3-2017

Bank Size, Returns to Scale and Cost Efficiency

Ayse Sapci
Colgate University, asapci@colgate.edu

Bradley Miles
Colgate University, bmiles@colgate.edu

Bank Size, Returns to Scale and Cost Efficiency

Bradley Miles,*and Ayse Sapci †

March 2017

Abstract

Since the passage of Dodd-Frank, government regulators have become more interested than ever in the significant increase of bank size in the U.S. financial sector. To shed light on the reasons of the bank size increase and its effects on banks, we study the dynamic interactions between size, cost efficiency and returns to scale. Using Fourier flexible form, we show that banks of all but the largest sizes exhibit increasing returns to scale. As banks grow, they tend to benefit from cost efficiencies more, but they lose returns to scale gains. Banks seem to exploit increasing returns to scale until they become too large; however, they continue to enjoy their cost efficiency. We also analyze the effects of regulations in the past 25 years to understand whether imposing (or removing) limits on the size of banks causes real economic costs. Our findings show that both restrictive and loose regulations help larger banks but hurt smaller banks by creating extra costs.

JEL Classification: C11, C14, G21, L11

*Department of Economics, Colgate University, Hamilton, NY 13346

†Corresponding Author, Department of Economics, Colgate University, Hamilton, NY 13346. email: asapci@colgate.edu

1 Introduction

The past 25 years have been characterized by heavy regulations (and deregulations) as well as significant increases in bank size. Assets of the five largest banks as a share of total commercial banking assets increased from 23 percent in 1996 to about 48 percent in 2014.[1] In this paper, we investigate the reasons and the effects of the bank size increase, by analyzing economies of scale and cost efficiency throughout the banking industry. We further study the economic costs of the regulatory environments in the past 25 years on differently sized banks.

Two of the most important reasons why banks are growing are returns to scale and cost efficiency gains. Although the earlier literature has mixed results on scale economies, recent literature has found increasing returns to scale for most banks.[2] For instance, Wheelock and Wilson (2012) and Hughes and Mester (2013) find significant economies of scale for even the largest banks. Thus, returns to scale can encourage banks to grow. Cost efficiency, i.e. the ratio of total noninterest expenses to total assets, might be another related factor driving the size increase in the banking sector. In fact, Kovner et al. (2014) find that every 10 percent increase in bank assets is associated with 0.3 to 0.6 percent decrease in noninterest expenses. We complement and contribute to this literature by analyzing the dynamic relationship between size, returns to scale, and cost efficiency through a Panel and Bayesian Panel VAR from 1992:3 to 2014:2.

Using Fourier flexible form model, we first show that all but the largest banks exhibit increasing returns to scale. Then we demonstrate that an increase in total assets leads to an increase in cost efficiency but a decrease in returns to scale. Banks seem to exploit increasing returns to scale until they become too large;

[1] Data are obtained from the Global Financial Development Database (GFDD), The World Bank.

[2] Schweitzer (1972), Noulas et al. (1990) and Hunter et al. (1990) find increasing returns to scale for all but the largest banks, which exhibit decreasing returns to scale. Similarly, McAllister and McManus (1993) and Wheelock and Wilson (2001) find increasing returns to scale for most banks but constant returns to scale for the largest banks.

however, they continue to benefit from cost efficiency gains.

The regulatory environment might also play an important role in the observed size increase of the banking sector. For example, regulations that gave more freedom to bank holding companies (such as Gramm-Leach-Bliley) likely resulted in cost synergies through an increase in merger activity, particularly for larger banks. Dodd-Frank, which is more restrictive, may have helped the banks by reducing the risk, but it may also have imposed costs on banks. To more effectively address whether bank size alters the costs arising from regulations, we use the cost decomposition data from the largest 198 commercial bank holding companies and account for each pertinent regulatory period. We find that both restrictive and loose regulations tend to benefit larger banks but hurt smaller banks.

2 Bank Regulations and Their Effects on Costs and Returns to Scale

The literature on the effects of regulations on bank costs is extensive, but many papers focus either on a country's general policies at a given point in time (such as Barth et al. (2004)) or focus on earlier regulatory periods (such as Demirguc-Kunt and Detragiache (2002)). These discussions help reinforce the connection between returns to scale and regulation; however, they provide little in helping to formulate expectations for how returns to scale and costs respond specifically to more recent regulatory periods in the United States. By focusing on the three banking regulations over the past 25 years, we can pinpoint the precise regulatory environment that affects returns to scale and bank costs. The following sections give details on these specific regulations and provide evidence on how they could affect bank costs and scale economies.

2.1 Riegle-Neal Interstate Banking and Branching Efficiency Act

The Riegle-Neal Act was passed in 1994 and is still in effect today. In September 1994, Congress enacted the Riegle-Neal Act, allowing bank holding companies to acquire banks across state lines regardless of individual state laws (this provision became active on September 30, 1995). In addition, on June 1, 1997, the bill permitted interstate mergers of pre-existing banks.

Cornett et al. (2006) analyze the effects of Riegle-Neal on bank profits. They measure a number of variables related to costs and revenues both before and after bank mergers completed during the period from 1990 to 2000. They find that majority of mergers occurred after the passage of Riegle-Neal. This conclusion echoes the findings of Chronopoulus et al. (2015) that bank sizes grew substantially after Riegle-Neal. Cornett et al. (2006), on the other hand, measure how the costs and revenues of banks changed after mergers following the Riegle-Neal Act. The authors conclude that increases in short-term profitability were higher in mergers that occurred in the period after Riegle-Neal was enacted due to cost synergies from consolidating operations.

2.2 Gramm-Leach-Bliley Act

The Gramm-Leach-Bliley Act was enacted on November 12, 1999, with the intention of repealing the Glass-Steagall Act. The Gramm-Leach-Bliley Act allowed banks to maintain both investment and commercial banking divisions. Chronopoulous et al. (2015) discuss the effects that the Gramm-Leach-Bliley Act had on bank profit persistence. Unlike the Riegle-Neal Act, the authors find that the Gramm-Leach-Bliley Act actually reduced competition by spurring consolidations that had previously been not allowed. This reduction in competition led to profit persistence increasing post-passage of the act. Chronopoulous et al. (2015) further show that the merger activity following the Act caused an increase in bank size. Barth et al. (2000 and 2004) discuss the potential effects of the Gramm-Leach-Bliley Act on bank costs. They argue that bank costs would be reduced, primarily from the cost synergies associated with combining investment and commercial banking facilities. For example, a commercial bank could leverage its current telecommunications and data processing divisions to include the sales of insurance and securities for low additional costs.

2.3 Dodd-Frank Wall Street Reform and Consumer Protection Act

The Dodd-Frank Act was signed into law on July 21, 2010, and is still in effect today. The Dodd-Frank was signed following the worst recessionary period in the recent U.S. history. This recession was spurred, in part, by loose regulations governing financial institutions' securitization businesses. Over 2,000 pages long, Dodd-Frank reform aims to limit the risks that banks take and helps minimize the severe results from bank failures.

