

The Cost Consequences of Breaking up Large Banks: Do Large Banks Enjoy Technological Cost Advantages?

Joseph P Hughes*

Department of Economics, Rutgers University, New Brunswick, NJ, USA

Abstract

Proposals to break up the largest banks seek to reduce the systemic risk they impose on the economy. However, if these banks experience scale economies that reduce the average cost of their financial products and services, breaking them into smaller institutions might reduce their ability to compete in global markets and provide them with incentives to evade break up and operate outside the regulated financial system - with the potential for new sources of systemic risk. Textbooks assert that large scale is associated with such cost economies, but the evidence for these economies is difficult to obtain. Is such evidence illusive or elusive? This paper explores some of the published evidence and the reasons why it is elusive, not illusive.

Keywords: Banking; Production; Risk; Scale economies; Too big to fail

JEL Codes: D20, D21, G21, L23

The Policy Debate

Fisher [1], the president of the Federal Reserve Bank of Dallas, has asserted, "Hordes of Dodd-Frank regulators are not the solution; smaller, less complex banks are. We can select the road to enhanced financial efficiency by breaking up TBTF banks - now." Bair [2], the former chairman of the Federal Deposit Insurance Corporation, expanded the assertion: "The public policy benefits of smaller, simpler banks are clear. It may be in the enlightened self-interest of shareholders as well." And Purcell [3], the former chief executive of Morgan Stanley, made the same point more emphatically: "Breaking these companies into separate businesses would double to triple the shareholder value of each institution." Each of these assertions suggests that the collection of smaller banks that would replace the largest financial institutions would operate more efficiently in some sense. However Tarullo [4], a governor of the Federal Reserve System, has warned that there may be a trade-off between systemic risk and financial efficiency: "An additional concern would arise if some countries made the trade-off by limiting the size or configuration of their financial firms for systemic risk reasons at the cost of realizing genuine economies of scope or scale, while other countries did not. In this case, firms from the first group of countries might well be at a competitive disadvantage in the provision of certain cross-border activities."

Mester [5] considers the policy implications of this competitive disadvantage: "... if policymakers do conclude that the costs of size outweigh the benefits, the existence of scale economies suggests that a strict size limit on banks is not likely to be an effective solution. Such limits work against market forces and do not align incentives. Given the potential benefits of size, strict limits would create incentives for firms to avoid these restrictions, and could thereby push risk-taking outside of the regulated financial sector, without necessarily reducing systemic risk."

Are Scale Economies at the Largest Banks Illusive or Elusive? The Textbook Case

Governor Tarullo [4] asks if such economies are genuine: "Generally, though, even where intuition suggests economies in some other areas - such as the breadth of securities distribution networks and the ability to provide all forms of financing in significant amounts - evidence for the existence of such economies is limited and mixed. Moreover, even where significant scale is necessary to achieve certain

economies, an important question will be what the minimum efficient scale - or, perhaps more realistically, the minimum feasible scale - actually is. It is possible that a firm would need to be quite large and diversified to achieve these economies, but still not as large and diversified as some of today's firms have become." He poses two fundamental questions: (1) are scale economies elusive or illusive and (2) if such economies exist, can they be achieved by institutions smaller than the ones we observe today? Former Federal Reserve Chairman Alan Greenspan has answered the first question negatively: "For years the Federal Reserve had been concerned about the ever larger size of our financial institutions. Federal Reserve research had been unable to find economies of scale in banking beyond a modest-sized institution." However, textbooks answer positively, and a growing literature based on new estimation techniques is finding evidence of economically significant scale economies even at the largest financial institutions. Studies using new techniques that find such confirming evidence include Hughes, et al. [6,7], Berger, et al. [8], Hughes, et al. [9,10], Hughes, et al. [11], Bossone, et al. [12], Wheelock, et al. [13], Feng, et al. [14], and Dijkstra [15]. Parenthetically, several of these authors are former or current researchers in the Federal Reserve System.

Economies of scale are defined by cost that increases less than proportionately with output so that the average cost of output declines. In the case of constant returns to scale, cost increases proportionately with output, and in the case of diseconomies of scale, more than proportionately so average cost increases with output. Textbooks often cite the spreading of overhead expenses, such as those associated with information technology, and network economies related to the payments system, such as on-us check clearing, as important sources of economies of scale.

