal has parameters μ and σ^2 obtained from the 2005 data set. However, the distribution fails the Kolmogorov-Smirnov test for goodness-of-fit. This is true for almost all the years in the data set, as well as for the other size measures.

The estimated lognormal distribution does a particularly poor job of fitting the right tail of the distribution. This is hard to see in Figure 4 because of the small number of large banks. The fit at the right tail is better seen if we use a rank-frequency, or Zipf plot. For a power law distribution plotted on a log scale, this type of graph has the valuable property that the slope will be linear. For example, let $x_r = cr^{-\alpha}$, where r is the rank of a variable. Taking the

Figure 4 Size Distribution of Banks in 2005



Notes: The parameters μ and σ are the mean and standard deviation, respectively, of the natural log of assets. The lognormal distribution is calculated using these parameters.

logarithm of both sides, we obtain the equation,

$$ln(x_r) = ln(c) - \alpha ln(r).$$

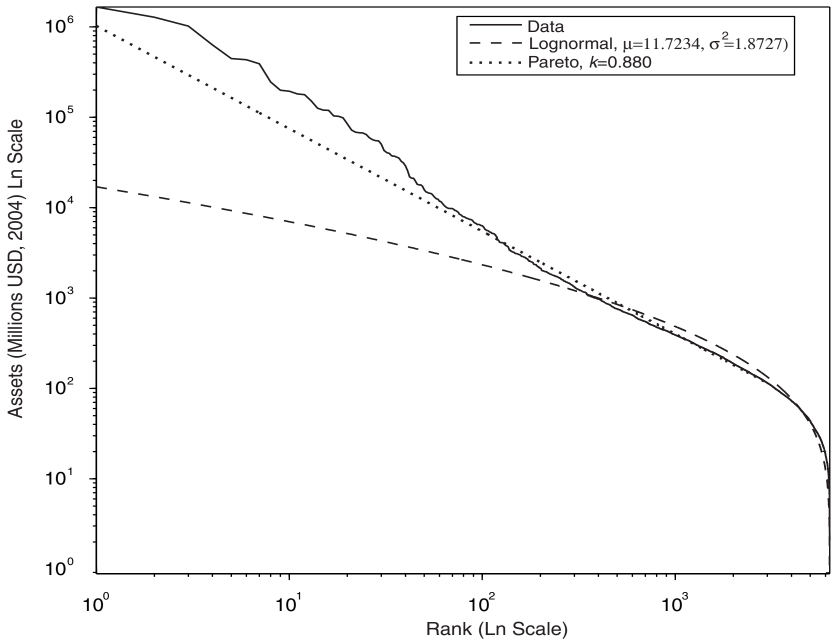
If $\alpha = 1$, then Zipf's Law holds.

Figure 5 is a Zipf plot for the bank size distribution in 2005.[9] It plots the data and the lognormal distribution using the sample mean and variance.

Figure 5 demonstrates that the lognormal distribution underestimates the density of the right tail (which is the left side in the Zipf plot). Indeed, if the plot is linear, we know that the right tail of the distribution of bank holding companies can be better approximated by the Pareto distribution. Note, however, that only the right tail of the distribution—which corresponds to the left side of the figure—appears to fit Zipf's Law. That is, the distribution of bank holding companies seems to be lognormally distributed with a Pareto-distributed tail.

[9] Zipf plots for different size measures look very similar.

Figure 5 Zipf Plot of Banks in 2005



Notes: The parameters μ and σ are the mean and standard deviation, respectively, of the natural log of assets. The lognormal distribution is calculated using these parameters. The estimate for the Pareto distribution is only for the 3,000 largest banks.

We can formally test whether the right tail of the distribution satisfies Zipf's Law. One common method is to estimate the slope by fitting a linear regression of the form,

$$ln(x_i) = \alpha_0 + \alpha_1 ln(r_i) + \epsilon_i, \tag{1}$$

where r is the rank of the bank. The coefficient α_1 is a power exponent in the Pareto distribution and if it is equal to –1, then Zipf's Law holds.

Ordinary least squares estimates of (1) will underestimate α_1 (see Gabaix and Ioannides 2004). For this reason, we use the maximum likelihood estimator in Newman (2005) to estimate the power coefficient in the Pareto distribution, or equivalently the slope α_1:

$$\alpha_1 = -n^{-1}[\sum_{i=1}^{n}(ln(x_i) - ln(x_{\min}))].$$

Table 1 Zipf's Law: Maximum Likelihood Estimates

	1960		1970		1980		1990		2005	
	$-\alpha_1$	SE	$-\alpha_1$	SE	$-\alpha_1$	SE	$-\alpha_1$	SE	$-\alpha_1$	SE
Assets	1.001	0.001	1.030	0.001	1.018	0.001	1.017	0.001	1.136	0.001
Deposits	0.998	0.001	1.016	0.001	0.983	0.001	0.982	0.001	1.092	0.001
Loans	1.028	0.001	1.051	0.001	1.007	0.001	1.082	0.001	1.187	0.001
Employees	–	–	1.022	0.065	1.031	0.073	1.005	0.084	1.052	0.096

Notes: The estimates are reported as $-\alpha_1$.

