Why the Volcker Rule Is a Useful Tool for Managing Systemic Risk

By Matthew Richardson

NYU Stern School of Business

Summary:

An optimal regulation policy for financial institutions would be to have those firms internalize the costs of systemic risk that they produce. An outcome of such regulation would lead to these firms organically choosing to be less leveraged and holding less risky asset positions. In lieu of this approach to financial regulation, a natural alternative would be to directly impose capital requirements and restrictions on asset holdings. The regulation enacted during the last major financial crisis in the 1930s, namely the Banking Act of 1933 ("Glass-Steagall" Act), followed this path. So too does the Dodd-Frank Act with its emphasis on capital regulation and the Volcker rule.

This white paper explains:

1. why from an economic point of view the Volcker rule makes sense as a regulatory tool.
2. why capital requirements are not a panacea for financial regulation and instead the need to supplement them with Volcker rule type restrictions. Specifically, certain types of asset holdings, such as those prone to regulatory arbitrage, carry trades and financial guaranty insurance, are ill suited for capital regulation.
3. why some of the recent arguments regarding the impact on capital market liquidity of the Volcker rule actually provide support, not opposition, for the Volcker rule.
4. why the argument that the Volcker rule limits diversification and increases systemic risk is a poor one.
5. why the Volcker rule needs to be principle-based with safe harbors as opposed to a strictly rule-based approach. The reason relates closely to the current debate on what constitutes principal trading versus market making.

Matthew Richardson, Charles E. Simon Professor of Applied Financial Economics, Director of the Salomon Center for the Study of Financial Institutions, NYU Stern School of Business, and consultant to Morgan Stanley. Some of the ideas for this white paper come from NYU Stern’s book on analyzing the Dodd-Frank Act of which I was one of the four co-editors, "Regulating Wall Street: The Dodd-Frank Act and the new Architecture of Global Finance,” in particular, Chapter 7 “Large Banks and the Volcker Rule” (co-authored with Roy Smith and Ingo Walter). This chapter describes the history of Glass-Steagall and how to consider the Volcker rule in this context. While I have been a consultant for a trading group within Morgan Stanley for 7 years, neither this white paper nor NYU Stern’s other work on the financial crisis has been reviewed or submitted to any officials at Morgan Stanley.
I. The Economics of the Volcker Rule

The economic theory of regulation is very clear. Regulate where there is a market failure. It is apparent that a major market failure in this crisis was the emergence of systemic risk. More concretely, systemic risk emerged when aggregate capitalization of the financial sector became low. The intuition for why this is a problem is straightforward. When a financial firm’s capital is low, it is difficult for that firm to perform financial services; and when capital is low in the aggregate, it is not possible for other financial firms to step into the breach. This breakdown in financial intermediation is the reason severe consequences occurred in the broader economy. When financial firms therefore ran aground during the crisis period, they contributed to the aggregate shortfall, leading to consequences beyond the firm itself. The firm has no incentive to manage the systemic risk. The negative externality associated with such risks implies that private markets cannot efficiently solve the problem, so government intervention is required.

The Dodd-Frank Act missed the opportunity to fix the negative externality of systemic risk by imposing the standard economic solution of taxing such an externality.\textsuperscript{2} In other words, there should be an additional fee, surcharge or premium that is tied specifically to the systemic risk of financial institutions as these systemic failures bear costs on the rest of the financial sector and the real economy (e.g., Acharya, Pedersen, Philippon and Richardson (2010)). In order to avoid these fees, these firms would organically become less systemic by choosing less leverage and holding assets that have less aggregate tail risk.

Instead, the Dodd-Frank Act seeks to regulate systemic risk by setting up a number of barriers, two of which are particularly important, namely enhanced capital requirements and the principal trading restriction of the Volcker rule. This way, the Dodd-Frank Act can be considered a second-best solution by hard wiring leverage constraints and restrictions on risk-taking activities.

Does history tell us anything about whether this is a reasonable approach to regulating the financial sector? In the earlier part of the century, e.g., during the Panic of 1907 and the various banking panics between 1930-1932 in the wake of the Great Depression, uncertainty and lack of information about which financial institutions were insolvent led to system-wide bank runs. In response to these systemic runs, the government created the Federal Reserve with its lender of last resort facility, the FDIC and deposit insurance, along with a number of banking and investment acts. Arguably, the most important part of the legislation was that depositors no longer had to run on the bank because the government guaranteed the funds.

Of course, it is well understood that this safety net creates a moral hazard, that is, an incentive for the bank to undertake greater risk than it would otherwise would without this insurance. Over time, regulators and policymakers therefore set up a number of countervailing barriers: (i) banks would have to pay to be a part of the deposit insurance system, so, at least, on an \textit{ex ante} basis, regulators took into account the cost of the insurance, (ii) the risk-taking activities of banks were

\textsuperscript{2} It is interesting to note that, in the House version of the precursor bill to Dodd-Frank, such a provision was included.
ring-fenced to the extent that there was a separation of the commercial and the more risky, investment banking activities, and (iii) enhanced supervision, and winding-down provisions, of individual banks, generally in the form of capital requirements and prompt corrective action was established.

One approach advocated by the Vickers Report in the U.K. would be to ringfence domestic retail banking services from global wholesale/investment banking with the ringfenced bank being legally and operationally separate with a higher capital requirement. Specifically, the Vickers Report (2011) concludes that “a number of UK banks combine domestic retail services with global wholesale and investment banking operations. Both sets of activities are economically valuable while both also entail risks – for example, relating to residential property values in the case of retail banking. Their unstructured combination does, however, give rise to public policy concerns, which structural reform proposals – notably forms of separation between retail banking and wholesale/investment banking – seek to address.”

