Martin Neil Baily
Douglas J. Elliott
The Brookings Institution
July 11, 2013

The Role of Finance in the Economy: Implications for Structural Reform of the Financial Sector

Executive Summary

The U.S. financial system is critical to the functioning of the economy as a whole and banks are central to the financial system. In addition to providing substantial employment, finance serves three main purposes:

Credit provision. Credit fuels economic activity by allowing businesses to invest beyond their cash on hand, households to purchase homes without saving the entire cost in advance, and governments to smooth out their spending by mitigating the cyclical pattern of tax revenues and to invest in infrastructure projects. Banks directly provide a substantial amount of credit in the U.S., but, unlike in almost any other economy, financial markets are the ultimate providers of most credit.

Liquidity provision. Businesses and households need to have protection against unexpected needs for cash. Banks are the main direct providers of liquidity, both through offering demand deposits that can be withdrawn any time and by offering lines of credit. Further, banks and their affiliates are at the core of the financial markets, offering to buy and sell securities and related products at need, in large volumes, with relatively modest transaction costs. This latter role is particularly important in the U.S., given the dominance of markets, but is often under-appreciated.

Risk management services. Finance allows businesses and households to pool their risks from exposures to financial market and commodity price risks. Much of this is provided by banks through derivatives transactions. These have gotten a bad name due to excesses in the run-up to the financial crisis but the core derivatives activities provide valuable risk management services.

Many argue that the U.S. financial system grew overly large in the bubble period and is still too large today. We agree that some of the activities that took place in the bubble period involved taking on excess amounts of risk, but it is extremely hard to determine the right size of the financial system based on well-grounded economic theories. In truth, it is very difficult to judge the right size of almost any industry and attempts at the use of central planning and other mechanisms to correct assumed problems of this nature have usually failed.

Nonetheless, it is reasonable to assume that a sector will be too large if there are unwarranted economic subsidies flowing to it. This does appear to have been the case in the bubble and may still be the case, although such subsidies have been much reduced by a series of actions to remove government support and to force the financial industry to operate more safely.

However, we suspect the excessive size in the bubble period was considerably less than many argue and we believe it is important to be cautious in drawing policy conclusions as it seems impossible to prove whether the sector was or is too large and by how much.

There are a number of important proposals to force major changes in the structure of the financial industry, including to:

- Eliminate Too Big to Fail banks by forcing their break-up or downsizing
- Limit the functions of banks à la Glass Steagall or the Volcker Rule

Banks that are central to our financial system, whether through sheer size or the critical nature of the services they provide, are perceived by many to benefit from an implicit government guarantee that ought to be eliminated. The main categories of proposals are:

- Break up the largest banks
- Mandate a size limit
- Push large banks to shrink voluntarily by imposing stiff costs for size
- Put in place a credible plan for resolving the largest institutions

We do not favor the proposals to break up the banks or force them to shrink dramatically. We believe that the best analysis indicates considerable economic benefits to size and scope and that these advantages are likely to grow further with increasing globalization, complexity, and improved information and management systems. America should have at least a few financial institutions with global scale, capable of providing a wide range of related commercial and investment banking services, operating on a scale in individual product lines that produces real efficiency.

This will almost certainly mean these firms are important enough to the economy that the government and regulators will need to watch them particularly carefully and may create need for special assistance, in extreme crisis situations of the level that are unlikely to occur more than once or twice a century. For this reason, we agree on the need to designate systemically important financial institutions and to require them to operate with higher safety margins.

We believe that the societal benefits of breaking up the large banks are over-stated. The recent financial crisis was much more about system-wide problems than about issues resulting from excessive size of financial institutions. A simple thought experiment illustrates this. If we had broken up the big banks a decade ago into 10 or 20 pieces each, they would likely all or virtually all have made the same mistakes. They would have over-invested in real estate-related products, taken excessive risks across the board, created opaque and risky securitizations and derivatives products, pushed accounting rules to their limits, etc. The other players in the financial system would presumably also have made the same mistakes, including the ratings agencies, governments, central banks, regulators, and families and businesses. It is difficult to presume that the disaster would have been much different. Indeed, there is a chance that the clean-up would have been more difficult without the ability to pull 17 key CEOs into a room and force them to accept the TARP arrangements.

The next financial crisis will almost certainly differ from the last, as every such crisis varies, but it remains difficult to see how a system of many mid-sized banks would be appreciably safer than one with some large banks as part of the mix.

We do favor ensuring that even the most important banks can be resolved effectively without the use of taxpayer funds, except perhaps for relatively short-term liquidity purposes and backed by solid collateral. Dodd-Frank goes a long way towards achieving this goal, but more could be done.

Activity limitations

U.S. commercial banks and their affiliates have always faced limitations on the business they are allowed to undertake, in order to reduce the risk of business disasters that would endanger their ability to fulfill their critical role at the heart of the economic system.

These limitations were considerably extended in the Great Depression. The Glass-Steagall Act was passed, making it illegal for a commercial bank to be affiliated with an investment bank. The former could undertake the types of activities we normally associate with banking, such as taking deposits and lending. The latter were principally involved in the securities business, through helping firms raise capital by selling stocks and bonds, assisting investors in buying and selling those securities, and trading them for the investment bank's own account.

The anti-affiliation provisions of Glass-Steagall were dramatically modified in the 1990's, allowing commercial and investment banks to be part of the same financial group, although there remain a number of important restrictions to limit dealings within the group.

There is a range of proposals to further limit the ability of banks to operate in the securities and derivatives businesses. Some call for a restoration of the anti-affiliation provisions of Glass-Steagall. Others want Glass-Steagall Lite, since they recognize that changing times make it difficult to simply turn back the clock. The Volcker Rule is intended to separate out proprietary trading completely from commercial banks *and* investment banks.

We do not favor any of the major proposals for further structural divisions between commercial banking and securities and derivatives activities. We believe that the U.S. capital markets are world leaders and that their strength is an important economic advantage for America. Those markets are underpinned by the role of major securities dealers that are closely affiliated with commercial banks. A major reason for the close linkages is the desire of corporate customers to be able to deal with financial firms that can provide a solid range of products from financial advice to loans to securities offerings to risk management via derivatives to purely operational products. The institutional knowledge and relationships that a banking group has in regard to its corporate customers is a valuable advantage both for the bank and for those customers.

Further, times have changed and will not change back. Glass-Steagall was based on a clear difference between a loan and a security, a difference that no longer exists now that most large loans are tradable among banks and also specialized investors. At this point, it is usually possible to structure a given transaction as a loan or a security or a derivatives transaction or often as insurance or another contractual arrangement.

Finally, any transition from the current system to an older-style system will create very considerable displacement of activities, with a real potential for problems. Some of this might occur through the divestiture of investment banking subsidiaries from banking groups, which would be the simplest approach; however even this would involve a large amount of change at a time when the U.S. economy remains in a fragile recovery that resulted in part from the disruption of the financial sector. Another source of displacement would result from striving mid-level securities firms grabbing market share. Although this could bring advantages, it also creates the danger of a repeat of a situation such as developed at MF Global, where the push for growth overcame proper risk management practices.

Introduction

The financial sector is a critical component of the economy. How well it works is a key factor in determining how the rest of the economy functions, as was clearly demonstrated when the recent financial crisis plunged economies into recession around the globe. The structure of the financial sector is under great scrutiny as a result of the crisis and some significant changes are being mandated. Others of potentially even greater import are under discussion, such as breaking up the largest banks or returning to some form of strict prohibition on the affiliation of banking and securities firms.

We believe it to be crucial for any major changes to be based on a careful analysis of the financial sector and its relationship to the "real economy." In many cases, there is a need for considerably more research and analysis and in other cases the existing state of knowledge is too frequently ignored or inconvenient realities played down. For these reasons, we convened a conference in December 2012, at the Brookings Institution that brought together many experts to discuss some key questions about how finance works now and how it should work in the future. Federal Reserve Board Governor Daniel Tarullo gave an excellent keynote address that buttresses our own belief that more research and more careful consideration of known facts is needed.[1]

This paper represents the chance to summarize our views on the purposes and current and optimal structures of the financial sector. It is necessarily not a definitive work, but provides an overview, with particular emphasis on some points that we believe receive too little consideration.

