The Evolution of Scale Economies in U.S. Banking

DAVID C. WHEELOCK          PAUL W. WILSON*

February 2017

Abstract

Continued consolidation of the U.S. banking industry and a general increase in the size of banks has prompted some policymakers to consider policies that discourage banks from getting larger, including explicit caps on bank size. However, limits on the size of banks could entail economic costs if they prevent banks from achieving economies of scale. This paper presents new estimates of returns to scale for U.S. banks based on nonparametric, local-linear estimation of bank cost, revenue and profit functions. We report estimates for both 2006 and 2015 to compare returns to scale some seven years after the financial crisis and five years after enactment of the Dodd-Frank Act with returns to scale before the crisis. We find that a high percentage of banks faced increasing returns to scale in cost in both years, including most of the 10 largest bank holding companies. And, while returns to scale in revenue and profit vary more across banks, we find evidence that the largest four banks operate under increasing returns to scale.

Forthcoming, Journal of Applied Econometrics

*Wheelock: Research Department, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166–0442; wheelock@stls.frb.org. Wilson: Department of Economics and School of Computing, Division of Computer Science, Clemson University, Clemson, South Carolina 29634–1309, USA; email pww@clemson.edu. This research was conducted while Wilson was a visiting scholar in the Research Department of the Federal Reserve Bank of St. Louis. We thank the Cyber Infrastructure Technology Integration group at Clemson University for operating the Palmetto cluster used for our computations. We thank three anonymous referees for comments, and Peter McCrory and Paul Morris for research assistance. The views expressed in this paper do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System. JEL classification nos.: G21, L11, C12, C13, C14. Keywords: banks, returns to scale, scale economies, nonparametric, regression.
1 Introduction

The financial crisis of 2007–08 raised new concerns about the size and complexity of the world’s largest banking organizations. Many of the largest banks are now considerably bigger than they were before the crisis. For example, on December 31, 2006, the largest U.S. bank holding company (Citigroup) had total consolidated assets of $1.9 trillion, while two others (Bank of America and JPMorgan Chase) also had more than $1 trillion of assets. By contrast, on December 31, 2015, the largest holding company (JPMorgan Chase) had $2.35 trillion of assets and three others had assets in excess of $1.7 trillion.

Are banks destined to become ever larger and, if so, is that cause for concern? The answer to this question depends, in part, on why banks have been getting larger. The policy implications are likely different if banks are growing larger to exploit technologically-driven scale economies than if government policies that encourage large size or excessive risk taking are driving bank growth. Of particular concern is the perception that regulators consider very large banks “too-big-to-fail” (TBTF), which would provide an implicit funding subsidy to banks that exceed a certain size threshold. The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 was intended to eliminate TBTF by establishing a formal process for resolving failures of large financial institutions, as well as by imposing a tighter financial regulatory regime. However, some economists and policymakers argue that Dodd-Frank does not go far enough to contain TBTF, and that banks should be subject to firm caps on their size (e.g., Fisher and Rosenblum, 2012). The imposition of size limits on banks could have a downside, however, if they prevent banks from achieving economies of scale (as noted, e.g., by Stern and Feldman, 2009). Hence, the extent to which there are scale economies in banking is an important question that has attracted renewed interest among researchers and policymakers.

This paper presents new estimates of returns to scale (RTS) for U.S. bank holding companies (BHCs) and independent (i.e., not BHC owned) commercial banks. The paper makes two main contributions. First, we report estimates for both 2006 and 2015 (as well as for 1986 and 1996) to provide a comparison of RTS for the largest banks some seven years after the financial crisis and five years after the enactment of Dodd-Frank with estimates for 2006 and earlier years. Second, whereas previous studies focus exclusively on scale economies in
terms of cost, we estimate RTS in terms of revenue and profit, as well as cost. Although estimates of RTS from a cost perspective indicate whether society’s resources are employed efficiently in providing banking services, economies of scale in revenue or profit are of concern to bank shareholders, as well as to policymakers interested in the forces driving industry consolidation.

Conventional wisdom, based largely on studies that use data from the 1980s and 1990s to estimate returns to scale from cost functions, holds that banks exhaust scale economies at low levels of output, e.g., $100–$300 million of total assets. However, several recent studies find evidence of increasing returns to scale (IRS) among much larger banks, including banks with more than $1 trillion of assets. Improved estimation methods and data could explain the difference in findings between older studies and more recent ones. However, recent advances in technology are often thought to have favored larger banks, and perhaps increased the size range over which banks could experience IRS (Berger, 2003; Mester, 2005). Wheelock and Wilson (2009) find that larger banks experienced larger gains in productivity over 1985–2004 than did smaller banks. Feng and Serletis (2009) find similar evidence for 1998–2005. These studies suggest that technological advances may have also generated IRS for banks. Indeed, using a variety of methodologies and datasets, several recent studies find more evidence of substantial economies of scale in banking, with some finding that even very large banks operate under IRS (e.g., Wheelock and Wilson, 2012; Hughes and Mester, 2013; and Kovner et al., 2014 for U.S. banks, and Becalli et al., 2015 for European banks). However, other studies are less conclusive (e.g., Feng and Zhang, 2014; Restrepo-Tobón and Kumbhakar, 2015) and questions remain.

Changes in regulation, notably the Dodd-Frank Act of 2010, might also have affected returns to scale by altering the environment in which banks operate. Most studies use data on banks from before the financial crisis of 2007–08 or just shortly thereafter. However, many of the largest U.S. banks have continued to grow even larger since the crisis, perhaps to the point of exhausting potential scale economies. Research using more recent data is

---

1 Studies of scale economies in banking from the 1980s and before typically relied on estimation of translog or other parametric specifications of bank cost functions. However, subsequent studies, including the present paper, find that the translog function is a misspecification of bank cost relationships and therefore can lead to erroneous estimates of returns to scale (RTS). See McAllister and McManus (1993) and Wheelock and Wilson (2001) for discussion and evidence on the bias introduced by estimating bank scale economies from a translog cost function.
thus required to determine whether earlier conclusions about the extent of scale economies in banking are still true.

In addition to providing an update to previous research on scale economies from the perspective of bank costs, we also examine scale economies in revenue and profit. Several studies have estimated revenue and profit relationships for banks to study such topics as revenue economies of scope (Berger et al., 1996), profit efficiency (e.g., Berger and Mester, 1997), and profit productivity (Berger and Mester, 2003). In addition to economies of scope, Berger et al. (1996) estimate revenue ray-scale economies for a sample of banks using data for 1978, 1984 and 1990. That study finds evidence of significant revenue scale economies in 1978 and 1984, especially for banks with less than $500 million of assets, but not in 1990. We are unaware of any other studies that examine revenue or profit scale economies for banks. Berger and Mester (2003) argue, however, that studies that ignore revenues when evaluating bank performance could be misleading. For example, Berger and Mester (2003) find that during the 1990s, banks became less productive in terms of cost (essentially that cost per unit of output rose after controlling for output quantities, input prices, and various environmental conditions), but more productive at generating profits. Berger and Mester (2003) attribute this finding to efforts by banks to increase profits by providing more or better quality services that raise their revenues by more than they increase costs. Similarly, an examination of the evolution of RTS from a revenue or profit perspective could provide a more complete picture of scale economies in banking than a focus solely on economies of scale in terms of cost.