Textbooks also contend that as banks grow in size, their larger base of deposits usually becomes more diversified. If reserves and other liquid

*Corresponding author: Joseph P. Hughes., Department of Economics, Rutgers University, New Brunswick, NJ 08901-1248, USA, Tel: (917) 721-0910; E-mail: jpHughes.@rci.rutgers.edu

Received September 16, 2013; Accepted November 18, 2013; Published November 20, 2013

Citation: Hughes JP (2013) The Cost Consequences of Breaking up Large Banks: Do Large Banks Enjoy Technological Cost Advantages? J Stock Forex Trad 3: 113. doi:10.4172/2168-9458.1000113

Copyright: © 2013 Hughes JP. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

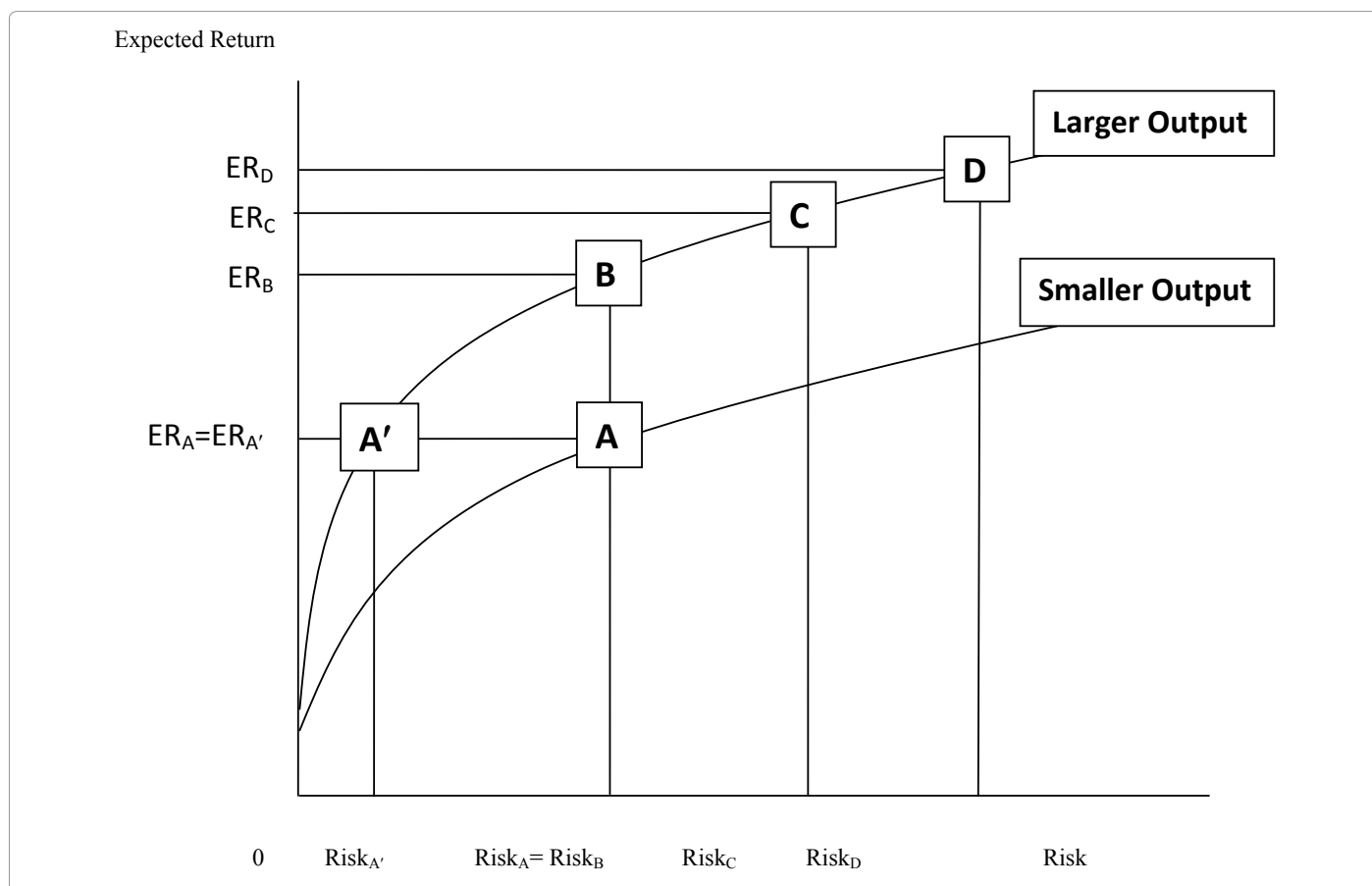
assets held to protect against liquidity risk increase proportionately with assets, liquidity risk is reduced by the better diversification. Alternatively, the better diversified bank can increase its reserves and liquid assets less than proportionately without increasing liquidity risk. Freeing some amount of liquid assets by a less than proportionate increase allows the bank to make more profitable investments in loans and other assets. And, the better diversification reduces the average cost of risk management.