Purposes of the financial sector

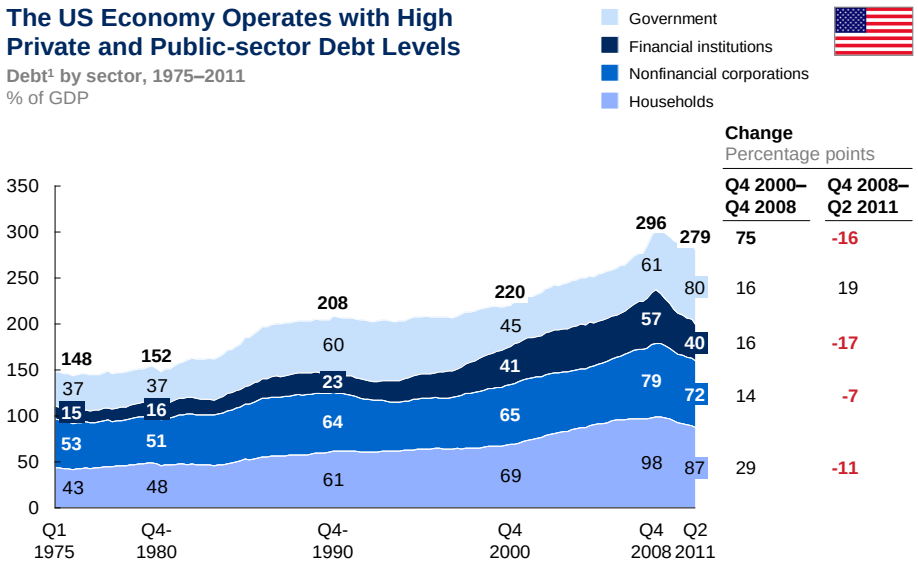
One way the financial sector's impact on the overall economy has been measured in the past is through its direct contribution to employment and GDP. For example, in 2006 there were 6.19 million people employed in the finance and insurance sector of the American economy, representing 5.4 percent of total nonfarm private sector payrolls, according to data from the Bureau of Labor Statistics. By 2010 employment in the financial sector had dropped to 5.76 million but total payroll employment was also down and so the share of employment in the sector remained comparable. By 2012 employment in the sector had risen only marginally, to 5.83 million, and the sector's share of employment was down to 5.2 percent. Nevertheless, the financial sector remains a large part of the economy and a major employer.

[1] A detailed summary of the points raised by the participants is available at http://www.brookings.edu/research/papers/2013/01/17-bank-restructuring-elliott A transcript and various of the presentations can be found at http://www.brookings.edu/events/2012/12/04-financial-industry-structure#ref-id=20121204_ES_keynote

In this section, however, our purpose is not to talk about the direct employment or GDP impact of the sector, but instead to emphasize the economic importance of the services the sector provides. A modern financial system exists primarily to provide three types of services to the rest of the economy:

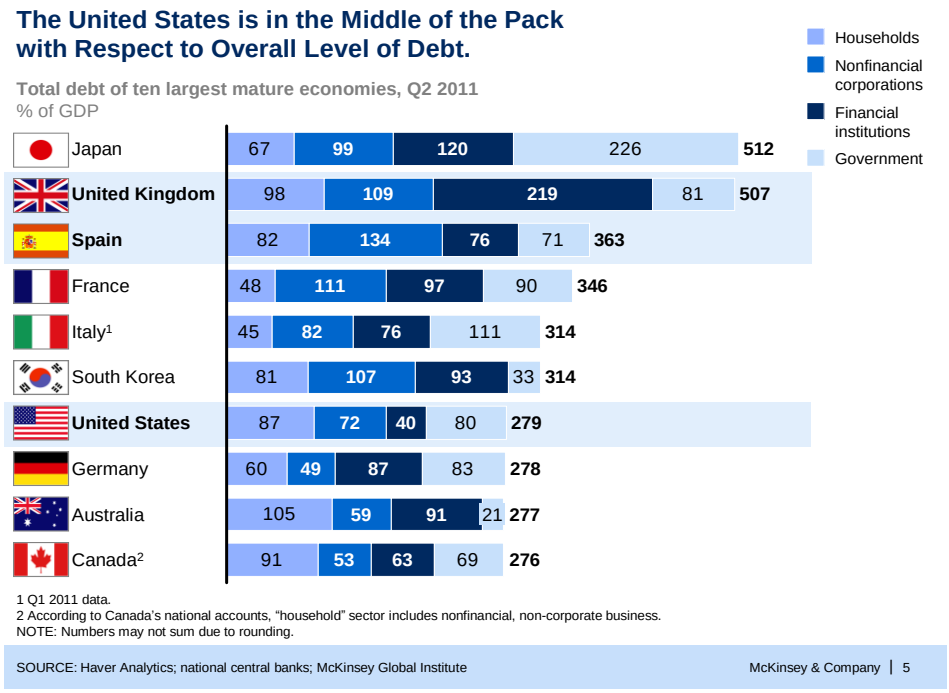
- Credit provision
- Liquidity provision
- Risk management services

Credit provision. The figure below[2] shows the U.S. economy is highly leveraged and hence very dependent on the availability of debt. At the peak, prior to the crisis, total debt outstanding equaled about 300 percent of GDP, split among households, government, financial corporations and non-financial corporations. Of course, the ratio of debt to GDP involves some "double counting" in the sense that households borrow from financial institutions but also own the assets of these institutions. However, the total amount of debt gives an indication of the scale of financial intermediation that is occurring. And, more than most other countries, foreign entities own a substantial fraction of the outstanding U.S. debt, including much of the government debt. The U.S. economy is very connected to global financial markets, providing pricing and liquidity to those markets and depending on foreigners to buy public and private debt instruments.



[2] This figure and the next one are both taken from *Debt and Deleveraging,* a study by the McKinsey Global Institute, part of McKinsey & Company.

One can certainly make the case that there was too much leverage at the peak of the bubble. The figure shows there has been some deleveraging since then, but overall leverage has been over 200 percent of GDP since the 1990s and is sure to remain above that level into the future. The United States is often pictured as a profligate borrower but as the figure below shows, we are only in the middle of the pack in terms of the ratio of debt to GDP, slightly below Korea and well below France and the UK. The United States is unique only because it is such a large economy and the fact that it is a net foreign borrower. The overall message from the figure is that the advanced economies all rely on large amounts of borrowing and lending as an important part of the operation of their economies.



Total U.S. commercial bank credit in 2011 was $9.4 trillion, with $6.9 trillion in the form of loans and leases outstanding. On top of the outstanding credit amounts, the banks have committed considerably more through contingent arrangements such as lines of credit that allow companies and individuals to know that funds would be available if needed. In addition to traditional lending, commercial and investment banks also take credit risk in many other fashions, particularly through derivatives exposures and the ownership of bonds and other financial instruments issued by companies and governments.

The Consequences of Interruptions in the Flow of Credit

The importance of credit provision to the larger economy was demonstrated by the recent financial crisis and ensuing sharp recession. Virtually all observers agree that the

difficulty of obtaining credit and the uncertainty as to its availability and cost were critical drivers of the recession. The economics literature has documented the impact of bank crises and the resulting loss of access to credit. A paper from researchers Luc Laeven and Fabian Valencia at the IMF updated their previous analysis of bank crises and concluded that for advanced economies the median cumulative loss of national income relative to trend from bank crises over the period 1970-2011 was nearly 33 percent, taking account of both larger and smaller crises. These same economies hit by crises also faced a median increase in their government debt levels of 21 percent and the median duration of the crisis was three years.[3] The most recent crisis is likely to have higher costs and longer duration than the historical values the IMF computed, given its severity and the fact that the financial crisis in Europe has now evolved into a sovereign debt crisis. Studies on financial conditions indices also show that credit conditions affect economic growth even in more normal times.

Most of the credit in the United States is ultimately provided by investors in bonds and other credit instruments, rather than residing on the balance sheets of banks. Only about a fifth of U.S. credit is held by banks, in contrast to Europe, where banks hold close to three quarters. U.S. equity investment is even more heavily skewed away from banks, with 99% of common stock held by non-bank investors.

The centrality of financial investors to the U.S. economy means that efficiently operating financial markets are crucial to the provision of credit and equity investment for American businesses and families.

Liquidity provision. Many of the major debt and equity investors care significantly about the degree of liquidity of their investments. Liquidity has two important and related aspects. First, there is the question of how easily one can buy or sell a position of the relevant size without moving the market price adversely. A large mutual fund, for example, will be less interested in owning thinly traded shares whose price will move up during the process of their buying shares. This is simply due to the pressure of those purchases, especially because the opposite would likely occur when they wished to liquidate their position. Illiquidity would therefore raise the average cost to investors of buying into a position and lower their average benefit from selling out at the other end of the transaction. In many debt instruments and some equities, there is the more extreme problem that it may take some time, potentially even days for a large position, to actually find investors who own the particular security and are interested in selling at something close to the market price. This creates risk.

The table below is based on FINRA TRACE data. It shows that the great majority of such bonds are traded rather infrequently. In fact, 43% of the bond classes traded in 2009

[3] Luc Laeven and Fabian Valencia, "Systemic Bank Crises Database: An Update," IMF Working Paper WP/12/163, Washington DC, June 2012.

did not even have one trade in the institutional market and 47% traded on 10 or fewer days, less than once a month. Even weighting the bonds by their dollar size shows a similar, though smaller, effect. $1.0 trillion of bonds (17%) had no institutional trades, while another $0.8 trillion (13%) traded on less than 10 days. In other words, most corporate bonds have low levels of liquidity. Another important feature is that bonds of Fortune 500 companies are much more frequently traded than the bonds of smaller companies. The majority (57%) of thinly traded bonds were issued by non-Fortune 500 companies. The bonds of the large companies are substantially more liquid than those of small companies and this is an important advantage for them.