The Volcker rule takes an alternative, yet related, approach to managing the risks of large complex banks by restricting certain types of trading activities. In this sense, the Volcker rule can be viewed as a sensible approach to mitigating moral hazard and preventing the systemic risk buildup across banking institutions and non-bank systemically important financial institutions (SIFIs) when these firms fall under the government safety net. In fact, Paul Volcker, the highly respected former Federal Reserve chairman, had long urged that the scope of any implicit federal guarantee be limited to a relatively small number of important banking institutions and to core banking functions, rather than extended across the spectrum of financial intermediaries and risky activities. In exchange for the banking safety net, Volcker recommended that banks be allowed to engage in the full range of commercial and investment banking functions as financial intermediaries, but not be permitted to engage in such nonbanking activities as proprietary trading, principal investing, commodity speculation, and hedge fund and private equity fund management. These activities could be spun off to nonbank asset management firms and would be subject to whatever regulation is necessary for those types of institutions. The legacy banks would be allowed to have no economic interest in the spun-off entities.

In other words, what sense does it make for financial institutions to be able to invest on their own account when these institutions can finance their activities at below-market rates due to the government safety net? These distortions occur not only at banks with access to FDIC insurance, but also with respect to the government sponsored enterprises like Fannie Mae and Freddie Mac and other too-big-to-fail large, complex, financial institutions.

A natural question is what is so special about principal trading and its contribution to systemic risk versus other activities housed at banking institutions? Indeed, given its aggregate risk component, a bank’s portfolios of loans is likely to do poorly in a crisis and add to the overall

---

5 The Volcker rule applies to nonbanking SIFIs by potentially subjecting these firms to additional capital requirements and quantitative limits on proprietary trading. For the rest of this white paper, I refer to “banks” but it should be clear that the arguments extend to SIFIs. Specifically, a number of the provided examples include non-bank SIFIs.
aggregate capital shortfall of the financial sector. These loan portfolios also contribute to systemic risk. That said, loans to households, small businesses and the like are core to banking and not easily replicable outside the banking sector (e.g. Fama (1985), Diamond (1991) and Petersen and Rajan (1994)). This is not true, however, of financial activities such as principal trading. Rather, these activities are performed in the capital markets at large by pension funds, mutual funds, hedge funds, sovereign wealth funds, etc... These funds are less levered and generally do not have access to the safety net. Therefore, because some non-core functions of banks can be performed elsewhere in the less systemically risky part of the financial system, the Volcker rule is an economically sensible way to manage the systemic risk of banks.

II. Capital Requirements Are Not Sufficient

Some analysts and academics have argued that systemic risk, irrespective of where it is housed in a bank, can be managed via capital requirements. For example, in a paper in direct response to the Volcker rule, Duffie (2012) argues instead for “rigorous capital and liquidity requirements for market makers, combined with effective supervisory monitoring, with the objective of ensuring that banks have abundant capital and liquidity to cover their market-making risks.” I disagree with this assessment. Unfortunately, capital regulation is not a panacea for managing systemic risk.

In trying to minimize systemic risk, the regulator can use the Dodd-Frank Act to effectively pull on two possible levers – either through greater capital requirements (i.e., less leverage) or through restrictions on asset holdings (e.g., using the Volcker rule). It is important to understand, however, that these levers are not perfect substitutes for each other. The regulator needs to pull both these levers, adjusting them for the type of financial activity the bank is engaged in. Specifically, certain types of asset holdings, such as those prone to regulatory arbitrage, carry trades and financial guaranty insurance, are ill suited for capital regulation.

Before discussing each of these activities in turn, generally, the reason why capital regulation fails is that it can be a highly inefficient way to manage systemic risk. To understand this point, note that the aforementioned activities generally produce small gains with a high probability against large losses with a low probability. In order to protect the financial sector against the emergence of systemic risk, regulators will require banks to hold sufficient capital to cover losses in the low probability state. Because these losses are generally large and occur only when aggregate risk is present, the necessary capital requirements will be large. But this would mean that banks would have to hold excess capital even though most of the time there is little or no probability of failure (e.g., Kashyap, Rajan and Stein (2008)). A more efficient approach to managing systemic risk, therefore, would be to pull on the Volcker rule lever and loosen the lever on capital requirements by effectively banning or restricting the level of these activities.

A. Regulatory Capital Arbitrage

Some of the examples provided below involve securities, such as governments and agencies, which are exempt from the Volcker rule. This does not change the structure of the economic argument in favor of the Volcker rule except to argue the rule may not have been broad enough.
One of the lessons of the financial crisis is the importance of capital requirements being consistently set across markets and institutions. In other words, if the risk of the underlying loans is the same, it should not matter how those loans are sliced and diced through securitization in terms of determining the required capital buffer of banking institutions. Unfortunately, however, Basel III and Dodd-Frank continue the risk-weights approach both within and across asset classes, as well as the internal ratings approach that many have forcefully argued against as a result of the crisis. A core problem of the financial architecture continues the focus on risk-weighted capital measures of individual firms as the main indicator. Thus, across-the-board higher capital requirements, as are being proposed for large banks and SIFIs, may actually exacerbate the problem. One of the problems that emerged in the financial crisis was the preferred capital treatment provided to certain asset-backed securities, such as AAA-rated mortgage-backed securities (MBS). With simply higher capital requirements for SIFIs, it cannot be ruled out that the preference for AAA-rated MBS or likewise securities will be even greater under a higher overall requirement, causing an even bigger buildup in aggregate systemic risk. Thus, capital requirements may be a particularly poor lever to deal with areas prone to regulatory capital arbitrage (e.g., Acharya, Nulkarni and Richardson (2010)).

With respect to the financial crisis, given large banks’ access to cheap financing, the banking sector ended up holding certain securities because, through regulatory loopholes, the warehousing of these securities required less regulatory capital and, as a result, the financial intermediaries were free to lever up to the hilt. These securities offered a spread over the financing rate precisely because they were less liquid and faced systemic risk.