We use a nonparametric, local-linear estimator to estimate cost, revenue and profit relationships from which we derive estimates of RTS. The nonparametric approach avoids the potential for functional-form specification error associated with parametric estimation. Although nonparametric estimators are plagued by the “curse of dimensionality,” i.e., slow convergence rates (compared to parametric estimators) that become exponentially slower with more model dimensions, we take steps to mitigate this problem. Specifically, we estimate our models using a large dataset consisting of over 800,000 observations on all U.S. banks for 1986–2015, and we employ principal components techniques to reduce the dimensions of our empirical models. Our estimation methodology is similar to that of Wheelock and Wilson (2012). However, Wheelock and Wilson (2012) focus exclusively on scale economies
in terms of cost and estimate RTS for U.S. banks in 2006. Here we extend the methodology to the estimation of RTS in terms of revenue and profit, report estimates for both 2006 and 2015 (as well as 1986 and 1996), and test whether changes in RTS between 2006 and 2015 are statistically significant.

We find that, despite the growth in size of many of the largest banks during and since the financial crisis, the very largest banks continued to face IRS in terms of cost in 2015. In fact, our estimates indicate that many of the largest banks experienced statistically significant increases in RTS between 2006 and 2015. Among all banks, approximately 35 percent of banks operated under IRS in 2006, while 43 percent faced IRS in 2015. Among all banks that were in existence in both 2006 and 2015, 27 percent more banks faced IRS in 2015 than in 2006, while only a few banks experienced decreasing returns to scale (DRS) in either period. Our results for revenue and profit economies are more mixed. While overall we find that fewer banks faced IRS in terms of revenue or profit than in terms of cost, we find evidence of IRS among a number of the largest banks in both 2006 and 2015, especially among the four largest U.S. banking organizations.

The next section describes the microeconomic specification of our cost, revenue, and profit functions and the statistics to measure RTS. Section 3 introduces the econometric specification, and Section 4 discusses the nonparametric methods we use for estimation and inference. Sections 5 and 6 present our empirical findings and conclusions. Additional details on data, estimation, and results are provided in separate Appendices A–E, which are available online.

2 Microeconomic Specification

To establish notation, let \( \mathbf{x} \in \mathbb{R}_+^p \) and \( \mathbf{y} \in \mathbb{R}_+^q \) denote column-vectors of \( p \) input quantities and \( q \) output quantities, respectively. Let \( \mathbf{w} \in \mathbb{R}^p \) denote the column-vector of input prices corresponding to \( \mathbf{x} \), and let \( \mathbf{r} \in \mathbb{R}^q \) denote the column-vector of output prices corresponding to \( \mathbf{y} \). Then variable costs are given by \( C := \mathbf{w}' \mathbf{x} \), which firms (banks) seek to minimize with respect to \( \mathbf{x} \), subject to \( h(\mathbf{x}, \mathbf{y}) = 0 \) where \( h(\cdot, \cdot) \) represents the product-transformation function that determines the possibilities for transforming input quantities \( \mathbf{x} \) into output quantities \( \mathbf{y} \). Solution to the constrained minimization problem yields a mapping \( \mathbb{R}_+^q \times \mathbb{R}^p \mapsto \mathbb{R}_+^p \).
\[ \mathbb{R}_+^n \text{ such that } x = x(y, w); \text{ substitution into } C = w'x \text{ yields} \]

\[ C = w'x = w'x(y, w) = C(y, w) \]  

(2.1)

where \( C(y, w) \) is the variable cost function.

The story so far is part of the standard microeconomic theory of the firm (e.g., see Varian, 1978). Under perfect competition in output markets, the same body of theory implies that banks maximize revenue \( R := r'y \) with respect to output quantities, again subject to \( h(x, y) = 0 \), yielding the solution \( y = y(r, x) \). Substitution then yields \( R = r'y(r, x) = R^s(x, r) \), i.e., a standard revenue function that maps input quantities and output prices to revenue. Fuss and McFadden (1978) and Laitinen (1980) describe the conditions on \( h(x, y) \) required for existence of the revenue (and profit) function(s).

Banking studies, however, often estimate alternative revenue or profit functions, where revenue (or profit) are functions of output levels and input prices. As discussed, for example, by Berger and Mester (1997), the alternative revenue and profit functions provide a means of controlling for unmeasured differences in output quality across banks, imperfect competition in bank output markets (which gives banks some pricing power), any inability of banks to vary output quantities in the short-run, and inaccuracy in the measurement of output prices.

Estimates of economies of scale from alternative revenue and profit functions provide information about the extent to which revenue (or profit) rises for a given increase in output, holding input prices constant. Berger et al. (1996) describe the assumptions underlying standard and alternative revenue functions, and the validity of those assumptions for banks. The standard form assumes that banks are price takers. The alternative form, by contrast, assumes that banks have some pricing power, and views banks as having greater on-going flexibility in setting output prices than output levels. Based on a review of available evidence, Berger et al. (1996) conclude that some two-thirds of bank revenues are associated with services that reflect a degree of price-setting behavior, and they proceed by viewing banks as negotiating prices and fees, where feasible, to maximize revenues and profits for given levels of output. They argue that this model better represents how banks actually operate than the perfectly-competitive model which underlies standard revenue and profit functions. Berger and Mester (1997, 2003) elaborate further on the advantages of the alternative form of the revenue and profit function. For example, they note that in addition to admitting
the possibility that banks have some degree of pricing power, the alternative form can be informative about bank performance when there are unmeasured differences in the quality of bank services across banks, when banks are unable to adjust their sizes quickly, or when output prices are not measured accurately. Indeed, bank input prices are, for the most part, more readily observed in bank call report data than output prices. The absence of output price information for the vast majority of banks means that standard revenue or profit functions cannot be estimated (unless outputs are aggregated to an even greater degree than they already are in our models).

Following Berger et al. (1996) and others, we assume that banks maximize revenue with respect to output prices $r$, subject to $g(y, w, r) = 0$, where $g()$ is an implicit function representing the bank’s opportunities for transforming given output levels $y$ and input prices $w$ into output prices $r$. Solution of this constrained optimization problem yields a mapping $\mathbb{R}_+^q \times \mathbb{R}^p \mapsto \mathbb{R}^q$ such that $r = r(y, w)$; then $R = r' y = r(y, w)' y = R(y, w)$, where $R(y, w)$ is the alternative revenue function introduced by Berger et al. (1996).

Turning to profits, let $P = [r' \ w]'$ and $Q = [y' - x']'$. Standard theory suggests that firms operating in perfectly competitive input and output markets maximize profit $\pi := P'Q$ with respect to $Q$, subject to $h(x, y) = 0$. Solution of the constrained optimization problem yields $Q = Q(P)$; substituting this back into the profit function $\pi = P'Q$ gives $\pi = P'Q(P) = \pi^*(w, r)$, i.e., the standard profit function that maps input and output prices into profit. Under imperfect competition in output markets, however, banks maximize profit with respect to input quantities $x$ and output prices $r$, subject to $h(x, y) = 0$ and $g(y, w, r) = 0$. The solution results in a mapping $\mathbb{R}_+^q \times \mathbb{R}^p \mapsto \mathbb{R}^p$ such that $x = x(y, w)$, and a mapping $\mathbb{R}_+^q \times \mathbb{R}^p \mapsto \mathbb{R}^q$ such that $r = r(y, w)$. Substituting these into the profit function gives

$$\pi = P'Q = [r(y, w)' \ w'] [y' - x(y, w)']' = \pi(y, w) \quad (2.2)$$

where $\pi(y, w)$ is the alternative profit function that maps output quantities and input prices to profit.

Note that the cost function $C(y, w)$ must be homogeneous of degree one with respect to input prices $w$ since the cost minimization problem implies that factor demand equations must be homogeneous of degree zero in input prices. However, there is no such requirement for the alternative revenue and profit functions. Without additional assumptions, the al-
ternative revenue and profit functions are neither homogeneous with respect to input prices \( w \) nor homogeneous with respect to output quantities \( y \). See Berger et al. (1996) and Restrepo-Tobón and Kumbhakar (2014) for discussion.