The maximum likelihood estimates for several different years are reported in Table 1.[10] We restrict the estimates to the tail by limiting the sample to the largest 3,000 banks for each year. Although they are not identical across years, they broadly support Zipf's Law in the upper tail, but the results are sensitive to the cutoff.

The worst fit in Table 1 is in 2005. This can be seen in Figure 5. The straight line in Figure 5 graphs the power distribution for the maximum likelihood estimate based on the 3,000 largest banks. The distribution is a straight line with slope –1.136. The slope is too high for Zipf's Law. Furthermore, the slope is high because the larger banks are larger than predicted by Zipf's Law, though this is less true for the largest.

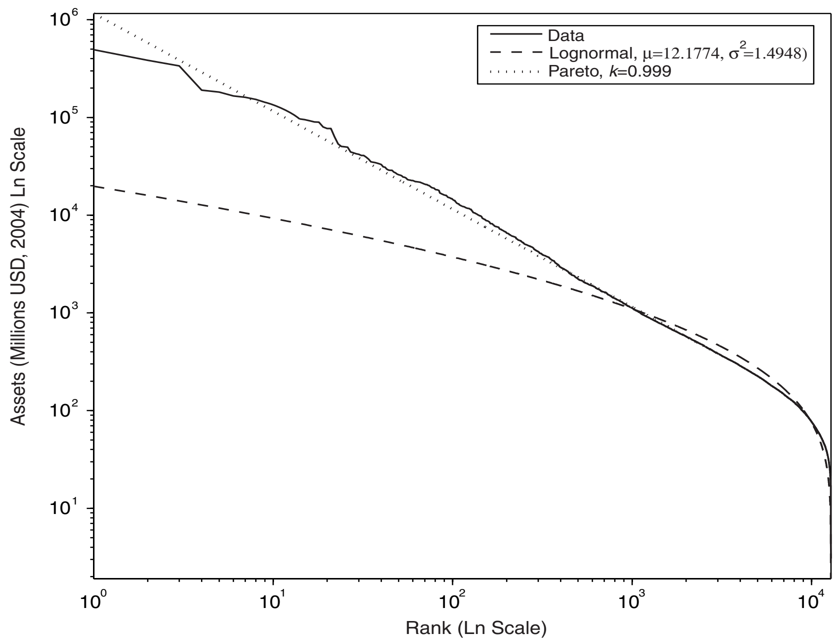
In contrast, the fit for assets in 1960 is excellent. Figure 6 is a Zipf plot for assets in 1960. The distribution is a straight line with slope –0.999, practically the same as in Zipf's Law.

The estimates for assets in 1970, 1980, and 1990 are similar, but this hides an important difference. The year 1970 is similar to 1960 in that many of the largest banks are smaller in size than predicted by the estimate. In 1980, this pattern changes to one where the largest banks are larger than predicted. The differences in the predictions for the largest banks are even larger in 1990 and look similar to that of 2005 (see Figure 5).

To summarize, only the right tail can reasonably be considered to be fitted by Zipf's Law, and the fit depends on the year. It does well in 1960, but starting in 1980 Zipf's Law predicts that the largest banks will be smaller than in the data. The lognormal distribution poorly fits the size of the largest banks but better fits the small banks.

[10] We calculate standard errors using the method outlined in Gabaix and Ioannides (2004).

Figure 6 Zipf Plot of Banks in 1960



Notes: The parameters μ and σ are the mean and standard deviation, respectively, of the natural log of assets. The lognormal distribution is calculated using these parameters. The estimate for the Pareto distribution is only for the 3,000 largest banks.

5. DOES GIBRAT'S LAW HOLD FOR BANKS?

Gibrat's Law states that firm size growth is independent of firm size. For our first test of this law, we fit the following linear equation,

$$ln(x_{it+1}) = \beta_0 + \beta_1 ln(x_{it}) + \epsilon_{it}, \tag{2}$$

where x_{it} is the size measure (assets, employees, etc.) of bank i at time t. A coefficient value of $\beta_1 = 1$ means that growth is independent of size.

We estimated (2) over each decade and over 2000–2005 using ordinary least squares. We only considered banks in the sample that were around at the beginning and end of the estimation period. Those that exited were dropped from the sample.[11] The estimates are reported in Table 2.

[11] Sometimes the literature includes exiting banks and sometimes it does not. See Sutton (1997) for more information. To reduce survivorship bias, we also estimated equation (2) over each year and found similar results to those reported in Table 2.

