CORE PROFITABILITY OF COMMUNITY BANKS: 1985–2015

Abstract

The relatively low profitability reported by community banks since the 2008 financial crisis has sparked concerns about the core profitability of the community banking model. This paper constructs an econometric model using 31 years of data to estimate the impact of macroeconomic shocks on industry average pretax return on assets (ROA). After accounting for macroeconomic factors, the remaining unexplained variation is considered to be the core component of profitability. Core return on assets is found to have been relatively stable between 1985 and 2015. It trended downward over the 1990s, but the effect of the financial crisis on industry composition has led to a reversal and a modest increase in core profitability. More than 80 percent of the post-crisis decline in profitability can be explained by negative macroeconomic shocks.

Community Bank Profitability, 1985 to 2015

Profitability across FDIC-insured institutions fell to record lows during the financial crisis of 2008 and the subsequent recession. In the years since, several measures of performance show the banking industry has rebounded. For example, by year-end 2015, noncurrent loans, loan-loss provisions, and net charge-offs had fallen to pre-crisis levels, and less than 5 percent of all institutions were unprofitable.

In contrast, profitability has stayed 20 to 30 basis points below pre-crisis levels, as measured by the average industry return on assets of 1.03 percent at year-end 2015. Banks continue to feel the strain of an economy marked by slow growth and low interest rates. The industry's net interest margin was just over 3 percent at year-end 2015, even as the share of longer-term assets with maturities over three years grew to just over one-third of total assets. In short, economic growth since the recession has helped the banking industry recover, but the weakness of the recovery has led to lackluster profitability.

Community banks—which accounted for 93 percent of all banks and 13 percent of total industry assets in 2015—have followed the same performance trends as the overall industry.[1] Their noncurrent loans, loan-loss provisions, net charge-offs, and percentage of unprofitable institutions have returned to pre-crisis levels, while profitability has remained below pre-crisis levels. Pretax ROA for community banks was more than a full percentage point higher in 2015 than the lows seen during the crisis but remained 20 to 30 basis points below the annual averages reported in the pre-crisis years.[2] This is even as the share of longer-term assets with maturities over three years grew to about one-half of total assets at community banks.

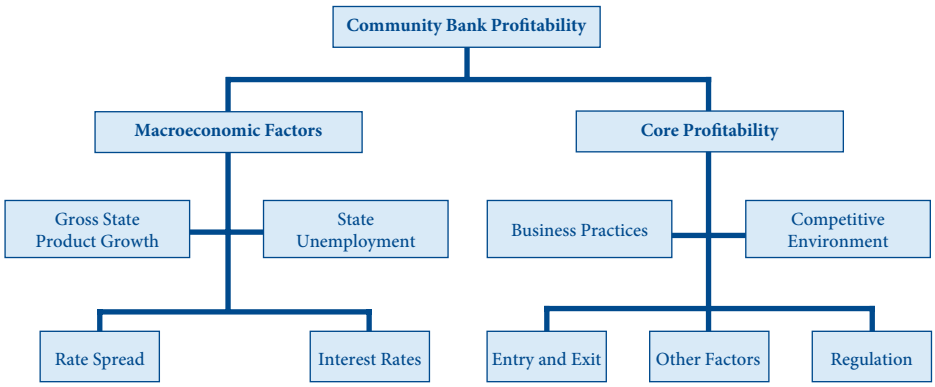
What accounts for the relatively low level of profitability among banks in the post-crisis period? Have macroeconomic factors that are external to the banking industry placed downward pressure on profits, or have structural factors within the industry—such as business practices and the regulatory environment—changed the intrinsic profitability of banks? Here we focus on the profitability of community banks and the factors that affect it. We use an econometric model to separate pretax ROA into two parts: one part is attributable to cyclical variations in pretax ROA caused by macroeconomic factors, and the second part is attributable to structural factors that reflect the operational environment of the banking industry and represents the core component of profitability.

Understanding the impact of macroeconomic factors on bank profitability has been a long-standing challenge. Several papers have investigated the relationship between business-cycle variables and bank profitability, but they do so to estimate the marginal effect of macroeconomic factors on individual bank profits rather than to estimate core profitability.[3] This

[1] We use the FDIC (2012) definition of community bank, which is a functional definition rather than a fixed-asset size definition.
[2] We evaluate pretax ROA for community banks since about one-third of community banks are pass-through Subchapter S corporations, which do not pay federal income taxes.
[3] For example, Demirgüç-Kunt and Huizinga (1999), Bikker and Hu (2002), Athanasoglou et al. (2006), and Beckmann (2007) use national GDP to examine the effects of macroeconomic conditions on bank profits. Other studies, such as Albertazzi and Gabacorta (2009), Tregenna (2009), Kanas et al. (2012), and Morris and Regehr (2014), examine the effects of bank-level characteristics and other structural factors on bank profits, including macroeconomic variables as control variables.