Dodd-Frank contains a number of clauses giving the Federal Deposit Insurance Corporation (FDIC) power in cases of perceived insolvency. More specifically, if regulators determine that a bank's imminent default could have severe economic implications, regulators may submit an appeal requesting control over the bank in question. The Dodd-Frank effectively gives the FDIC the ability to navigate the troubled bank through the default and liquidation processes. The FDIC, being more concerned with controlling economic shocks than with returning equity to shareholders, should cause fewer shocks in the economy than would the banking heads, were they to remain in control of their company. However, with the FDIC presumably putting the interests of shareholders second, the bank could have a more difficult time raising capital. Thus, this clause could actually destabilize the banking industry due to reduced liquidity from limited access to capital markets. Dodd-Frank also aims to reduce risk by placing regulations on institutions and on the instruments that can be traded.

In order to reduce institutional risk, Dodd-Frank focuses primarily on the bank holding companies that are more likely to cause significant economic shocks (those banks with over $50 billion in assets). The regulation imposes a number of controls on the size and quality of reserves that these large institutions are required to hold. The increase in held reserves creates a significant buffer, preventing liquidity shocks that had previously damaged bank holding companies. Dodd-Frank regulation also prohibits a single bank holding company controlling more than 10 percent of total liabilities of all financial institutions which effectively imposes a limit on the size of banks and restricts bank mergers and acquisitions. Wheelock and Wilson (2012) and Kovner et al. (2014) demonstrate that limiting size of banks causes significant increases in bank costs, primarily because banks exhibit increasing returns to scale. Berger and Hannan (1998), however, contradict this point with a discussion of the "quiet life" hypothesis. Under this hypothesis, banks in highly concentrated markets will collude and have little incentive to compete and cut costs. In other words, if the size of market concentration is restricted, bank costs are likely to be reduced. Berger and Hannan (1998) test this hypothesis and find that reducing market concentration decreased bank operating costs by 8 to 32 percent. Based on these interpretations, we can conclude that the effects of Dodd-Frank on bank costs and returns to scale are not binary. Despite the lack of a conclusion, a dichotomy between Dodd-Frank and the other acts is obvious here: While the Gramm-Leach-Bliley and Riegle-Neal Acts gave banks more freedom, Dodd-Frank added more restrictions.

3 Data

Our dataset consists of 15 operating expenses for 198 commercial bank holding companies from 1992:3 to 2014:2. The dataset captures the largest publicly traded bank holding companies with assets valued at over $300 million as ranked by the December 2007 Federal Reserve Board Report.[3] Individual bank income statements and balance sheets are obtained from the Mergent Online database. In order to assess costs more efficiently, we divide these operating expenses into three general categories: Fixed, quasi-fixed, and variable costs. Fixed costs include occupancy, supplies, printing, software, and equipment expenses. Quasi-fixed costs include personnel and employee benefits. Variable costs include marketing, telecommunications, litigation, data processing, loan processing, professional fees, postal and courier, and other noninterest expenses.[4]

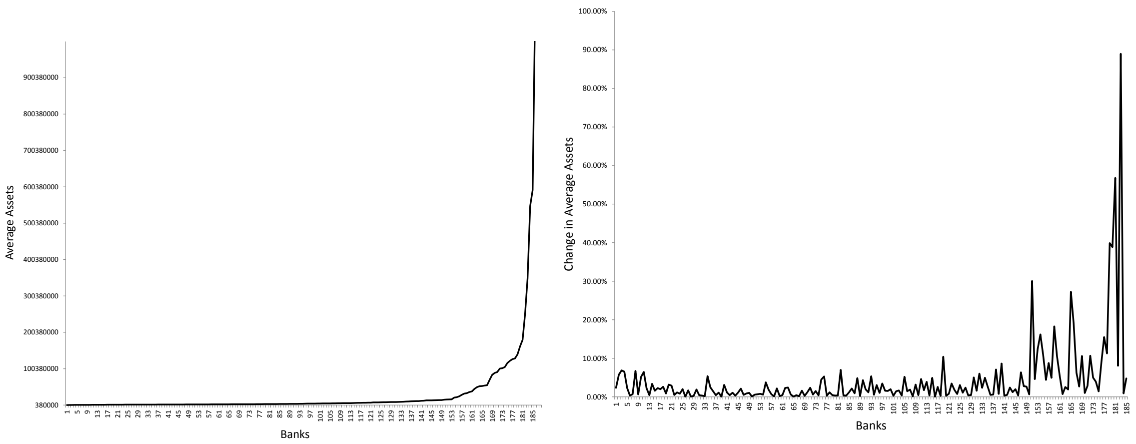
Following Jaremski and Sapci (2014), we control for macroeconomic variables that could affect the bank size. For instance, we control for real GDP and industrial production, representing the overall health of the economy. We also control for the inflation rate (CPI inflation) and money supply (M2) for the effects of monetary policy. Last, we include the Dow Jones Industrial Average to control for the health of the financial markets and the Case-Shiller U.S. National Home Price Index to control for house prices.

Figure 1a plots each bank's average assets for the period, whereas Figure 1b displays the percentage change in average asset size across banks. Using the average assets size of a bank compared to their nearest neighbor, we divide the banks into four groups: Too-big-to-fail (TBTF henceforth), large, medium and small banks. In particular, bank holding companies with an average asset size of over $20 billion made up the TBTF group, between $2.5 billion and $20 billion

[3] We use the rankings before the Great Recession to account for the banks that did not survive after the turmoil.

[4] Please refer to Appendix A.1 for a detailed description of the dataset.

made up the large group, between \$500 million and \$2.5 billion made up the medium group, and under \$500 million made up the smallest bank category. Tables 4 through 7 in Appendix A.1.2 show the summary statistics of cost decomposition data for the different bank groups.



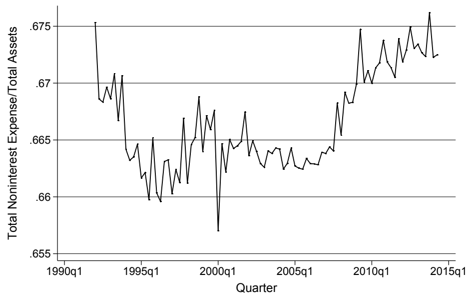
(a) Average Bank Size

(b) Banks' Size Compared to Nearest Neighbors

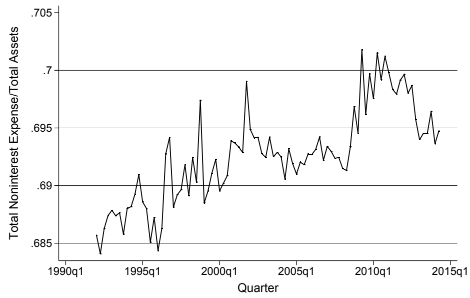
Figure 1: Bank Size Groups

To assess the range of cost efficiency over time and across bank groups of different sizes, Figure 2 plots the average value of total noninterest expenses divided by total assets for the 1992:3 to 2014:2 period.[5] Interestingly, small and medium commercial bank holding companies appear to become more cost efficient over time, whereas the TBTF and large commercial bank holding companies experienced declines in their cost efficiency since mid 1990s. This analysis shows that there is some heterogeneity among banks in terms of cost efficiency gains.