Table 2 Gibrat's Law: Estimates

	1960–1969		1970–1979		1980–1989		1990–1999		2000–2005	
Assets										
β_0	0.143	(0.031)	0.574	(0.034)	0.043	(0.057)	–0.187	(0.074)	0.015	(0.061)
β_1	**0.987**	(0.003)	**0.953**	(0.003)	**1.011**	(0.005)	**1.024**	(0.006)	**0.998**	(0.005)
R^2	0.929		0.915		0.847		0.823		0.869	
Deposits										
β_0	0.435	(0.030)	1.004	(0.033)	0.131	(0.058)	0.654	(0.075)	0.379	(0.063)
β_1	**0.971**	(0.003)	**0.929**	(0.003)	**0.999**	(0.005)	**0.972**	(0.007)	**0.970**	(0.005)
R^2	0.923		0.902		0.826		0.787		0.847	
Loans										
β_0	0.545	(0.037)	1.368	(0.039)	0.000	(0.071)	0.635	(0.074)	0.444	(0.069)
β_1	**0.960**	(0.003)	**0.901**	(0.003)	**0.993**	(0.006)	**0.964**	(0.007)	**0.962**	(0.006)
R^2	0.889		0.870		0.767		0.782		0.812	
Employees										
β_0	–	–	–0.021	(0.012)	0.240	(0.016)	0.284	(0.022)	0.101	(0.017)
β_1	–	–	**1.006**	(0.003)	**1.001**	(0.004)	**1.008**	(0.006)	**0.994**	(0.004)
R^2	–		0.897		0.862		0.829		0.900	

Notes: The table provides ordinary least squares estimates of equation (2) for each sample period. Standard errors are in parentheses. Numbers in bold are the slope estimates of the effect of firm size on growth. An estimate close to one is consistent with Gibrat's Law.

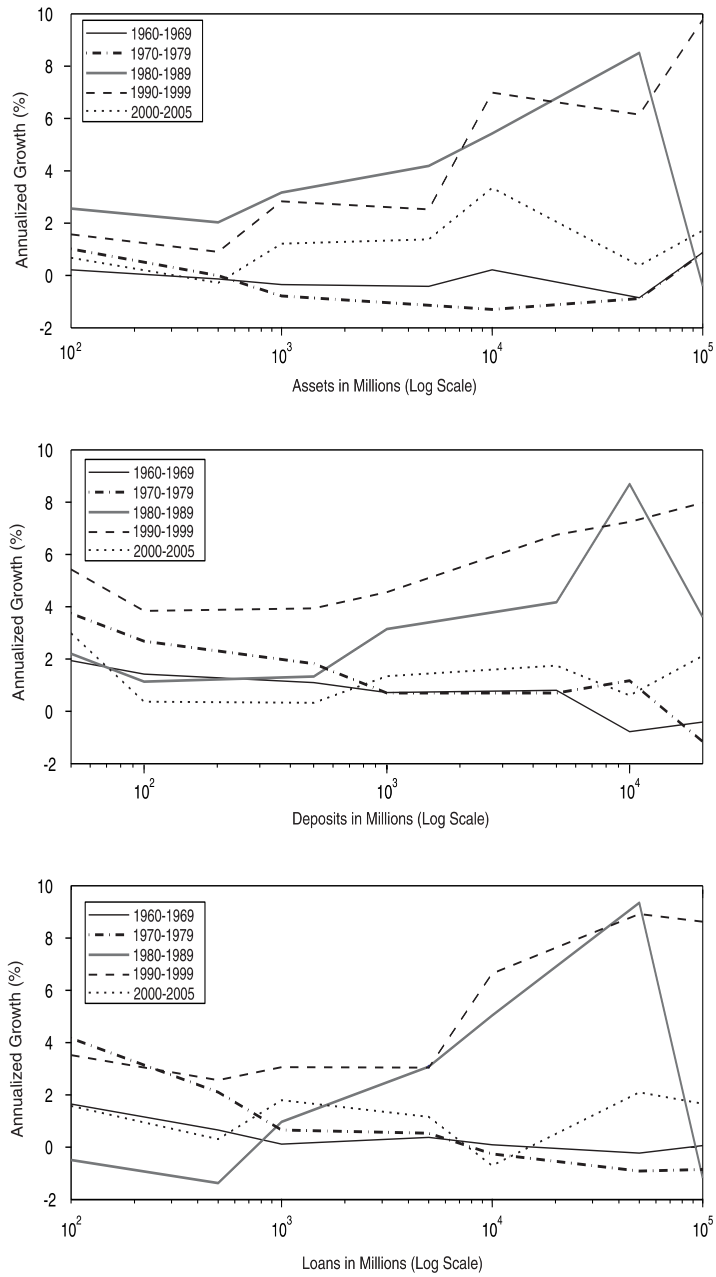
The estimates in Table 2 are close to one for all the decades and variables. While broadly supportive of Gibrat's Law, these estimates put a great deal of emphasis on small banks because they comprise most of the sample. For this reason, we broke the sample into different size categories and then calculated the annualized growth rates over the same periods for banks in each category. As before, we only considered banks that survived. Figure 7 reports the growth rates for assets, deposits, and loans.[12]

Just as in Table 2, growth measures do not vary much with size in the 1960s and 1970s.[13] In the 1980s and 1990s, however, the numbers in Figure 7 present a different picture than the estimates in Table 2. As demonstrated by the figure, the largest banks clearly grew faster than the small banks. This is also true for the 2000–2005 period, but the effect is less pronounced.

