

The Journal of Financial Perspectives

Ernst & Young Global Financial Services Institute

March 2013 | Volume 1 - Issue 1

The Ernst & Young Global Financial Services Institute brings together world-renowned thought leaders and practitioners from top-tier academic institutions, global financial services firms, public policy organizations and regulators to develop solutions to the most pertinent issues facing the financial services industry.

The Journal of Financial Perspectives aims to become the medium of choice for senior financial services executives from across banking and capital markets, asset management and insurance, as well as academics and policy-makers who wish to keep abreast of the latest ideas from some of the world's foremost thought leaders in financial services. To achieve this objective, a board comprised of leading academic scholars and respected financial executives has been established to solicit articles that not only make genuine contributions to the most important topics, but that are also practical in their focus. *The Journal* will be published three times a year.

gfsi.ey.com

The articles, information and reports (the articles) contained within *The Journal* are generic and represent the views and opinions of their authors. The articles produced by authors external to Ernst & Young do not necessarily represent the views or opinions of EYGM Limited nor any other member of the global Ernst & Young organization. The articles produced by Ernst & Young contain general commentary and do not contain tailored specific advice and should not be regarded as comprehensive or sufficient for making decisions, nor should be used in place of professional advice. Accordingly, neither EYGM Limited nor any other member of the global Ernst & Young organization accepts responsibility for loss arising from any action taken or not taken by those receiving *The Journal*.

Editorial

Editor

Shahin Shojai
Ernst & Young LLP

Advisory Editors

Dai Bedford
Ernst & Young LLP
Carmine DiSibio
Ernst & Young LLP
David Gittleson
Ernst & Young LLP

Shaun Crawford
Ernst & Young LLP
Ratan Engineer
Ernst & Young LLP
Bill Schlich
Ernst & Young LLP

Special Advisory Editors

Paul Feeney
Old Mutual Wealth
Paolo Fiorentino
UniCredit

Robert W. Jenkins
Bank of England
Thomas C. Wilson
Allianz

Editorial Board

Viral V. Acharya
New York University
John Armour
University of Oxford
Tom Baker
University of Pennsylvania
Law School
Philip Booth
Cass Business School and IEA
Kalok Chan
Hong Kong University of Science
and Technology
J. David Cummins
Temple University
Allen Ferrell
Harvard Law School
Thierry Foucault
HEC Paris
Roland Füss
University of St. Gallen
Giampaolo Gabbi
SDA Bocconi
Boris Groysberg
Harvard Business School
Scott E. Harrington
The Wharton School
Paul M. Healy
Harvard Business School
Jun-Koo Kang
Nanyang Business School
Takao Kobayashi
Aoyama Gakuin University
Howard Kunreuther
The Wharton School
Deborah J. Lucas
Massachusetts Institute of Technology

Massimo Massa
INSEAD
Patricia A. McCoy
University of Connecticut
School of Law
Tim Morris
University of Oxford
John M. Mulvey
Princeton University
Richard D. Phillips
Georgia State University
Patrice Poncet
ESSEC Business School
Michael R. Powers
Tsinghua University
Andreas Richter
Ludwig-Maximilians-Universitaet
Philip Rawlings
Queen Mary, University of London
Roberta Romano
Yale Law School
Hato Schmeiser
University of St. Gallen
Peter Swan
University of New South Wales
Paola Musile Tanzi
SDA Bocconi
Rolf Tilmes
EBS University
Marno Verbeek
Erasmus University
Ingo Walter
New York University
Bernard Yeung
National University of Singapore

Reworking the global financial architecture: is universal banking the best way forward?