This point is not just relevant for the financial crisis. A quick look at the 2009 balance sheets of the four largest banks - JPMorgan Chase, Bank of America, Citigroup, and Wells Fargo - shows holdings of $1.1 trillion worth of available-for-sale securities. While banks will argue these holdings are necessary for liquidity, if this were in fact the case, then they would be holding short-term Treasuries or cash. Instead, many of these available-for-sale securities are long-term governments and asset-backed securities funded using overnight repos. However, in the unlikely event that bad times occur and liquidity and market risk surface, these securities would lose value. Since there is little or no capital underlying these positions - and bank-type levered entities would already be facing trouble from loan losses - systemic risk emerges.

Commingling systemically risky security holdings with economically important financial intermediation at banks and other large financial institutions was one of the main causes of the recent crisis. This is why finance theory argues persuasively that the business model of securitization never intended asset-backed securities to be held on banks’ balance sheets, and especially not to skirt capital requirements.

B. Carry Trades

Financial firms, especially those with access to below-market borrowing rates via the government safety net, like to engage in so-called carry trades. Carry trades involve buying high yielding financial securities financed by a lower (usually short-term) yielding security. Examples of carry trades are riding the yield curve (e.g., buying long-term government-backed securities), taking small amounts of credit risk (e.g., buying AAA-rated securities), holding liquidity (e.g.,
buying less liquid securities with little or no credit risk), investing in tail risk (e.g., market risk that gets realized very rarely) and buying high interest rate currencies financed by low interest rate currencies. These trades are usually low risk in nature, paying off the yield spread in most states. Common to all carry trades, however, is a low probability of a systemically large loss.

To better understand these trades, consider as an example the business model of Fannie Mae and Freddie Mac. These firms invested approximately $1.5 trillion in bank-originated pools of mortgages at a very high leverage ratio. Because of the implicit government guarantee of their debt, these firms were able to take these bets at a low financing cost. It is now widely recognized that this model was a recipe for disaster, since it combined private profit taking with socialized risk. The banking sector during the financial crisis, however, looked almost identical to Fannie and Freddie. A Lehman Brothers report from April 2008 shows holdings of residential mortgage-backed securities of U.S. banks and thrifts. These holdings included $901 billion of agency securities and $483 billion of subprime AAA-rated securities, versus $741 billion and $308 billion held by Fannie and Freddie. And like Fannie and Freddie, these positions held by the banks and thrifts were funded at a lower cost of capital than the underlying risk because of either explicit government guarantees of bank deposits or implicit too-big-to-fail guarantees (via repo financing of the trading books of large, complex banks).

As another example of a carry trade, consider the recent financial crisis in the euro-zone area. Leverage to the hilt, many European financial institutions are broadly exposed to the debt of countries facing a significant probability of default. Because investors are aware of the euro-area banks’ leverage and sovereign debt exposure, depositors and other bank creditors are running towards safer countries (i.e., to Germany and non-Euro areas). In response, the European Central Bank (ECB) has provided three year “cheap” financing to most large Euro-area banks. These banks, however, are using the funds – encouraged by preferential capital treatment – to further invest in high yielding government debt, earning a positive carry spread. The problem is that, from a systemic point of view, we most worry about a sovereign debt default and a partial breakup of the euro-zone area. Of course, this outcome would simultaneously impose large losses on the banking sector’s carry trade on sovereign debt, resulting in even greater systemic risk. This illustrates the problem with banks taking on large-scale carry trades.

C. Writing Financial Guaranty Insurance

There are other financial products that mimic the properties of carry trades. The most obvious one is the writing of out-of-the-money put options on aggregate risk. Writers of put options receive a small premium each period in return for promising to payout when an aggregate risk is realized. Out-of-the-money puts simply mean that the aggregate risk is an unlikely event. Similar to carry trades, these out-of-the-money puts require large amounts of capital to be set aside because if the event occurs, the payout is large. It is suboptimal to house this type of financial guaranty insurance at large banks not only because of the higher capital requirements involved but also because undercapitalized losses can bleed into the rest of the institution.

A case in point. Banks can use OTC derivatives to tailor their own risk-taking and leverage build-up, since some of these positions are not reflected on their balance sheets, either from a regulatory or statutory disclosure perspective. Consider the following illustration. On page 122
of the quarterly filings of AIG from early February 2008, well before they were brought down, AIG describes their now infamous $527 billion dollar of credit default swap (CDS) positions by their AIG financial products group (AIGFP). AIGFP’s CDSs were written on AAA-rated securities and thus could only ever fail if extreme aggregate risk was realized. Why did AIGFP write such contracts? As stated in the document, “approximately $379 billion of the $527 billion in notional exposure of AIGFP’s super senior CDS portfolio as of December 31, 2007 represents derivatives written for financial institutions, principally in Europe, for the purpose of providing them with regulatory capital relief rather than risk mitigation.” If financial institutions held AAA-rated securities and bought protection on those securities from a AA- or AAA-rated insurance company, then these institutions could hold zero capital. Such rules possibly explain the huge leverage positions of UBS, ABN Amro and investment banks, like Merrill Lynch, among others.

Of course, this is not to argue financial guaranty insurance should never be written. If financial institutions outside the safety net and not systemically risky write such contracts, and buyers of these contracts recognize the counterparty risk, then such trades should be subject to the discipline of the markets. That said, the likely outcome would be that, given sufficient demand for such insurance, the insurance would be offered by highly capitalized institutions. In fact, in the financial crisis of 2007-2009, it became known that Berkshire Hathaway had written put options on the S&P 500, the FTSE 100, the Euro Stoxx 50, the Nikkei 225 and other indices. Over the years, Berkshire had collected billions in premiums albeit with over $35 billion in gross exposure 10 to 20 years out. When stock markets collapsed in the late fall of 2008 and winter of 2009, on a market valuation basis, Berkshire was exposed to heavy losses on these derivatives. Berkshire’s equity value fell partly to reflect these derivative losses. However, unlike the banking sector and other SIFIs, Berkshire was extremely well capitalized, thus producing much less systemic risk for the financial sector and overall economy.