[12] A figure showing employee growth rates is excluded because the growth rates jump around significantly and do not show any clear patterns.

[13] We have also calculated annual growth rates over each period and then averaged them to get a similar figure. (Over the 1960–1969 period, for example, this meant calculating the 1960–1961, 1961–1962, etc., growth rates and then averaging them.) The results are similar.

Figure 7 Annualized Growth Rates by Size Categories



Notes: The size measure is broken into seven size categories and then average annualized growth rates are reported for each category.

Gibrat's model that yields the law of proportionate effect assumes that growth rates are not persistent over time. To check the validity of this assumption, we calculate the correlation of growth rates for surviving banks in each decade. We find that correlations of growth rates for all variables are very low. In the 1960s, it is 0.0610; in the 1970s, it is 0.0331; in the 1980s, it is 0.0691; in the 1990s, it is 0.0722; and after 2000, it is 0.1617. Correlation coefficients of growth in the remaining variables are of similar magnitude.

To conclude, growth rates appear to be independent of size in the 1960s and 1970s, but they are positively related to size in the 1980s and 1990s. In the 2000–2005 period, the growth rates are also higher for the largest banks, but less so than in the previous two decades. It appears that the 1960s and 1970s were relatively stationary periods, but the 1980s and 1990s were a long transition period, no doubt due to the legal, regulatory, and technological changes of the period. Finally, the 2000–2005 period appears to be the end of the transition, as the size dynamics seem to be returning slowly to the numbers of the 1960s and 1970s. Of course, this conclusion is tentative because trends calculated from five years of data can easily be transitory.

6. ENTRY AND EXIT

Despite the large number of banks that have exited the industry over the last 45 years, there has been a consistent flow of new bank entries. The number of entries and exits (including mergers) expressed as a fraction of the banking population is reported in Figure 8.

Visually, it is apparent that the flow of new banks is a relatively constant fraction of the banking population. To check this, we estimated a linear time trend of the number of entries as a fraction of the number of banks operating. Specifically we estimated the equation,

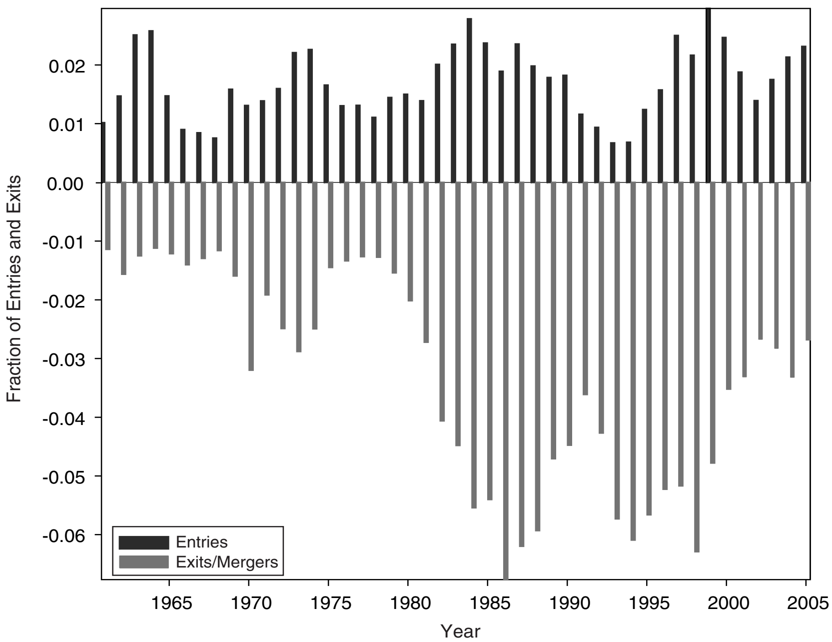
$$y_t = \gamma_0 + \gamma_1 t + \epsilon_t,$$

where y_t is the fraction of entries per year and t is the time trend. The ordinary least squares estimates are $\gamma_0 = 0.0145$ (0.0018), $\gamma_1 = 0.0001$ (0.0001) with R^2=0.0612. Standard errors are in parentheses. There is no time trend in the flow of new banks, though there is a significant amount of cyclical variation.

The fraction of banks that exit varied a great deal over time. There was a significant increase starting in the 1980s and, except for a short dip in the early 1990s, the high level continued through the late 1990s. No doubt much of this exit was due to mergers, particularly those that occurred in the 1990s, but our data does not allow us to distinguish between these two sources of exit. It is only in the last five years that the rate of exit seems to slow down.

It is striking that despite the huge number of bank exits starting in the 1980s, entry remained strong throughout the entire period. Interestingly, it is virtually uncorrelated with exit. For example, the correlation between exit

Figure 8 Fraction of Banks that Enter and Exit by Year



Notes: The chart reports the gross flow of banks that enter and exit expressed as a fraction of banks in each year.