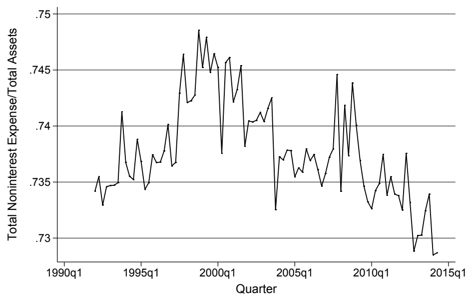
[5]It is important to note that total noninterest expenses divided by assets is a measure of cost inefficiency. Throughout the paper, lower values always denote higher cost efficiency.



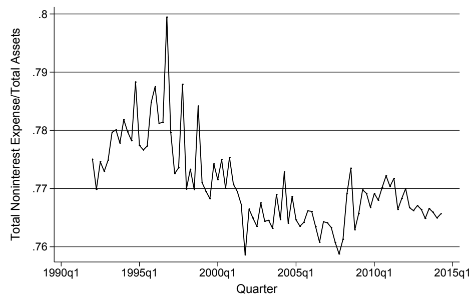
(a) Small Banks



(b) Medium Banks



(c) Large Banks



(d) Too-Big-To-Fail Banks

Figure 2: Cost Efficiency Across Time for Different Bank Groups

4 Empirical Models

4.1 Fourier Flexible Form

The Fourier flexible form represents a semi-nonparametric approach which is useful when the true functional form of the cost function is unknown. Because sine and cosine functions are orthogonal within the 0 to 2π range, an infinite series of sines and cosines with varying frequencies can accurately represent any continuously differentiable function. Because of computational and dimensionality limits, an infinite Fourier series (which would be a fully nonparametric estimate) is not feasible. A finite Fourier series, however, is semi-nonparametric and (unlike, for instance, the Translog cost function) is a global approximation of the cost function.[6] The Fourier flexible form used in this paper is presented below:

$$\ln C \approx \beta_o + \beta_Q \ln Q + \sum_m \beta_m \ln p_m + \sum_n \tau_n \ln p_n \ln Q + \frac{1}{2}\sum_m \sum_n \delta_{mn} \ln p_m \ln p_n$$
$$+ \frac{1}{2}\gamma \ln Q^2 + \sum_{i=1}^{N} [\zeta_i \sin(k_i V) + \phi_i \cos(k_i V)] + \epsilon A_t + u_i + e_{i,t} \quad (1)$$

Where C is total costs, Q is the vector of outputs, p is the vector presenting input prices, k_i is a vector of integer values and V is a vector of the logged input and output quantities and A_t is the vector containing macroeconomic controls as well as bank regulation dummy variables.[7] Lastly, u_i is a vector of bank-fixed effects. In the case of bank holding companies, a firm's total loans will serve as a measure of a bank's output. The inputs for a firm's production function are labor, capital, and deposits. The amount of labor is specified by the sum of employee benefits and personnel costs. The amount of capital is a sum of the supplies and printing, software, occupancy, and equipment expenses.

In order for the Fourier flexible form to accurately measure returns to scale, several criteria must be fulfilled. In particular, the input and output variables must be transformed so that they vary within the interval [0,2π]. In terms of choosing k_i, there is no substantial literature on it; however, we follow the criteria laid out in Skolrud (2013) and use a 6fh order Fouries series expansion. To ensure orthogonality between the cosine and sine terms, all values within the k_i vectors must be integers. In order to maintain linear homogeneity, the sum of all the β_m coefficients must be 1. Put simply, this restriction ensures that, should each operating cost be multiplied by some scalar constant, the total cost will change by the same constant. In other words: $\lambda C = C(Q, \lambda \mathbf{p})$; the cost function must

[6]We also analyzed returns to scale of banks using Translog cost function. Both Translog and Fourier provide similar conclusions. Translog cost function results are available upon request.

[7]Note that the dummy variables take the value of zero before a regulation was enacted and one in all periods thereafter.

be homogeneous of degree one. From the arguments of linear homogeneity and symmetry, it follows that the τ and δ coefficients must also sum to zero. The addition of the vector V means restrictions must also be imposed on the values of k_i. In particular, the sum of k_i integers, which are multiplied with input prices, must equal zero. This restriction ensures that $\lambda C = C(Q, \lambda\mathbf{p})$ is still satisfied when using the Fourier flexible form.

Returns to scale measures can be obtained from the Fourier flexible form by taking a partial derivative with respect to the natural log of output:

$$\frac{\partial \ln C}{\partial \ln Q} = \beta_Q + \sum_n \tau_n \ln p_n + \gamma \ln Q + \sum_{i=1}^{N} k_{i,Q} \left[\zeta_i \cos(k_i V) - \phi_i \sin(k_i V)\right] \qquad (2)$$

This partial derivative gives the change in cost resulting from an increase or decrease in output. If this number is one, then cost is perfectly correlated with output and commercial bank holding companies exhibit constant returns to scale. Along the same argument, a value higher than one yields decreasing returns to scale, and less than one yields increasing returns to scale.

4.2 Panel Vector Autoregression (PVAR)

VAR, or vector autoregressive models, are those that project a given variable's current value as a function of its lagged values and lagged values of other variables. VAR models are particularly useful because all input variables are treated as endogenously determined and interdependent. With purely time-series data, the general equation for a vector autoregressive model can be written as

$$X_t = \beta_0 + \Gamma_1 X_{t-1} + \Gamma_2 X_{t-2} + ... + \Gamma_m X_{t-m} + \epsilon_t \qquad (3)$$