III. The Impact on Liquidity of the Volcker Rule

One of the main criticisms of the Volcker rule relates to the rule’s impact on the liquidity of financial markets. Liquidity of secondary markets is clearly important. If issuers and investors know that there is an active secondary market, then the liquidity premium is reduced, providing for better pricing (e.g., lower rates for fixed income products). Skeptics of the Volcker rule have argued that an implemented Volcker rule will hamper liquidity either by restricting market making activities of banks or reducing trading volume by banning the principal trading of banks. The argument is flawed for three reasons.

First, and foremost, the issue of liquidity is most relevant during times of crisis. In normal times for active markets, liquidity is not as pressing an issue. Rather, the level of liquidity premiums and the most extreme form of market liquidity risk occurs when dealers shut down during severe times of stress (e.g., Scholes (2000) and Pedersen (2008)). This point actually provides support, not opposition, for the Volcker rule.
Second, some of the discussion regarding liquidity impact seems to imply that all trading by large banks and SIFIs reflect market making activities. We should not confuse the difficulty in measuring trading intent versus the fact that trading intent does exist, that is, some trades are driven through market making versus some through proprietary bets. Section V below describes an approach to implementing the Volcker rule which is less dependent on identifying precise differences between these activities. In particular, if the financial institution stays within the safe harbors of the Volcker rule, and chooses to manage its gross and net inventory of reasonably liquid assets within a fraction of its capital, then there would be enough leeway for the affected financial institutions to make active markets in most securities.

Third, some analysts and academics have argued that the Volcker rule would reduce the quality and capacity of market making services. For example, Duffie (2012) writes “non-bank providers of market-making services would fill some of the resulting void in market making capacity, but with an unpredictable impact on the safety and soundness of financial markets.” In the long run, it is not so clear that financial markets would not adjust to a new regime in which banks played a smaller role as broker-dealers. The evolution of the finance system post the enactment of Glass-Steagall provides an historical lesson which puts in doubt this Volcker rule criticism. Moreover, it is reasonable to question whether the dominance of banks as broker-dealers is due to their access to cheaper funding via the safety net as opposed to greater efficiency on their part. The former can induce excess liquidity and distort market prices.

In order to elaborate on these three reasons, as an example, I consider a recent report by Oliver Wyman (2011) (denote OW (2011)). The report argues that the Volcker rule will have a large negative impact on the liquidity of the corporate bond market. Their prime argument relies on research by Dick-Nielsen, Feldhutter and Lando (2011) (denote DFL (2011)). DFL (2011) find that the liquidity premium of corporate bonds increases dramatically with the onset of the subprime crisis, and, most important, these bonds become less liquid when financial distress hits a lead underwriter. OW (2011) extrapolates these liquidity effects to what might happen if the Volcker rule restricts market making in the corporate bond market.

First, OW (2011) misinterprets the real meaning of the DFL (2011) study and previous investigations of liquidity effects during crises. DFL (2011) show that large financial institutions in distress can no longer offer financial intermediation such as market making services during crises. Because these institutions, as lead underwriters, are key dealers in certain markets and securities, there is a huge liquidity impact on these markets and securities. For example, DFL (2011) write that “the liquidity of bonds underwritten by Lehman was close to the liquidity of an average bond in the market up until August 2008, but this changed when the ‘illiquidity gap’ between Lehman underwritten bonds and average market bonds increased strongly in response to Lehman filing for bankruptcy on September 15.”

Lehman did not fail because of their underwriting or market making activities. Instead, it was due to the combination of (i) $70 billion in commercial real estate and residential mortgage-related assets and private equity investments that had dropped in value, and (ii) high leverage with a heavy reliance on short-term, especially repo, funding. (See, for example, the FDIC’s (2011) study of the potential for an orderly liquidation of Lehman Brothers.) This shows the danger of concentrating key financial intermediation activities within large complex financial
institutions. Losses related to principal trading can bleed into other parts of the business causing financial disintermediation and, in this case, large liquidity effects. The Volcker rule would limit the losses from principal trading and thus make it less likely a Lehman Brothers would occur. Even if the losses were at a bank and had come from their loan portfolio, it shows the risks associated with having key financial functions performed by just a few financial institutions. By restricting principal trading, the Volcker rule would limit the reliance on any particular dealer by the market at large.

The size and power of large banks is worrisome. For example, in 2009, the world’s five largest wholesale banks were responsible for the origination of nearly 60 percent of all capital market transactions and the six largest U.S. banks (in order, Bank of America, JPMorgan Chase, Citigroup, Wells Fargo, Goldman Sachs, and Morgan Stanley) accounted for $8.97 trillion of assets, or approximately 55 percent of all assets held in the entire U.S. banking system. They operate aggressively because the financial marketplace is extremely competitive, and mandates are won or lost based less on the ideas proposed than on the tightness of the pricing and the willingness to bear risk. Their big balance sheets allow for diversification of risk, but only as long as risks do not become highly correlated (as they now tend to be in moments of panic that engender liquidity crises). The Volcker rule partially solves these problems by restricting the amount of principal trading, and some would argue the scale of market making activities, by large banks (and SIFIs). This will lessen the liquidity impact of the systemic risk of these institutions on the broader financial markets and real economy.

Hence, I interpret the DFL (2011) analysis as supporting, not contradicting, the Volcker rule.

Second, OW (2011) effectively assumes that all trading transactions by large banks (and SIFIs) are in fact market making activities. There is a long history of trading losses at banks (and SIFIs) which suggests this is not the case. Examples of singular bets gone wrong include $0.5 billion losses in interest-only and principal mortgage-backed securities at Merrill Lynch in 1987; $3.5 billion losses in copper futures at Sumitomo corporation in 1996; $7 billion in losses in European index futures at Societe Generale in 2008; and $8.5 billion in losses in subprime CDOs at Morgan Stanley in 2008, among many others. Fixed income departments housed within large banks and SIFIs suffered large losses during the summer of 1998 when the hedge fund LTCM failed; high yield trading departments such as Morgan Stanley’s endured losses in 2000 during the collapse of the internet bubble; widespread losses occurred within mortgage-related departments during the recent financial crisis (see Section V.A below); and recently many European banks, Dexia in particular, have experienced catastrophic losses on exposure to the sovereign debt of the peripheral Euro-zone countries.