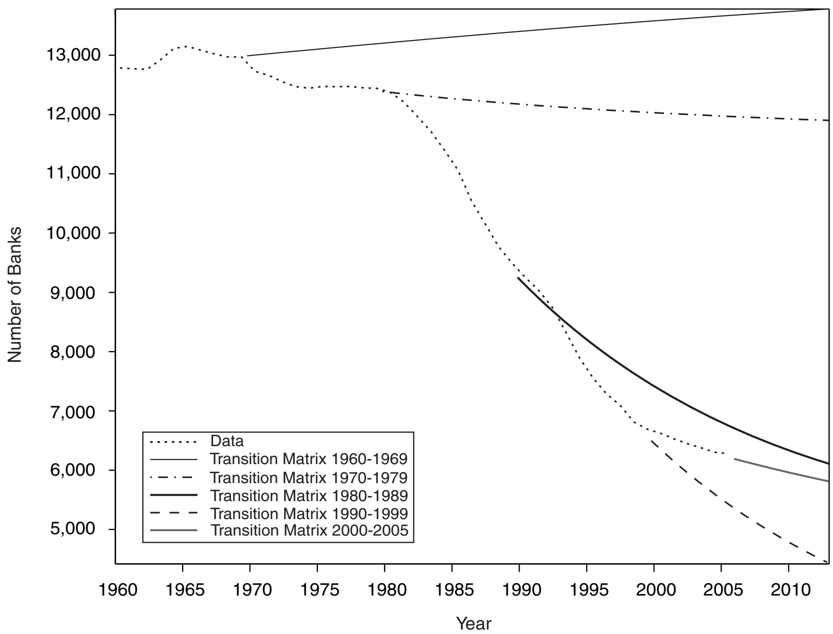
and entry for the 1985–2005 period is only –0.07. This evidence means that a theory of the bank size distribution needs to be able to generate robust entry, even when the total number of banks is declining.

7. THE DYNAMICS OF THE SIZE DISTRIBUTION

In this section, we use the information previously documented—the size distribution at a point of time, entry and exit rates, and bank dynamics—to make forecasts of what will happen to the size distribution. We also perform counterfactual experiments such as what would have happened to the bank size distribution if the dynamics had *not* changed. The purpose of these exercises is to develop a sense of how changes in size dynamics have mattered for the size distribution.

To make these calculations, we do not use growth rates at the individual bank level. Instead, we do something similar by constructing a simple Markov chain model following Adelman (1958). A Markov chain model splits the

Figure 9 Projected Number of Banks



Notes: Each projection is made by taking the transition matrix estimated over a given period and then calculating the total number of banks that would operate after the last year in that given period.

banks into a finite number of size categories. Probabilities of moving between size categories are summarized with a Markov, or transition, matrix. A Markov matrix is a square matrix, P, where element P_{ij} specifies the probability a bank that starts in size category i will move into size category j in the next period.

A Markov model allows for straightforward predictions about changes to a size distribution. For example, if the size distribution at time t is s_t, then the size distribution at time $t + n$ is

$$s_{t+n} = P^n s_t.$$

If a Markov model has the property that a bank starting in any category has a positive probability of moving to any other size category in a finite number of periods, then several useful theorems apply. First, there exists a

Table 3 Transition Matrix Size Categories (Scaled Dollars)

Size Categories	Assets
1	<100m
2	100m-500m
3	500m-1b
4	1b-5b
5	5b-10b
6	10b-50b
7	50b<

Notes: A list of size categories used in the Markov model. In all years, data are converted into market share numbers and then multiplied by the total quantity of assets in 2004.

stationary distribution, that is, there exists s such that $s = Ps$.[14] Second, the stationary distribution is unique and independent of the initial distribution of banks. Therefore, regardless of the initial distribution, if the transition matrix, P, is repeatedly applied to the distribution, then the size distribution will approach the unique stationary distribution.

We construct seven different size categories. These are listed in Table 3. All the data are scaled as we discussed earlier to make it possible to compare across years. We also include an eighth category that represents banks that are inactive. New banks come from this category and exiting banks move into it. We calculate the number of banks in each of the seven active categories. For the inactive banks, we assume that there is a large pool of 100,000 potential entries.

Our transition matrix is calculated by counting the fraction of banks that move from size category i to j each year over the specified time period. Entry rates are calculated so that a constant fraction of potential banks enter.

For our first exercise, we use the transition matrices estimated over several ranges of time periods to forecast the aggregate number of banks in the industry. We estimate several transition matrices and make predictions to 2013. The results are illustrated in Figure 9. For each period, we calculate the transition matrix and then forecast the change in the total number of banks as if the transition probabilities had not changed from that time period forward. This exercise is similar to one in Jones and Critchfield (2005), although our methodology is very different. An advantage of our Markov chain model is that we have information on the entire distribution at each point in Figure 9.

As is clear from the earlier analysis, as well as from Figure 9, the size dynamics changed significantly over this period. Nevertheless, the exercise

[14] Let s^j be the fraction of banks in size category j. The stationary distribution s is the solution to the set of equations $s = Ps$ and $\sum_j s_j = 1$.