Anthony Saunders

John M. Schiff Professor of Finance, Stern School of Business, New York University

Ingo Walter

Vice Dean of Faculty, and Seymour Milstein Professor of Finance, Corporate Governance and Ethics, Stern School of Business, New York University¹

Abstract

Consolidation has been a fact of life in the wholesale financial services sector, resulting in fundamental change in the financial architecture and public exposure to systemic risk. The underlying drivers include advances in transactions and information technologies, regulatory changes, geographic shifts in growth opportunities, and the rapid evolution of client requirements, which in combination have obliged financial firms to rethink their roles as intermediaries. Moreover, financial sector reconfiguration has accelerated as a result of the global market turbulence that began in 2007, with governments either forcing or encouraging combinations of stronger and weaker financial firms in an effort to stem the crisis and improve systemic robustness.

In the process, financial firms that are “systemic” in nature and had a major role in creating the crisis have come out of it with even larger market shares and greater systemic importance. Given the episodic socialization of risk in the form of widespread use of public guarantees to firms judged too big or too interconnected to be allowed to fail, the role of systemically important financial institutions (SIFIs) is central to the financial architecture and the public interest going forward.

This survey paper considers the sources of systemic gains, losses, and risks associated with SIFIs in historical context, in the theoretical and empirical literature, and in public policy discussions – i.e., what is gained and what is lost as a result of the available policy options to deal with the dominant role of SIFIs in the financial architecture?

¹ Paper adapted for the *Journal of Financial Perspectives* from “Systemic risk in financial architecture,” *Financial Markets and Portfolio Management*, March 2012 and “Architecture financière, risque systémique et les banques universelles,” *Revue d’Economie Financière*, April 2012.

Introduction

A quarter century ago we formed a small group of financial economists to re-examine the relative merits of specialized versus universal banking with respect to efficiency and innovation as well as safety and soundness in financial intermediation [Walter (1985)]. At the time in the United States, financial architecture had been characterized by specialized financial intermediation dating from the Glass-Steagall provisions of the Banking Act of 1933 – a structure that also characterized Japan after the American occupation following World War II and the British tradition of designated clearing banks alongside independent merchant banks, brokers, dealers, and insurers. Meanwhile, virtually all other countries followed some form of universal banking. This divergence suggested a “natural experiment” based on available evidence and stylized facts. Which system seemed to have delivered the best results based on a set of social welfare benchmarks, notably some composite of efficiency, growth, and stability?

Our conclusion at the time was that financial architecture ought to be essentially market-driven, based on a level regulatory playing field, and that market forces and the economics of industrial organization in financial intermediation should be allowed to determine optimum firm structures [Saunders and Walter (1994)]. We argued that market-based drivers would harvest both scale and scope economies for the benefit of the public interest, and that market discipline and appropriate corporate governance would adequately promote safety and soundness in the financial system.

The arguments we put forward in the 1980s eventually carried the day. Gradual liberalization undertaken by the key regulatory agencies in the U.S. allowed progressively greater functional and geographic latitude for financial intermediaries, culminating in the 1999 Gramm-Leach-Bliley Act. Parallel liberalization occurred in the U.K. with the “Big Bang” in 1986 and a regulatory shift from the Bank of England to the Financial Services Authority, and in due course in Japan as well. With the disappearance of many specialized financial intermediaries into the major universal banks and financial conglomerates of the time, it seemed that the market had provided an answer – the forces of scale and scope would ultimately produce a better (more efficient and more stable) financial system, one dominated by a monoculture of national and global financial behemoths.

But would the apparently inexorable progression to exploit scale and scope necessarily be in the public interest? Since the dominant institutions were likely to be too big, too complex, and too interconnected to fail, would the evolving architecture, based on size and scope, generate negative externalities not priced in the market, so that what appeared to be a market-driven institutional result would turn out to be based on badly distorted signals? Do externalities rooted in scale and scope inevitably produce “privatization of returns and socialization of risks,” so that the institutional evolution we were observing would actually be wealth-redistributing rather than wealth-creating? Like the dog that caught the bus, what you wish for may not be exactly what you get.