Another contention of textbooks focuses on assets: as banks' assets increase, their loan portfolio becomes better diversified. If equity capital, which absorbs loan losses and protects the bank from insolvency, increases proportionately with assets, better diversification implies that insolvency risk will decrease. Alternatively, the bank can increase its equity capital less than proportionately and, thanks to better diversification, maintain the same insolvency risk. The less than proportionate increase in equity allows the bank to increase its return on equity. Moreover, the better diversified loan portfolio reduces the average cost of risk management.

Better diversification of larger banks implies they experience a better risk-expected-return trade-off than smaller, less diversified banks. In Figure 1, assume the smaller bank operates on the lower

frontier with the investment strategy at point A. The larger bank's better diversification of loans and deposits gives it a higher frontier. One of its investment strategies is to continue to operate with the same ratio of reserves and liquid assets to total deposits and the same ratio of equity to assets - point A' represents the same expected return as the smaller bank but less risk due to better diversification. As noted above, another investment strategy involves reducing the ratio of reserves and liquid assets to deposits and the ratio of equity to assets to increase the expected return while maintaining the same risk as the smaller, less diversified bank - point B.

In general, though, larger banks may be expected to take more risk for even higher expected return by adopting investment strategies to the right of point B at such points as C and D. The incentive to take more risk results in part from lower marginal and average costs of risk management and from the cost-of-funds subsidy of mispriced explicit and implicit deposit insurance. The largest banks especially are thought to benefit from too-big-to-fail policies that lower the cost of their uninsured borrowed funds. While smaller banks may pursue more conservative investment strategies to protect their charter made valuable by growth opportunities and market power, larger banks, operating in more competitive markets with less valuable growth



Source: Hughes and Mester (forthcoming)

Figure 1: Risk-Expected-Return Frontiers for Smaller and Larger, Better Diversified Banks. Larger banks whose size improves their diversification experience a better risk-expected-return frontier than smaller banks. Suppose smaller banks produce with the investment strategy at point A. If larger banks adopt the same investment strategy, their better diversification reduces the strategy's risk and is represented by point A'. The increase in cost from A to A' is proportionately less than the increase in outputs due to better diversification, which implies economies of scale characterize production. However, this cost-saving effect of better diversification may be obscured when larger banks take more risk, say the strategies at points B, C, and D. When the extra risk is costly to take, the increase in cost from A to B may appear proportionate to the increase in output - constant returns to scale - and from A to C or D as more than proportionate - diseconomies of scale. The appearance of proportionate or more than proportionate increase in costs obscures the underlying scale economies due to better diversification that improve banks' risk-expected-return trade-off.

opportunities and benefiting from federal safety-net subsidies, tend to adopt more aggressive investment strategies¹. To the extent that additional risk-taking involves additional cost, these additional costs may bias estimates of scale economies - the response of cost to a proportional increase in outputs - and obscure the cost-saving effects of better diversification.

Are Scale Economies Illusive or Elusive? Disentangling the Effects of Risk-Taking on Cost

Consider two output vectors, one belonging to the smaller bank and one, to the larger bank. Let the larger bank's output vector simply be the smaller bank's mix of outputs where the quantity of each output is increased equi-proportionately. In addition, let's interpret the risk-expected-return frontier of the smaller bank in Figure 1 as the risk-return menu of this given vector of outputs and the frontier of the larger bank as the menu for the larger bank's mix of outputs scaled up equi-proportionately in quantity. Scale economies, then, are defined by how cost varies as we proportionately increase the outputs of the smaller bank - in this case, from the lower frontier to the higher frontier. The cost of the larger output on the higher frontier depends on the risk-return strategy that is adopted. If additional risk-taking is costly, then the cost of the larger output increases from A' to B to C to D. At point A' where the larger output is produced with the investment strategy of the smaller output at A, *but with better diversification*, cost increases *less than proportionately* with output, as the textbooks explain. However, if the larger output is produced with a more risky, costly investment strategy, the measure of scale economies will be reduced. For example, the reduced measure may yield an equi-proportional increase in cost at point B - *constant returns to scale* - and at points C and D, a more than proportional increase in cost - *diseconomies of scale*. The technology evolves from various sources of scale economies which improve the risk-expected-return frontier as outputs are increased. However, *detecting those scale economies requires controlling for risk-taking* - the investment decision along the improved frontier. If larger banks on average adopt the strategy around point A', the estimation of cost will yield scale economies. However, if larger banks adopt more risky strategies such as those around B, C, and D and the estimation technique does not control for endogenous risk-taking, the estimation will likely mistakenly find constant returns to scale or diseconomies of scale, the commonly found result that Greenspan [16] described. Thus, the estimation fails to detect the underlying economies of scale that characterize banking technologies.