	Trading Frequency		**Size of Original Issuance** [(1)]		
Number of days traded in 2009	Number of Bonds	Percent of Total	Principal Value (MM)	Dollar-Weighted Percent	Avg. Issue Size (MM)
0	11807	43%	1,046,200	17%	89
1-10	7424	27%	806,061	13%	109
>10	8258	30%	4,277,034	70%	518
Total	**27489**		**6,129,295**		

Source: Bloomberg, FINRA
Notes: 1. 5,282 of the 27,489 bonds had no data from Bloomberg, possibly because they have already been redeemed

Second, the transaction costs for thinly traded securities are markedly higher than for more liquid instruments. In particular, bid/ask spreads are usually substantially wider for illiquid securities. The bid/ask spread is the difference between the price at which a dealer is willing to bid for a security and the price it asks to sell that same security. Bid/ask spreads exist for two principle economic reasons. For one, the spread determines the average amount that a dealer will earn from buying a security from one investor and selling it to another, assuming that price changes between the time of purchase and sale average out to be neutral. This is the main economic mechanism by which financial institutions generate the revenues to support the infrastructure of employees, computers, real estate, etc. necessary to be a securities dealer, as well as the related capital and other financial requirements.

Additionally, the bid/ask spread compensates dealers for the risk they may be forced to hold a security in inventory for some time while looking for another investor willing to buy the security. There is at least a modest financing cost to holding securities and, more importantly, there is the risk of a price decline. The danger of a fall in the price is exacerbated with illiquid securities because a modest amount of selling may be enough to push the price down for some time. For this reason, dealers are sometimes willing to

narrow the bid/ask spread for a good client who assures them that they are using only that dealer for the given transaction. Otherwise, there is a risk that other dealers are also being loaded up with inventory on that security as the investor sells through multiple outlets, depressing the price further.

(All of the logic about liquidity holds true as well when investors buy a security from a dealer, who may be forced to "sell short" by borrowing that security from someone else, with the intent of buying the security in the near future when it becomes available at a reasonable price somewhere. However, it tends to be intuitively easier to understand the economic and market principles when the explanations focus on an investor selling and a dealer taking the security into inventory.)

Thus, investors care about liquidity because it affects:

- Their direct transaction costs of getting into and out of a trade
- The average price they pay or receive if their transaction is large enough to move the market for that security
- The length of time it takes to execute a transaction, potentially forcing purchases or sales to be spread over days or weeks.

The net effect of these liquidity issues is to make investors less willing to own illiquid securities and therefore to necessitate a lower price or higher return to lure them in. There are large classes of investors who will not put their money into illiquid securities because they deal in such large volumes that it would not be sensible to make the effort for investments which would take long periods to buy into or sell out of. There are still other investors who might be willing to buy, but only if they are offered a higher average return to make up for the illiquidity. Both these factors reduce the average price investors are willing to pay, either by reducing the sources of demand or changing the price at which that demand exists.

For debt instruments, where the market interest rate shows explicitly what return investors need, analysts refer to the "liquidity premium" as the portion of the demanded interest rate that investors want as compensation for liquidity risks. This liquidity premium is nearly zero for short-term government bonds, but can be very high for long-term and highly illiquid corporate bonds.

Academic Literature on Banks and Liquidity

There is an extensive academic literature on the issue of liquidity with perhaps the most frequently cited theoretical treatment being by Diamond and Dybvig (1983). There is a less technical summary of this article in Diamond (2007), which also cites several

important and more recent contributions to the literature.[4] These authors lay out in a careful and rigorous analysis the framework in which banks create liquidity and they also show the economic basis for the demand for liquidity in the marketplace, both by consumers and businesses. Diamond (2007) draws the following conclusions from this analysis:

> "Banks create demand deposits to provide investors with liquid assets. When there is a demand for more liquid assets from investors or entrepreneurs, demand deposit contracts serve as a means for quick access to liquidity. Demand deposits work very well when investors forecast that banks will survive, but can cause severe damage if investors lose faith in banks. There is scope for banks to write more refined contracts, such as deposits with suspension of convertibility of deposits to cash. In addition, there may be a role for government policies to eliminate self-fulfilling runs on banks. The government plays a role because its taxation authority is not available to private firms." (page 199).

Although, in a sense, Diamond and Dybvig's findings simply repeat the common understanding of banks that has been around since Bagehot, their modeling, based on a solid base of economic theory, is valuable in providing a clear analysis of the role of banks in the economy. Banks provide an essential service to the economy that works very effectively in normal times but is subject to "tail risk" if some economic event triggers depositors or other bank creditors to wish to withdraw their funds all at the same time. Bank crises have occurred regularly in U.S. economic history, including in the early years of the 20th century, which led to the formation of the Federal Reserve System in 1913-14. Demand deposit insurance was instituted as a result of the Great Depression of the 1930s to stop runs by retail depositors. The provision of liquidity is central to the function of banks but also makes them prone to instability. Conservative economist Milton Friedman faulted the Federal Reserve in the Great Depression for not doing more to support the liquidity of the banking system and hence contributing to the massive disruption of the economy.[5]

There is also valuable academic literature that explores the importance of liquidity from an empirical perspective. A paper by Dick-Nielsen, Feldhütter and Lando[6] describes and summarizes prior analyses and proposes new estimates of the liquidity premium based on combining a number of different measures of liquidity. These include the impact of

---

[4] Douglas W. Diamond and Philip H. Dybvig, 1983,"Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91 (5): 401–19; Douglas W. Diamond, "Banks and Liquidity Creation: A Simple Exposition of the Diamond-Dybvig Model," Federal Reserve Bank of Richmond, *Economic Quarterly 93 (2),Spring 2007: 189–200*

[5] Milton Friedman and Anna Schwartz,*A Monetary History of the United States 1867-1960*, (Princeton, NJ: Princeton University Press, 1963).

[6] Jens Dick-Nielson, Peter Feldhütter and David Lando (2012), "Corporate bond liquidity before and after the onset of the subprime crisis," *Journal of Financial Economics*, 103:471-492.

trades on asset prices, an estimate of the bid/ask spread and the turnover or volume of trading activity. Although the number of days when a security does not trade ("zero trading" days) has been used in prior work to measure liquidity, Dick-Nielson et al. do not find it to be useful. These authors use principal components analysis, a statistical method, to combine several possible liquidity metrics into an overall measure or index of liquidity. They use the FINRA TRACE data for their analysis and we refer the reader to the paper itself for a discussion of the detailed methodology and the various controls that are introduced into their regressions. Prior to the crisis the liquidity premium was small for highly rated investment grade bonds, as small as 0.8 bps for AAA ratings, up to 3.9 bps for BBB. Despite its modest size, the premium is always found to be highly statistically significant. For non-investment grade bonds the premium was much higher, averaging 57.6 bps. They also find that the liquidity premium increases with the maturity of the security, reaching 83.9 bps for non-investment grade bonds with maturities of 5 to 30 years.

How are these numbers to be interpreted? Different bonds are generally compared by looking at the yield "spread" of the bond, given by the interest yield on the bond relative to the yield on a risk-free bond with equivalent maturity. Often, U.S. Treasuries are used to determine the risk free rate, but in this study the authors use swap data that they say have been found to be a better way to estimate the risk-free rate. Their findings are estimated separately for different classes of bonds. The finding for AAA bonds means that a bond with this rating that has average liquidity carries a yield that is 0.8 bps higher than a very liquid bond with the same rating, controlling for other factors. Basically, this tells us that liquidity differences were not a very large determinant of interest rate spreads among AAA corporate bonds over the pre-crisis period. These bonds were issued by well-established and usually large companies. For lower rated bonds and long maturity bonds, however, the yield difference between average and high liquidity bonds is much greater.

The study also looks at liquidity premiums in the period following the credit crisis and the rise in the premiums is very striking. The premium for an average liquidity AAA bond rises after the crisis to 4.9 bps compared to a high liquidity bond, a premium six times the level prior to the credit crisis period. The liquidity premium rises to nearly 200 bps for non-investment grade securities on average, two percentage points on the interest rate. It seems likely that the liquidity effects going forward will fall roughly midway between the pre-crisis and crisis figures. The pre-crisis figures are from a period of very low risk-aversion and "bubble thinking" that led investors to place too little value on liquidity. At the same time, the value of liquidity always shoots up in crises and then falls, so one would not expect those levels to be sustained over longer periods.

The bottom line is that liquidity matters and it matters much more for lower rated corporate bonds and non-investment grade bonds, which are often issued by smaller companies and those without a strong credit history. These are the companies that are having the hardest time recovering from the financial crisis. Liquidity premiums were

generally low in the years prior to the crisis but have risen substantially since then. Providing liquidity is an essential service that the financial sector must make available to companies large and small in order to maintain the flow of corporate bond funding needed to sustain economic growth. Even small increases in interest rate spreads can be significant. An increase in spread of 10 bps on 2011's issuance of $1.2 trillion of corporate debt would mean an additional $1.2 billion a year in cost for the life of those bonds. If the average life of a corporate bond is five years, this would accumulate with each year's issuance to an eventual increase in corporate interest costs of about $6 billion a year.

Risk management. Financial institutions and markets help businesses and individuals to protect themselves against many forms of financial risk, generally through various forms of "derivatives". This is the term used for financial instruments whose value is derived from the value of some underlying asset. One of the oldest forms of derivatives is a futures contract on a crop, such as wheat. The concept has since expanded to financial instruments as well as many non-agricultural commodities.