To measure RTS using the cost function, we define

\[
E_{C,i} := \left( \delta C(y_i, w_i) - C(\delta y_i, w_i) \right) \left( \delta C(y_i, w_i) \right)^{-1}
\]  

(2.3)

where \( \delta > 1 \) is a constant and \( y_i \) is the observed vector of output quantities produced by the \( i \)th bank facing observed input prices \( w_i \). Clearly, \( E_{C,i} < 1 \). The statistic \( E_{C,i} \) measures expansion-path scale economies as the difference between \( \delta \) times the cost of producing output quantities \( y_i \) and the cost of producing output quantities scaled by the factor \( \delta \). The difference is normalized by dividing by \( \delta C(y_i, w_i) \). If \( E_{C,i} > (=, <) 0 \) then bank \( i \) faces IRS (CRS, DRS) in terms of cost.

To interpret the magnitude of \( E_{C,i} \), rearrange terms in (2.3) to obtain

\[
\eta_{C,i} := \delta (1 - E_{C,i}) = C(\delta y_i, w_i)/C(y_i, w_i).
\]

(2.4)

Hence firm \( i \) increases its output by a factor \( \delta > 1 \), its cost increases by a factor \( (1 - E_{C,i})\delta \). For example, if \( \delta = 1.1 \) and \( E_{C,i} = 0.05 \), then firm \( i \) incurs a 4.5-percent increase in cost when it increases its output level by 10 percent since \( 1.1 \times (1 - 0.05) \approx 1.045 \). The measure \( \eta_{C,i} \) defined in (2.4) can be interpreted as a “pseudo elasticity.” For \( \delta = 1.1 \) (i.e., for a 10-percent increase in output levels), costs increase by \( (\eta_{C,i} \times 100)\)-percent, and the firm faces IRS (CRS, DRS) if \( \eta_{C,i} < (=, >) 1.1 \).

As in many empirical studies, the revenue measure introduced below in Section 3 consists of *net* revenues and can take negative values. Of course, profits can also be negative. Therefore, to measure RTS from the revenue and profit functions, we define

\[
E_{R,i} := \left( R(\delta y_i, w_i) - \delta R(y_i, w_i) \right) \left( \delta |R(y_i, w_i)| \right)^{-1}
\]

(2.5)

and

\[
E_{\pi,i} := \left( \pi(\delta y_i, w_i) - \delta \pi(y_i, w_i) \right) \left( \delta |\pi(y_i, w_i)| \right)^{-1}.
\]

(2.6)

\footnote{The measure defined in (2.4) has an additional interpretation. Some algebra reveals that \( E_{C,i} (>, =, <) 1 \) iff \( E_{C,i} (<, =, >) 1 - \delta^{-1} \). For \( \delta = 1.1 \), \( (1 - \delta^{-1}) \approx 0.09091 \). Hence values of \( E_{C,i} \) less than 0.09091 indicate that a 10 percent increase in output levels results in an increase in total (variable) cost, whereas values of \( E_{C,i} \) greater than 0.09091 indicate that a 10 percent increase in output reduces cost. Of course, it is probably unlikely, but perhaps not impossible, for an increase in output to reduce total cost.}
In the definitions of $\mathcal{E}_{R,i}$ and $\mathcal{E}_{\pi,i}$, the constant factor $\delta$ multiplies output levels $y$ in the first term of the numerator, in contrast to the definition of $\mathcal{E}_{C,i}$ in (2.3), where $\delta$ multiplies output levels $y$ in the second numerator term. Similarly, $\delta$ multiplies the second term of the numerators of $\mathcal{E}_{R,i}$ and $\mathcal{E}_{\pi,i}$, rather than the first term as in (2.3). These differences reflect the fact that banks attempt to maximize revenue and profit but minimize cost. In addition, the denominators in (2.5) and (2.6) involve absolute values to account for the possibility that measured revenue or profit can be negative. Clearly, $\mathcal{E}_{R,i} > (=, <) 0$ implies IRS (CRS, DRS) and similarly for values of $\mathcal{E}_{\pi,i}$. Moreover, following the logic in footnote 2, it is easy to show that a 10 percent increase in output levels (i.e., $\delta = 1.1$) increases revenue or profit whenever $\mathcal{E}_{R,i}$ or $\mathcal{E}_{\pi,i}$ is greater than $-0.09091$ (although revenue or profit might increase by less than 10 percent).

To facilitate interpretation by providing a pseudo elasticity measure for revenue and profit analogous to the one given for cost in (2.4), consider the scale measure $\mathcal{E}_{\pi,i}$ defined in (2.6) (similar reasoning applies to the scale measure $\mathcal{E}_{R,i}$ defined in (2.5)). Suppose $\pi(y_i, w_i) > 0$ and $\pi(\delta y_i, w_i) > 0$, which is the most common (by far) scenario. Then (2.6) can be rearranged to define

$$\eta_{\pi,i} := (1 + \mathcal{E}_{\pi,i}) \delta = \pi(\delta y_i, w_i) / \pi(y_i, w_i).$$

Clearly, in this case $\mathcal{E}_{\pi,i} \geq -1$. Increasing output levels by a factor $\delta$ leads to an ($\eta_{\pi,i} \times 100$) percent change in profits. Hence, for $\delta = 1.1$, values $\eta_{\pi,i} > (=, <) 1.1$ indicate IRS (CRS, DRS). Moreover, increasing output levels by a factor $\delta > 1$ leads to an increase in profit whenever $\eta_{\pi,i} > 1$. Using similar reasoning, we define $\eta_{R,i} := (1 + \mathcal{E}_{R,i}) \delta = R(\delta y_i, w_i) / R(y_i, w_i)$, whose interpretation is analogous to $\eta_{\pi,i}$.

The next section describes the models we estimate to obtain the predicted values needed to estimate the returns-to-scale measures defined in (2.3)–(2.6). Subsequently, Section 4 explains how we estimate the models and make inferences.

3 Econometric Specification

To obtain estimates of the returns-to-scale measures $\mathcal{E}_{C,i}$, $\mathcal{E}_{R,i}$, and $\mathcal{E}_{\pi,i}$, we must specify versions of the cost function $C(y, w)$, revenue function $R(y, w)$, and profit function $\pi(y, w)$

---

3 See Appendix A for details about the interpretation of the returns-to-scale measures $\mathcal{E}_{R,i}$ and $\mathcal{E}_{\pi,i}$ when revenue or profit are negative.
for estimation. We define response and explanatory variables in the present section, and discuss our fully nonparametric estimation methods in Section 4.