Chart 1

Profitability Is the Net Effect of Macroeconomic and Structural Factors



Source: FDIC.

paper further differs from past studies by considering macroeconomic variables in terms of growth rates, evaluating profitability across all community banks in aggregate, and controlling for the econometric bias introduced by entry and exit of banks in our sample over time.[4]

Our econometric model estimates the impact of macroeconomic factors on pretax ROA across community banks. The remaining variation in pretax ROA is attributable to structural factors, and we refer to this as core profitability. Core profitability is the intrinsic earning capacity of a bank, after controlling for the impact of macroeconomic factors. It reflects the net impact of the structural factors, which could include, for example, business practices, entry and exit of banks, the competitive environment, and the regulatory environment (see Chart 1). Our econometric model does not estimate the contributions of individual structural factors on profitability separately; instead, it estimates their net effect.

Our results show that community bank profitability from 1985 through 2015 may be divided into three distinct periods: the savings and loan (S&L) crisis years from 1985 to 1990, the economically strong years from 1991 to 2007, and the financial crisis and recovery years from 2008 to 2015. We find that relatively low profitability during the S&L crisis was the result of structural factors within the industry and was largely independent of the macroeconomic environment. Structural changes following the S&L crisis resulted in a sharp increase in profitability. During the second period, profitability was relatively high largely due to the exceptionally strong economy; however, profitability trended down slowly over this period as the strong economy was able to sustain increasingly less efficient institutions.

We find that the sharp decline in profitability during the 2008 financial crisis and subsequent recovery are largely the result of adverse macroeconomic conditions, and that structural factors played only a modest role. After controlling for macroeconomic factors, we find that core profitability has been above its long-run average over much of the post-crisis period, which has been obscured by the strong economic headwinds affecting observed ROA. These findings suggest that the core earnings model of community banks remains sound, despite the challenging post-crisis economy.

[4] Recent work by Adams and Gramlich (2016) takes a similar approach. They estimate the contribution of nonregulatory factors on de novo charters and find that nonregulatory factors explain 75 percent of the recent decline in de novo charters. Other papers on factors affecting bank profitability include DeYoung and Roland (2001) and Boyd and Gertler (1993).

Econometric Approach

Our econometric model uses time series panel analysis to estimate the impact of four macroeconomic factors on community bank pretax ROA: economic growth, unemployment rate, interest rate, and interest-rate spread. The data are in panel form, meaning we follow each bank every year during the sample period. This allows us to control for bank fixed effects, which are characteristics unique to each bank that do not vary over time; however, entry and exit of banks over time will distort or bias our results. We control for this bias using established econometric methods.[5] Our model also includes one lag for each macroeconomic factor to capture the dynamic effect that macroeconomic conditions in a given year may have on bank performance in subsequent years.[6]

The macroeconomic variables used in the model are measured in deviations from the mean. Consequently, when we interpret our model results, we are measuring the impact of our macroeconomic factors on the average ROA.

Data

We use 31 years of annual Call Report data from 1985 through 2015 for all FDIC-insured community banks. These data consist of 20,335 unique community banks, of which there were 15,957 in 1985 and 5,874 in 2015. There were 4,368 community banks that existed over the entire 31-year period.[7] This large dataset with many banks over a long period of time gives our model significant statistical power to determine the impact of the macroeconomic factors on profitability. Table 1 lists summary statistics for the data.

Chart 2 shows that the average annual pretax ROA for community banks varied from a high of 1.54 percent in 1993 to a low of 0.01 percent in 2009. Pretax ROA was relatively low from 1985 until 1990, averaging just 0.58 percent. It increased sharply during the early 1990s and was above 1.25 percent from 1992 until 2005. Following its low in 2009, ROA rebounded at the end of our sample period. Chart 2 also shows loan-loss provisions as a percentage of total assets, which move counter cyclically with ROA.

Table 1

Summary Statistics Reveal Significant Variance Within Factors

	Mean (Percent)	Std. Dev.
Return on Assets	0.99	2.18
Unemployment	5.84	1.83
Gross State Product Growth	5.15	3.08
Spread	1.48	0.96
Interest Rate	6.11	2.25