As outlined in section V below, underwriting and market making are important financial intermediation activities. Some of these activities may best be performed at banks. Therefore, there is no reason why these activities cannot continue within banks, especially those allowing for limited inventory holdings and making markets in reasonably liquid assets. But because of both the systemic risk produced by these trades and the importance of market making, it does not follow that all of these activities must be executed within banks.
Another argument, one made especially by foreign policymakers, is that banks help create liquidity in the markets simply through trading, irrespective of whether these activities are considered market making or proprietary. But is this a good thing?

With access to the safety net, these firms distort market prices and arguably produce excess liquidity. One would prefer the discipline of the market to choose the pricing of these securities and the amount of liquidity. If liquidity cannot be reached through straightforward means, then it suggests more about the underlying security’s viability than anything else.

Third, both Duffie (2012) and OW (2011) argue that the Volcker rule will restrict market making activities by banks, causing short-term liquidity problems for sure and likely long-term liquidity problems as well. There is no doubt that there will be a transition stage as financial markets find a new equilibrium of market making inside and outside of banks. But the argument that this rule should be watered down because banks are currently the major players is weak. The argument should rely on what system is best in the long run. Moving some of these activities away from the large banks will lower systemic risk. The financial crisis of 2007-2009 is a perfect example of the problems associated with having these activities concentrated within just a few firms.

Some have expressed concern that market making activities will move to the “shadows” of the banking system. Specifically, this system of financial institutions, so-called shadow banks, performs functions like banks but take the form of other financial firms or entities. These financial institutions borrow short-term in rollover debt markets, leverage significantly, and lend and invest in longer-term and illiquid assets. As has been documented, the growth of shadow banking over the last 25 years has been extraordinary relative to the growth in deposits and, for the most part, was unregulated leading into the crisis. But this is not an argument for subsidizing and concentrating these activities within the systemically risky banking sector. Instead, to the extent these functions get taken on by shadow banks, shadow banks should be subject to likewise supervisory regulation.

In terms of how the banking sector might evolve, history provides an important lesson. Specifically, the Glass-Steagall Act forced the dissolution of the universal banks - for example, the breakup of J.P. Morgan into the Morgan Bank and Morgan, Stanley & Company. Continental Europe, in contrast, engaged in no such functional separation and largely continued with the universal banking tradition. The United Kingdom went its own way with a commercial banking structure centered on a short list (determined by the Bank of England) of publicly listed clearing banks and a long tradition in the securities sector of single-capacity jobbers (dealers), brokers, and merchant banks. Without access to the markets for deposits and commercial loans, but protected from competition by commercial banks, U.S. investment banks’ share of financial intermediation grew rapidly as financial flows progressively shifted to the financial markets. They in turn had a great deal to do with accelerating this process. Commercial paper markets, high-yield securities, asset securitizations, money market mutual funds, and similar innovations were in part the products of investment banks’ successful incursions into the market share of credit institutions.

By the 1980s, the U.S. financial system had become heavily market dominated while other financial systems remained dominated by universal banks. For example, local banks in
continental Europe were strongly resistant to cannibalization of profitable business at home. While this structural difference may have had something to do with a persistently higher U.S. rate of economic growth during the 1980s and 1990s, the so-called Anglo-Saxon financial architecture was arguably more efficient, more disciplined, and more innovative than the bank-dominated system of continental Europe. If true, then the Glass-Steagall legislation may have paid handsome growth dividends for over half a century, dividends that might have been forgone if the United States had persisted with a universal banking model after 1933. Internationally as well, a consequence of Glass-Steagall may have been the progressive dominance of U.S. investment banks in rapidly evolving global capital markets. Had universal banking remained in place in the United States after 1933, the lack of competitive pressure across very different strategic cohorts might well have involved significant opportunity costs for the U.S. economy.

This is not to suggest that history will repeat itself except that there is a strong argument structural reforms of the sort provided by the Volcker rule may in effect lead to greater innovation and efficiency.

IV. Diversification Benefits Are Minimal

One of the popular arguments put forth with respect to the Volcker rule’s restrictions on principal trading is that it will actually increase risk because it will not allow banks to diversify their risks. In terms of systemic risk, the diversification argument seems particularly weak. The fundamentals of modern finance tell us that there are two types of risk: idiosyncratic or firm-specific risk, which is diversifiable, and systematic or market-wide risk, which is not. While it is certainly true that the expansion of financial firms into multiple business lines may reduce the volatility of their overall asset portfolios, this is not necessarily what society most cares about. Because an economic crisis is the realization of market-wide risk, the problem society really cares about is whether banks - large and small - can withstand such risk and continue to perform critical intermediation functions.

When the economy craters, banks’ loans become impaired, the value of their securities holdings falls, their underlying investment banking business produces far less revenues, and the value of their asset management business plummets. Wagner (2009) argues that, while diversification makes individual bank default (and distress costs) less likely, it actually increases the likelihood of systemic risk. (See also Freixas, Loranth, and Morrison 2007.) Recent empirical work supports this theory. For example, De Jonghe (2009) documents that the tail betas of diversified financial institutions are higher and therefore these firms create more systemic risk. In a series of papers, DeYoung and Roland (2001), Stiroh (2004, 2006), and Stiroh and Rumble (2006) find that movement away from traditional banking activities toward other financial services increases the volatility and market risk of the firms. This work argues that the costs more than outweigh the benefits of diversification.