The importance of futures contracts to the economy and economic growth is illustrated by the sudden change in the energy environment in the United States. Energy experts such as Daniel Yergin and Philip Verleger believe that newly developed technologies can safely extract huge amounts of natural gas and increased amounts of oil from shale deposits, resulting in plentiful cheap energy for America for many years to come. The technologies were developed primarily by small wildcat drilling companies and these companies were dependent on futures contracts to stay in business. Natural gas and oil prices go through wild swings up and down creating risks that small drillers cannot absorb. By buying futures contracts the drillers were able to lock in pricing to reduce their risks and stay in business, something that has paid off for them and for the whole economy.[7]

The other main form of derivatives is an option to buy or sell a commodity or a security at the fixed price for a given period of time. For example, rather than committing to a specific transaction at a future date, and relinquishing any potential upside, a farmer may simply wish to protect against the crop price moving too far down. Options can be used like an insurance purchase, where one hopes that the event does not occur, even though that means the insurance would be "wasted." Businesses and families are willing to take an expected loss on insurance purchases in order to avoid the risk of outcomes that would be ruinous to them, such as fire destroying an uninsured home.

[7] Yergin is the founder of Cambridge Energy Research Associates and author of *The Prize: The Epic Quest for Oil Money and Power.* He has talked on CNBC and in other forums about the impact of the expansion of U.S. energy supply. Verleger is the President of PKVerleger LLC, the author of "The Amazing Tale of U.S. Energy Independence," *The International Economy* (Spring 2012) and is a visiting fellow at the Peterson Institute where he is completing a book on the benefits of low-cost energy.

Futures and options transactions now extend well beyond commodities to cover a wide range of financial risks, since firms and individuals are exposed to very significant dangers from potential adverse movements in interest rates, stock prices, or exchange rates. Other market participants use derivatives to take credit risk or other investment positions more efficiently or even as speculation. Although speculation tends to be viewed negatively, it is an important element of efficient market functioning, since speculators are necessary when there is an imbalance in supply and demand from other participants.

Derivatives have gotten a bad name with much of the public, since the financial industry developed some derivatives prior to the financial crisis that were excessively complex and too far removed from any real world transactions. These derivatives helped fuel some of the speculative excess that led to the financial crisis and resulted in major losses for some parties. However, it is important to remember that derivatives are a key risk management tool for companies and other "real world" parties to manage their risks from future economic and financial market developments.

It is difficult to quantify the economic value of risk management products, but the academic literature has long recognized that people do place great value on risk reduction. As a starting point, there is a strong consensus among economists, and most other observers, that humans and the institutions they run are risk-averse. This implies that they would willingly pay for protection or give up potential value in exchange for increasing certainty. This analysis goes back to at least the writings of Daniel Bernoulli in 1738 and has been thoroughly incorporated into economic thought, particularly the analysis of financial markets.

In 1995 the Harvard Business School published a pioneering analysis of the financial system and the chapter on "The Allocation of Risk" was written by leading HBS professor Scott P. Mason.[8] He said that there are three generic ways to manage risk: hedging, diversification and insurance. Hedging allows an individual or organization to eliminate a risk through the spot sale of the risk or through a transaction in an instrument that represents an obligation to sell the risk in the future. Diversification reduces risk-bearing through the combining of less than perfectly correlated risks. And insurance refers to a set of securities or contracts that limit risk in exchange for the payment of a premium. He argues that all risk allocation activities activity can be analyzed in terms of this taxonomy.

[8] Dwight B. Crane et al., *The Global Financial System: A Functional Perspective (*Harvard Business School, 1995). Chapter 5 is "The Allocation of Risk," by Scott Mason who, unfortunately, died in 1998. Leading college textbooks covering this material are Zvi Bodie, Robert C. Merton and David L. Cleeton, *Financial Economics,* second edition 2009 and Frederic S. Mishkin, *Economics of Money, Banking and Financial Economics*, tenth edition 2012.

Mason also talks about the benefits of risk management products to the economy. He describes the basic contribution in terms of improving the welfare of participants in the economy that are exposed to risk and are risk-averse. In financial markets this shows up in the strong preference exhibited by many market participants for assets that carry little or no risk of default or that have low price volatility. Retirees in particular choose safe assets with assured rates of return. In order for financial institutions to match the desire of the holders of financial assets, they must strip the risks out of the assets that are generated in the market before selling them. These institutions cannot concentrate all the risks on their own books and so they make use of derivatives to spread the risks to other institutions, a strategy that can reduce risk through diversification. Alternatively, they pass the risks on to specialized institutions such as hedge funds that have the knowledge and the resources to take on the risks and profit from them.

One aspect of risk management that Mason does not deal with fully involves the desire of managers of nonfinancial companies to avoid taking risks they do not understand and may not be able to absorb. One example of this was given earlier—wildcat drillers used futures contracts to protect against oil and gas price swings. This practice is very widespread, however, and even large companies engage in hedging or risk management strategies. For example, a U.S. company that is exporting to Europe wants to focus on what it does best, designing and making a product, selling it and servicing it. The company does not want to find that its profit margin on the export contract has disappeared because the value of the dollar changed in relation to the euro. Similarly, a European exporter selling to the United States also wants protection against exchange rate risk. Financial institutions can arrange exchange rate swaps that protect both parties; indeed, it may be able to arrange a swap with no residual risk held by the financial institution—a win-win situation. More generally, managers do not want to guess about future prices but would rather spend their time and expertise on their companies' core competencies and pay financial institutions to take care of the risk management. If they are forced to hold risk, they may curtail their activities instead. The innovations in derivatives creation and pricing have been a tremendous boon in allowing people to specialize in the activities they know and do well. This is an extension of Adam Smith's view of the benefits of the division of labor, or of David Ricardo's analysis of comparative advantage.

Another important expert on risk management is Yale Professor Robert Shiller who is well-known for having predicted both the stock market decline that resulted from the dot com bubble and the collapse of housing prices. He did pioneering studies of the stock market showing that it is prone to swings in value that cannot be accounted for by the variability of corporate earnings. Importantly, Shiller has responded to this understanding by proposing that more innovative financial products be created to help people and institutions handle risk. He made this case in two important books, the first a largely theoretical treatment published in 1993 and the second a more popular

version in 2003.[9] Shiller notes the importance people place on the ability to buy insurance on their assets, including their investment portfolios through futures and options, but he points out that they cannot yet buy insurance against a decline in the price of their house. He proposes the creation of derivatives that provide that type of insurance. One prerequisite for such derivatives is objective, accurate price indexes for metropolitan areas. The Case-Shiller price indexes are now reported regularly in the press. In summary, Shiller is an economist who does not believe that expectations are always rational or that markets are always perfect, but he does see the tremendous value in risk management products and believes society needs more of them.

As the Nobel Prize-winning economist Kenneth J. Arrow points out, derivatives are quite similar to insurance, about which there is a wider and deeper pool of economic thought. He indicates "[t]he purest form [of contracts to share risk] is insurance. However, risk-bearing can, in fact, be shifted by many contractual arrangements, of which common stock is a good example and derivatives securities of all kinds even better illustrations."[10]

One economist who has attempted a comprehensive evaluation of the value of financial innovation is our former Brookings colleague and financial market expert Robert E. Litan. In a paper published on the Brookings website February 17, 2010, entitled "In Defense of Much, But Not All, Financial Innovation," Litan looks at a broad range of financial innovations and concludes that most of them have added to society's welfare.[11] Unfortunately, he was not able to provide precise quantification of the value of these innovations but he is able to give ratings for risk management products based on their contributions along three dimensions: Access, Convenience and GDP (or productivity). Scores range from double minus for the most negative contributions up to double plus. His first category is "options, futures exchanges and pricing" and these get single pluses on access and convenience and a double plus on GDP contribution. The second category is "interest, currency swaps" and these receive double pluses on access and convenience and between a single and double plus on GDP contribution. The final category is "credit default swaps" and these are rated single plus for all three types of contribution. In summary, risk management products get an endorsement from Litan's careful and balanced assessment.

---

[9] Robert Shiller *Macromarkets: Creating Institutions for Managing Society's Largest Risks* (Oxford: Clarendon Press, 1993); *The New Financial Order: Risk in the 21st Century* (Princeton and Oxford: Princeton University Press, 2003).

[10] Kenneth J. Arrow, "The Theory of Risk-Bearing: Small and Great Risks", *The Journal of Risk and Uncertainty* (May 1996): 1.

[11] http://www.brookings.edu/~/media/research/files/papers/2010/2/17%20financial%20innovation%20litan/0217_financial_innovation_litan.pdf.

How are the banking and securities industries structured to provide these products?

The previous section demonstrated the importance to the "real economy" of the financial system's provision of credit and liquidity, as well as its role in risk management. Investment and commercial banks are the central providers of these functions.

Credit provision. Banks provide $11 trillion of credit to the U.S. economy by holding loans and debt securities on their balance sheets, representing 20% of the total outstanding. The proportion they provide of the debt of non-financial corporations is even higher at 29%. However, the role of these banks in credit provision goes considerably beyond this.

They are the major originators of loans to businesses and families. A significant majority of loans are put together by these banks. Much of the credit risk is retained by them, but a great deal is also passed on to end-investors through loan participations, securitizations, derivatives, and other means.