Table 4 Stationary Distribution of Banks by Assets

Size	1960–1969		1970–1979		1980–1989		1990–1999		2000–2005	
Categories	1969	SS	1979	SS	1989	SS	1999	SS	2005	SS
1	0.333	0.357	0.379	0.439	0.379	0.471	0.440	0.523	0.480	0.540
2	0.504	0.516	0.492	0.485	0.488	0.435	0.436	0.377	0.399	0.363
3	0.085	0.076	0.064	0.044	0.067	0.049	0.061	0.051	0.060	0.050
4	0.060	0.042	0.046	0.021	0.043	0.027	0.044	0.037	0.043	0.033
5	0.008	0.004	0.007	0.003	0.008	0.005	0.006	0.004	0.008	0.006
6	0.009	0.005	0.010	0.005	0.010	0.007	0.007	0.005	0.006	0.005
7	0.002	0.000	0.002	0.003	0.005	0.007	0.005	0.002	0.005	0.003

Notes: Columns under a year list the actual size distribution in that year. Columns under "SS" list the stationary distribution as calculated from the estimated transition probabilities for that period.

is interesting because it illustrates how the dynamics matter for the size distribution. For example, the 1960s and 1970s look relatively stable. If these dynamics had not changed, there would not have been much change in the total number of banks. In the 1980s and 1990s, the dynamics predicted continued substantial declines in the number of banks. However, the transition probabilities changed significantly in the 2000–2005 period and as a consequence, the decline in the number of banks leveled off.

Table 4 reports stationary size distributions calculated from the estimated transition matrices for several periods in the data. We compare them with the actual distribution at the end of each period in order to illustrate whether the existing distribution was close to the stationary distribution. Interestingly, for every period, we find that the fraction of banks that are in the smallest category is higher in the stationary distribution than in the final year of the period. We also find that for every other period, the fraction of banks in each of the other size categories is less in the stationary distribution than in the final year (the only exceptions are during Category 2 [1960–1969] and Category 7 [1970–1979] and [1980–1989]).

Table 4 has several implications. First, even in the relatively stable decades of the 1960s and 1970s, the size distribution was not at a stationary point and under the estimated transition probabilities, the distribution would have continued to change. Second, there will continue to be large numbers of small banks, even if the fraction of assets they hold is not large. Third, we can see that the dynamics for the 1990s imply even more concentration than the dynamics from the 1980s. This was also suggested by Figure 9. Finally, there will continue to be a large number of mid-size banks. The most recent merger wave led some commentators to speculate that the bank size distribution would take a "barbell" shape with only small banks and large banks. The small banks would

Table 5 Distribution of Bank Assets

	2005		Stationary	
Size Categories	Fraction of Banks	Fraction of Assets	Fraction of Banks	Fraction of Assets
1	0.480	0.014	0.540	0.022
2	0.399	0.050	0.363	0.062
3	0.060	0.024	0.050	0.027
4	0.043	0.047	0.033	0.050
5	0.008	0.034	0.006	0.035
6	0.006	0.068	0.005	0.083
7	0.005	0.763	0.003	0.721

Notes: Fraction of assets for the stationary distribution was calculated by assuming that the mean assets in each size category are the same as in 2005.

survive because of their comparative advantage at small business lending and the big banks would take advantage of their scale economies. Based on the stationary distribution calculated from the 2000–2005 transition probabilities, there will continue to be more mid-size banks than large banks. Indeed, the 2000–2005 stationary distribution has a higher fraction of Category 5 banks than the stationary distributions of the 1960s and 1970s.

Table 5 reports the fraction of assets held by each size category in 2005 and in the stationary distribution based on the 2000–2005 data. Interestingly, in the stationary distribution, only the largest size category holds a smaller percentage of assets than it does in 2005. This striking finding demonstrates the importance of the transition probabilities for determining the size distribution.

Finally, we report in Table 6 the estimated transition matrix for the 2000–2005 period. The first row lists the probability of a new bank forming and starting in each size category. The first column lists the probability a bank of each size category exits (and for an inactive bank, stays inactive). For the given time period, all entering banks enter into the smallest size category. The first column shows the probability of exit, either by failure or merger, from the industry. Exit is most common in the largest category and represents mergers. We see that banks in any size category are most likely to remain in the same bin as denoted by the diagonal entries. Finally, the matrix shows that banks gradually change in size. Except for exits, banks almost always stay within one size category of the previous year.

8. CONCLUSION

In this paper, we documented the large changes in the size distribution of banks that occurred starting in the 1980s. We found that the lognormal distribution poorly fits the right tail of the size distribution. Zipf's Law fits better, but for some size measures in some periods, this distribution does not fit the largest banks that well.

We also documented some differences and similarities in the size dynamics over time. First, we found that new banks are a constant fraction of the total number of banks. Second, we also found that Gibrat's Law is a good approximation for the 1960s and 1970s, before deregulation, but does not describe the 1980s and 1990s. In these decades, the large banks grow the fastest. The last five years of data suggest that the dynamics are returning to the earlier, more stable period. Of course, five years of data are not enough to make a strong prediction.