Our specification of right-hand-side (RHS) explanatory variables closely follows Wheelock and Wilson (2012) and much of the banking literature. We define four inputs and five outputs that, with one exception (the measure of off-balance sheet output), are those used by Berger and Mester (2003). Specifically, we define the following output quantities: consumer loans ($Y_1$), real estate loans ($Y_2$), business and other loans ($Y_3$), securities ($Y_4$), and off-balance sheet items ($Y_5$) consisting of net non-interest income.\footnote{Of the commonly used measures of off-balance sheet output, net non-interest income is the most consistently measurable across banks and over time. However, as a net, rather than gross measure of income, it is potentially a biased measure of off-balance sheet output because losses would appear to reduce off-balance sheet output. Data that would permit calculation of a gross measure of non-interest income are not available. See Clark and Siems (2002) for discussion of alternative measures of off-balance sheet activity.}

We define three variable input quantities: purchased funds and core deposits, consisting of the sum of total time deposits, foreign deposits, federal funds purchased, demand notes, trading liabilities, other borrowed money, mortgage indebtedness and obligations under capitalized leases, and subordinated notes and debentures ($X_1$); labor services, measured by the number of full-time equivalent employees on payroll at the end of each quarter ($X_2$); and physical capital ($X_3$). The first input quantity, $X_1$, captures non-equity sources of investment funds for the bank.\footnote{Wheelock and Wilson (2012) treat core deposits (i.e., total deposits less time deposits of $\$100,000$ or more) and other funding liabilities as a separate inputs. Here, we combine them into a single input due to reporting differences in the FR Y-9C call reports for bank holding companies and the FFIEC call reports for commercial banks prior to 2001.}

We measure the corresponding prices ($W_1, \ldots, W_3$) by dividing total expenditure on the given input by its quantity. We include financial equity capital ($EQUITY$) as a quasi-fixed input, which controls somewhat for differences in risk across banks (see Berger and Mester (2003) for details).\footnote{We define $EQUITY$ as the sum of the book values of common and preferred stock, surplus, and retained earnings, which are items RCFD3210 and BHCK3210 from the FFIEC and FR Y-9C call reports, respectively.}

As an additional control for differences in bank risk, we also include a measure of non-performing assets ($NPER$) consisting of (i) total loans and lease financing receivables past due 30 days or more and still accruing, (ii) total loans and lease financing receivables not accruing, (iii) other real estate owned, and (iv) charge-offs on past-due loans and leases.\footnote{We thank a referee for pointing out that adding charge-offs to past-due and nonaccrual assets eliminates bias caused by differences in charge-off strategies across banks.} With the exception of labor input (which is mea-
sured as full-time equivalent employees) and off-balance sheet output (which is measured in terms of net flow of income), our inputs and outputs are stocks measured by dollar amounts reported on bank balance sheets, consistent with the widely used intermediation model of Sealey and Lindley (1977).

In addition to the variables defined above, we index quarters 1986.Q4 through 2015.Q4 by setting \( T = 1 \) for 1986.Q4, \( T = 2 \) for 1987.Q1, \ldots, \( T = 117 \) for 2015.Q4. Although \( T \) is an ordered, categorical variable, we treat it as continuous since it can assume a wide range of possible values. The regulatory environment and the production technology of banking changed a great deal over the 30 years covered by our data; including \( T \) as an explanatory covariate allows functional forms to change over time. Two features of our estimation strategy allow a great deal of flexibility. First, because we use a fully nonparametric estimation method, we impose no constraints on how \( T \) might interact with other explanatory variables. Second, the local nature of our estimator means that when we estimate cost at a particular point in time, observations from distant time periods will have little or no effect on the estimate. Typical approaches that involve estimation of a fully parametric translog cost functions by OLS or some other estimation procedure are not local in the sense that when cost is estimated at some point in the data space, all observations contribute to the estimate with equal weight. Moreover, the typical approach requires the imposition of a specific functional form a priori for any interactions among explanatory variables.\(^8\)

Turning to the response variables, we define our cost variable \( C \) as the sum of expenditures on purchased funds and core deposits, labor, and physical capital so that \( C := W_1X_1 + W_2X_2 + W_3X_3 \). We define our revenue variable, \( R \), similarly to Berger and Mester (2003); i.e., \( R := \text{total interest income} + \text{total non-interest income} + \text{realized gains (losses) on held-to-maturity securities} + \text{realized gains (losses) on available-for-sale securities} - \text{provision for loan and lease losses} - \text{provision for allocated transfer risk reserves} \). Finally, we measure profit (\( \pi \)) as the difference between revenue and cost; i.e., \( \pi := R - C \).

Our cost, revenue, and profit functions include as right-hand side (RHS) variables the vector \( \mathbf{y} := [Y_1 \ Y_2 \ Y_3 \ Y_4 \ Y_5] \) of output quantities defined above. When estimating the cost function, we also include \( \mathbf{w}_1 := \left[ \frac{W_2}{W_1} \quad \frac{W_3}{W_1} \quad T \quad \text{EQUITY} \quad \text{NPER} \right] \) with the price of purchased funds (\( W_1 \)) serving as the numeraire (we also divide cost on the left-hand side (LHS) by \( W_1 \)).

\(^8\) The local nature of our estimator is discussed in more detail below in Section 4 and in Appendix D.
to ensure homogeneity with respect to input prices). As discussed in Section 2, we do not
impose linear homogeneity when estimating the revenue and profit functions. Consequently,
we include on the RHS (in addition to \( y \)) \( w_2 := [W_1 \ W_2 \ W_3 \ T \ EQUITY \ NPER] \) when
estimating revenue and profit functions.

Our cost, revenue and profit functions are each of the form

\[
\mathcal{Y} = m(y, w) + \varepsilon \tag{3.1}
\]

where \( \mathcal{Y} \) represents one of our dependent variables (i.e., either \( C \), \( R \) or \( \pi \)), \( w \) represents
either \( w_1 \) or \( w_2 \) (depending on the LHS variable), and \( \varepsilon \) is a stochastic error term with \( E(\varepsilon \mid y, w) = 0 \) so that so that \( m(\cdot, \cdot) \) is a conditional mean function. In addition, we assume that
the densities of the continuous RHS variables are twice continuously differentiable at each
point where the conditional mean function is estimated, but otherwise make no functional
form assumptions regarding \( m(\cdot, \cdot) \). Consistency of our estimator requires that the dependent
variable \( \mathcal{Y} \) must be continuous at \((y, w)\) when the conditional mean function is estimated
at \((y, w)\), and that \( E(|\mathcal{Y}|^{2+\nu} \mid y, w) \) exists for some \( \nu > 0 \). One may view the conditional
mean functions as either parametric but of unknown form, or nonparametric (i.e., infinitely
parameterized). We provide details on estimation and inference below in Section 4 and in
Appendix D.

Given a set of RHS variables, our minimal assumptions on the response function \( m(y, w) \)
and inclusion of the time variable \( T \) allow far more flexibility than any parametric model.
In banking and other industry studies, it has become fashionable in recent years to spec-
ify parametric models that allow (to some degree) technological heterogeneity across firms
(examples include Orea and Kumbhakar, 2004 and Poghosyan and Kumbhakar, 2010). Al-
though we maintain an assumption of continuity, our nonparametric specification and local
estimation method means that \( m(y, w) \) can be quite different for different firms. In addi-
tion, the interaction of time \( T \) in the response function is left unspecified, allowing far more
flexibility than in typical parametric specifications.

We estimate the models using a dataset comprised of consolidated balance sheet and
income statement observations for all U.S. bank holding companies (BHCs) for 1986.Q3–
2015.Q4. We include in our dataset observations for commercial banks that are not owned
by holding companies. We use the seasonally adjusted, quarterly gross domestic product
Implicit price deflator to convert all dollar amounts to constant 2015 dollars.\textsuperscript{9}

Using data at the level of holding companies (where relevant) permits more accurate tallying of inputs and outputs than is possible at the level of individual commercial banks, for example by accounting for interbank transfers among subsidiaries of a single holding company, as well as expenses incurred at the holding company level. Moreover, our primary interest is in the largest institutions in the industry, and these are typically holding companies. After pooling data across 117 quarters and deleting observations with missing or implausible values, 847,299 observations remain for estimation. Summary statistics are provided in Tables B.2–B.6 of the separate Appendix B.

4 Details on Estimation and Inference

Various approaches exist for estimating conditional mean functions such as those in the models described above in Section 3. A common approach is to specify a fully parametric translog functional form for the conditional mean function and then estimate the parameters via least-squares methods. However, our data easily reject the translog specification using specification tests similar to those used by Wheelock and Wilson (2001, 2012); see Appendix C for details.