They are the dominant securities dealers, as well as substantial holders of securities in their roles as investors. Without them, financial markets would operate considerably less efficiently, as described below, and would not be as effective in providing credit.

They arrange the financing for major categories of investors that operate with financial leverage, including hedge funds. These investors rely on the ability to borrow against securities in order to obtain funding at relatively low cost to support their investments. Without such funding, these investors would cut back substantially on their credit provision.

Some banks provide key elements of infrastructure for the credit and risk management markets, allowing these markets to operate efficiently, through the provision of clearing services, securities lending, and other activities that allow markets to provide credit smoothly to the wider economy.

Liquidity provision. Securities dealers increase the liquidity of many securities sharply by holding those instruments in inventory for sale to interested investors or by being willing to temporarily take substantial long or short positions in order to accommodate the buying or selling demand of investors. Their willingness to do this is partially dependent on the existence of a "wholesale" network of other large securities dealers to whom they can lay off their own exposure quickly. Ultimately, they and the other dealers transfer their positions to the "retail" market, but this can take considerable time. ("Retail" in this report refers to non-dealers, rather than to small investors as is sometimes true in other contexts.) In the interim, it is important for the initial dealers to

reduce their own direct exposure through the wholesale market in order to hold down their risk.

The analogy to wholesale and retail markets in the real economy is quite apt for significant parts of the activities undertaken by the major dealers. For example, auto dealers hold a large number of cars in inventory so that retail customers can walk in and buy cars off the lot if they want. Similarly, one of the key ways that securities dealers provide liquidity is by owning an inventory of specific securities that they believe might be of interest to their customers.

Structural Change in the Provision of Liquidity: Money Market Funds

In a prior section of this paper, we discussed the nature of liquidity provision and the theoretical and empirical analysis that has explored this issue. One important structural change in financial markets that has taken place in the past thirty years has been the growth of money market mutual funds that provide liquidity at both the retail and wholesale level and that experienced substantial stress in the financial crisis. These funds grew up almost accidentally as a result of financial regulation, specifically Regulation Q that limited the interest that banks could pay on deposits. Regulation Q aimed to provide additional stability to banks by preventing them from competing for funds by offering higher interest rates to depositors. This, it was feared, would lead them to make riskier loans in order to be able to pay the higher rates.

As interest rates rose in the 1970s and 80s, interest-sensitive depositors became unhappy with the low rates they were getting from banks and looked elsewhere for greater returns. Money market funds were started to satisfy this demand and, in an important regulatory change, they were allowed to provide checking accounts linked to these funds. Money funds were essentially providing demand deposits backed by holdings of commercial paper and this pulled deposits away from the commercial banks and savings and loans. Regulation Q was subsequently repealed and banks and S&Ls started to compete for deposits. The money market mutual fund industry continued to grow, however, and held $2.7 trillion in assets in 2012, with about 65 percent of the deposits held being from wholesale sources, largely corporate treasurers managing the inflow and outflows of money for their companies. These wholesale deposits are very sensitive to interest rate differentials and, as revealed in the crisis, will run very quickly if a fund threatens to “break the buck” when it is unable to redeem deposits dollar for dollar.

Although money market funds have effectively taken over a portion of the deposit base that traditionally resided in bank there is still a relationship between banks and money funds where the latter buy short-term securities issued by the banks. Short-term wholesale funding from money funds and other sources provided the means to purchase CDOs or other mortgage-backed assets in the build-up to the financial crisis.

Banks and other financial institutions still make use of money market funds as a source of financing.

In the crisis, the intermediation role played by money market funds created a serious dilemma for the authorities. When the first signs of trouble appeared, wholesale depositors started to run. $400 billion was withdrawn from these funds within weeks after the Lehman bankruptcy. Fearing that the run would escalate and eventually impact retail deposits, the Treasury and the Federal Reserve pledged to guaranty money market fund deposits 100 percent. Most wholesale and retail depositors had believed money market funds were safe and, in the event, they were because the government stepped in. Congress did not like this, however, and passed a law saying the Treasury and Federal Reserve could not enact such a policy in the future, effectively ruling out this expedient in future crises. In order to avoid future problems, the SEC has required money market funds to hold liquid assets that can cover large unexpected withdrawals and has recently proposed rules whereby prime funds (serving wholesale depositors) would not try to maintain a fixed net asset value and that there may be restrictions on withdrawals in the event that a fund is depleted.

The Industry and Regulatory Challenge in Liquidity Provision

The traditional role of financial intermediaries is to undertake maturity transformation, borrow short and lend long. They provide liquidity to the rest of the economy to satisfy the huge demand for assets that can be withdrawn almost instantaneously with no risk of capital loss. In periods of financial stress, as Diamond and Dybvig point out, the government is the only entity with the resources that can guaranty vast amounts of liquid assets. And in the case of small countries, like Iceland and Ireland, even the government cannot provide the guaranty.

In the wake of the financial crisis, policymakers around the developed world have been working to limit the responsibility placed on central banks and ultimately taxpayers to make sure depositors remain whole. One approach is to limit the ability of banks to carry out maturity transformation and provide liquid assets. The Basel III negotiators have developed liquidity rules forcing institutions to rely more on longer term funding. In Europe, the Cyprus crisis resulted in large uninsured bank deposits taking a haircut, a strategy that has been endorsed by the European Union. The danger in this strategy is that large depositors may run in the future. In the United States, Dodd-Frank and the regulators who are implementing its provisions are seeking to limit short-term funding by banks and to impose increased regulation on money market funds. In a recent speech, Federal Reserve Governor Daniel K. Tarullo talked of the dangers of short-term funding:

“Most importantly, relatively little has been done to change the structure of wholesale funding markets so as to make them less susceptible to damaging runs.” And later he

argues that there may be "a 'maturity rat race', in which each creditor acts to shorten the maturity of its lending so as to facilitate quick and easy flight…."[12]

For money market funds, the SEC has already mandated that money funds must hold 10 percent of their assets in Treasury securities and have 30 percent in a form that can be very quickly redeemed. Additional regulations are being considered to require such funds to hold capital or to limit the speed of deposit redemption. With Regulation Q long gone, there is a question of whether the money market fund industry would be able to survive if some of the proposed regulations are enacted.

A Comprehensive Review of Liquidity Provision Is Needed

The aim of Dodd-Frank was to seek out areas where the crisis had revealed stress points and enact rules to make these parts of the system safer. A motivation for this paper is to look at the functions of the financial system and how the structure of the industry has evolved to provide these functions. To the extent that Dodd-Frank rules affect these functions, it will trigger a restructuring of the industry and perhaps a re-pricing and a reduced level of provision of the services that are now subject to tighter regulation. The market shows there is a high demand for liquid assets, although of course that could change if pricing changes. There is also a tendency for market participants to believe that certain assets are perfectly safe, even if there is no explicit government guaranty for them, an assumption that has in practice been validated by *ex post* policy actions. The importance of liquidity in the economy strongly suggests that regulators should proceed cautiously in order to see how industry structure changes as a result of implementation and determine how drastically any cutback in the provision of liquid assets impacts real economic activity.

The interaction between capital requirements and liquidity provision is something where our understanding of the issues is evolving. Influential academics and policymakers have argued that banks should be required to hold much higher levels of capital as a buffer against adverse shocks. The counterargument is that much higher capital levels will force a reduction in lending and other financial activities. In an important, but as yet untested paper, DeAngelo and Stulz (2013) argue that banks should be highly levered because of their importance in the provision of liquidity.[13] As long as the market values liquidity, it is optimal for banks to be highly levered. The argument makes intuitive sense and they make a good case that the traditional

[12] Daniel K. Tarullo, "Evaluating Progress in Regulatory Reforms to Promote Financial Stability." Speech at the Peterson Institute for International Economics, May 3, 2013. Available at the Federal Reserve website. The quotations are from Page 4 of the speech, as printed.

[13] Harry DeAngelo and Rene M. Stulz, "Why High Leverage is Optimal for Banks," Working Paper, Ohio State University, April 12, 2013.

Modigliani-Miller theorem does not fully apply to banks because of their unique position as providers of liquidity. If they are correct, higher capital will be costly to banks and hence to the process of liquidity creation. Time will tell if their argument is validated, but their findings reinforce the conclusion that regulators should be measured as they take steps to raise capital requirements. In a recent paper, Elliott (2013)[14] sets out other reasons to believe that funding with equity capital is more expensive for banks than debt funding, including the differential taxation of debt and equity and the agency and signaling problems associated with issuing new equity.

Risk management products. Major banks are the largest providers of risk management products to businesses. For comparison, the $27 trillion of gross market value of derivatives globally in 2011 is higher than the roughly $15 trillion of insurance reserves globally[15]. (These figures are not directly comparable, but they make clear that financial risk management using derivatives is a very important activity, of a broadly similar scale to the insurance industry.) Of this $27 trillion of derivative value, a large percentage resides with the largest U.S. banks.