Finding evidence of diversification also involves disentangling the effects of endogenous risk-taking on a bank's risk exposure. A naïve interpretation of the textbook's assertion that larger banks are better diversified would characterize larger banks to be on average less risky than smaller banks - in effect, to be operating at point A' in Figure 1. However, if larger banks tend to adopt more risky strategies such as those at points C and D, the extra risk of these strategies will obscure the inherent reduction in risk due to better diversification observed

¹ Banks must obtain a charter to operate. Marcus [18] shows that restrictions on entry into banking and safety-net subsidies create dichotomous value-maximizing investment strategies. Valuable investment opportunities make episodes of financial distress especially costly: liquidity crises may arise, regulators may intervene, and the charter may be revoked and investment opportunities lost. Banks with valuable investment opportunities pursue relatively safe strategies to protect their charters (and ability to pursue their investment opportunities) from financial distress. Banks with less valuable opportunities adopt more risky strategies to exploit cost-of-funds subsidies of explicit and implicit deposit insurance. Mid-range risk strategies do not maximize value. Keeley [21], McConnell, et al. [22], and Hughes, et al. [23] find evidence of these dichotomous strategies. Using stochastic frontier estimation, the latter paper and Hughes and Mester [24] find that larger banks have less valuable investment opportunities.

directly at point A'. That is to say, the larger bank's better diversification improves its risk-expected-return trade-off. It does not necessarily mean the larger bank will be less risky. Demsetz, et al. [17] demonstrate this point by using estimates of firm-specific risk of commercial banks from several asset pricing models and regressing these measures of risk on bank size and then on bank size and a number of controls for bank risk-taking. When they do not control for risk-taking, they find only a small negative relationship with asset size, but when they control for endogenous risk-taking, they find a much more negative relationship with size. A similar research strategy much be adopted to disentangle the effects of risk-taking on the relationship between cost and output.

Estimating Scale Economies with Controls for Endogenous Risk-Taking

In terms of Figure 1, the theory underlying the standard cost function would identify a minimum cost of producing the smaller output on the lower frontier and the larger output on the higher frontier, but the risk-expected-return choice on the frontier does not enter into the choice of inputs or the resulting minimum cost. Thus, the theoretical definition of the standard minimum cost function does not account for endogenous risk-taking. The econometric estimation of this cost function would consider cost as a function of the quantities of outputs, the quantity of equity capital, and the prices of the variable inputs without controls for endogenous risk-taking. In failing to control for endogenous risk-taking, it would identify the increase in cost from the smaller output at point A to the larger output at point C or D as more than proportional to the increase in outputs - *diseconomies of scale*.

Hughes, et al. [6,7] propose a technique for estimating a cost function that controls for endogenous risk-taking and disentangles its effect on cost from the underlying potential scale economies of the banking technology. With such controls, the effect on cost of an equi-proportionate increase in the quantities of all outputs from the smaller output vector on the lower frontier at point A to the larger output vector on the higher frontier would be gauged from the cost of the risk-return strategy at point A to that at A', not that at C or D. Recall that A' results when the increase in outputs is effected holding constant the ratio of liquid assets to total assets and the ratio of equity capital to total assets; hence, the expected return is held constant and better diversification reduces liquidity risk and insolvency (credit) risk. By controlling for the endogenous choice of risk and expected return at A, it captures the diversification of a larger output that puts the bank at A', given its initial choice of A. Hughes, et al.[10] refer to the effect on cost of increasing output in the neighborhood of point A and holding sources of risk-taking constant -- a move from point A to A' -- as the *diversification effect*, and the effect on cost of taking on additional risk -- the move from point A' to C or D -- as the *risk-taking effect*. Endogenous risk-taking moves the bank from point A to point C or D as scale increases, and cost increases more than proportionately with scale if the extra risk is costly. Thus, to uncover the effects of better diversification on risk and cost and to detect correctly the underlying scale economies, the estimation of banking cost must control for this risk-taking. Controlling for endogenous risk-taking requires modeling production decisions where risk influences managers' ranking of production plans.