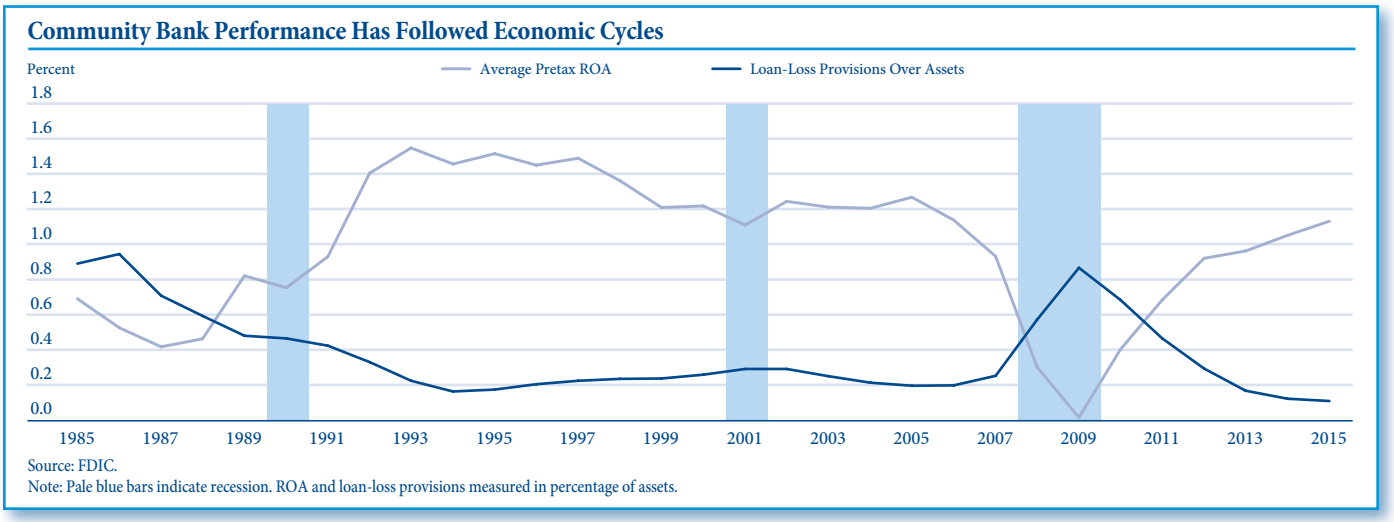
Sources: FDIC, Bureau of Labor Statistics, Bureau of Economic Analysis, and Federal Reserve Board (Haver Analytics).

[5] Specifically, we use a standard Heckman selection-type model using bank-level variables to first predict the likelihood of a bank entering or exiting the panel, and then include the inverse Mills ratio from the first stage as an explanatory variable in a second-stage regression. A common alternative approach to account for entry and exit is to include only banks that exist over the full sample period; however, this would cut three-fourths of the community banks from our sample, which would likely bias our estimation results in other ways, particularly if the factors affecting entry and exit are correlated with the factors affecting bank profitability.

[6] Since each macroeconomic factor is correlated with its own prior value, a single lag term per factor captures the dynamic impact of macroeconomic conditions in all preceding periods on the current period's ROA. The addition of multiple lags confirms that no switching effect exists—for example, a crisis doesn't lower the following year's ROA but then raise the third or fourth year's ROA. Rather, a crisis depresses each subsequent year's ROA, though the effect diminishes as time passes. See Achen (2001) for further discussion.

[7] This includes institutions insured by the Federal Savings and Loan Insurance Corporation prior to 1990.