There is a debate about the benefits of large complex financial institutions as it pertains to economies of scale. The argument for synergies has some grounding. For example, many analysts would argue that it is important for firms that are active in the primary market for securities (i.e., underwriting) to be important participants in the secondary market (i.e., dealers). Nevertheless, the empirical evidence remains decidedly mixed. Notably, Laeven and Levine
(2007) report evidence that contradicts the existence of wide-scale synergies in large banks in the banking sector. They argue that there is a financial conglomerate discount; in other words, the whole is worth less than the sum of the parts. See also DeLong (2001), who performs an event study on diversifying bank mergers. In a study that goes beyond banks and looks at all financial intermediaries, Schmid and Walter (2009) document similar evidence. Interestingly, they find a premium for the very large firms, indicating that there is most likely a too-big-to-fail guarantee that supports the market value of these firms. Therefore, the reason for the growth in SIFIs may simply be due to the below-market cost of financing through the central bank or public guarantees. From a societal point of view, the benefits to SIFIs of a too-big-to-fail guarantee are clearly not a valid reason to oppose reinstatement of some form of Glass-Steagall such as the Volcker rule, since too-big-to-fail standing encourages moral hazard.

In contrast, Baele, De Jonghe, and Vander Vennet (2007) and Elsas, Hackethal, and Holzhauser (2009) provide evidence that the large bank model does improve bank profitability, and generally argue that these gains are due to economies of scale. The reasons for the different findings can be attributed to both different data sources and different methodologies. In this white paper, I am not going to be able to resolve this current debate. Indeed, the recent studies mirror the findings of the survey article by Berger and Humphrey (1997) some 15 years earlier, which argued there was no predominance of evidence either for or against economies of scale in the financial sector.

What is less controversial in the literature, however, is that the expansion to multiple functions, the large, complex bank model, produces greater systemic risk. As noted above, there is now a plethora of research—including Acharya, Brownlees, Engle, Farazmand and Richardson (2010), DeYoung and Roland (2001), Stiroh (2004, 2006), Stiroh and Rumble (2006), De Jonghe (2009), and even papers loosely in support of the large bank model such as Baele, De Jonghe, and Vander Vennet (2007)—that finds the large bank model more risky. This creates a strong economic case for some form of return to Glass-Steagall and functional separation, the Volcker Rule being one such example.

V. What Form Should the Volcker Rule Take?

Paul Volcker has argued for a much simpler bill than is currently being proposed. His basic idea is that the Volcker rule would ban proprietary trading and make the chief executive and firm’s Board of Directors responsible for compliance. If the banks did not comply with the “spirit” of the rule, then there should be strong regulatory action. I believe Paul Volcker is right.

It would be a mistake to make the Volcker rule into a complex set of specific rules that banks must follow. The reasons are straightforward. First, once the set of rules are written, it would be straightforward for banks and SIFIs to circumvent these rules over time. Financial markets are dynamic with continuous financial innovations. One only has to look at the way the ever more complex capital regulation rules of Basel II were exploited in the most recent financial crisis. Basel II was not set up with this in mind, yet this was the outcome. Second, a strictly rules-based approach provides false cover for regulators. Regulators can “check” the box that the rule has been adhered to rather than supervising the financial institution more broadly. Numerous examples from the most recent financial crisis exist, including regulators providing a clean bill of
health to Fannie Mae and Freddie Mac shortly before being brought into conservatorship; the large Belgian bank Dexia scoring high on stress tests just before failing; European banks' build up of exposure to AIGFP’s CDS for purely regulatory capital reasons; Lehman Brothers use of repo 105 transactions to show lower leverage; and so forth.

Paul Volcker’s comments suggest that he had a more principles-based rule in mind; in fact, when asked during Senate hearings how he would identify such trading, Volcker replied that “it’s like pornography – you know it when you see it.” The intuitive definition is that proprietary trading constitutes any trading conducted by the firm for its own account.

The Dodd-Frank Act states that: The term “proprietary trading,” when used with respect to a banking entity or nonbank financial company supervised by the Board, means engaging as a principal for the trading account of the banking entity or nonbank financial company supervised by the Board in any transaction to purchase or sell, or otherwise acquire or dispose of, any security, any derivative, any contract of sale of a commodity for future delivery, any option on any such security, derivative, or contract, or any other security or financial instrument that the appropriate Federal banking agencies, the Securities and Exchange Commission, and the Commodity Futures Trading Commission may . . . determine.

While exemptions for excluding customer-related trading and hedging are logical, they create a gray area for implementing the rule. For example, when a large bank acts as intermediary between buyers and sellers, especially for less liquid securities, the firm will often be exposed to one side of the transaction. In fact, a number of normal market- and client oriented transactions, such as trading in foreign exchange, fixed-income securities, and derivatives, as well as services like bridge financing, prime brokerage, and the like, might result in the firm technically trading on its own account but doing so to serve client needs. This gray area also invites manipulation. What is to prevent a bank from accumulating a large exposure in a given security or derivative in expectation of an eventual customer demand for the asset? How are regulators to distinguish between identical trades where the intent of one is clearly customer-driven and the intent of the other is proprietary? Should there be a time limit set on holdings related to customer-related trading? Should there be the requirement that the aggregate market exposure associated with these holdings be hedged? How can such holdings be differentiated from those related to pure trading bets in the real world?

From my perspective, a principles-based approach would be the best way to address these issues.

First, as detailed in Section II above, certain types of asset holdings, such as those prone to regulatory arbitrage, carry trades and financial guaranty insurance, should be banned for banks and certainly restricted for SIFIs. If a bank wanted to engage in these activities for customers and/or hedging of the bank’s own risks, then the onus would be on the bank to convince the regulator that (i) it was necessary and economically important, and (ii) it did not impose unnecessary systemic risk.

Second, to the extent there is a fine line between principal trading and market making, the best approach would be to describe safe harbors to the Volcker rule, in other words, proprietary and hedge activities always permitted within the rule. Any activity outside these safe harbors would
require permission by the bank’s (or non-bank SIFI’s) regulator. Again, my view is that this safe harbor-approach is most in line with original intent of the Volcker rule.

What would some of these safe harbors look like in practice? Here are some examples:

1. As long as the firm’s aggregate gross (i.e., unhedged) inventory of assets is below a specified fraction (denote $K\%$) of the firm’s capital, there would be the presumption that the trading positions are market making. If the firm wishes to hold inventory above the fraction, then the onus would be on the bank to convince the regulator that the entire inventory (i) is necessary and economically important for market making, and (ii) does not impose unnecessary systemic risk.