The standard cost function characterizes banking technology by giving the minimum cost associated with any given vector of outputs, y , quantity of equity capital, k , and prices, w , of variable inputs, x . Technology defines feasible production plans, (y, x, k) , as those where the output vector can be produced with the input vector and equity

capital. For any given state of the world, s , managers associate an amount of realized profit, π , with each production plan. Hence, the realization of profit is a function of the production plan and its interaction with the state of the world: $\pi = g(y, x, k; s)$. Given managers' expectation of the probability distribution of states of the world and how the realization of the state interacts with the production plan to yield profit, a subjective probability distribution of profit can be associated with any production plan: $f(\pi; y, x, k)$. Under certain restrictive conditions, this probability distribution can be represented by its first two moments, the expected profit, $E(\pi; y, x, k)$, and its variance, $S(\pi; y, x, k)$. *When managers maximize profit, which necessarily entails minimizing cost, they rank production plans or, equivalently, probability distributions of profit by their first moment, expected profit.* Thus, they adopt the plan with the highest expected profit, which implies its associated output is produced with the inputs and equity capital that minimize its cost. In fact for any output, the standard cost function would give the minimum cost of producing it.

Risk matters to production decisions in banking because higher risk investment strategies can result in a liquidity crisis, regulatory intervention, and even insolvency and the loss of the valuable charter. On the other side of the coin, higher risk strategies exploit explicit and implicit deposit insurance and increase expected profit². As noted above Marcus [18], smaller banks with more valuable growth opportunities maximize value by pursuing relatively low risk investment strategies to protect their charters from financial distress while larger banks in more competitive markets with poorer growth opportunities maximize value by pursuing relatively more risky strategies to exploit deposit insurance. Thus, these managers will rank production plans, not just by their first moment, expected profit, but also by higher moments that characterize the plans' risk.

When managers rank production plans by their risk as well as their expected profit, they may not maximize profit or minimize cost. Maximizing a bank's market value, which is the discounted value of its stream of expected profits, involves accounting for the market priced risk that determines the discount rate as well as the expected profits. A riskier production plan that results in higher expected profit and a discount rate sufficiently higher that market value falls would be ranked lower than some plans with less expected profit. In the case of banking, these calculations are made more complicated by contrasting risk-taking incentives due to the potential costs of financial distress and by deposit insurance and too-big-to-fail policies.

Hughes, et al. [6,7] develop a model of production where managers rank production plans, or equivalently, subjective probability distributions of profit, by higher moments as well as by the first moment, expected profit. For example, as Marcus [18] suggests, a smaller bank might rank plans with less risk and less expected profit higher than those with more risk and more expected profit while larger banks rank

in the opposite order. Generalized profit and cost functions for banking are obtained from this model of managerial utility maximization and include the standard profit and cost functions as special cases where higher moments of the subjective probability distributions of profit do not influence the ranking. The conditions for this restrictive case can be tested to determine if the data are consistent with standard model or if they require accounting for endogenous risk-taking. When this cost function is estimated with banking data, the estimated relationship of cost to output will depend on how risk-taking varies with output and affects cost.

Hughes, et al. [10] estimate the standard minimum cost function described above for top-tier U. S. bank holding companies in 2003, 2007, and 2010 and find on average constant returns to scale for 2007 and 2010 and decreasing returns to scale for 2003. Specifically, for 2007 and 2010, a 10 percent increase in all outputs would result in an average 10 percent increase in cost while, for 2003, an average 10.70 percent increase in cost. By contrast, they use the technique of Hughes, et al. [6,7] to estimate a cost function that controls for endogenous risk-taking, which, in terms of Figure 1, means that scale economies are measured for any given output vector from the equivalent point A to A' rather than A to C or D. In doing so, this technique uncovers evidence of economically and statistically significant scale economies that increase with the size of the bank³. For the full sample in 2003, a 10 percent increase in all outputs results in an average 8.45 percent increase in cost; in 2007, an average 8.78 percent increase in cost; and in 2010, an average 8.00 percent increase in cost. In 2010, for the smallest banks with consolidated assets less than \$800 million, a 10 percent increase in outputs yields an average 8.15 percent increase in cost; for banks in the range of \$2 billion to \$10 billion in assets, an average 7.54 percent increase in cost; while, for the largest banks with assets greater than \$100 billion, an average 7.00 percent increase in cost. Thus, when the effect on cost of endogenous risk-taking is disentangled from the basic technology's underlying scale effects on cost, evidence shows that scale economies are large and increase with the size of the bank. The largest banks experience the largest scale economies.