Some derivatives, generally the most standardized and liquid futures and options, are traded on exchanges. However, many derivatives historically have been traded on a bilateral basis, where at least one side of the transaction is a derivatives dealer. For example, an auto company may wish to hedge a financial risk, such as the possibility that the euro will appreciate at the same time that European interest rates will rise, squeezing profitability of its operations over there. There is no exchange-traded derivative that would easily accommodate this combined risk management issue, but a dealer would be willing to construct a customized derivative to meet this need. Customers may also transact on a bilateral basis, even on a standard contract, in cases where they felt they could get a better price through a dealer. This is particularly likely if they wish to transact in a large enough volume that it would affect the price on an exchange were it executed there.

The major banks are important participants in the exchange-traded derivatives markets as well. Much of this is simply executing activity from their clients, in cases where an exchange-traded contract is the best option. Beyond that, some of this activity is for

[14] Douglas J. Elliott, "Higher Bank Capital Requirements Would Come at a Price," The Brookings Institution, February 2013, http://www.brookings.edu/research/papers/2013/02/20-bank-capital-requirements-elliott.

[15] There are approximately $4.7 trillion of reserves on insurance policies in the U.S., according to the ACLI 2011 Life Insurer Fact Book and Morgan Stanley Equity Research for the property/casualty industry. North America makes up about 30% of the world insurance market according to figures from CEA (Insurance Europe) and the U.S. is the vast bulk of the North American market.

their own risk management purposes as large financial institutions and some of it is for laying off of risks taken on for customers. For example, the transaction with the auto maker described above might be passed along in part to the exchange-traded markets by buying futures or options on the euro and/or European interest rates.

What is the right size for the financial sector?

Most policymakers and analysts believe that the financial sector grew to be too large and this certainly matches public sentiments. We have some sympathy with this view. However, caution is in order, as it is extremely hard to determine the right size of the financial system based on well-grounded economic theories. Developed nations generally rely primarily on straightforward market forces to determine the size of any industry. This effectively assumes that whatever size develops from the interactions of the businesses and individuals in the economy will be appropriate in the long run, even if swings in sentiment may cause sectors to over-expand or contract too far in the short run. Experiments in economies with more central planning have ranged from sub-optimal to disastrous.

Nonetheless, there remain reasonable grounds to argue that finance grew too large. At the intuitive level, it is easy to view the big increases in the relative importance of finance over recent decades, and particularly in the last decade, as evidence that the sector over-expanded. Since the economy as a whole functioned reasonably well in those earlier times, one might conclude that the expansion was unnecessary. On an even more basic level, many people simply cannot believe that we need one worker in twenty to be involved in an activity that seems ancillary to "real" activities.

Arguments of this nature have been made throughout history to attack changes in the composition of the economy. For example, it was inconceivable to many that manufacturing could grow to employ far more people than farming, given the central importance of food to life. Going back further, Confucian theory put merchants well below peasants and other classes in their importance to society, since they produced nothing, but merely shifted ownership. Today, of course, we view merchants and other businesspeople as essential to the economy.

A more analytical extension of the intuitive arguments is the contention that a large portion of the growth of finance consisted of activities that are not "socially useful." Again, we agree that some activities that were created or substantially expanded during the bubble times were socially useless or harmful. Again, though, it is very tricky to prove this and particularly to sort out the wheat from the chaff. We simply do not have solid enough theories of the financial sector to nail down the value added or subtracted from each type of transaction. It may be akin to the famous saying that "half of all advertising dollars are wasted; we just can't tell which half." That said, there are certainly extreme cases, such as some of the wildly complex and opaque securitization

structures that grew up in the bubble, where few would argue in their defense. The clear cases, though, have all disappeared now and therefore have less relevance.

The best theoretical arguments that the financial sector grew too large revolve around policy distortions. To the extent that government policy provided a subsidy to financial activity, one would fully expect those activities to expand beyond what would occur without the subsidy. Therefore, any such subsidy that we view as inappropriate, such as implicit and free government guarantees of financial institutions that were viewed as Too Big to Fail, pushes the financial sector to be overly large. However, there are two problems with concluding that distortions make the sector too large. First, there is considerable subjective argument about which policies represent a "distortion" and which are appropriate for some societal purpose. For example, government guarantees of student loans expanded the financial sector, but also served a social purpose. Second, there are also government actions that increase the cost and decrease the attractiveness of conducting financial activities, in particular in the regulatory realm. One has to balance out the effects in each direction in order to conclude what the overall impact is of non-market factors such as government policy.

In sum, we find it easy to believe that the financial sector grew too large in the bubble period and possibly starting even earlier. However, we suspect the excess was considerably less than many argue and we believe it is important to be cautious in drawing policy conclusions as it seems impossible to prove whether the sector was or is too large and by how much.

Key structural proposals

There are a number of important proposals to force major changes in the structure of the financial industry. The remainder of this paper gives our views on proposals to:

- Eliminate Too Big to Fail banks by forcing their break-up or downsizing
- Limit the functions of banks á la Glass Steagall or the Volcker Rule

Proposed remedies for Too Big to Fail financial institutions

Banks that are central to our financial system, whether through sheer size or the critical nature of the services they provide, are perceived by many to benefit from an implicit government guarantee. The actions that governments around the world took to preserve their financial systems during the recent financial crisis give these views added weight.

We agree that the most important financial institutions in the U.S. significantly benefitted during the financial crisis from the clear intent of the government to protect against a major bank failure that might spread panic, having seen the disastrous effects

of the Lehman Brothers failure. Further, it seems likely that the major firms benefitted before the crisis from a perception that such a safety net might exist, although the overall level of concern about risk in the financial system was so low that the impact was likely relatively small. (One of the problems in trying to measure the subsidy value of implicit government guarantees of this nature is that it is difficult to differentiate between over-optimism in the markets with resulting underpricing of risk, a belief that the diversification and professional management of major banks made them low risk, and assumed government support.) There is considerable controversy as to whether such implicit subsidies still exist, how much they are worth on a gross basis, and the extent to which they are offset by the regulatory and other disadvantages that they now encounter as a result of their size. The industry largely argues that the regulatory burdens now outweigh any remaining guarantee benefits, while those who want to break up or limit the banks maintain that implicit subsidies are still big.

There are four main remedies proposed by those who believe that subsidies for Too Big to Fail banks remain large and are distorting the markets.

Break up the largest banks. Some argue for forcing the largest banks to break up into several smaller pieces. The transition details would be very complex, but the core idea is simple.

Mandate a size limit. Federal Reserve Board Governor Daniel Tarullo, among others, has offered for consideration ideas for placing size limits on the largest banks. A number of the proposals would essentially cap growth by these firms, while others would combine a cap with a milder form of break-up, where the maximum size would not be much lower than the current size of the largest banks. It should be noted that there were already long-standing limits on bank acquisitions and mergers that, with certain exceptions, block banks from having more than 10% of the national deposit market, except by achieving this through organic growth, as opposed to M&A activity. Dodd-Frank extended these limits to a wider range of liabilities than just deposits.

One of the problems with size limits is that they would actually make it much harder to resolve failing institutions of a certain type. The FDIC has developed over time a process by which failing banks are restructured and then sold to another bank. Moderately-sized regional banks have been through this process many times and it happened several times in the crisis. What would happen to a failing bank that was around the size limit set by the Federal Reserve? The FDIC would not be able to sell the bank to another entity and would be stuck with trying to support the bank until it could be restarted, or else they would have to close the bank with potential disruption or systemic effects.

Push large banks to shrink voluntarily by imposing stiff costs for size. The newly proposed Brown-Vitter bill in the Senate would impose much higher capital requirements on the biggest banks, which industry analysts generally believe would

effectively force the largest banks to break up or shrink dramatically. Other proposals have been floated, including the global agreement to impose a capital "surcharge" on Systemically Important Financial Institutions (SIFIs).

Put in place a credible plan for resolving the largest institutions. The three remedies considered above all address the problem of banks being too big but another alternative is to find a credible way for them to fail. If large global banks can be put through the FDIC's resolution process or, better yet, be placed into bankruptcy, with the costs of resolution borne fully by the equity and debt holders, and without the systemic disruptions that occur with a disorderly liquidation process, then any subsidy they may have at present will disappear.

Pioneering work by the FDIC and the Federal Reserve has opened up this possibility by utilizing the single point of entry approach (SPOE) to achieve such an orderly resolution.[16] The Financial Regulatory Reform Initiative of the Bipartisan Policy Center (BPC) released a report on May 14, 2013, which lays out a path to deal with failing large global banks (G-SIFIs) building on the FDIC's SPOE approach.[17] The first key is to require that G-SIFIs hold an adequate amount of going- and gone-concern loss absorbing capacity in the form of equity and long-term structurally subordinated debt at the bank holding company (BHC) level. These equities and liabilities of the bank then provide a loss-absorbing capacity. If the institution gets into trouble, the FDIC is appointed receiver under Title II and places the top tier holding company into receivership and then transfer most if not all of the holding company's assets (including operating subsidiaries) to a new "bridge holding company (the single point of entry) and removes the management deemed responsible for the failure and all the debt and equity held at that level. All the losses of the company will be borne by these debt and equity holders and because the debt is of at least a year's maturity, the debt holders will not be able to run from the bank at the first sign of trouble.