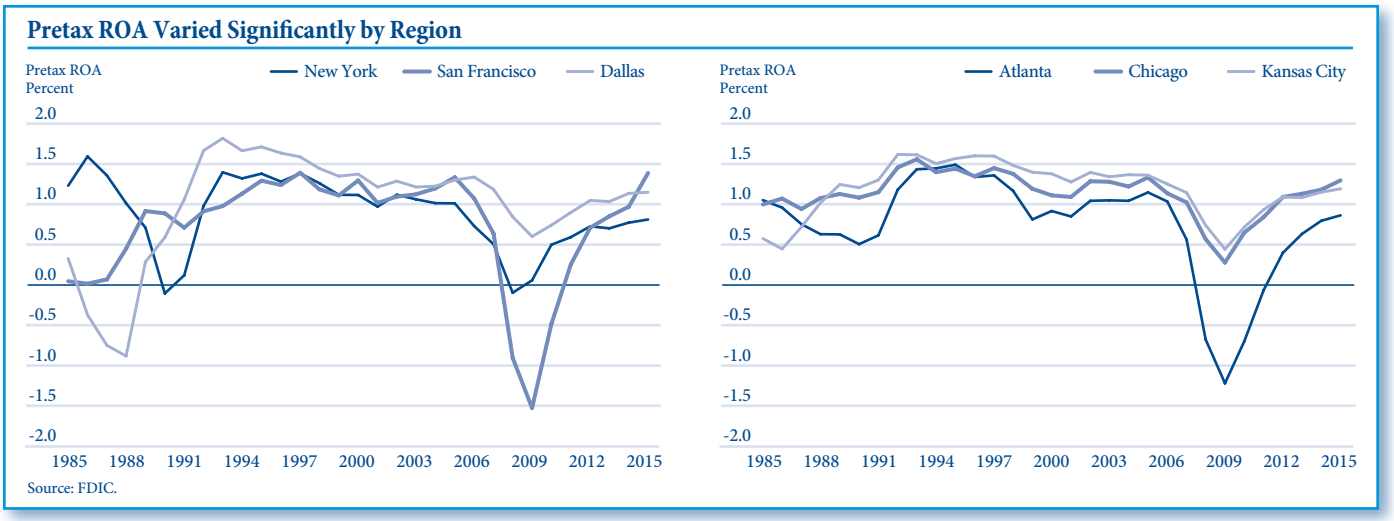
Chart 2



Because community banks are heavily influenced by local economic conditions, ROA varies significantly by region across the country. For example, Chart 3 shows that community banks in the Dallas region saw their average pretax ROA approach –1 percent in the late 1980s, while average pretax ROA for community banks in the New York region rose above 1.5 percent. In contrast, banks in the Dallas region weathered the recent financial crisis relatively well. Relative to other regions, their average pretax ROA dipped only slightly to 0.6 percent, whereas at the same time the average pretax ROA of community banks in the San Francisco region fell sharply to below –1.5 percent.

To capture regional variation in macroeconomic factors facing community banks, we use nominal gross state product (GSP) and state-level unemployment rates. GSP data are from the Bureau of Economic Analysis, and state-level unemployment rates are from the Bureau of Labor Statistics. We also include the interest rate, as measured by the return on a ten-year Treasury note, and the interest-rate spread, as measured by the difference in return on the ten-year and one-year Treasury notes. By including both the interest rate and the rate spread, we capture the effects of the level and slope of the yield curve as they change over time. Interest-rate data are from the Federal Reserve Board, as reported through Haver Analytics.

Chart 3



Estimating the Impact of Macroeconomic Factors on Pretax Return on Assets

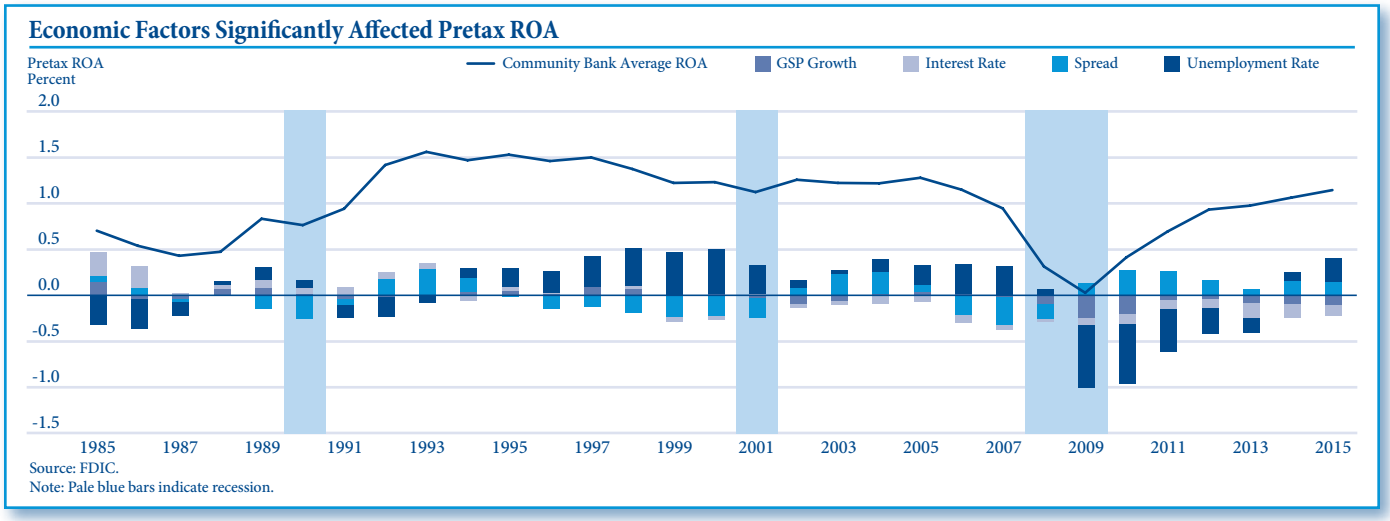
Our model estimates find that the four macroeconomic factors—economic growth, unemployment rate, interest rate, and interest-rate spread—together explain a significant part of the variation in pretax ROA across community banks over time. On average, macroeconomic factors accounted for more than half (57 percent) of the total variation in ROA; however, this varied significantly over time. Structural industry factors explain nearly all of the observed deviation from the period average core ROA in the late 1980s, as macroeconomic factors played a negligible role in determining ROA for community banks during the S&L crisis. During the stronger, more economically stable years of the 1990s and early 2000s, macroeconomic factors explain 76 percent of the variation in ROA. Finally, since the financial crisis in 2008, macroeconomic factors explain 80 percent of the variation in community bank ROA during the exceptionally weak, post-crisis economic expansion.