2. If the firm’s aggregate gross inventory of assets is below $K\%$ of the firm’s capital, then there would be an additional restriction that the firm’s aggregate net (i.e., hedged) inventory of assets also lie below an even smaller fraction (denote $K^{**}\%$) of the firm’s capital. In other words, if gross inventory were above $K^{**}\%$ of capital, the residual portion of the firm’s inventory must be hedged. If the firm wishes to hold unhedged inventory above the fraction, then the onus would be on the bank to convince the regulator that the hedge is unnecessary.

3. Hedging risk of the bank’s loan portfolio or risk of its inventory of assets is provided safe harbor as long as those hedges fit into one of two types. The first type is a hedge of aggregate risk using one of a pre-specified set of highly liquid derivative contracts amongst futures, options, swaps, or indices. This hedge must be justified in terms of a maximum allowed tracking error. The second type is a hedge of specific risk of an asset held by the firm using a tightly pre-specified derivative contract on the underlying (e.g., corporate bond and a CDS on the underlying bond, equity and an option on the underlying equity, etc...). If the firm wishes to hedge aggregate risk or hedge a specific risk using a non-specified contract, then the onus would be on the bank to explain to the regulator why the use of such a contract is unnecessary.

4. Even if the above safe harbors are satisfied, there would be an additional restriction that the effective holding period for the economic risk of the assets in inventory must be less than a given period of time (e.g., three months). If the financial firm expects to hold the assets longer than the specified time length, or goes beyond this time length ex post, then the onus would be on the bank to explain to the regulator why the additional time is (or was) needed.

5. The firm would be restricted from holding in inventory any type of level 3 assets. Level 3 assets are generally highly illiquid assets whose fair value cannot be determined using either market prices or models. If the financial firm wishes to make markets in level 3 assets or otherwise hold these assets in inventory, then the onus would be on the bank to explain to the regulator why the holding of the level 3 assets is necessary, and, in conjunction with their other holdings, does not impose unnecessary systemic risk.

As can be seen from the above discussion, a principle-based rule with safe harbors creates a much tighter regulatory regime between the bank (and/or SIFI) and the regulator. While at first glance this regime might seem onerous to the financial institution, it need not be so. If the financial institution stays within the safe harbors of the Volcker rule, and chooses to manage its
gross and net inventory of reasonably liquid assets within a fraction of its capital, then the financial firm would be free to make markets and take limited proprietary bets. Some large banks will choose to organize their business along these lines. Other banks might choose to make markets in more illiquid assets and/or go beyond their inventory limits. These banks would be subject to greater regulatory oversight and may or may not get permission to expand their activities beyond the boundaries of the Volcker rule.

A. The Volcker Rule, If Applied Prior to the Financial Crisis of 2007-2009

It has been argued that the Volcker rule would not have been useful in the run-up to the financial crisis of 2007-2009 in that principal trading was not a root cause of the crisis. While this argument regardless does not imply that the Volcker rule is not a useful tool for managing systemic risk, the argument also happens not to be true. Considerable losses of large complex banks and other SIFIs occurred in their holdings of asset-backed securities, especially those tied to residential mortgages. It is true that, for many of these firms, their initial exposure related to their underwriting and warehousing of these securities. This business model, however, morphed into these firms effectively holding a large volume of these securities either on- or off-balance sheet. The evidence shows that some of this behavior was clearly willful and simply a carry trade in disguise. However, even if these financial firms were stuck holding these securities because there was little demand in capital markets for them, why should financial firms under the safety net be creating excess liquidity and systemic risk? The principles-based Volcker rule with safe harbors would have greatly restricted these activities either through an outright ban on a recognized carry trade or having the regulator question the risk of these activities. That is, financial firms would have needed to seek permission to go beyond the safe harbor and greatly expand their inventory holdings. Such a request would have been a red flag as the firm would be increasing their balance sheet with little or no additional capital. Much of the analysis below is taken from Clementi, Cooley, Richardson and Walter (2009).

In the summer of 2005, one of the major players in subprime mortgage collateralized debt obligations (CDOs), UBS, ramped up its CDO warehouse business. In this business, UBS would purchase residential mortgage backed securities (RMS) primarily made up of subprime mortgages, house them in its CDO warehouse, prepare them for securitization, and then sell the multi-tranche CDOs in the marketplace. UBS’s CDO desk received structuring fees on the notional value of the deal ranging from 30bps to 150bps, depending on the credit quality of the tranche. Because this process from start to finish took 2-4 months, the CDO warehouse was an important component of UBSs value at risk and UBS recognized this as such. In 2005, the CDO business, albeit a risky one, worked as intended. UBS faced short-term holding risk during the securitization process but was compensated by being paid considerable fees. The credit risk that would normally be held by banks or mortgage lenders was transferred to the capital market.

Starting in 2006, however, UBS began to hold the so-called AAA rated, super senior tranches of the CDOs rather than sell them. These tranches have the highest priority within the CDO and thus are somewhat protected by the junior tranches. The senior tranches are only hit if there are substantial defaults and low recoveries. That said, the super senior tranches were structured to

---

5 This account is taken from UBS’s “Shareholder Report on UBS’s Write Downs” prepared for the Swiss Federal Banking Commission.
hold as much of the pool of subprime loans as possible and still maintain the AAA rating given by rating agencies. From holding almost none of these securities in February 2006, the CDO desk was holding over $50 billion in September 2007. The main reason for retaining these tranches on their books was that these securities offered a yield above UBS’s internal funding rate which hovered around LIBOR, yielding an immediate ongoing profit.

Moreover, because these securities were rated AAA, they barely registered on UBS’s value at risk or stress tests even when totally unhedged. Thus, the excess yield was treated as pure alpha. As a result of this “pure alpha”, there were no aggregate notional limits placed on the CDO warehouse. Thus, every extra dollar of CDOs retained increased the desk’s “profit”. Moreover, because the UBS compensation structure did not differentiate between profits derived from a low cost of funding versus the generation of true excess return (i.e., alpha), the desk’s compensation was directly linked to the size of the CDO’s mortgage book. There was no liquidity premium charged the group. That is, there was little or no differentiation between liquid and illiquid assets even though there are many examples of almost identical securities offering different yields in the markets (e.g., off-the-run versus on-the-run treasuries).