All of the assets of the holding company and the subsidiaries are then placed under the control of a bridge holding company, or Newco. Since Newco has no obligation to pay the debt and equity holders of the original holding company, it will be fully capitalized except under the most extreme circumstances. The subsidiaries are not in general placed into bankruptcy, and this includes foreign subsidiaries, which greatly simplifies

[16] The Clearing House in its Banking Brief White Paper Series, *Ending "Too Big to Fail": Title II of the Dodd-Frank Act and the Approach of "Single Point of Entry" Private Sector Recapitalization of a Failed Financial Company,* explains how the SPOE approach of the FDIC would operate to resolve financial institutions.

[17] Financial Reform Initiative, Bipartisan Policy Center, "Too Big to Fail: The Path to a Solution," May 14, 2013, Washington DC. The principal authors of the report are Randall Guynn, John Bovenzi and Thomas Jackson. Available on the BPC website: http://bipartisanpolicy.org/library/report/too-big-fail-path-solution

cross-border issues. Since the subsidiaries are still open for business the morning after the BHC has been placed into resolution, the impact on the financial system will be minimized. However, it is very likely that Newco would have trouble securing private liquidity financing immediately (debtor-in-possession or DIP financing) and so the government, through the Orderly Liquidation Fund established by Title II or through the Federal Reserve lending window, would have to provide short-term secured funding to ensure the continued operation of the subsidiaries.[18] A restructured institution would over time be returned to the private sector. The debt and equity holders of the original institution would be able to receive any residual value from the estate of Newco. They would be better off under this procedure than under the usual result of bankruptcy because the going-concern value of the company would have been preserved. Of course, the restructuring process could include selling off subsidiaries and/or closing some down in an orderly fashion.

The BPC report also argues that this same procedure could be followed by a bankruptcy judge with specialized knowledge of the financial sector. The FDIC and the Federal Reserve would have to be a party to the bankruptcy of the holding company because of their role as an insurer of deposits and to approve the provision of liquidity funding. There are some revisions to the bankruptcy code that should be made to deal with G-SIFIs, but even the current code would allow the SPOE process. Section 363 of the code allows for an immediate sale of the asset of a bankrupt holding company to a bridge Newco. The report argues that the Orderly Liquidation Authority in Title 2 of Dodd-Frank should be brought into alignment with the bankruptcy provisions in Title 1 and the living will process should provide the blueprint for either approach. The normal resolution path should be through bankruptcy, with the Orderly Liquidation Authority used only in extreme circumstances.

This brief description of the proposed resolution process for G-SIFIs does not do justice to the complexity of the problem, but at the same time it points to a way forward to solve the Too Big to Fail problem without breaking up the large banks or forcing them to meet size limits. Title II provides the tools necessary to resolve a U.S.-based G-SIFI without imposing costs on taxpayers. Moreover, the bankruptcy code could be amended to facilitate a SPOE approach and could then be just as effective as Title II.

[18] The OLF is for liquidity purposes only and funds advanced through OLF would not be used for recapitalization so there is no government bailout; in fact, title II specifically prohibits taxpayer funded bail out. Liquidity would be secured against the assets of Newco, would be subject to haircuts according to the riskiness of the collateral and would require a penalty interest rate on the loan. Such provision of temporary liquidity is consistent with historical role of the government as lender of last resort in times of extreme crisis in order to mitigate against panic, as described, for example, by Bagehot.

Economies of Scale and Scope

Proponents of breaking up the big banks or forcing them to shrink sharply generally believe that there is little economic reason for a financial institution to be very large, other than to take advantage of unfair benefits of size, such as the potential to be viewed as Too Big to Fail and thereby benefit from an implicit government guarantee. There are two fundamental ways in which any firm may benefit from large size. First, they may find "economies of scale", meaning that it is cheaper per unit of activity to do a lot of something than a little. For example, there may need to be a very considerable cost to create a computer system or to build an industrial plant of an efficient size, but ramping up production can spread the fixed costs over more units. There are areas in banking that clearly work in this manner, particularly transactional activities that require major investments in computer systems. But, the benefits can show up outside of information systems. It takes a certain scale of activity before it becomes efficient to hire an expert on a certain market or particular type of loan or investment, and loan officers are more efficient if they have more customers in a specific industry or geography.

Second, there can be benefits from conducting a range of activities, known as "economies of scope." Gas stations have become convenience stores because of the additional sales possibilities, just as airports have become malls. In finance, there are clear advantages to being able to offer a wide range of services, which are easiest to illustrate with regard to corporate customers. For instance, one reason that banks value mergers and acquisitions assignments so much is that they produce a wide range of related transactions which the bank is then in the best position to provide. There will be a need for financing, perhaps foreign exchange transactions, risk management products such as derivatives, and various operational services. It is most convenient for everyone to have the bank provide many of these services in combination.

The original academic work on economies of scale in banking suggested that the benefits topped out at about $100 billion in assets, much smaller than the size of the largest U.S. financial institutions. More recent academic work has been considerably more sophisticated and generally finds economies of scale going much higher. The Clearing House Association, an industry body, issued a study that surveyed existing research, concluding that large banks in the U.S. provide a series of important economic benefits that would not be achievable to the same extent by smaller banks. [19] The study found that the existence of big banks created $20 to $45 billion a year in benefits from economies of scale, figures they reached by looking in detail at various areas of banking where one would expect scale to matter. Similarly, benefits of a wide scope of product offerings at these banks and their affiliates produced another $15 to $35 billion a year of economic benefits. Finally, benefits that big banks provide through encouragement of

[19] The Clearing House Association, "Understanding the Economics of Large Banks," November 7, 2011.

financial innovation came to another $15 to $30 billion, bringing the total to $50 to $110 billion a year. The study argued that these findings were relatively conservative, pointing to a finding by Wheelock and Wilson of the Federal Reserve Bank of St. Louis[20] that even capping banks at $1 trillion in size, a relatively high limit compared to some proposals, would result in a loss of $79 billion in benefits.

The Wheelock and Wilson study, and a Federal Reserve of Philadelphia working paper by Hughes and Mester[21], find economies of scale going up to much larger sizes than earlier studies. This is partly because the newer studies are substantially more sophisticated and partly because it appears that economies of scale in U.S. banking have actually grown as technology and its application have improved and geographic and other legal barriers have been reduced. (The earlier studies are mostly a decade or so older.)

However, both these newer studies and the older ones understate the underlying economies of scale by treating compensation costs as mandatory, whereas there is an element of profit-sharing in them. Ronald Anderson attempted to estimate the economic "rents" captured by the managers of financial institutions and concluded that they have been substantial.[22] Some of the benefits of economies of scale and scope have presumably been siphoned off by managers, but they are nonetheless economic benefits that need to be considered. The various economic pressures being placed on the financial industry in response to the financial crisis are likely to squeeze much of these rents back out, reallocating them to shareholders and customers.

Our Views

We do not favor the proposals to break up the banks or force them to shrink dramatically. We believe that the best analysis indicates considerable economic benefits to size and scope and that these advantages are likely to grow further with increasing globalization, complexity, and improved information and management systems. America should have at least a few financial institutions with global scale, capable of providing a wide range of related commercial and investment banking services, operating on a scale in individual product lines that produces real efficiency.

---

[20] David C. Wheelock and Paul W. Wilson, "Are U.S. Banks too Large," Federal Reserve Bank of St. Louis Working Paper, December 2009.

[21] Joseph P. Hughes and Loretta J. Mester, "Who Said Large Banks Don't Experience Scale Economies? Evidence From a Risk-Return-Driven Cost Function," Federal Reserve Bank of Philadelphia Working Paper, July 2011.

[22] Ronald W. Anderson and Karin Joeveer, "Bankers and bank investors: Reconsidering the economies of scale in banking", September 2012, available at http://personal.lse.ac.uk/ANDERSOR/bankscale%20posting%20130912.pdf

This will almost certainly mean these firms are important enough to the economy that the government and regulators will need to watch them particularly carefully and may create need for special assistance, in extreme crisis situations of the level that are unlikely to occur more than once or twice a century. For this reason, we agree on the need to designate systemically important financial institutions and to require them to operate with higher safety margins.

We believe that the societal benefits of breaking up the large banks are over-stated. The recent financial crisis was much more about system-wide problems than about issues resulting from excessive size of financial institutions. A simple thought experiment illustrates this. If we had broken up the big banks a decade ago into 10 or 20 pieces each, they would likely all or virtually all have made the same mistakes. They would have over-invested in real estate-related products, taken excessive risks across the board, created opaque and risky securitizations and derivatives products, pushed accounting rules to their limits, etc. The other players in the financial system would presumably also have made the same mistakes, including the ratings agencies, governments, central banks, regulators, and families and businesses. It is difficult to presume that the disaster would have been much different. Indeed, there is a chance that the clean-up would have been more difficult without the ability to pull 17 key CEOs into a room and force them to accept the TARP arrangements.

The next financial crisis will almost certainly differ from the last, as every such crisis varies, but it remains difficult to see how a system of many mid-sized banks would be appreciably safer than one with some large banks as part of the mix.