Where X_t is an $(nx1)$ vector of endogenous variables at some time t. The vectors on the right-hand side of the equation represent lagged values of the $(nx1)$ vector. The Γ's represent (nxn) coefficient matrices and the ϵ term is an i.i.d. $(nx1)$ vector of error terms. The number of lags is equal to m, the value of which we determine by Akaike and Schwartz information criteria. Written in a more accessible format below is an example of an n-variable, two-lag vector autoregression:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{bmatrix} = \beta_0 + \begin{bmatrix} \gamma^1_{1,1} & \gamma^1_{1,2} & \cdots & \gamma^1_{1,n} \\ \gamma^1_{2,1} & \gamma^1_{2,2} & \cdots & \gamma^1_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^1_{n,1} & \gamma^1_{n,2} & \cdots & \gamma^1_{n,n} \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ \vdots \\ x_{n,t-1} \end{bmatrix} + \begin{bmatrix} \gamma^2_{1,1} & \gamma^2_{1,2} & \cdots & \gamma^2_{1,n} \\ \gamma^2_{2,1} & \gamma^2_{2,2} & \cdots & \gamma^2_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^2_{n,1} & \gamma^2_{n,2} & \cdots & \gamma^2_{n,n} \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ \vdots \\ x_{n,t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{n,t} \end{bmatrix} \quad (4)$$

It is straightforward to see that the above equation can be extrapolated to any number of lags given the addition of more matrices. Presented thus far have only been vector autoregressions for use with times series data. Given that our data consist of 198 commercial bank holding companies across 24 years, we must extrapolate our time series model to work with a panel dataset. The econometric application of a PVAR model is slightly more complicated, but the theoretical extension to a panel dataset can be written compactly as

$$X_{i,t} = \beta_{i,0} + \Gamma_{i,1} X_{i,t-1} + \Gamma_{i,2} X_{i,t-2} + ... + \Gamma_{i,m} X_{i,t-m} + \epsilon_{i,t} \quad (5)$$

For our purposes, the most useful information from the PVAR comes in the form of impulse response functions. Impulse response functions measure the responses of current and future values of each variable to a shock, defined as a unit increase in the current value of one of the variables. For our VAR model, the shock will be a change in row n of the $(nx1)$ ϵ column vector. Because of the interpendencies that characterize a PVAR, this shock will likely affect all variables in X_t. The response over time of the X_t variables creates the impulse response functions. In general, the ϵ_n shock imposed is the size of the standard deviation of variable x_n. In addition, all variables are typically normalized to have a value equal to zero prior to the shock.

For commercial bank holding companies, PVAR is particularly interesting because the parameters that characterize an impulse response function will show the effects that an increase in assets have on cost efficiency and returns to scale.

4.3 Bayesian Panel VAR

All forms of regression analysis contain bias based on a researcher's prior beliefs. In order to develop a strong model, the researcher will select variables and, in some cases, even the form of the model (when the model is parametric). While this is a typical frequentist approach in estimation, it does impose the researcher's biases on the model. The Bayesian form of estimation helps fix this issue by assuming that model parameters are random quantities. Bayesian analysis is contrasted by frequentist analysis, which instead assumes that the data are a random sample of the global data but that the parameters are fixed and unknown. By assuming that the model's parameters are random, Bayesian analysis allows the researcher to incorporate prior knowledge in the estimation process. These priors are a researcher's best estimate of the distribution of the model's parameters before any regressions have been run. These prior distributions are then updated in a series of simulations until a final posterior distribution is determined. This form of estimation gives a more robust result than frequentist analysis because Bayesian estimation uses both data on hand and prior knowledge. Bayesian analysis will serve as a robustness check for our results obtained through a frequentist PVAR.

In Bayesian analysis, the posterior has two components: a likelihood function, which contains information about the parameters from observed data, and a prior, which contains information about the parameters from other sources. The posterior distribution generally cannot be solved for analytically, as it typically does not have a closed-form solution. Thus, to find the posterior solution one must use Markow Chain Monte Carlo sampling to approximate the distribution. Despite the differences in their theoretical basis, the form of the Bayesian PVAR is equivalent to that of Equation (5) used in PVAR, the only difference being the additional reliance on priors and the determination of a solution through estimation.

5 Results

5.1 Measurements of Returns to Scale

First, we use the Fourier flexible form to determine the returns to scale for each banks size group. The estimates of $\partial \ln C/\partial \ln Q$ are presented in Table 1.

Table 1: Returns to Scale (RTS) Measurements from the Fourier Flexible Form

Size	**RTS**
Too-Big-To-Fail Banks	2.013
Large Banks	0.106
Mid-Size Banks	0.093
Small Banks	0.526

Notes: RTS higher than one yields decreasing returns to scale, and less than one yields increasing returns to scale.

Table 1 shows that the largest commercial bank holding companies in our study exhibit decreasing returns to scale. Our results are consistent with the previous research which find increasing returns to scale up to a particular size limit. Although each size category exhibits increasing returns to scale, the medium commercial bank holding companies exhibit the largest returns to scale compared to large and small commercial bank holding companies showing a non-monotonic distribution of returns to scale.

Table 2: The Effects of Different Regulations on Total Cost

Variables	Fourier Flexible Form		
	RN	GLB	DF
TBTF	-0.858***	0.313	0.242
	(0.262)	(0.261)	(0.234)
Large	0.039	-0.263***	-0.051
	(0.099)	(0.094)	(0.082)
Medium	0.020	0.029	-0.033
	(0.031)	(0.028)	(0.027)
Small	0.064	0.071*	0.166***
	(0.049)	(0.039)	(0.035)

Notes: Here ***, **, and * denote the 5, 1, and 0.1 percent levels of significance, respectively. Standard deviations are in parenthesis.

Next, we turn to assessing the role that the banking regulations play in total costs banks incur. Interestingly, we find that the total costs of small commercial bank holding companies increased significantly as a result of Dodd-Frank legislation. This result is surprising because the Dodd-Frank legislation was marketed as a way to reduce banking risk, specifically for the largest U.S¿ commercial bank holding companies. Although the coefficients are insignificant, the total costs for TBTF has increased as expected, yet it decreased for large and medium commercial bank holding companies after the Dodd-Frank. The largest commercial bank holding companies (TBTF and large) benefited the most from legislation that granted commercial bank holding companies more freedom (Riegle-Neal and the Gramm-Leach-Bliley Act) while smaller bank holding companies hurt most in terms of cost increases. We can conclude that restrictive bank regulations, even those which are targeted towards the largest commercial bank holding companies, have the most significant negative effect on the smallest commercial bank holding companies.

5.2 Panel VAR (PVAR) Results

PVAR allows the study of the interaction between bank size, cost efficiency, and returns to scale. Since each of these variables are interrelated, and thus endogenous, PVAR is the best estimation model for this dynamic relationship. The variables used in our PVAR are cost efficiency, bank size (measured as total assets), and returns to scale. Although we can observe shocks to each endogenous variable, we are most interested in how a bank's returns to scale and cost efficiency respond to a growth in total assets. This analysis is particularly important, as Wheelock and Wilson (2012) as well as others find that returns to scale differ for firms of different sizes, while Kovner et al. (2014) conclude similarly for the cost efficiency. Because we use a panel dataset, we address the issues relating to bank fixed effects by using Helmert transformation. The use of vector autoregressive models in returns to scale and bank cost efficiency research is novel, and it represents our contribution to the field.