Chart 4 shows the contributions of the four macroeconomic factors to pretax ROA over time. Together the macroeconomic factors raised pretax ROA by 2 basis points from 1985 to 1990; changes in structural factors were the primary forces affecting community bank ROA during the S&L crisis. Macroeconomic factors increased pretax ROA by 15 basis points from 1991 to 2007, peaking at 33 basis points in 1998. In 2009, macroeconomic factors had a large negative impact, reducing pretax ROA by 87 basis points during the financial crisis and severe economic recession. The drag from macroeconomic factors gradually declined as the post-recession recovery progressed. In 2014, macroeconomic factors began to lift pretax ROA, and by 2015 they increased pretax community bank ROA by 18 basis points.

The unemployment rate is the dominant macroeconomic factor affecting community bank ROA across the sample period and in most years. This is not surprising given that community banking is focused on relationship lending and a strong local job market boosts demand for loans, while a weak job market may raise delinquency rates. From 1994 to 2008, low unemployment boosted community bank ROA by an average of 25 basis points annually; however, the sharp increase in unemployment during the financial crisis was associated with a large decline in community bank profitability, reducing pretax ROA by 68 basis points in 2009 and 64 basis points in 2010. The subsequent decline in unemployment first lessened the drag on community bank profitability and eventually contributed to profitability in 2014 and 2015.

Gross state product growth is the least influential of the four macroeconomic factors. This likely stems from the fact that unemployment, interest rates, and interest-rate spreads affect community bank profitability directly, while economic growth affects banks indirectly and often through the other macroeconomic factors.

Chart 4



The interest rate generally had a relatively small impact on community bank profitability over the sample period; however, it had its largest positive impact of 25 basis points in 1985 and was one of the largest contributors to profitability in the first two years of the sample when interest rates were at their highest. The interest rate had its largest negative impact on profitability during post-crisis years, when central banks maintained a zero-interest-rate policy. The impact of low interest rates became progressively more adverse to community bank profitability through 2013, when it reduced ROA by 17 basis points. By 2015, the impact began to weaken as the extension of loan maturities coincided with an improvement in the macroeconomic environment.

The impact of the interest-rate spread on community bank profitability generally is directly related to the size of the spread and generally moved independently of the other three macroeconomic variables. The spread had a small positive impact in the first two years, a negative impact when the yield curve inverted prior to the 1990 to 1991 recession, and a large positive impact when the yield curve steepened sharply in the first half of the 1990s. The spread again had a negative impact in the mid-1990s through the recession of 2001, when the yield curve flattened before inverting in 2000. From 1995 through 1998, the spread was the only macroeconomic factor that had a negative impact on profitability. The pattern repeated itself following the 2001 recession, with a steep yield curve boosting community bank profits before gradually flattening and finally inverting just before the 2008 financial crisis.

It is interesting that the spread still had a relatively large positive impact on profitability in 2014 and 2015 as the yield curve flattened. This likely reflects a search for yield, as community banks raised the percentage of loans that mature in over three years from about 40 percent in 2012 to about 50 percent in 2014.