Conversely, are those who today argue for breaking-up or constraining “systemic” financial conglomerates in order to constrain the system’s exposure to negative externalities, seriously underestimating the offsetting benefits of scale and scope in financial intermediation?

In this survey paper, we review the available evidence regarding financial architecture in the perspective of the systemic turbulence that surfaced in 2007 and grew to encompass virtually all major financial markets and institutions, inflicted massive financial losses on market participants, and went on to contaminate the real economy and public finances – and ultimately tested the limits of public policy. Section 2 of the paper traces the pathology of regulation, deregulation, and the evolution of the financial architecture from watershed events in the 1930s to the systemic failures of the recent past. Section 3 surveys the industrial economics of financial intermediation and explains why market forces could produce system-critical financial institutions as a dominant institutional form, identifying what would be lost and gained if they were to be broken-up. Section 4 considers the public-interest benchmarks against which institutional design in financial intermediation should be calibrated, and how scope-related regulation would be viewed in this context. Section 5 concludes.

Legacies of the past

Counterfactuals have become a favorite technique of modern historians. Suppose Napoleon had won a resounding victory at Waterloo in 1815. Suppose the British and the Germans had formed an alliance after Archduke Ferdinand’s assassination in

Sarajevo in 1914. What would the world have looked like in the ensuing years and decades? The imagination is allowed free rein – immune to definitive rejection, since the argumentation is based on some key assumptions.

Here we can develop a counterfactual on the functional separation of “banking” and “commerce” in 1933, marking the end of decades of universal banking in the U.S. and a key part of serious regulatory reform in the wake of debilitating financial and economic turbulence that took hold in 1929. What might the U.S. and global financial system have looked like in the absence of forced break-up of the legacy American universal banks? There are some available comparisons, after all, since the countries of continental Europe specifically did not follow the American approach as they struggled through the Depression, and continued with the universal banking model to the present day.

The so-called Glass-Steagall Act actually consisted of four provisions of the Banking Act of 1933, arguably the most substantial piece of U.S. financial legislation to emerge from the Great Depression – legislation that also encompassed deposit insurance and other reforms designed to restore and maintain financial stability. One of the key provisions of the Act mandated a virtually complete separation of investment banking (commerce) from deposit-taking activities (banking). It thus eliminated involvement by firms with a commercial banking charter in the securities business – specifically, underwriting and dealing in corporate debt and equity securities and municipal revenue bonds (obligations not guaranteed by the full faith and credit of municipal governments). That business, which gained dramatically in volume during the 1920s, was dominated by an amalgam of universal banks such as JP Morgan and National City Corporation and broker-dealer banks such as Goldman Sachs and Lehman Brothers. The former were listed companies engaged in a full array of commercial and investment banking activities, and the latter were private partnerships engaged mainly in securities underwriting and trading, and investing their partners’ capital.

Contemporary critics of the U.S. universal banking model feared that bank involvement in securities underwriting had directly and indirectly led them to ramp-up (warehouse) their holdings of long-term financial instruments – funded in part by customer deposits – thereby exposing themselves, their clients, and the system to

potentially dangerous market, credit, and liquidity risk. When this risk materialized with a vengeance after 1929, it was thought to have contaminated the entire U.S. financial system by triggering the collapse of banks nationwide, which in turn had disastrous consequences for the real economy. About 40% of all U.S. banks failed during this period, undermining their role as financial intermediaries and cutting off the air supply to the real economy. Little wonder, in the heat of the moment, that investment banking might have been considered “too risky” for commercial banks.

The fact is that the big universal banks did increase their holdings of equities and long-term debt securities during the 1920s, probably because the demand for short-term corporate loans declined substantially. But there is little evidence that the quality of bank securities holdings was responsible for the cascading bank failures of 1930-1933. Rather, the Federal Reserve System’s unwillingness to provide liquidity to a banking system beset by depositor runs was largely responsible for the banking collapse of the time. Under the circumstances, most of the banks that failed would have collapsed even if they had held no long-term bonds at all. Evidence that commercial banks’ securities activities somehow directly caused the Great Depression has remained elusive, although the indirect causality remains arguable.