These facts meant that the CDO desk had the incentive to grow the balance sheet for CDOs as large as possible because, by construction, their bonuses were tied to instant profits with no recognition of any risk. This growth continued even during the first half of 2007 as subprime lenders were going bankrupt and hedge funds were reporting losses. In fact, UBS shut down one of its own operations, Dillon Read Capital Management, in May 2007 for losses in their subprime investment portfolio. In March 2007, the Treasury group within UBS, alarmed at the tremendous growth of UBS’s balance sheet especially in relatively illiquid ABSs, argued for a limit on illiquid assets, a haircut funding model (in which illiquid assets would no longer get short term funding), and an overall freeze on their balance sheet. This call went unanswered.

Putting aside the issue of whether these securities were truly “AAA” in quality, there is no doubt that their underlying risk was very asymmetric. That is, the securities would pay a premium above LIBOR in most states of nature, but in the rare event that there were substantial defaults and low recoveries, they would get hit. Historically, this rare event would arise only if the underlying collateral (i.e., house prices) fell dramatically or there was a sharp economic downturn (i.e., as in previous recessions). In finance terms, due to the priority structure of the claims, the holders of the senior tranche were essentially invested in a risk-free asset, like LIBOR, while simultaneously writing a way out-of-the-money put option on the market.

If the aforementioned description of the governance (or lack thereof) of the CDO desk at UBS was an aberration, then there would be less of an issue. Unfortunately, it seemed to be the norm during this period and endemic to many firms. Some further examples:

- First, over the period 2005-2007, Citigroup was one of the largest CDO issuers (ranked first in 2005 with $28 billion, third in 2006 with $33 billion, and first again in 2007 with

---

6 A majority of the super senior tranche holdings were partially hedged and treated as having zero effect on the firm's value at risk.
7 Coval and Jurek (2008).
8 The following description is based on a series of articles in the New York Times called “The Reckoning” which covered, among other firms, Citigroup, Merrill Lynch, AIG, Fannie Mae and Freddie Mac.
$40 billion). There are remarkable similarities to UBS: (i) by the third quarter, Citigroup had accumulated over $55 billion of AAA-rated higher tranche mortgages, (ii) there was no clear separation between risk management and trading in the fixed income group, leading to misaligned incentives, (iii), according to the NY Times, despite their size, yet using the rating agencies as justification, these securities do not show up on the firm’s Value at Risk analysis⁹, and (iv) the CDO group was one of the highest paid in the firm.

- Second, during this same period, Merrill Lynch also poured into this sector, ranking second in CDO issuance in 2005 with $27 billion, first in 2006 with $54 billion and second in 2007 with $38 billion. Before 2005, like UBS and Citigroup, Merrill would hold onto the AAA-rated tranches and get them insured via a credit default swap (CDS) with the insurance giant, A.I.G. In 2005, After A.I.G. decided not to insure subprime-backed CDOs any longer Merrill decided to continue to issue CDOs and hold them, essentially unhedged. With risk management taking a back step to the carry profit of the AAA-tranches (i.e., the premium minus Merrill’s short-term funding), Merrill held over $70 billion of these securities by the time the financial crisis started.¹⁰

- A third example involves AIG (see also Section II above). One of the more profitable groups within the A.I.G. financial empire was A.I.G. Financial Products (representing 17% of A.I.G.’s operating income in 2005).¹¹ As one of its many investments, the group ended up writing a staggering $500 billion worth of CDSs on mostly the AAA-rated tranches of CDO-like structured products on mortgages, corporate bonds and loans. Like the banks above, risk management was not a separate function. The model was essentially identical to that of UBS, Citigroup and Merrill, that is, one of writing out of the money puts on the underlying, in this case, many defaults. In fact, in August 2007, Joseph Cassano (the former head of the group) stated, “It is hard for us, without being flippant, to even see a scenario within any kind of realm of reason that would see us losing one dollar in any of those transactions.” Because these CDSs were claims on the upper priority of portfolios of loans that could only get hit upon a large systematic shock, it essentially meant that A.I.G. would receive fees most of the time. These fees were booked as income, resulting in huge compensation packages for the group (e.g., compensation hovered around $500 million for a staff of 300 or so). Of course, there was no accountability to this group for the rare event which would result in essentially all CDSs being hit, and a significant fraction of $500 billion having to be paid out, bringing down the firm.

While the three examples above are arguably the most extreme cases in the U.S. financial sector during the crisis, similar problems also show up for Bear Stearns, Morgan Stanley, Fannie Mae, Freddie Mac, Lehman, and most likely other undocumented cases. As outlined above, the Volcker rule would have been able to address these issues.

Conclusion

¹¹ The facts are taken from the article, “A Spreading Virus” by Gretchen Morgenson, New York Times, September 27, 2008.
As part of an effort to seriously address excessive systemic risk, I find the logic of limiting
government guarantees to core banking activities and segregating nonbanking risk-taking
businesses to be fundamentally sound and in the public interest. This approach is akin to that of
the 1930s, but adapted to conduct proprietary activities in a way that can be effectively regulated.
It is a development that would be in line with the public interest as well as common sense, and
one that is unlikely to trigger significant social costs in terms of financial efficiency and
innovation. Indeed, based on a careful reading of the unintended consequences of the Glass-
Steagall restrictions of 1933, quite the opposite could be the case.
References:


The Orderly Liquidation of Lehman Brothers Holdings Inc. under the Dodd-Frank Act

http://bankingcommission.independent.gov.uk/
Independent Commission on banking Final report


Pedersen, Lasse Heje, 2008, Liquidity Risk and the Current Crisis, VoxEU, 15 November.


UBS’s “Shareholder Report on UBS’s Write Downs