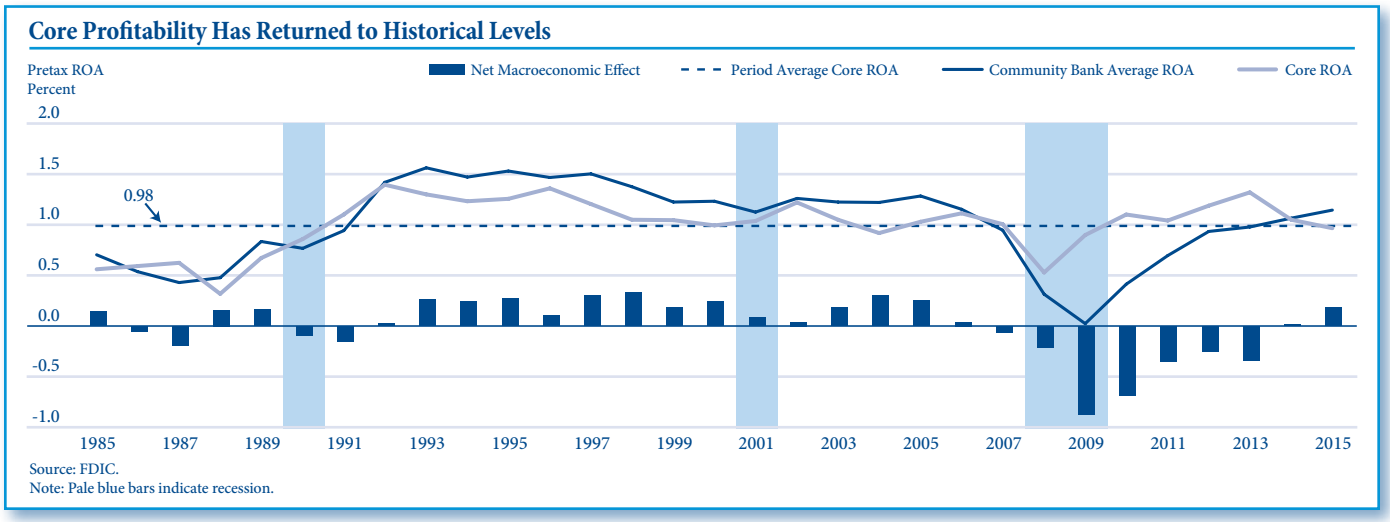
Trends in Core Profitability

One explanation put forward for the decline in profitability among banks is the impact of new regulations put into place following the financial crisis. Among these are a range of regulations mandated by the Dodd-Frank Act of 2010 and the Basel III capital standards introduced in 2013; however, regulation is just one among many noneconomic factors that may contribute to structural change in community bank profitability. Other structural factors may include the rise of nonbank lending, competition from larger banks, and changes in loan portfolios and other business practices. Given that macroeconomic factors explain 80 percent of the post-crisis variation in ROA, the net effect of structural factors combined explains the remaining 20 percent of post-crisis ROA variation.

Core profitability is the intrinsic earning capacity of a bank, after controlling for the impact of macroeconomic factors. It is a measure of the impact of structural factors on pretax ROA. Our econometric model estimates the impact of macroeconomic factors on ROA, and then the remaining variation in ROA is attributed to structural factors. Chart 5 shows that core ROA averaged 0.98 percent from 1985 through 2015. Core ROA and observed ROA generally evolve together, but core ROA is more stable with less variability around the average. Note that the difference between core ROA and observed ROA each year is the net effect of the model's macroeconomic factors.

Core profitability among community banks was at its lowest during the S&L crisis of the late 1980s. It improved through the early 1990s, possibly as the competitive environment evolved following the failure of more than 1,700 banks and savings and loans, which eliminated many less-profitable institutions. After peaking in 1992, core profitability declined gradually over the following decade as a strong economy helped boost earnings, but enabled less profitable banks to operate. Bank responses to the recession of 2001 abated the downward trend in core profitability for two years, after which core ROA resumed its downward trend. Core profitability reached its lowest level since the S&L crisis during the financial crisis in 2008, falling to 0.51 percent. The subsequent failure of 440 banks in 2009 through 2012, which eliminated many underperforming banks, combined with other structural changes to the competitive environment, reversed the long-term downward trend and resulted in a marked upturn in core profitability.

Chart 5



Core profitability has been relatively strong throughout the post-crisis period, remaining at or above its historical average. The sharp decline in observed ROA during the financial crisis and subsequent recession is largely attributable to the severity of the downturn in macroeconomic factors, primarily the unemployment rate. At their extreme, macroeconomic factors reduced community bank profitability by 87 basis points in 2009. From 2010 onward, the slow pace of macroeconomic recovery and the persistence of historically low interest rates continued to be a drag on profitability, although by less in each successive year. By 2014, macroeconomic factors were no longer a headwind to ROA and were essentially profit-neutral.

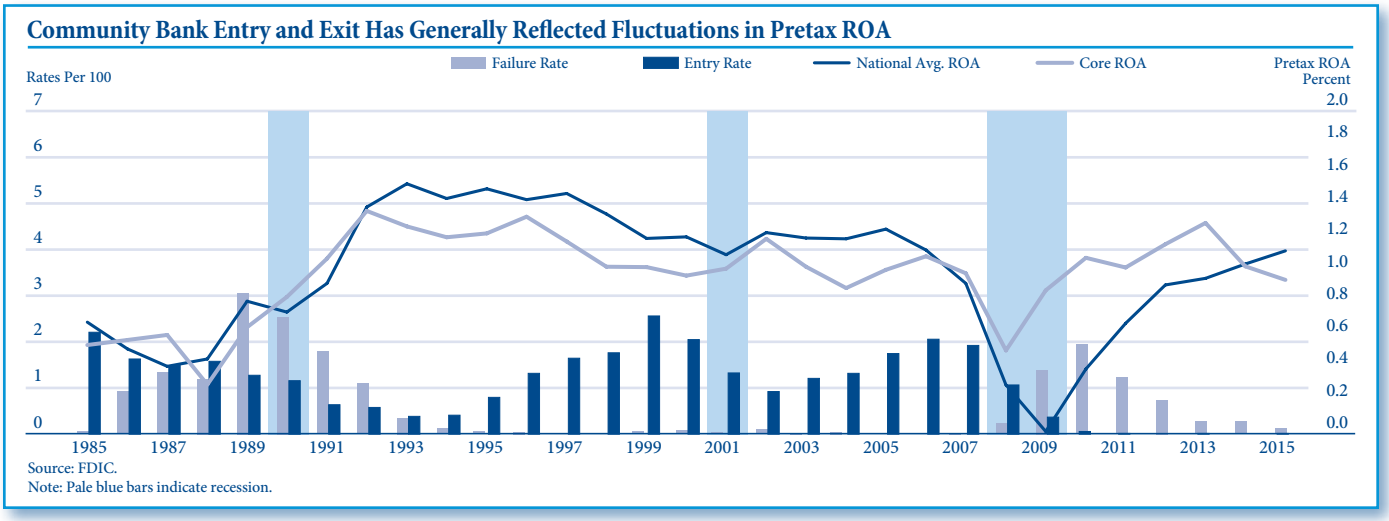
The Potential Effect of Entry and Exit on Core Profitability Trends