We also performed a simple forecasting exercise, using the transition probabilities taken from different time periods. Again, the relative constancy of the number of banks in the 1960s and 1970s suggests that this period was relatively stable. The projected rapid decline in the number of banks using 1980s and 1990s transition probabilities is evidence of the rapid changes that occurred in the banking industry during that time. Finally, the 2000–2005 transition probabilities predict a leveling off in the number of banks.[15] If that trend continues, then we will be returning to a relatively stable period in banking, at least as measured by the number of banks. The size dynamics imply that the U.S. banking structure will continue to have large numbers of small banks and a decent number of mid-size banks.

As illustrated by the transition probability analysis, the size distribution depends, ultimately, on the size dynamics. Therefore, a theory of the changes in bank size distribution needs an explanation of why the size dynamics changed and by how much. The data demonstrate that these changes started in the 1980s as deregulation proceeded, so the natural place to start is with an understanding of how removals to growth and size limits change the growth rates of different size banks. A successful theory would also need to account for the robust entry over this period, despite the large number of banks that exited.

[15] Jones and Critchfield (2005) make a similar prediction, using a different forecasting method.

Table 6 Transition Probability Matrix for Bank Assets: 2000–2005

Size Categories	Inactive	1	2	3	4	5	6	7
Inactive	**0.999**	**0.001**	0.000	0.000	0.000	0.000	0.000	0.000
1	**0.024**	**0.927**	**0.049**	0.000	0.000	0.000	0.000	0.000
2	**0.030**	**0.043**	**0.908**	**0.018**	0.000	0.000	0.000	0.000
3	**0.040**	0.000	**0.076**	**0.827**	**0.058**	0.000	0.000	0.000
4	**0.050**	0.000	0.003	**0.048**	**0.879**	**0.020**	0.001	0.000
5	**0.028**	0.000	0.000	0.000	**0.060**	**0.827**	**0.086**	0.000
6	**0.053**	0.000	0.000	0.000	0.000	**0.068**	**0.849**	**0.030**
7	**0.061**	0.000	0.000	0.000	0.000	0.000	**0.020**	**0.920**

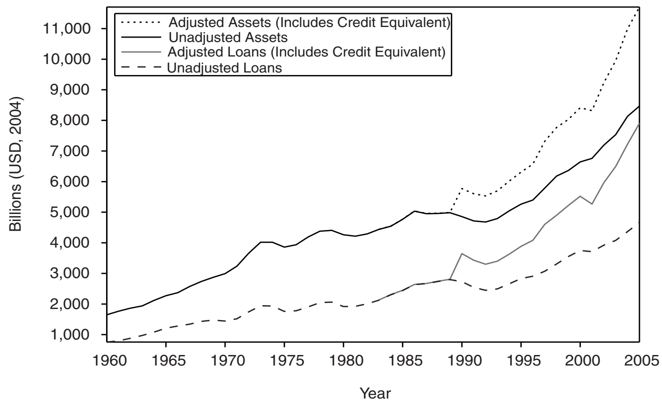
Notes: This matrix has the property that a bank starting in any size category will reach with positive probability any other size category in a finite number of periods. Probabilities that are significantly different than zero are highlighted in bold.

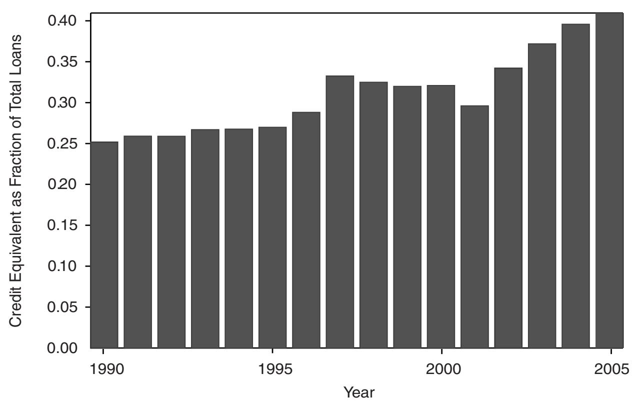
APPENDIX: OFF-BALANCE-SHEET ITEMS

Banks can make commitments that are not directly measured by a traditional balance sheet. For example, a loan commitment is a promise to make a loan under certain conditions. Traditionally, this kind of promise was not measured as an asset on a balance sheet. As documented by Boyd and Gertler (1994), providing this and other off-balance-sheet items have become an important service provided by banks, which means that traditional balance sheet numbers did not accurately report some of the implicit assets and liabilities of a bank.

We account for loan commitment and other off-balance-sheet items such as derivatives by converting them into *credit equivalents* and then adding them to on-balance-sheet assets and loans. This method is similar to the "Basel Credit Equivalents" series found in Boyd and Gertler (1994). Off-balance-sheet items are weighted by a credit conversion factor to create credit equivalents. We make these adjustments starting in 1989 because it is only from this year that we have the complete data to make them. Both panels of Figure 10 demonstrate the importance of the adjustment by plotting aggregate assets and loans with and without the adjustment, as well as by plotting credit equivalents as a share of total loans. A detailed list of off-balance-sheet items and credit equivalent weights is found in Table 7. These weights are used by federal regulators to determine credit equivalents for regulatory capital purposes.