Rejection of the translog functional form is hardly surprising. The translog function is merely a quadratic in log-space, which limits the variety of shapes the conditional mean function is permitted to take. Further, the translog is derived from a Taylor expansion of the cost (or revenue, or profit) function around the means of the data; one should not expect it to fit well data that are highly variable and highly skewed, as is the case with U.S. banking data.\textsuperscript{10}

Several studies have noted that the parameters of a translog function are unlikely

\textsuperscript{9} BHC data are from Federal Reserve report FR Y-9C, which we downloaded from the website of the Federal Reserve Bank of Chicago. Data for independent commercial banks are from the Federal Financial Institutions Examination Council (FFIEC 031 and 041 reports). The reports record expenses and other flow variables (as opposed to stocks of deposits, etc.) from January 1 to the end of each quarter (March 31, June 30, September 30 and December 31). Hence for quarters 2, 3 or 4 of a given year, the previous quarter’s call report must be used to first-difference flow variables in order to obtain expenses for a particular quarter. Although we use data from the 1986.Q3 reports for this purpose, our final data represent quarters from 1986.Q4 through 2015.Q4.

\textsuperscript{10} The summary statistics for banks’ total assets given in Tables B.2–B.6 in Appendix B reveal that the distribution of banks’ sizes is heavily skewed to the right. In fact, estimates of Pearson’s moment coefficient of skewness for total assets in each of 117 quarters range from 27.49 to 49.03. Moreover, skewness is increasing over time, despite the consolidation in the industry over the years covered by our data. Regressing the
to be stable when the function is fit globally across units of widely varying size; see, for example, Guilkey et al. (1983) and Chalfant and Gallant (1985) for Monte Carlo evidence, and Cooper and McLaren (1996) and Banks et al. (1997) for empirical evidence involving consumer demand, Wilson and Carey (2004) for empirical evidence involving hospitals, and McAllister and McManus (1993), Mitchell and Onvural (1996), and Wheelock and Wilson (2001, 2012) for empirical evidence involving banks. Similarly, Hughes and Mester (2013, 2015) estimate a nonstandard profit function and input demand equations that allow banks to trade profits for reduced risk. Their system reduces to the translog form when parameter restrictions are consistent with profit maximization and cost minimization, but their tests of these restrictions reject the translog function, implying that banks trade profits for lower risk.

We use fully nonparametric methods to avoid likely specification errors. Although nonparametric methods are less efficient than parametric methods in a statistical sense when the true functional form is known, nonparametric estimation avoids the risk of specification error when the true functional form is unknown, as in the present application. We use local-linear estimators described by Fan and Gijbels (1996) to estimate our cost, revenue, and profit functions. Both the local-linear estimator as well as the Nadaraya-Watson kernel regression estimator (Nadaraya, 1964; Watson, 1964) are examples of local order-$p$ polynomial estimators with $p = 1$ and 0, respectively. For a locally-fit polynomial of order $p$ used to estimate a conditional mean function, going from an even value to an odd value of $p$ results in a reduction of bias with no increase in variance (e.g., see Fan and Gijbels, 1996 for discussion). Hence, we use a local-linear estimator to estimate conditional mean functions, resulting in lower asymptotic mean square error than one would obtain with the Nadaraya-Watson estimator.

Nonparametric regression models can be viewed as infinitely parameterized; as such, any parametric regression model (such as an assumed translog functional form) is nested within a nonparametric regression model. Clearly, adding more parameters to a parametric model affords greater flexibility. Nonparametric regression models represent the limiting outcome of adding parameters, and may be viewed as the most general encompassing model that a

---

skewness coefficients for each quarter on the time variable $T$ yields a positive estimate of the slope coefficient, $0.07302$ that is significantly different from zero at .1 significance.
particular parametric specification might be tested against.\footnote{Several methods for nonparametric regression exist. Cogent descriptions of nonparametric regression and the surrounding issues are given by Fan and Gijbels (1996, chapter 1), Härdle and Linton (1999), and Henderson and Parmeter (2015). Härdle and Mammen (1993) propose a test of a parametric regression against a nonparametric alternative where the test statistic is an estimate of the integrated square difference between the two regressions. Although we do not implement the Härdle and Mammen test in order to avoid computational expense, it seems almost certain the test would reject the translog parametric model in view of the results from our simple specification tests discussed above and in the separate Appendix C.}

Most nonparametric estimators suffer from the “curse of dimensionality,” i.e., convergence rates fall as the number of model dimensions increases. The convergence rate of our local-linear estimator is $n^{1/(4+d)}$ where $d$ is the number of distinct, continuous RHS variables, and there are $d = 10$ RHS variables in our cost function and $d = 11$ RHS variables in our revenue and profit functions. The slow convergence rate of our estimator means that for a given sample size, the order (in probability) of the estimation error we incur with our nonparametric estimator will be larger than the order of the estimation error one would achieve using a parametric estimator in a correctly specified model with the usual parametric rate $n^{1/2}$. However, our nonparametric estimation strategy avoids specification error that would likely render meaningless any results that might be obtained using a misspecified model. We adopt the view of Robinson (1988), who argues that parametric models are likely misspecified and should be viewed as root-$n$ inconsistent instead of root-$n$ consistent.\footnote{Convergence results for nonparametric estimators are often expressed in terms of order of convergence in probability. Briefly, for a sequence (in $n$) of estimators $\hat{\theta}_n$ of some scalar quantity $\theta$, we can write $\hat{\theta} = \theta + O_p(n^{-a})$ when $\hat{\theta}$ converges to $\theta$ at rate $n^a$, and we say that the estimation error is of order in probability $n^{-a}$. This means that the sequence of values $n^a|\hat{\theta}_n - \theta|$ is bounded in the limit (as $n \to \infty$) in probability. See Serfling (1980) or Simar and Wilson (2008) for additional discussion.}

To mitigate the curse of dimensionality in our application, we (i) use a large sample with more than 800,000 observations and (ii) employ a simple dimension-reduction method. Multicollinearity among regressors is often viewed as an annoyance, but here we use the multicollinearity in our data to reduce dimensionality, thereby increasing the convergence rate of our estimators and reducing estimation error. We do this by transforming the continuous RHS variables in each model to principal components space. Principal components are orthogonal, and eigensystem analysis can be used to determine the information content of each principal component. In each model we estimate, we sacrifice a small amount of information by using only the six principal components of the continuous RHS variables that correspond to the six largest eigenvalues, hence reducing the number of continuous...
RHS variables in our regressions from 10 or 11 to six. The six principal components account for 92.86 percent of the independent linear information among the RHS variables in our cost function, and 89.70 percent and 88.52 percent in the revenue and profit functions. The transformation to principal-components space can be inverted, and the interpretation of the estimators of the conditional mean functions in each model based on six principal components of the (continuous) RHS variables is straightforward because our estimator is fully nonparametric. Additional details about our principal components transformation and nonparametric estimation strategy are provided in Appendix D.

To implement the local-linear estimator we must select a bandwidth parameter to control the smoothing over the continuous dimensions in the data. We use least-squares cross-validation to optimize an adaptive, $\kappa$-nearest-neighbor bandwidth. In addition, we employ a spherically symmetric Epanechnikov kernel function. This means that when we estimate cost, revenue or profit at any fixed point of interest in the space of the RHS variables, only the $\kappa$ observations closest to that point can influence estimated cost, revenue or profit. In addition, among these $\kappa$ observations, the influence that a particular observation has on estimated cost, revenue or profit diminishes with distance from the point at which the response is being estimated. Our estimator is thus a local estimator, and is very different than typical, parametric, global estimation strategies (e.g., OLS, maximum likelihood, etc.) where all observations in the sample influence (with equal weight) estimation at any given point in the data space. Moreover, because we use nearest-neighbor bandwidths, our bandwidths automatically adapt to variation in the sparseness of data throughout the support of our RHS variables.