Applying this technique to 1990 data drawn from U. S. commercial banks, Hughes, et al. [7] find similar patterns of scale economies that are large and increase with bank size. And using 1994 data on top-tier U. S. bank holding companies, Hughes, et al. [11] obtain the same qualitative results.

Dijkstra [15] applies their technique to data drawn from banks in 12 countries of the European Monetary Union during the period 2002-2011 and finds estimates of scale economies that are similar to Hughes, et al. [10] for the U. S. in this period. In 2002, a 10 percent increase in outputs is associated with an average 8.62 percent increase in cost. By 2011, the increase in cost is an average of 7.87 percent. Like U. S. banks, European banks experience scale economies that increase with size. The largest banks obtain the largest scale economies. On the other hand, the estimation for European banks of the standard cost function that fails to account for endogenous risk-taking yields little evidence of these scale economies. Dijkstra [15] finds that in 2002 a 10 percent increase in outputs is associated with a 9.89 percent increase in cost and in 2011, a 10.02 increase in cost. This evidence given by the standard cost function suggests approximately constant returns to scale.

Notably, in all these studies carried out by Hughes, and Mester as well as those with their coauthors, the test of consistency of the data with the assumptions of cost minimization and profit maximization strongly rejects cost minimization and profit maximization. Thus, managers rank production plans by higher moments characterizing risk as well as by the first moment. Thus, estimation of the standard

² Tufano [25] discusses reasons why risk-neutral managers would find managing risk a value-maximizing strategy.

³ Smaller and larger banks differ in how they produce and evaluate information dealing with the credit risk of loan applicants. Smaller banks rely on the relationship with the customer, often a small business, and knowledge of the local economy to produce "soft" information that cannot easily be captured in "hard" data. Larger banks rely on transactional data that is readily verifiable and, thus, is easy to transfer in the organization. Soft data is not easy to transfer. Consequently, Nakamura [26] and Brickley, et al. [27] contend that the differences in information between large and small banks imply that their ownership structure and the markets they serve will be different. Moreover, Dijkstra [15] hypothesizes that the transactional nature of information in larger banks is more likely to generate scale economies than the relationship nature of information in smaller banks. This difference in information and the type of customer served by smaller and larger banks suggest that, despite the greater scale economies of larger banks, smaller banks will not be supplanted by larger banks.

model of cost minimization with these data yields biased estimates of scale economies.

Do Scale Economies Result from Too-Big-To-Fail Subsidies in the Cost of Funds?

Textbooks contend that scale economies result from technological advantages conferred by large size. However, many suggest that large size also confers a subsidy in the cost of funds due to the too-big-to-fail doctrine. Is it possible that the estimated scale economies of these banks results from such a subsidy? As a theoretical matter, the cost function is defined for a given set of input prices. Hence, input prices are held constant as the response of cost is gauged for a equi-proportional increase in all outputs. Hughes, et al. [10] offer two robustness tests to supplement this theoretical assurance. First, using 2007 data, they recalculate scale economies for the largest banks whose assets exceed \$100 billion by substituting the median interest rates paid on borrowed funds by smaller banks that do not obtain a cost-of-funds subsidy for the interest rates paid by these too-big-to-fail banks⁴. They obtain essentially the same estimated scale economies for these largest banks. Using the actual interest rates paid by the largest banks with assets greater than \$100 billion (in 2007), they estimate that a 10 percent increase in outputs is associated with an average 7.50 percent increase in cost. Using the median interest rates paid by smaller banks, a 10 percent increase in the outputs of the largest banks would result in an average 7.43 percent increase in cost. Second, they drop all banks whose consolidated assets exceed \$100 billion from the sample and re-estimate the model. They use the parameter estimates obtained from this sample that excludes the largest banks to compute the scale economies for these too-big-to-fail banks, and they obtain essentially the same results: a 10 percent increase in outputs occasions an average 7.42 percent increase in cost compared with a 7.50 increase in cost when the largest banks are included in the estimation of the cost function. They conclude that technology, not too-big-to-fail subsidies, account for the scale economies of the large financial institutions.