Figure 10 Total Credit Equivalents in Assets and Loans





Notes: Top panel reports assets and loans both unadjusted and adjusted for off-balance-sheet items. Bottom panel reports the fraction of adjusted loans that are due to the credit equivalent adjustment. Credit equivalents are based on the weights used by regulators to determine regulatory capital requirements.

Table 7 Off-Balance-Sheet Items and Credit Equivalents

Item	Conversion Factor
Financial Standby Letters of Credit	1.00
Performance and Standby Letters of Credit	0.50
Commercial Standby Letters of Credit	0.20
Risk Participations in Bankers' Acceptances	1.00
Securities Lent	1.00
Retained Recourse on Small Business Obligations	1.00
Recourse and Direct Credit Substitutes	1.00
Other Financial Assets Sold with Recourse	1.00
Other Off-Balance-Sheet Liabilities	1.00
Unused Loan Commitments (maturity >1 year)	0.50
Derivatives	–

Notes: Conversion factors used by regulators for determining credit equivalents of off-balance-sheet items. The source is FFIEC 041 Schedule RC-R retrieved from: www.ffiec.gov/forms041.htm (accessed on November 10, 2005).

REFERENCES

Adelman, Irma G. 1958. "A Stochastic Analysis of the Size Distribution of Firms." *Journal of the American Statistical Association* 53: 893–904.

Alhadeff, David, and Charlotte Alhadeff. 1964. "Growth of Large Banks, 1930–1960." *Review of Economics and Statistics* 46: 356–63.

Axtell, Robert L. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science* 293: 1818–20.

Berger, Allen N., Anil K. Kashyap, and Joseph M. Scalise. 1995. "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been." *Brookings Papers on Economic Activity* 2: 55–201.

Boyd, John H., and Mark Gertler. 1994. "Are Banks Dead? Or Are the Reports Greatly Exaggerated?" Federal Reserve Bank of Minneapolis *Quarterly Review* 18 (Summer): 2–23.

Ennis, Huberto M. 2001. "On the Size Distribution of Banks." Federal Reserve Bank of Richmond *Economic Quarterly* 87 (Fall): 1–25.

Gabaix, X. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114: 739–67.

Gabaix, X., and Yannis Ioannides. 2004. "The Evolution of City Size Distributions." In J. Vernon Henderson and Jacques-Francois Thisse (eds.) *Handbook of Regional and Urban Economics IV*. North-Holland.

Gibrat, Robert. 1931. *Les Inégalités Économiques*. Paris, France: Librairie du Recueil Sirey.

Goddard, John A., Donald G. McKillop, and John O. S. Wilson. 2002. "The Growth of U.S. Credit Unions." *Journal of Banking and Finance* 26: 2327–56.

Hart, P. E., and S. J. Prais. 1956. "The Analysis of Business Concentration: A Statistical Approach." *Journal of the Royal Statistical Association* 119: 150–91.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (September): 1127–50.

Jayaratne, Jith, and Philip E. Strahan. 1997. "The Benefits of Branching Deregulation." Federal Reserve Bank of New York *Economic Policy Review* 3 (December): 13–29.

Jones, Kenneth D., and Tim Critchfield. 2005. "Consolidation in the U.S. Banking Industry: Is the 'Long, Strange Trip' About to End?" FDIC *Banking Review* 17: 31–61.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50 (May): 649–70.

Kane, Edward J. 1996. "De Jure Interstate Banking: Why Only Now?" *Journal of Money, Credit and Banking* 28: 141–61.

Mengle, David L. 1990. "The Case for Interstate Branch Banking." Federal Reserve Bank of Richmond *Economic Review* 76 (November/December): 3–17.

Newman, M. E. J. 2005. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics* 46: 323–51.

Rhoades, S. A., and A. J. Yeats. 1974. "Growth, Consolidation and Mergers in Banking." *Journal of Finance* 29: 1397–1405.

Robertson, Douglas D. 2001. "A Markov View of Bank Consolidation: 1960–2000." OCC Economic and Policy Analysis Working Paper 2001-4.

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2006. "Establishment Size Dynamics in the Aggregate Economy." Mimeo, Princeton University.

Saunders, Anthony, and Ingo Walter. 1994. *Universal Banking in the United States*. New York, NY: Oxford University Press.

Simon, Herbert, A., and Charles P. Bonini. 1958. "The Size Distribution of Business Firms." *American Economic Review* 48: 607–17.

Spong, Kenneth. 2000. *Banking Regulation*. 5th ed. Kansas City, MO: Federal Reserve Bank of Kansas City.

Sutton, John. 1997. "Gibrat's Legacy." *Journal of Economic Literature* 35 (March): 40–59.

Tschoegl, Adrian E. 1983. "Size, Growth, and Transnationality Among the World's Largest Banks." *Journal of Business* 56: 187–201.