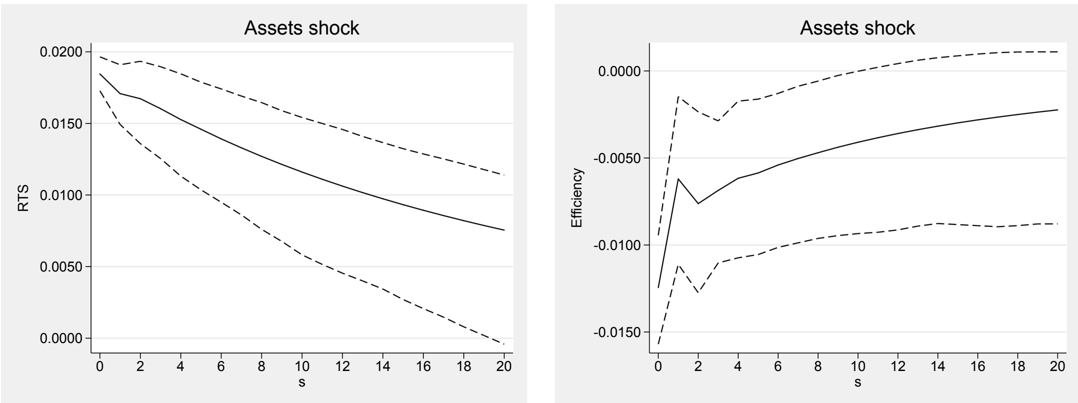


Figure 3: PVAR IRFs: Responses of Cost Efficiency and Returns to Scale After a Standard Deviation Increase in Assets

A shock to a firm's size both increases cost efficiency (decreases cost to asset ratio) and decreases returns to scale (increases in the partial derivative of costs relative to output). While the returns to scale findings match the previous literature, the cost efficiency findings are consistent with Kovner et al. (2014), where the authors find that an increase in firm size is associated with an increase in cost efficiency.

5.3 Bayesian Panel VAR Results

We obtain posterior distributions for the Bayesian PVAR using Metropolis-Hastings MCMC sampling. The Metropolis-Hastings algorithm requires initial prior distributions for each coefficient in the Bayesian PVAR. The posterior distribution of these coefficients is then updated through a series of iterations.[8] To deal with the issue of bank fixed effects, we again apply a Helmert transformation to each of the variables used.[9]

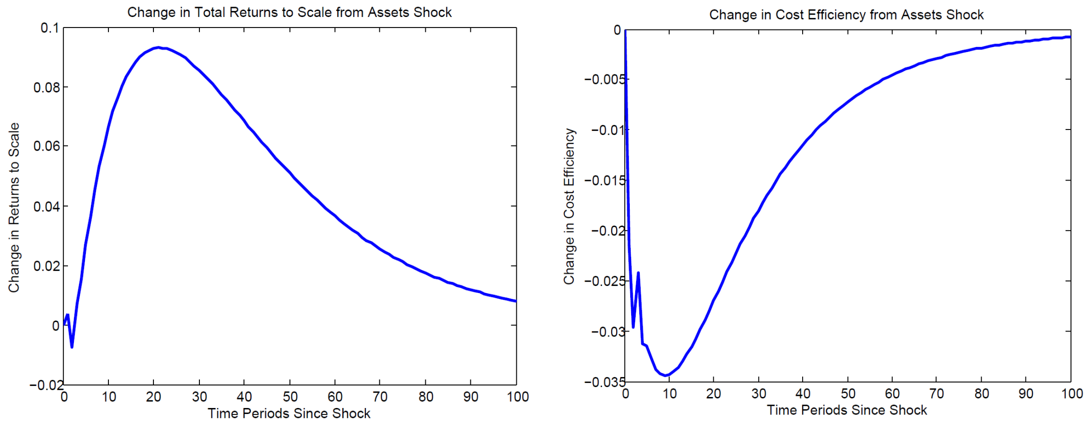


Figure 4: Bayesian PVAR IRFs: Responses of Cost Efficiency and Returns to Scale to a Standard Deviation Increase in Assets.

The conclusions drawn from the Bayesian PVAR impulse responses are consistent with those of the PVAR IRFs. An increase in total assets increases cost efficiency while decreasing returns to scale.

[8]The posterior distribution was based on 1 million MCMC iterations (of which the first 500,000 were discarded). As few as 100,000 iterations were also tested, and in all cases (with iterations between 1,000,000 and 100,000) all convergence criteria were met.

[9]Please refer to A.2 for the prior and posterior coefficient distributions for each of the nine variables included in the Bayesian PVAR.

5.4 Robustness Checks

First, we test our findings using a balanced panel dataset for the period of 1998:1 to 2014:2 to understand whether the unbalanced nature of the dataset affects the results. We find that returns to scale and PVAR results are very similar under a balanced dataset as seen in Figure 5. This analysis shows that failed banks in our period do not seem to skew our findings.

Next, we use the recovery period between the 2001 and 2007 recessions, i.e. the period from 2002:1 to 2007:3, to determine whether recessionary periods are leading to any biases. We find similar conclusions in terms of returns to scale and PVAR results. Using data exclusively from 2002:1 to 2007:3, Figure 6 shows that our conclusions are robust regardless of the inclusion or exclusion of the recessionary periods.

As was previously mentioned, once banks are placed in a given size category, they do not leave that group regardless of whether they grow larger or smaller over time. Using bank average real assets, we establish groups which banks could enter or leave across time. Allowing banks to change groups does not change the main conclusions of the paper.

Recognizing that our results could be influenced by the controls that we select, we run both the PVAR with and without macroeconomic controls as can be seen in Figure 7. The conclusions drawn from running the PVAR without controls are the same as when the PVAR was run with controls.

While our model is based on the most conservative ordering, we changed the Cholesky ordering in our PVAR estimation to account for ordering effects.[10] We note that the model is robust to ordering: a change in the variable order has no significant effects on the IRFs.

[10]For our model, the most conservative order is total assets, cost efficiency, and returns to scale.

6 Conclusion

In this paper we try to understand the underlying reasons why banks have been growing larger for 25 years. In particular, we ask the following two questions: If a bank grows larger in size, would it benefit from scale economies and/or cost efficiency gains? Do dissimilar regulatory environments affect banks differently by lessening (or increasing) costs which could in turn encourage (or discourage) banks to grow? To answer these questions, we study the effects of an increase in bank size on returns to scale and cost efficiency gains.

Knowing that banks are heterogeneous, we first divide our dataset into four groups based on bank asset sizes. Using the Fourier flexible form, we find that all but the largest banks exhibit increasing returns to scale. Next, we analyze the dynamic relationship between bank size, returns to scale, and cost efficiency through Panel VAR and Bayesian Panel VAR analyses. Both methods show that an increase in bank size decreases the chances that a bank can exploit returns to scale. On the other hand, cost efficiency gains increase as size increases. Thus, banks seem to exploit increasing returns to scale up until they become too large; however, they continue to benefit from cost efficiency gains even when they are the largest size.