For statistical inference about our estimates of RTS, we use the wild bootstrap introduced by Härdle (1990) and Härdle and Mammen (1993), which allows us to avoid making specific distributional assumptions. We estimate confidence intervals using methods described in Wheelock and Wilson (2011, 2012). Although our estimators are asymptotically normal, the asymptotic distributions depend on unknown parameters; the bootstrap allows us to avoid the need to estimate these parameters, which would introduce additional noise.\textsuperscript{13}

\textsuperscript{13} Additional details about our inference methods are given in the separate Appendix D.
5 Empirical Results

We estimate the cost, revenue and profit specifications described in Section 3 using the methods described in Section 4 to obtain estimates of cost, revenue, and profit. We substitute these into the RTS measures defined in Section 2 to obtain for each bank $i$ estimates $\hat{\mathcal{E}}_{C,i}$, $\hat{\mathcal{E}}_{R,i}$ and $\hat{\mathcal{E}}_{\pi,i}$ as well as estimates $\hat{\eta}_{C,i}$, $\hat{\eta}_{R,i}$ and $\hat{\eta}_{\pi,i}$ of the pseudo elasticities defined in Section 2. We use the bootstrap methods discussed in Section 4 and described in Appendix D to make inference about the corresponding true values of the RTS measures and corresponding pseudo elasticities.

Table 1 provides an overview of results from our estimation for 1986.Q4, 1996.Q4, 2006.Q4 and 2015.Q4. The table reports the number of banks for which we reject CRS (at .05 significance) in favor of IRS or DRS, or for which we cannot reject CRS for cost, revenue and profit in each quartile of total assets in each period.\footnote{See Tables E.3 and E.5 in the separate Appendix E for similar counts at .1 and .01 significance levels.}

In each quartile, we find that a large majority of banks faced either CRS or IRS in cost in each period, even though the distribution of banks’ sizes in terms of total assets shifted rightward over time.\footnote{Figure B.1 in Appendix B shows kernel estimates of the density of total assets for 1986.Q4, 1996.Q4, 2006.Q4, and 2015.Q4, where the rightward shift is apparent.} Even among banks in the fourth quartile (the largest 25 percent of banks by assets), we reject CRS in favor of IRS for a substantial number of banks, and in favor of DRS for very few banks. Our results are thus similar to other recent studies finding that even many large banks operate under increasing returns to scale (e.g., Wheelock and Wilson, 2012; Hughes and Mester, 2013; Kovner et al., 2014 and Becalli et al., 2015).\footnote{We estimated returns to scale for two alternative cost specifications, including one that treats physical capital as quasi-fixed, rather than as a variable input, and one that uses total cost, rather than the sum of expenditures on the variable inputs, as the dependent variable. Results of those models, which are reported in the separate Appendix E, are qualitatively very similar to those reported in Table 1.}

With regard to revenue economies, we are unable to reject CRS for a majority of banks in each quartile and period. However, we reject CRS in favor of IRS for more banks than we reject in favor of DRS.\footnote{Results are qualitatively very similar for an alternative specification, reported in the separate Appendix E, in which we define revenue as total unadjusted revenue.} Similarly, for profit economies, we also fail to reject CRS for more banks than not, but generally we reject CRS in favor of IRS for more banks than we reject in favor of DRS. The two exceptions are for banks in the largest-size quartile in 2006.Q4 and 2015.Q4, where we reject in favor of DRS for 420 and 378 banks, respectively, but in favor
of IRS for only 235 and 206 banks.\footnote{As with cost and revenue economies, results are robust to alternative measures of profit, constructed from different measures of cost and revenue as described in the preceding footnotes. See the separate Appendix E for specific results.}

Because the distribution of bank asset sizes is quite skewed, the largest asset-size quartile represents a much larger range than the other three quartiles. Further, because much of the interest in economies of scale pertains to the very largest banks, we report estimates of returns to scale for the 10 largest banks in each period. Specifically, Tables 2–3 report estimates of the pseudo elasticities defined in Section 2 for each of the 10 largest banks in each period. For the cost model, pseudo elasticity estimates that are significantly less than 1.1 indicates IRS, whereas for the revenue and profit models, estimates that are significantly greater than 1.1 indicate IRS.

For cost economies, the results in Tables 2–3 indicate that we reject CRS in favor of IRS in nearly every case (34 estimates out of 40) among the 10 largest banks in each period. Moreover, we reject CRS in favor of IRS for each of the four largest banks in 1986.Q4, 1996.Q4, and 2015.Q4 (and two of the four largest in 2006.Q4). In no case do we reject CRS in favor of DRS. Thus, the evidence suggests strongly that even the very largest U.S. banks faced increasing returns to scale throughout the sample period.\footnote{A referee suggested that there might be a break in RTS around $50 billion of assets due to regulatory and enforcement differences for banks beyond that threshold. Inspection of pseudo elasticity estimates for the largest 100 banks in each quarter, which are reported in Tables E.6–E.9 in the separate Appendix E, reveal no obvious break at $50 billion of assets.}

In contrast with the substantial evidence that the very largest banks face IRS in cost, the estimates shown in Tables 2–3 indicate that the largest banks mostly face DRS in revenue. We reject CRS in favor of DRS in revenue in 33 of 40 cases, and reject in favor of IRS in only two cases (Citigroup and Wells Fargo in 2015.Q4). However, all but one (Wells Fargo in 2006.Q4) of the revenue pseudo elasticity estimates are greater than 1.0, indicating that even though the largest banks face DRS in revenue, their revenues would still rise with an increase in output levels (though by a less than proportionate amount).

Finally, the estimates of returns to scale in profit reveal the relative importance of cost and revenue economies in determining profit economies for each bank in each period. For 1986.Q4, only two of the profit pseudo elasticity estimates in Table 2 are significantly different from 1.1, and in both cases CRS is rejected in favor of DRS. However, in 1996.Q4, nine
estimates are significantly less than 1.1, indicating DRS. Nonetheless, these nine estimates are greater than 1.0, indicating that profits increase with size for these banks, albeit by a less than proportionate amount.

Whereas we find no indication that any of the 10 largest banks faced IRS in revenue or profit in either 1986.Q4 or 1996.Q4, results reported in Table 3 indicate that the three largest banks (JPMorgan Chase, Bank of America, and Citigroup) operated under IRS in profit in 2006.Q4, and all four of the very largest banks did so in 2015.Q4. Further, the results indicate that both Citigroup and Wells Fargo also faced IRS in revenue in 2015.Q4. By contrast, among the remaining banks in the top 10, we either fail to reject CRS, or reject in favor of DRS for both revenue and profit in both 2006.Q4 and 2015.Q4. Nonetheless, the pseudo elasticity estimates for 2015.Q4 are all greater than 1.0 for these banks, indicating that profit would increase with an increase in output (but by a less-than proportionate amount).

Evidence on the extent to which changes in RTS over time were statistically significant is shown in Table 4. Specifically, for the 10 largest banks in 2015.Q4 that were also in existence in 2006.Q4, we report whether the change in the bank’s pseudo elasticity between those two periods was statistically significant (at the .05 level) and, if so, the direction of the change. Upward arrows indicate significant changes in the direction of greater returns to scale, whereas downward arrows indicate statistically significant decline in returns to scale, and the absence of an arrow indicates that the change in pseudo elasticity between the two periods is not statistically significant. As the table shows, we find statistically significant gains in RTS in terms of cost for seven banks, and a significant decline for only one bank. All of the banks that experienced a significant increase in RTS except Citigroup and JPMorgan Chase, for which we do not reject CRS in 2006.Q4, already faced IRS in 2006.Q4.