Scale Economies at the Largest Banks and the Implications of Breaking Them into Smaller Banks

The evidence of the largest scale economies at the largest institutions suggests that a policy to break them up into smaller institutions would substantially increase their average costs and raises the possibility that they would no longer be globally competitive. Hughes, et al. [10] use their estimation of costs in 2007 to obtain evidence on this question. Following the work of Brewer, et al. [19], they identify banks holding consolidated assets greater than \$100 billion as too big to fail and consider their total cost if each of the 17 banks in this group is reduced in size by one half to create 34 banks. The 17 banks largest U. S. banks incur total costs of \$410 billion while the 34 banks half the size of these 17 banks incur combined costs of \$506 billion. These banks hold \$9.1 trillion in consolidated assets. The average cost per dollar of assets is 4.5 percent for the 17 largest banks and 5.6 percent for the 34 banks scaled back by one half in size. Some of these banks, even scaled back in size by one half would still be considered too big to fail.

⁴ Davies, et al. [28] use Moody's ratings data to compute for each bank an interest rate on borrowed funds that assumes no government assistance. They substitute this rate for the observed rate and then re-estimate the cost function. Re-estimating the cost function is problematical because the observed cost is regressed on pseudo prices that did not give rise to this cost. The observed cost results from the actual prices, not the pseudo prices. Hence, the results obtained from this re-estimation do not result from the underlying technology. Instead, the actual prices and cost should be used in the estimation, and the estimated parameters can then be used to compute scale economies at the pseudo prices. Another important problem with their estimation is that they do not control for endogenous risk-taking.

Using the benchmark of \$100 billion to define too-big-to-fail, Hughes, et al. [10] also consider breaking the 17 largest U. S. institutions into a number of institutions holding \$100 billion in assets so that the total assets of this group equals the \$9.1 trillion of the 17 banks. In this case, the combined predicted cost of the resulting smaller institutions would total \$1.48 trillion which yields an average cost per dollar of assets of 16.3 percent.

Between these two cases where banks are reduced in size by one half and where they are reduced to \$100 billion in consolidated assets - an increase in predicted average cost from 4.5 percent to an average cost between 5.6 and 16.3 percent - the global competitiveness of these largest financial institutions would likely be significantly compromised. In these two cases the quantities of financial products and services offered by each bank are scaled down proportionately so the smaller version of the bank produces the same product mix as the larger bank. At the smaller sizes, the product mix of the larger bank will likely be uneconomical so that the broken-up banks will produce a different mix. How the demand for the products and services that can no longer be produced economically by the broken-up domestic banks will be accommodated is a key question.

Boyd, et al. [20] use the estimates of scale economies obtained by Hughes, et al. [11] to calculate the benefits of bank scale, which they contrast with their calculation of the social costs of the recent financial crisis. Attributing these costs to the scale of banks, they draw negative conclusions: "Our calculations indicate that the cost to the economy as a whole due to increased systemic risk is of an order of magnitude larger than the potential benefits due to any economies of scale when banks are allowed to be large. When distributional and intergenerational transfer issues are taken into account, the potential benefits to economies of scale are unlikely to ever exceed the potential costs due to increased risk of financial crisis."

The conclusions of Boyd, et al. [20] implicitly assume that imposing a smaller scale on banks by regulation would ameliorate the systemic risk entailed by larger institutions. In contrast, Mester [5] suggests, "Such limits work against market forces and do not align incentives. Given the potential benefits of size, strict limits would create incentives for firms to avoid these restrictions, and could thereby push risk-taking outside of the regulated financial sector, without necessarily reducing systemic risk."

Conclusions

Textbooks claim that banking technology involves economies of scale that reduce the average costs of outputs. Many empirical studies have not found evidence of such economies. These studies, by using the standard theory of cost that ignores endogenous risk-taking, fail to account the effects on cost of increased risk-taking associated with a larger scale of output. While a larger scale of output improves diversification and reduces the average cost of producing the larger output with the same investment strategy (e.g., ratio of equity capital to assets and liquid assets to assets), larger banks, which have incentives to increase risk-taking as their scale increases, can experience an increase in cost that is more than proportional to the increase in output when extra risk-taking is costly. This risk-taking effect can obscure the cost-saving effect of better scaled related diversification. Thus, the estimation of banking technology must control for endogenous risk-taking to disentangle the effect on cost of scale related risk-taking from the effect of scale related diversification. Studies that account for endogenous risk-taking find evidence of economically significant scale economies in banking that increase with the size of the bank. The largest banks experience the largest scale economies. And these economies suggest

that proposals to scale back the size of these banks to reduce their systemic risk would compromise their global competitiveness and give them the incentive to reorganize outside the regulated financial sector - with the potential for new sources of systemic risk.