The number of community banks has fallen steadily from a high of 15,957 in 1985 to 5,874 at the end of our sample in 2015. One structural factor that is measureable and offers a potential explanation of trends observed in core profitability is entry and exit of community banks, which affects core ROA by changing the composition of community banks over time.[8] New entries may increase competitive pressures on existing community banks and may lower overall core earning potential. Conversely, the failure or merger of less-productive banks may cause both observed ROA and core ROA to rise, as underperformers are removed from the sample.

Chart 6 shows that bank entry and exit correlate closely with overall trends in ROA: A rise in failures corresponds to an increase in core ROA (and observed ROA), whereas higher entry and fewer failures correspond to a decline in core ROA. The rate of de novo entry follows a clear cyclical pattern—rising in expansions and falling in recessions—while the rate at which banks exit follows the opposite pattern. The period since the financial crisis in 2008 and subsequent recession is exceptional in that de novo charters have not increased as the economy has expanded.

[8] The compositional effect from entry and exit is independent of any econometric bias introduced by entry and exit. Banks exiting the sample results in a measurement bias if the same unobserved factors that influence profitability also affect failure rates. We use a standard econometric technique to identify and correct for this bias: an F-test shows that the coefficient on the inverse Mills ratio—the attrition bias correction factor—is statistically significant, confirming the presence of attrition bias. Analysis suggests that if banks had not outgrown the sample in the boom years, average ROA would have been up to 12 basis points higher in some years. Conversely, if banks had not failed out of the sample during the recent crisis, in some years average ROA would have been 11 basis points lower.

Chart 6



During the S&L crisis period from 1985 to 1990, when entry rates were falling and failure rates were growing, core ROA rose at an average rate of 4 basis points per year. During the economically strong period from 1991 to 2007, when many banks entered and few banks failed, core ROA fell at an average rate of 1.8 basis points per year. Finally, during the financial crisis, recession, and subsequent economic recovery, the high rate of failures and lack of de novos corresponded to a strong upward trend in core ROA of 5.5 basis points per year.

The correlation between bank entry and exit and community bank core profitability is strong; however, this is not to suggest that compositional effects from entry and exit explain all of the variation in core profitability. Changes in other structural factors such as business practices, competitive environment, and regulation would also play a role; however, unlike entry and exit, other structural factors are difficult to measure reliably.

What Can We Conclude About Core Profitability Among Community Banks?

Understanding the evolution of core profitability among community banks requires an econometric approach that distinguishes the impact of macroeconomic factors from structural factors on observed profitability. Our model estimates the impact of the macroeconomic factors on pretax ROA, and then the remaining variation in pretax ROA that is attributable to structural factors. We refer to the impact of structural factors on pretax ROA as core profitability.

Over our sample period from 1985 through 2015, our model finds that core profitability rose sharply from a low in the late 1980s during the S&L crisis to a high in the early 1990s, trended down slowly through the mid-2000s before falling sharply to a low during the financial crisis in 2008, and then returned to pre-crisis levels during the weak economic recovery in the years following the financial crisis. The model finds that macroeconomic factors are largely responsible for actual profitability being so low during and after the financial crisis, and that core profitability generally has been at or above its long-run level since 2009. These findings suggest that the fundamental earnings model of community banks remains sound, despite the challenging post-crisis economic environment.

Author:
Jared Fronk
Financial Economist
Division of Insurance and Research

Appendix: Econometric Model and Regression Results

$$ROA_{it} = \beta_0 + \beta_1 Unemployment_{k,t} + \beta_2 Unemployment_{k,t-1} + \beta_4 GSP\ Growth_{k,t} + \beta_5 GSP\ Growth_{k,t-1} + \beta_7 Spread_t + \beta_8 Spread_{t-1} + \beta_{10} Interest\ Rate_t + \beta_{11} Interest\ Rate_{t-1} + \beta_{13} Attrition\ Correction_{i,t} + \gamma_i + \varepsilon_{i,t}$$

where i indicates bank, t indicates year, and k indicates state. The attrition correction term is the inverse Mills ratio, which econometrically accounts for bias introduced by entry and exit of banks in the sample.[9] Bank fixed effects are given by γ_i. All variables are measured as deviations from the mean.[10]

The econometric model estimates core profitability by identifying the degree to which macroeconomic factors affect observed ROA. The model does not include direct measures of core profitability, as they are not directly observed in the data. Rather, core profitability is inferred as the component of ROA that is not explained by macroeconomic factors. Attrition is corrected for using a standard Heckman correction. Table A1 presents results from the regression model.