We further analyze the role of different regulatory environments and study whether the regulations impose (or lessens) costs on banks. In particular, we concentrate on the past 25 years and cover more freeing regulations such as Riegle-Neal Interstate Banking and Branching Efficiency Act and Gramm-Leach-Bliley Act as well as a more restrictive regulation like the Dodd-Frank Wall Street Reform and Consumer Protection Act. Both type of regulations seem to benefit larger banks but hurt smaller banks in terms of increasing operating costs. In particular, the largest commercial bank holding companies benefit more from the most flexible regulations, whereas the smaller commercial bank holding companies are generally hurt more by the most restrictive regulations. This is an interesting result because the most restrictive regulation (Dodd-Frank) is generally focused on the largest commercial bank holding companies. However, Dodd-Frank seemed to have a significant influence on the costs of the smallest commercial bank holding companies. Our results suggest that regulators must consider the auxiliary implications of regulations on commercial bank holding companies that may not be in their target group.

References

[1] Barth, J., Brumbaugh, R., & Wilcox, J. (2000). Policy Watch: The Repeal of Glass-Steagall and the Advent of Broad Banking. *Journal of Economic Perspectives*, 14(2), 191-204.

[2] Barth, J., Caprio, G., & Levine, R. (2004). Bank Regulation and Supervision: What Works Best? *Journal of Financial Intermediation*, 13(2), 205-248. doi:10.1016/j.jfi.2003.06.002

[3] Berger, A., & Hannan, T. (1998). The Efficiency Cost of Market Power in the Banking Industry: A Test of the "Quiet Life" and Related Hypotheses. *Review of Economics and Statistics*, 80(3), 454-465. doi:10.1162/003465398557555

[4] Chronopoulos, D., Mcmillan, F., & Wilson, J. (2015). The Dynamics of US Bank Profitability. *The European Journal of Finance*, 21(5), 426-433.

[5] Cornett, M., Mcnutt, J., & Tehranian, H. (2006). Performance Changes Around Bank Mergers: Revenue Enhancements versus Cost Reductions. *Journal of Money, Credit, and Banking*, 38(4), 1013-1050.

[6] Demirgüç-Kunt, A., & Detragiache, E. (2002). Does deposit insurance increase banking system stability? An empirical investigation. *Journal of Monetary Economics*, 49(7), 1373-1406.

[7] Hughes, J. P., & Mester, L. J. (2013). Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation*, 22, 559–585.

[8] Hunter, W., Timme, S., & Yang, W. (1990). An Examination of Cost Subadditivity and Multiproduct Production in Large U.S. Banks. *Journal of Money, Credit and Banking*, 22(4), 504-525. doi:10.2307/1992434

[9] Jaremski, M., & Sapci, A. (2014). Understanding the Cyclical Nature of Financial Intermediation Costs.

[10] Kovner, A., Vickery, J., & Zhou, L. (2014). Do Big Banks Have Lower Operating Costs? *Economic Policy Review*, 20(2), 1-27.

[11] McAllister, P., & McManus, D. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking Finance*, 17(2-3), 389-405. doi:10.1016/0378-4266(93)90039-G

[12] Noulas, A., Ray, S., & Miller, S. (1990). Returns to Scale and Input Substitution for Large U. S. Banks. *Journal of Money, Credit and Banking*, 22(1), 94-108. doi:10.2307/1992130

[13] Schweitzer, S. (1972). Economies of Scale and Holding Company Affiliation in Banking. *Southern Economic Journal*, 39(2), 258-266. doi:10.2307/1056596

[14] Skolrud, T. (2013). Reducing Approximation Error in the Fourier Flexible Form.

[15] Wheelock, D., & Wilson, P. (2001). New evidence on returns to scale and product mix among U.S. commercial banks. *Journal of Monetary Economics*, 47(3), 653-674. doi:10.1016/S0304-3932(01)00059-9

[16] Wheelock, D., & Wilson, P. (2012). Do Large Banks have Lower Costs? New Estimates of Returns to Scale for U.S. Banks. *Journal of Money, Credit and Banking*, 44(1), 171-199. doi:10.1111/j.1538-4616.2011.00472.x

A Appendices

A.1 Data

A.1.1 Data Definition and Measurement

Mergent Online obtains their data from FR Y-9C filings, which must be filled out quarterly by bank holding companies with over 500 million USD in assets (more information on these requirements can be found in the Bank Holding Company Act, Regulation Y, and the Homeowners Loan Act). In FR Y-9C filings, Schedule HI, item 7 contains a list of possible noninterest expenses a bank holding company could incur: salaries and employee benefits, expenses of premises and fixed assets, goodwill impairment, amortization expenses, other noninterest expenses, and total noninterest expenses. In addition, Schedule HI, memoranda item 7 contains several more possible expenses: data processing, advertising and marketing, director's fees, printing and supplies, postage, legal fees and expenses, FDIC deposit insurance assessments, accounting and auditing expenses, consulting and advisory expenses, interchange fees, and telecommunications expenses. Under this memoranda item, bank holding companies are also able to create additional accounts. The Federal Reserve Microdata Reference Manual lays out the definition of costs reported in FR Y-9C filings. Some of the largest operating costs definitions that we use in this paper are shown in Table 3.

Many of the bank holding companies in our dataset contain several quarters where data is unavailable for certain expenses. Often, missing data are a result of a company switching accounting procedures. If a missing quarter occurs between quarters containing data, the missing entry is replaced with an average of the quarters before and after the entry as long as the missing entries do not exceed three for the entire sample.

Since banks do not provide individual financial statements for their fourth-quarters, we obtained fourth quarter noninterest expenses by summing the three available quarters and subtracting the total from the similar cost presented in the annual reports. We paid special attention to matching the quarterly financial statements with annual statements for every bank and each year. When we fail to do so, we obtained missing fourth-quarter values by averaging the preceding third-quarter and following first-quarter values. If this adjustment is made for more than three years, however, we remove the bank from our dataset.

The cost decomposition data used in this paper are far more detailed than the reports that are regulated by the SEC.[11] Thus, sometimes, bank holding companies had slightly different methods of reporting their detailed financial information. Two such categories that required the most careful analysis were the personnel and occupancy expenses. For personnel, some banks report only salaries paid (breaking employee benefits into a separate account), whereas other holding companies would report only one general category. Similarly, some companies choose to group their occupancy costs with their equipment costs. Because the personnel and occupancy data were the largest cost categories, we combined the equipment and occupancy into one category and personnel and employee benefits costs into another for each bank holding company in our dataset.