Among banks that experienced a significant change in RTS in revenue, as reflected by a statistically significant change in pseudo elasticity, four experienced a significant increase in RTS, while two had a significant decline. Among the banks that experienced significant gains, both Citigroup and Wells Fargo went from facing DRS in 2006.Q4 to IRS in 2015.Q4.

Finally, for profit, three banks, including two of the top four, experienced significant gains in RTS and one had a significant decline, while the change for six banks was not statistically significant. Among the largest four banks in 2015.Q4, the pseudo elasticity estimates in
Table 3 indicate that Citigroup, Bank of America, and JPMorgan Chase already faced IRS in 2006.Q4, whereas we are unable to reject CRS for Wells Fargo, which was the smallest of the four banks in that period. Both Citigroup and Wells Fargo experienced statistically significant gains in RTS between 2006.Q4 and 2015.Q4, and in the latter period we reject CRS in favor of IRS for Wells Fargo.

On the whole, our results are consistent with earlier studies finding that even the largest U.S. banks face IRS in terms of cost. Our findings indicate that this remains true some eight years after the financial crisis and after substantial changes in bank regulation. Further, our results indicate that while many banks face IRS in cost, many fewer operate under IRS in revenue or profit. However, substantial numbers of banks, including the four largest U.S. banks, do appear to face CRS or IRS in revenue and profit. Further, we find that, if anything, the very largest banks faced greater returns to scale in terms of revenue and profit in 2015 than they had in 2006, before the financial crisis and introduction of a new regulatory regime.²⁰

6 Conclusions

As the number of banks has declined since 1986, many banks have grown considerably in size. Despite the growth in bank size, we find considerable evidence that the largest U.S. banks continue to operate under increasing returns to scale in terms of cost, as they did in 2006 and even earlier. It is perhaps not surprising that large banks faced increasing returns in earlier years, given that institutions grew larger, but it is interesting that even in 2015, the largest institutions had not exhausted scale economies in terms of cost.

The evidence for returns to scale in revenue and profit is more mixed. Still, our estimates suggest that relatively few banks with total assets below the largest 25 percent face decreasing returns to scale, while the rest face constant or increasing returns. We find that somewhat

²⁰ In the separate Appendix E, Table E.12 gives counts of firms that experienced a statistically significant change in RTS between 2006.Q4 and 2015.Q4. In addition, Tables E.16–E.24 give transition matrices showing the numbers of institutions facing IRS, CRS, or DRS in 2006.Q4 versus 2015.Q4. The number of banks appearing in our sample in both 2006.Q4 and 2015.Q4 is 4,148. For the cost model described in Section 3, at .05 significance, there are 3,064 significant changes, with 1,686 gains and 1,378 declines in RTS. For the revenue model, there are 2,194 significant changes, with 1,033 increases and 1,161 decreases in RTS. For the profit model, there are 1,210 significant changes, with 493 increases and 717 decreases in RTS. A majority of significant changes in cost RTS are gains, while the numbers of changes in revenue RTS are almost even between increases and decreases, and changes in profit RTS are more often downward.
fewer banks in the largest size quartile operate under increasing returns. Among the largest 10 banks, we find that some operated under increasing returns in revenue and, especially, profit in 2006 and 2015, but others faced constant or decreasing returns. In particular, the largest four U.S. banks—all of which are substantially larger than the next largest banks—faced increasing returns to scale in profit, as well as in cost, in 2015. Thus, it appears that the turmoil of 2007-08 and subsequent changes in regulation have not lessened returns to scale in terms of cost, revenue or profit for most U.S. banks. And, if anything, the largest four banks have seen significant increases in returns to scale since 2006, suggesting that scale economies still provide an impetus to become even larger.
References


<table>
<thead>
<tr>
<th>LHS</th>
<th>Period</th>
<th>1st quartile</th>
<th>2nd quartile</th>
<th>3rd quartile</th>
<th>4th quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IRS</td>
<td>CRS</td>
<td>DRS</td>
<td>IRS</td>
</tr>
<tr>
<td>Cost</td>
<td>1986.Q4</td>
<td>1138</td>
<td>1439</td>
<td>5</td>
<td>734</td>
</tr>
<tr>
<td></td>
<td>1996.Q4</td>
<td>669</td>
<td>1211</td>
<td>18</td>
<td>524</td>
</tr>
<tr>
<td></td>
<td>2006.Q4</td>
<td>630</td>
<td>932</td>
<td>9</td>
<td>526</td>
</tr>
<tr>
<td></td>
<td>2015.Q4</td>
<td>528</td>
<td>662</td>
<td>20</td>
<td>481</td>
</tr>
<tr>
<td>Revenue</td>
<td>1986.Q4</td>
<td>287</td>
<td>2142</td>
<td>153</td>
<td>333</td>
</tr>
<tr>
<td></td>
<td>1996.Q4</td>
<td>245</td>
<td>1549</td>
<td>104</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>2006.Q4</td>
<td>201</td>
<td>1226</td>
<td>144</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>2015.Q4</td>
<td>204</td>
<td>822</td>
<td>184</td>
<td>229</td>
</tr>
<tr>
<td>Profit</td>
<td>1986.Q4</td>
<td>637</td>
<td>1914</td>
<td>31</td>
<td>463</td>
</tr>
<tr>
<td></td>
<td>1996.Q4</td>
<td>620</td>
<td>1252</td>
<td>26</td>
<td>444</td>
</tr>
<tr>
<td></td>
<td>2006.Q4</td>
<td>528</td>
<td>1024</td>
<td>19</td>
<td>416</td>
</tr>
<tr>
<td></td>
<td>2015.Q4</td>
<td>327</td>
<td>827</td>
<td>56</td>
<td>269</td>
</tr>
</tbody>
</table>
Table 2: Returns to Scale for Largest Banks by Total Assets, 1986.Q4 and 1996.Q4

<table>
<thead>
<tr>
<th>Name</th>
<th>Assets</th>
<th>Cost</th>
<th>Revenue</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>—1986.Q4—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CITIBANK</td>
<td>275</td>
<td>1.0371***</td>
<td>1.0147***</td>
<td>0.9834***</td>
</tr>
<tr>
<td>BK OF AMER</td>
<td>204</td>
<td>1.0928**</td>
<td>1.0558***</td>
<td>1.1622</td>
</tr>
<tr>
<td>CHASE MHTN BK</td>
<td>150</td>
<td>1.0618***</td>
<td>1.0820</td>
<td>1.0616</td>
</tr>
<tr>
<td>MANU. HAN</td>
<td>139</td>
<td>1.0855***</td>
<td>1.0686***</td>
<td>1.1819</td>
</tr>
<tr>
<td>MORGAN GNTY TC</td>
<td>130</td>
<td>1.0306***</td>
<td>1.0381***</td>
<td>1.0603</td>
</tr>
<tr>
<td>SECURITY PACIFIC</td>
<td>113</td>
<td>1.0982</td>
<td>1.0593***</td>
<td>1.0774</td>
</tr>
<tr>
<td>CHEMICAL NY</td>
<td>109</td>
<td>1.0923***</td>
<td>1.0574***</td>
<td>1.0995</td>
</tr>
<tr>
<td>BANKERS TR NY</td>
<td>100</td>
<td>1.0461***</td>
<td>1.0354***</td>
<td>1.0542***</td>
</tr>
<tr>
<td>FIRST INTRST BC</td>
<td>100</td>
<td>1.0883*</td>
<td>1.0705***</td>
<td>1.0586</td>
</tr>
<tr>
<td>WELLS FARGO &amp; CO</td>
<td>81</td>
<td>1.0897</td>
<td>1.0488***</td>
<td>1.0622</td>
</tr>
<tr>
<td>—1996.Q4—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHASE MHTN</td>
<td>469</td>
<td>1.0557***</td>
<td>1.0334***</td>
<td>1.0401***</td>
</tr>
<tr>
<td>CITICORP</td>
<td>394</td>
<td>1.0368***</td>
<td>1.0376***</td>
<td>1.0426***</td>
</tr>
<tr>
<td>BK OF AMER</td>
<td>352</td>
<td>1.0412***</td>
<td>1.0469***</td>
<td>1.0485***</td>
</tr>
<tr>
<td>NATIONSBANK</td>
<td>266</td>
<td>1.0773***</td>
<td>1.0442***</td>
<td>1.0395***</td>
</tr>
<tr>
<td>MORGAN GNTY TC</td>
<td>245</td>
<td>1.0539***</td>
<td>1.0126***</td>
<td>0.9942***</td>
</tr>
<tr>
<td>FIRST UNION</td>
<td>195</td>
<td>1.0907**</td>
<td>1.0201***</td>
<td>0.9906***</td>
</tr>
<tr>
<td>WELLS FARGO &amp; CO</td>
<td>155</td>
<td>1.0630***</td>
<td>1.0839***</td>
<td>1.1083</td>
</tr>
<tr>
<td>FIRST NBD</td>
<td>150</td>
<td>1.0663***</td>
<td>1.0392***</td>
<td>1.0440***</td>
</tr>
<tr>
<td>BANC ONE</td>
<td>143</td>
<td>1.0653***</td>
<td>1.0594***</td>
<td>1.0600***</td>
</tr>
<tr>
<td>FLEET FNCL GROUP</td>
<td>123</td>
<td>1.0587***</td>
<td>1.0817***</td>
<td>1.0781***</td>
</tr>
</tbody>
</table>