References

1. Fisher R, Rosenblum H (2012) How huge banks threaten the economy. Wall Street Journal.
2. Bair S (2012) Why it's time to break up the "too big to fail" banks. Fortune.
3. Purcell P (2012) Shareholders can cure too big to fail. Wall Street Journal.
4. Tarullo DK (2011) Industrial organization and systemic risk: an agenda for further research. Conference on the Regulation of Systemic Risk, Federal Reserve Board, Washington, DC, USA.
5. Mester LJ (2010) Scale economies in banking and financial regulatory reform. The Region: Banking and Policy Issues Magazine, Federal Reserve Bank of Minneapolis, USA.
6. Hughes JP, Lang W, Mester LJ, Moon C-G (1996) Efficient banking under interstate branching. *Journal of Money, Credit, and Banking* 28: 1045-1071.
7. Hughes JP, Lang W, Mester LJ, Moon C-G (2000) Recovering risky technologies using the almost ideal demand system: an application to U.S. banking. *J Financ Serv Res* 18: 5-27.
8. Berger AN, Mester LJ (1997) Inside the black box: what explains differences in the efficiencies of financial institutions? *Journal of Banking and Finance* 21: 895-947.
9. Hughes JP, Mester LJ (1998) Bank capitalization and cost: evidence of scale economies in risk management and signaling. *Rev Econ Stat* 80: 314-325.
10. Hughes JP, Mester LJ (forthcoming) Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *J Financ Intermed*.
11. Hughes JP, Mester LJ, Moon C-G (2001) Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production. *Journal of Banking and Finance* 25: 2169-2208.
12. Bossone B, Lee J-K (2004) In finance, size matters: the 'systemic scale economies' hypothesis. *IMF Staff Papers* 51: 19-46.
13. Wheelock D, Wilson P (2012) Do large banks have lower costs? New estimates of returns to scale for US banks. *Journal of Money, Credit, and Banking* 44: 171-199.
14. Feng G, Serletis A (2010) Efficiency, technical change, and returns to scale in large US banks: Panel data evidence from an output distance function satisfying theoretical regularity. *Journal of Banking and Finance* 34: 127-138.
15. Dijkstra M (2013) Economies of scale and scope in the European banking sector 2002-2011. Working Paper 2013-11, Amsterdam Center for Law & Economics, University of Amsterdam, The Netherlands.
16. Greenspan A (2010) The crisis. *Brookings Papers on Economic Activity*.
17. Demsetz RS, Strahan PE (1997) Diversification, size, and risk at bank holding companies. *Journal of Money, Credit, and Banking* 29: 300-313.
18. Marcus AJ (1984) Deregulation and bank financial policy. *Journal of Banking and Finance* 8: 557-565.
19. Brewer E, Jagtiani J (2009) How much did banks pay to become too-big-to-fail and to become systemically important? Federal Reserve Bank of Philadelphia Working Paper 09-34.
20. Boyd JH, Heitz A (2012) The social costs and benefits of too-big-to-fail banks: a bounding exercise. Working Paper. University of Minnesota, USA.
21. Keeley MC (1990) Deposit insurance, risk, and market power in banking. *Am Econ Rev* 80: 1183-1200.
22. McConnell JJ, Servaes H (1995) Equity ownership and the two faces of debt. *J Financ Econ* 39: 131-157.
23. Hughes JP, Lang W, Moon C-G, Pagano M (1997,2004) Measuring the efficiency of capital allocation in commercial banking. Federal Reserve Bank of Philadelphia Working Paper 98-2.
24. Hughes JP, Mester LJ (2013) A primer on market discipline and the governance of financial institutions for those in a state of shocked disbelief. *Efficiency and Productivity Growth: Modelling in the Financial Services Industry*, John Wiley and Sons.
25. Tufano P (1996) Who manages risk? An empirical examination of risk management practices in the gold mining industry. *J Financ* 50: 1097-1137.
26. Nakamura L (1994) Small borrowers and the survival of the small bank: is mouse bank Mighty or Mickey? Federal Reserve Bank of Philadelphia Business Review.
27. Brickley J, Linck J, Smith C (2003) Boundaries of the firm: evidence from the banking industry. *J Financ Econ* 70: 351-383.
28. Davies R, Tracey B (forthcoming) Too big to be efficient? The impact of too-big-to-fail factors on scale economies for banks. *Journal of Money, Credit, and Banking*.