Moreover, modern studies of financial practices preceding passage of the Banking Act of 1933 suggest that the large universal banks did not in fact exploit conflicts of interest that arose from their role as commercial lenders and securities underwriters – for example using their privileged insight as lenders to prompt clients to issue securities and use the proceeds to repay commercial loans, thereby shifting losses from themselves to investors. Securities underwritten by (potentially conflicted) universal banks did not underperform those underwritten by (non-conflicted) investment banks [See, for example Gande et al. (1997), Gande et al. (1999), Puri (1994), and Puri (1996)]. Most of the questionable transactions that appeared to have occurred between banks and their securities affiliates in the 1920s would have been illegal under the financial laws and regulations that followed – the Securities act of 1933, the Securities and Exchange Act of 1934, and Investment Company Act of 1940.

The goal of financial reformers in the 1930s was to prevent a recurrence of the financial hurricane that had just passed,

rather than to seek retribution and punishment for those held to be responsible. Rightly or wrongly, legislation was vigorously debated and passed in good faith, courageously and under enormous political pressure in the “fog of war.” It forced the dissolution of U.S. universal banks. This included the breakup of JP Morgan & Co. into the Morgan Bank, which in 1959 merged with the Guaranty Trust Company to form the Morgan Guaranty Trust Company of New York, the nation’s premier wholesale commercial bank, and what later became Morgan Stanley, one of the top investment banks.

Continental Europe, in contrast, engaged in no such functional separation and largely continued with the universal banking tradition, complemented by a few independent securities and advisory firms, mainly in Italy (notably Mediobanca), France (the banques d’affaires), and the Netherlands. The U.K. went its own way with a commercial banking structure centered on a short list (determined by the Bank of England) of listed clearing banks and a long tradition in the securities sector of “single capacity” jobbers (dealers), brokers, and merchant banks, which was changed in Britain’s 1986 “Big Bang” reforms to allow the formation of broker-dealers, integrated investment banks and investment banking divisions of universal banks, and financial conglomerates. In 1948, Japan was forced under the U.S. occupation to adopt a version of Glass-Steagall in the form of Article 65 of the Japan Securities and Exchange Law, which contained strict separation of commercial and investment banking.

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 from the balance sheets of commercial banks and other credit institutions to the financial markets. In turn, the investment banks 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 a product of investment banks’ successful incursions into the credit institutions’ intermediation market share, aided by the substantially lighter regulatory burdens they had to bear (notably capital requirements and liquidity requirements). By the 1980s, the U.S. financial system had become heavily market-dominated while financial systems

in continental Europe remained universal bank-dominated, with the major banks understandably resistant to cannibalization of profitable businesses at home via capital market instruments. Without the disruptive pressure exerted by independent broker-dealers, European equity and corporate bond markets were slow to develop, and banks took significant shareholdings and control positions in nonfinancial companies, creating close and stable client relationships but arguably eroding the role of market discipline in capital allocation and corporate governance.

If the so-called Anglo-Saxon financial architecture of the time was indeed more efficient, more disciplined, and more innovative than the bank-dominated system of continental Europe, then the Glass-Steagall legislation may have paid handsome economic growth dividends for over half a century – dividends that might have been forgone if the U.S. had persisted with a universal banking model after 1933. Such collateral gains could then be added to the high degree of stability achieved by the U.S. financial system for almost 70 years. Other than the S&L debacle of the late 1980s and early 1990s – which saw 747 thrift failures (out of 3,234 institutional) but had nothing to do with failures of individual systemic financial institutions – the U.S. suffered no material financial setbacks during the separated-banking era.²

Globally as well, a consequence of Glass-Steagall may have been the progressive dominance of U.S. investment banks in rapidly evolving offshore and onshore capital markets worldwide. Surviving American broker-dealers, whose competitiveness was enhanced by the disappearance of fixed brokerage commissions in the SEC’s “Mayday” financial reforms in 1974, began a sustained offensive in financial markets abroad.