**NOTE:** For cost model, estimates of \((1 - \varepsilon_{C,i}) \delta\) are reported (\(\delta = 1.1\)). For revenue and profit models, estimates of \((1 + \varepsilon_{R,i}) \delta\) and \((1 + \varepsilon_{\pi,i}) \delta\) are given. For cost model, values **less than** 1.1 indicate increasing returns to scale, while for revenue and profit models, values **greater than** 1.1 indicate increasing returns to scale. Statistical significance (difference from 1.1) at the ten, five, or one percent levels is denoted by one, two, or three asterisks, respectively. Assets are given in millions of constant 2015 dollars.
Table 3: Returns to Scale for Largest Banks by Total Assets, 2006.Q4 and 2015.Q4

<table>
<thead>
<tr>
<th>Name</th>
<th>Assets</th>
<th>Cost</th>
<th>Revenue</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2006.Q4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CITIGROUP</td>
<td>2082</td>
<td>1.101</td>
<td>1.0808</td>
<td>1.1355***</td>
</tr>
<tr>
<td>BK OF AMER</td>
<td>1672</td>
<td>1.039</td>
<td>1.0930</td>
<td>1.1519***</td>
</tr>
<tr>
<td>JPMORGAN CHASE &amp; CO</td>
<td>1543</td>
<td>1.102</td>
<td>1.0982</td>
<td>1.2066***</td>
</tr>
<tr>
<td>WACHOVIA</td>
<td>726</td>
<td>1.046</td>
<td>1.0229</td>
<td>1.0663</td>
</tr>
<tr>
<td>WELLS FARGO &amp; CO</td>
<td>554</td>
<td>1.019</td>
<td>0.9996</td>
<td>1.0134***</td>
</tr>
<tr>
<td>U S BC</td>
<td>250</td>
<td>1.058</td>
<td>1.0493</td>
<td>1.0656***</td>
</tr>
<tr>
<td>COUNTRYWIDE</td>
<td>225</td>
<td>1.101</td>
<td>1.0013</td>
<td>0.9950***</td>
</tr>
<tr>
<td>SUNTRUST BK</td>
<td>210</td>
<td>1.073</td>
<td>1.0757</td>
<td>1.0865</td>
</tr>
<tr>
<td>HSBC BK USA</td>
<td>191</td>
<td>1.046</td>
<td>1.0096</td>
<td>0.9928***</td>
</tr>
<tr>
<td>NATIONAL CITY</td>
<td>160</td>
<td>1.069</td>
<td>1.0178</td>
<td>0.9778***</td>
</tr>
<tr>
<td><strong>2015.Q4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPMORGAN CHASE &amp; CO</td>
<td>2378</td>
<td>1.015</td>
<td>1.1007</td>
<td>1.1249***</td>
</tr>
<tr>
<td>BK OF AMER</td>
<td>2145</td>
<td>1.014</td>
<td>1.1030</td>
<td>1.1592***</td>
</tr>
<tr>
<td>CITIGROUP</td>
<td>1765</td>
<td>1.037</td>
<td>1.1337</td>
<td>1.1842***</td>
</tr>
<tr>
<td>WELLS FARGO &amp; CO</td>
<td>1764</td>
<td>1.034</td>
<td>1.1170</td>
<td>1.1180***</td>
</tr>
<tr>
<td>U S BC</td>
<td>418</td>
<td>0.965</td>
<td>1.0091</td>
<td>1.0811***</td>
</tr>
<tr>
<td>BK OF NY MELLON</td>
<td>384</td>
<td>1.069</td>
<td>1.0451</td>
<td>1.0348***</td>
</tr>
<tr>
<td>PNC FNCL SVC GROUP</td>
<td>359</td>
<td>0.963</td>
<td>1.0168</td>
<td>1.0283***</td>
</tr>
<tr>
<td>STATE STREET</td>
<td>246</td>
<td>1.156</td>
<td>1.0042</td>
<td>1.0117***</td>
</tr>
<tr>
<td>T D BK</td>
<td>243</td>
<td>1.052</td>
<td>1.0666</td>
<td>1.0717</td>
</tr>
<tr>
<td>BB&amp;T</td>
<td>209</td>
<td>1.048</td>
<td>1.0795</td>
<td>1.1012</td>
</tr>
</tbody>
</table>

**NOTE:** For cost model, estimates of \((1 - \mathcal{E}_{C,i})\delta\) are reported \((\delta = 1.1)\). For revenue and profit models, estimates of \((1 + \mathcal{E}_{R,i})\delta\) and \((1 + \mathcal{E}_{\pi,i})\delta\) are given. For cost model, values less than 1.1 indicate increasing returns to scale, while for revenue and profit models, values greater than 1.1 indicate increasing returns to scale. Statistical significance (difference from 1.1) at the ten, five, or one percent levels is denoted by one, two, or three asterisks, respectively. Assets are given in millions of constant 2015 dollars.
### Table 4: Significant Changes in RTS from 2006.Q4 to 2015.Q4 for 10 Largest Banks in 2015.Q4 (.05 Significance)

<table>
<thead>
<tr>
<th>Bank</th>
<th>Model Cost</th>
<th>Model Revenue</th>
<th>Model Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPMORGAN CHASE &amp; CO</td>
<td>↑</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>BK OF AMER</td>
<td>↑</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CITIGROUP</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>WELLS FARGO &amp; CO</td>
<td>—</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>U S BC</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>PNC FNCL SVC GROUP</td>
<td>↑</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>STATE STREET</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>T D BK</td>
<td>↑</td>
<td>↓</td>
<td>—</td>
</tr>
<tr>
<td>BB&amp;T</td>
<td>↑</td>
<td>↑</td>
<td>—</td>
</tr>
<tr>
<td>SUNTRUST BK</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

**NOTE:** Upward arrows indicate a significant increase in RTS pseudo-elasticity from 2006.Q4 to 2015.Q4. Downward arrows indicate significant decrease in RTS pseudo-elasticity from 2006.Q4 to 2015.Q4. Horizontal dashes indicate no significant change.