Table 3: Cost Decomposition and Descriptions

Variable	**Description**
Personnel	salaries and benefits for all officers and employees of the bank and its consolidated subsidiaries.
Occupancy	all noninterest expenses related to the use of premises, equipment, furniture, and fixtures. Premises and fixed assets are defined net of rental income. In addition to rental deductions, income from assets that indirectly represent premises, equipment, furniture, or fixtures included in "Premises and Fixed Assets" are also deducted.
Advertising and Marketing	advertising, production, agency fees, direct mail, marketing research, public relations, seminars, and customer magazines.
Professional Fees	sales training by consultants, public accountants' fees, management services, consulting fees for economic surveys, and other special advisory services.

[11]Please refer to the Tables from 4 to 7 for detailed cost breakdown

Other Noninterest Expenses	Other noninterest expenses is a category intended to include items not required to be reported individually in Schedule HI, item 7. The Federal Reserve Microdata Reference Manual lists 31 unique costs that should be included in other noninterest expenses. Some of these costs include civil penalties and fines as well as costs of gifts given to depositors.

A.1.2 Summary Statistics

Below are our tables showing the different summary statistics for TBTF, large, medium, and small commercial bank holding companies.

Table 4: Summary Statistics for Too-Big-Too-Fail Banks

Variable	Obs	Mean	Std. Dev.	Min	Max
Insurance Exp	0	.	.	.	.
Supp. and Print	0	.	.	.	.
Software Exp	32	106406	31925	53000	157000
Occupancy Exp	412	710902	612148	51200	2406000
Marketing Exp	221	321462	246331	10000	926000
Data Processing Exp	66	447955	223402	107000	856000
Loan Processing Exp	0	.	.	.	.
Prof. Services Exp	263	492191	507096	6424	2109000
Litigation Exp	0	.	.	.	.
Telecommunications Exp	301	411710	382001	32700	1646000
Travel Exp	0	.	.	.	.
Postal and Courier Exp	23	36820	13125	19324	67000
Card Processing Exp	0	.	.	.	.
Personnel Exp	435	3400753	2671402	108800	1.02e+07
Other Noninterest Exp	415	1611499	1744373	51000	1.31e+07
Total Noninterest Exp	526	6437326	5209732	253472	2.72e+07
Total Deposits	522	3.53e+08	3.48e+08	22000	1.32e+09
Gross Loans	465	3.22e+08	2.95e+08	475000	9.77e+08
Total Assets	546	7.68e+08	6.98e+08	2.12e+07	2.52e+09
Number of Banks	7				

Table 5: Summary Statistics for Large Banks

Variable	Obs	Mean	Std. Dev.	Min	Max
Insurance Exp	0	.	.	.	.
Supp. and Print	205	4779	2761	826	11872
Software Exp	171	47077	54013	3523	190000
Occupancy Exp	1673	77899	67253	681	447000
Marketing Exp	770	54175	79668	1060	511142
Data Processing Exp	315	58086	53740	4140	233000
Loan Processing Exp	48	49583	21035	26000	111000
Prof. Services Exp	857	54415	73099	220	518000
Litigation Exp	0	.	.	.	.
Telecommunications Exp	504	54993	59672	4126	266200
Travel Exp	0	.	.	.	.
Postal and Courier Exp	68	47027	31196	7432	81000
Card Processing Exp	70	115500	51452	44000	193000
Personnel Exp	1722	324001	260935	3430	1404000
Other Noninterest Exp	1113	116330	140158	1222	1218000
Total Noninterest Exp	1839	662793	589071	5520	7273350
Total Deposits	1685	4.95e+07	4.49e+07	452201	2.76e+08
Gross Loans	1504	4.87e+07	4.26e+07	1734832	2.44e+08
Total Assets	1832	7.54e+07	6.67e+07	2047633	3.89e+08
Number of Banks	25				

Table 6: Summary Statistics for Medium-Sized Banks

Variable	Obs	Mean	Std. Dev.	Min	Max
Insurance Exp	108	2791	1725	107	8589
Supp. and Print	432	2046	1729	366	6292
Software Exp	36	5646	3804	1704	11087
Occupancy Exp	3144	12366	12195	114	120213
Marketing Exp	1515	3131	2943	80	24870
Data Processing Exp	885	5605	5396	117	29071
Loan Processing Exp	208	4686	4342	577	40786
Prof. Services Exp	1453	4331	4083	242	29905
Litigation Exp	43	4008	2751	855	12806
Telecommunications Exp	632	3916	4406	129	31000
Travel Exp	114	2766	2138	622	10194
Postal and Courier Exp	400	3434	3669	165	15960
Card Processing Exp	71	5733	3913	1686	16018
Personnel Exp	3469	42467	39139	1028	362340
Other Noninterest Exp	2694	14170	13255	156	113300
Total Noninterest Exp	3665	81385	76686	1410	754678
Total Deposits	3553	7382313	5338367	301673	4.65e+07
Gross Loans	3389	6297472	4576498	43801	2.80e+07
Total Assets	3565	9973001	7160884	386737	5.00e+07
Number of Banks	54				

Table 7: Summary Statistics for Small Banks

Variable	Obs	Mean	Std. Dev.	Min	Max
Insurance Exp	156	637	610	11	3330
Supp. and Print	886	303	154	58	885
Software Exp	83	789	239	379	1575
Occupancy Exp	6204	2526	2616	41	42967
Marketing Exp	2477	657	564	0	10977
Data Processing Exp	2700	1052	1463	0	23161
Loan Processing Exp	38	281	204	61	1044
Prof. Services Exp	2232	1081	1060	6	10816
Litigation Exp	189	493	353	28	2001
Telecommunications Exp	1052	1300	2970	51	26122
Travel Exp	18	191	42	128	305
Postal and Courier Exp	561	480	300	25	1340
Card Processing Exp	268	730	582	90	2746
Personnel Exp	6709	10530	17476	67	395936
Other Noninterest Exp	5237	3530	3500	0	69911
Total Noninterest Exp	7140	19561	26843	1	482944
Total Deposits	6872	1792173	1289990	4801	1.20e+07
Gross Loans	6461	1573544	1158160	341	1.09e+07
Total Assets	6944	2332048	1725824	10161	1.53e+07
Number of Banks	101				

A.2 Bayesian Panel VAR Priors and Posteriors

Table 8: Distributions Where Total Assets is Dependent Variable

Variable	Prior (μ, σ^2)	Posterior (μ, σ^2)
L1.Total Assets	N(0,10xσ)	(0.863, 1.97e-04)
L2.Total Assets	N(0,10xσ)	(0.096, 3.49e-04)
L3.Total Assets	N(0,10xσ)	(0.007, 1.92e-04)
L1.RTS	N(0,10xσ)	(0.019, 1.81e-04)
L2.RTS	N(0,10xσ)	(0.000, 3.18e-04)
L3.RTS	N(0,10xσ)	(-0.030, 1.71e-04)
L1.Cost Efficiency	N(0,10xσ)	(-0.002, 2.72e-05)
L2.Cost Efficiency	N(0,10xσ)	(-0.012, 2.67e-05)
L3.Cost Efficiency	N(0,10xσ)	(0.009, 2.71e-05)