For further support of the econometrical approach, see Wooldridge (2010) and Petersen (2009).

Table A1

Macroeconomic Variables Have a Significant Impact on Industry Profits

	Pretax Return on Assets
Unemployment	-0.2461***
	(0.0066)
GSP Growth	0.0229***
	(0.0015)
Spread	0.0666***
	(0.0061)
Interest Rate	-0.0309***
	(0.0044)
Unemployment, t-1	0.0173***
	(0.0062)
GSP Growth, t-1	0.0241***
	(0.0014)
Spread, t-1	0.1347***
	(0.0051)
Interest Rate, t-1	0.064***
	(0.004)
Attrition Correction	-20.9338***
	(1.8218)
Constant	0.9679***
	(0.002)
Bank Fixed Effects	Yes
Observations	304,948
Unique Banks	19,872
Within R-Squared	0.0595
Between R-Squared	0.1821

Source: FDIC.

Note: Standard errors are clustered at the bank level and are reported in parentheses.

*** $p < 0.01$

[9] The method behind the calculation of the inverse Mills ratio is available from the author upon request.

[10] Measuring variables as deviations from the mean has no effect on estimated coefficients or standard deviations. It affects only the constant term β_0 such that it is easier to interpret, since it measures deviations from the average. See Vella (1998).

References

Achen, Christopher. 2001. "Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables." Working Paper 1001. University of Michigan.

Adams, Robert, and Jacob Gramlich. 2016. "Where Are All the New Banks? The Role of Regulatory Burden in New Bank Formation." *Review of Industrial Organization* 48: 181–208.

Albertazzi, Ugo, and Leonardo Gambacorta. 2009. "Bank Profitability and the Business Cycle." *Journal of Financial Stability* 3: 393–409.

Athanasoglou, Panayiotis, Sophocles Brissimis, and Matthaios Delis. 2006. "Bank-Specific, Industry-Specific and Macroeconomic Determinants of Bank Profitability." *Journal of International Financial Markets, Institutions, and Money* 18: 121–136.

Beckmann, Rainer. 2007. "Profitability of Western European Banking Systems: Panel Evidence on Structural and Cyclical Determinants." Deutsche Bundesbank Discussion Paper, Series 2, No. 17/2007, Deutsche Bank.

Bikker, Jacob, and Haixia Hu. 2002. "Cyclical Patterns in Profits, Provisioning and Lending of Banks and Procyclicality of the New Basel Capital Requirements." *BNL Quarterly Review* 221: 143–175.

Bureau of Labor Statistics, U.S. Department of Labor. 2015. "Local Area Unemployment Statistics." November 17. www.bls.gov.

Boyd, John, and Mark Gertler. 1993. "U.S. Commercial Banking: Trends, Cycles, and Policy." NBER Working Paper No. 4404, National Bureau of Economic Research.

Demirgüç-Kunt, Asli, and Harry Huizinga. 1999. "Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence." *The World Bank Economic Review* 13: 430–455.

DeYoung, Robert, and Karin Roland. 2001. "Product Mix and Earnings Volatility at Commercial Banks: Evidence From a Degree of Leverage Model." *Journal of Financial Intermediation* 10: 54–84.

Federal Deposit Insurance Corporation (FDIC). 2012. *FDIC Community Banking Study*. Washington, DC, FDIC.

Kanas, Aanglos, Dimitros Vasiliou, and Nikolaos Eriotis. 2012. "Revisiting Bank Profitability: A Semi-parametric Approach." *Journal of International Financial Markets, Institutions, and Money* 22: 990–1005.

Morris, Charles, and Kristen Regehr. 2014. "What Explains Low Net Interest Income at Community Banks?" *Federal Reserve Bank of Kansas City Economic Review* 2:59–87.

Petersen, Mitchell. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies* 22: 435–480.

Tregenna, Fiona. 2009. "The Fat Years: The Structure and Profitability of the US Banking Sector in the Pre-crisis Period." *Cambridge Journal of Economics* 33: 609–632.

U.S. Bureau of Economic Analysis. 2015. "Annual Gross Domestic Product by State." November 17. http://www.bea.gov/regional.

Vella, Francis. 1998. "Estimating Models with Sample Selection Bias: A Survey." *Journal of Human Resources* 33 (1): 127–169.

Wooldridge, Jeffrey. 2010. "Censored Data, Sample Selection, and Attrition." In *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., 837–845. Cambridge, MA: MIT Press.