Limitations on mixing commercial banking and securities and derivatives business

U.S. commercial banks and their affiliates have always faced limitations on the business they are allowed to undertake, in order to reduce the risk of business disasters that would endanger their ability to fulfill their critical role at the heart of the economic system. For example, it is relatively uncontroversial in America that banks are allowed to undertake only quite limited business activities that are not clearly “financial” in nature. This differs from some other countries, where banks have owned significant manufacturing, retail, or other businesses.

In the Great Depression, the Glass-Steagall Act was passed, making it illegal for a commercial bank to be affiliated with an investment bank. The former could undertake the types of activities we normally associate with banking, such as taking deposits and lending. The latter were principally involved in the securities business, through helping

firms raise capital by selling stocks and bonds, assisting investors in buying and selling those securities, and trading them for the investment bank's own account.

The anti-affiliation provisions of Glass-Steagall were dramatically modified in the 1990's, allowing commercial and investment banks to be part of the same financial group, although there remain a number of important restrictions to limit dealings within the group. These are intended to insure the commercial bank does not directly engage in most investment banking activities and does not enter into transactions with its investment banking affiliate except under arms-length terms and sometimes not even in that case. The idea is to prevent the commercial bank from acting as a giant piggy bank for the rest of the group, exposing it to excessive risk.

There is a range of proposals to further limit the ability of banks to operate in the securities and derivatives businesses. Some call for a restoration of the anti-affiliation provisions of Glass-Steagall. Others want Glass-Steagall Lite, since they recognize that changing times make it difficult to simply turn back the clock. The Volcker Rule is intended to separate out proprietary trading completely from commercial banks *and* investment banks, with the view that it is so pernicious that it ought not to be undertaken by important financial intermediaries. In the U.K., the Vickers Commission has recommended a "ring-fencing" approach that would function somewhat like the present version of Glass-Steagall in the U.S., although it will probably have more limitations on interactions within the financial groups than is true in the U.S. In the European Union, the Liikanen High-Level Expert Group has recommended its own version of ring-fencing.

Our Views

We do not favor any of the major proposals for further structural divisions between commercial banking and securities and derivatives activities. We believe that the U.S. capital markets are world leaders and that their strength is an important economic advantage for America. Those markets are underpinned by the role of major securities dealers that are closely affiliated with commercial banks. A major reason for the close linkages is the desire of corporate customers to be able to deal with financial firms that can provide a solid range of products from financial advice to loans to securities offerings to risk management via derivatives to purely operational products. The institutional knowledge and relationships that a banking group has in regard to its corporate customers is a valuable advantage both for the bank and for those customers.

Further, times have changed and will not change back. Glass-Steagall was based on a clear difference between a loan and a security, a difference that no longer exists now that most large loans are tradable among banks and also specialized investors. At this point, it is usually possible to structure a given transaction as a loan or a security or a derivatives transaction or often as insurance or another contractual arrangement.

Finally, any transition from the current system to an older-style system will create very considerable displacement of activities, with a real potential for problems. Some of this might occur through the divestiture of investment banking subsidiaries from banking groups, which would be the simplest approach; however even this would involve a large amount of change at a time when the U.S. economy remains in a fragile recovery that resulted in part from the disruption of the financial sector. Another source of displacement would result from striving mid-level securities firms grabbing market share. Although this could bring advantages, it also creates the danger of a repeat of a situation such as developed at MF Global, where the push for growth overcame proper risk management practices.

Conclusions

In the aftermath of the financial crisis there was a perceived urgency around enacting reforms that would ensure that such a crisis would never happen again. In that spirit the Administration and legislators looked at the markets and institutions that had experienced difficulty and sought ways to regulate them so they would not pose a threat in the future. Regulators have sought to implement the provisions of Dodd-Frank following the intent of Congress, perhaps erring too much on the side of caution.

Inevitably, such a big change in the rules of the game will bring about substantial structural change in the large and diverse financial industry. In anticipation of this, in a conference held at Brookings on December 4, 2012, a distinguished group of academics and practitioners gathered to look at how the industrial organization of the industry might be affected by Dodd-Frank. Governor Daniel Tarullo, the lead regulator at the Federal Reserve, spoke about his ideas on this topic and called for further study. This paper is intended to contribute to that research goal. We have looked at the different functions performed by financial institutions and how they may be changed as a result of regulatory reform. One major conclusion is that we do not support the case for breaking up the large banks, for several reasons. We do not think that size was an important cause of the crisis, which was instead the result of risky lending and buying risky assets by both large and small institutions, most of the bad assets tied to real estate. We think there are significant advantages to having both large and small financial institutions in the mix of the financial sector. And finally, we judge that substantial progress has been made in figuring out how to resolve or put into bankruptcy the largest institutions without triggering a broader crisis and without placing a burden on taxpayers.

This paper represents only a small step in the larger debate that should take place around the impact of financial regulation on the financial sector. This debate should be guided by the principle of balancing the crucial importance for economic growth of the provision of financial services with the vital importance of preserving the safety of the system.

Appendix

How are credit exposures created in providing these products?

Banks take on credit exposures in a wide variety of ways as they provide the economic functions of credit provision, liquidity provision, and risk management. This section outlines the major ways in which the exposures arise.

Long-term credit extension. Banks are major holders of loans and debt securities directly on their balance sheets, creating credit exposures. While some of these represent inventory for securities dealing, substantial amounts represent deliberate holdings taken on because the bank likes the economics of the underlying credit provision. In addition, committed lines of credit also produce exposures.

Facilitation of initial credit extension. A bank will often agree to take on a much larger credit risk to a customer than it is willing to retain, as an accommodation to that customer or in order to make additional revenues by handling the full amount. If the credit risk is passed on to others through purchases of credit protection, rather than an outright sale of that portion of the loan, then credit exposures will exist.

Securities inventories held to provide liquidity to customers. The inventories of securities banks hold to facilitate customer transactions create credit exposure. In addition, a bank, particularly an investment bank which has less access than commercial banks to cheap deposit funding, may choose to finance these positions by entering into repurchase agreements (repos), which produce credit exposures, usually to other large financial institutions.

Securities held for a bank's own liquidity management. Commercial and investment banks need to own a significant amount of short-term, highly creditworthy securities in case they need to raise cash quickly to cover deposit withdrawals or for other purposes. These securities create credit exposures.

Facilitation of customer securities sales exceeding desired inventory levels. Credit exposures can be created when a dealer buys more from a customer than it is comfortable keeping in inventory. The dealer may purchase credit protection and may finance the holdings through repos, both of which create credit exposures to the relevant counterparty.

Facilitation of customer purchases of securities exceeding existing inventory. Sometimes customers want to buy securities that are not held in inventory or are not held in sufficient quantity. The natural first response would be to locate the additional securities elsewhere and buy them. However, there may be a portion, or all, of the trade

which can only be accommodated by borrowing securities from another party and using them to "sell short" to the customer. Securities borrowings create credit exposures.

Facilitation of securities offerings, including government bond issues. Another important function of investment banks is to aid businesses, governments, and other entities to sell debt, equity, and other securities. Generally a bank will buy the securities initially, perhaps in a syndicate with other banks, and then distribute them to interested buyers. The initial direct effect is to create gross credit exposure on the securities inventories. This exposure is quite likely to be hedged through transactions such as credit default swaps that transfer credit exposure to the protection provider.

Most governments around the world rely on large banks as the initial purchasers of a major share of government bond offerings. In the United States, 21 "primary dealers" are offered the opportunity to bid directly when the Treasury conducts its securities auctions. In exchange, primary dealers are expected to participate in all auctions and to meet informal targets for their purchase volumes in order to retain their status as primary dealers. Thus, a key role for the dealers is to take large amounts of U.S. Treasury securities onto their books, distributing them over the course of time, usually within a few days. These dealers rely on the ability to hedge and to finance these temporary inventories, with short selling, derivatives purchases, and repo transactions, all of which create credit risk, usually to other large financial institutions.

Securities lending. Securities lending is quite helpful in increasing credit provision and market liquidity in various ways. It has a role in supporting securities inventories at the banks and supplying funding for hedge funds and similar credit investors. In addition, securities lending assists overall market liquidity by creating the ability for banks and other market participants to sell short, since this requires the ability to borrow the shorted security from someone. Banks may also act as an agent for a securities lender, which often requires them to provide a guarantee of the credit risk of the borrower. All of these uses of securities lending create credit exposures.

Custodial services. A small number of banks perform a variety of critical custodial and related services for the securities markets. Their fundamental job is to ensure that securities are safely held in the accounts of their owners and that securities and cash move as promised when trades, repurchase agreements, securities lending, and similar transactions are entered into. Thus, they are involved in very large flows, which can require them to bear credit risk for very short periods of time. Much of this activity is intraday and is explicitly exempted for that reason in many regulatory calculations of credit exposures. However, portions of the activity spill over beyond one day, either intentionally or unintentionally, creating credit exposure that fall under regulatory constraints.

Offerings of risk management products. Most of the risk management products offered by banks are in the form of derivatives transactions. Traditionally, banks have

primarily dealt on a bilateral basis, which leaves them with the credit risk that the counterparty will not perform as promised. An increasing share of derivatives transactions will be taking place using a Central Clearing Party (CCP), sometimes known as a "clearinghouse." Thus there are credit exposures to CCPs similar to derivatives exposures to a counterparty in a bilateral trade, since they are conceptually identical.