Table 9: Distributions Where RTS is Dependent Variable

Variable	Prior (μ, σ^2)	Posterior (μ, σ^2)
L1.Total Assets	N(0,10xσ)	(0.008, 2.04e-04)
L2.Total Assets	N(0,10xσ)	(-0.030, 3.58e-04)
L3.Total Assets	N(0,10xσ)	(0.048, 1.93e-04)
L1.RTS	N(0,10xσ)	(0.905, 1.89e-04)
L2.RTS	N(0,10xσ)	(0.189, 3.26e-04)
L3.RTS	N(0,10xσ)	(-0.163, 1.74e-04)
L1.Cost Efficiency	N(0,10xσ)	(0.007, 2.73e-05)
L2.Cost Efficiency	N(0,10xσ)	(-0.002, 2.64e-05)
L3.Cost Efficiency	N(0,10xσ)	(-0.010, 2.70e-05)

Table 10: Distributions Where Cost Efficiency is Dependent Variable

Variable	Prior (μ, σ^2)	Posterior (μ, σ^2)
L1.Total Assets	N(0,10xσ)	(-0.046, 0.001)
L2.Total Assets	N(0,10xσ)	(-0.011, 0.002)
L3.Total Assets	N(0,10xσ)	(0.028, 0.001)
L1.RTS	N(0,10xσ)	(-0.031, 0.001)
L2.RTS	N(0,10xσ)	(0.061, 0.002)
L3.RTS	N(0,10xσ)	(-0.015, 0.001)
L1.Cost Efficiency	N(0,10xσ)	(0.276, 1.70e-04)
L2.Cost Efficiency	N(0,10xσ)	(0.337, 1.66e-04)
L3.Cost Efficiency	N(0,10xσ)	(0.159, 1.68e-04)

A.3 Robustness Checks

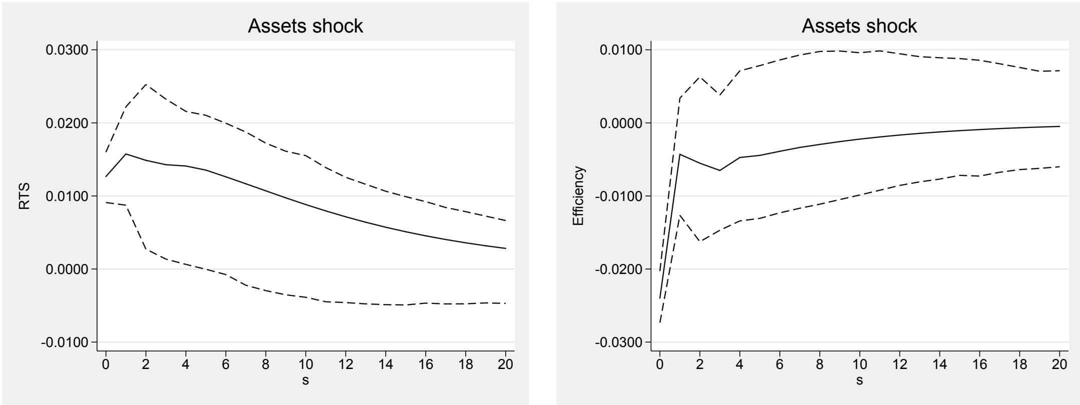


Figure 5: PVAR IRFs measuring cost efficiency and returns to scale after a total asset shock with data for 2002:1-2007:3, excluding recessions.

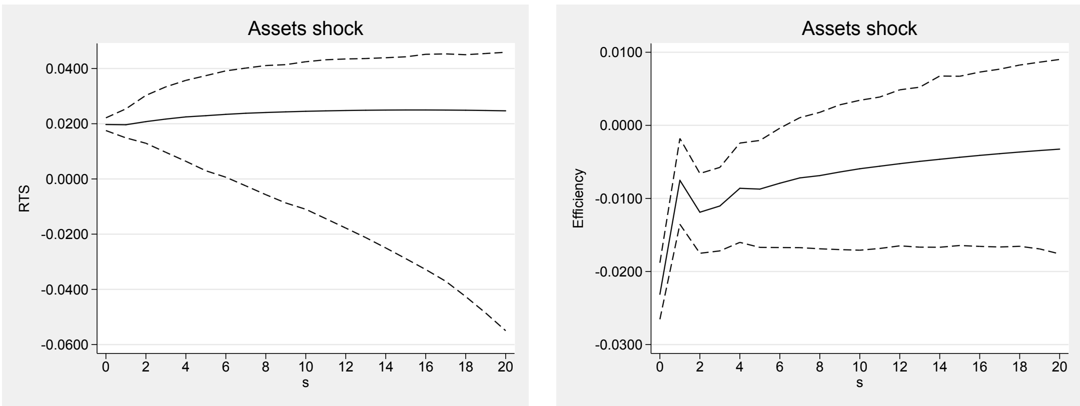


Figure 6: PVAR IRFs measuring cost efficiency and returns to scale after a total asset shock with a balanced dataset and data from 1998:1-2014:2.

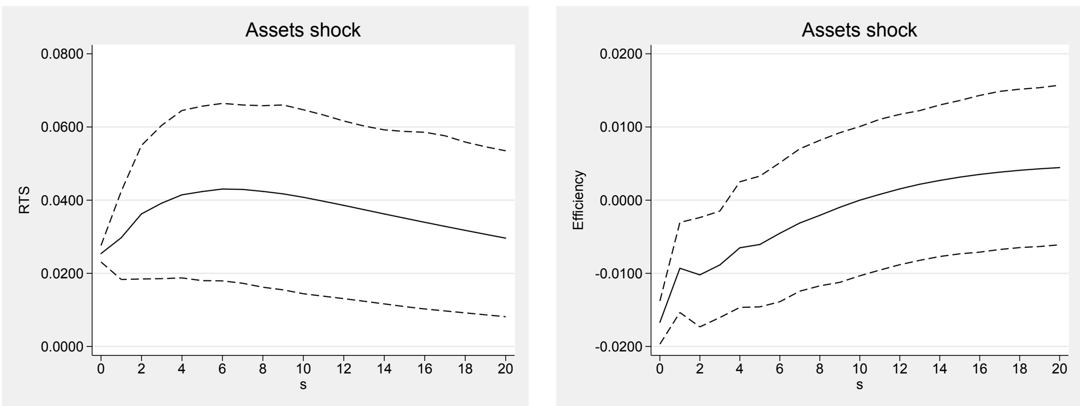


Figure 7: PVAR IRFs measuring cost efficiency and returns to scale after a total asset shock with the entire dataset, including macroeconomic controls.