Penetrating the fortresses of universal banking in one country after another, they mounted a sustained 20-year attack to wean European and later Asian corporates and governments from their reliance on intimate relationships with domestic universal

² An exception was the 1984 failure of Continental Illinois Bank and Trust Company, which required resolution by the Federal Reserve and the Federal Deposit Insurance Corporation in a U.S.\$4.5 billion rescue that wiped-out its shareholders, saw management and directors replaced, and led to eventual acquisition by Bank of America - but which successfully avoided any contamination of the banking system as a whole. Arguably the role of the Federal Reserve in the resolution of Long Term Capital Management in 1998 was another exception, with Fed-sponsored intervention and liquidation of the firm by a group of its prime-broker commercial banks intended to contain possible systemic impact on financial markets. Fed intervention came despite the existence of a commercial offer to restructure the firm in a private transaction led by Warren Buffet.

banks, offering lower funding costs and innovative financings. Meanwhile, they cultivated the global buy-side of the market – insurance companies, pension funds, and other institutional investors – with new investment alternatives and ideas for improving portfolio efficiency.

The U.S. investment banks' global offensive was so successful that virtually all the major universal banks in Europe launched vigorous efforts to develop investment banking divisions of their own, but without their U.S. rivals' advantages of a captive domestic market and having been battle-tested in global competition. By the early 1990s, American investment banks basically dominated their industry worldwide, with a global market share approaching 75% in terms of transactions volume. As a consequence, investment banking developed into one of the top U.S. export industries. Had universal banking remained in place in the U.S. after 1933, the lack of competitive pressure across very different strategic cohorts might well have had a very different outcome in terms of global competitiveness. So, with the passage of time, Glass-Steagall may well have had important positive unforeseen consequences.

In sum, it seems plausible that separated banking substantially delivered on its promise of efficiency, stability, and competitiveness for well over half a century, against the counterfactual that the same or even better results would have been achieved if U.S. universal banks and financial conglomerates had been allowed to persist.

Who knows? But revealed preference suggests otherwise, as the major U.S. wholesale commercial banks – notably Morgan Guaranty Trust Company, Bankers Trust, Chase Manhattan, and Citicorp – began to agitate vigorously by the early 1980s for reinstatement of universal banking powers to redress what they had come to regard as a debilitating competitive disadvantage. The grass was obviously much greener on the other side of the Glass-Steagall fence. While they could and did compete vigorously with investment banks in government bond, foreign exchange, and other traded markets as well as corporate advisories, they were hamstrung in the high-growth “Glass-Steagall illegal” sectors of the market that were critical for competing in high-growth financial intermediation channels.

Indeed, by the late 1980s commercial banks had gained the limited right to sell investment and insurance products to retail customers, as well as the right to operate separately capitalized, size-constrained wholesale securities subsidiaries (so-called Section 20 subsidiaries) under various safeguards, to prevent the commercial banking arm from contamination in the event of investment banking losses. Liberalization came in the form of administrative rulings on the part of the regulators, not legislative change. Perhaps a dozen of the major wholesale commercial banks took early advantage to build significant securities subsidiaries, especially in the bond business, to complement their powerful wholesale commercial banking and government bond activities and their emerging presence in corporate advisory work.

One important area in which the commercial banks made little headway was equities, a highly profitable growth market that was far removed from their traditional expertise in debt finance, and in which they had little sales and trading expertise and few natural relationships with companies undertaking IPOs. Moreover, lack of a market presence in equities seriously hampered their ability to build a competitive fee-based corporate finance business. This gap in their product line-up lent even more urgency to removal of the remaining Glass-Steagall restrictions through legislative action.

So, in the 1980s the commercial banks mounted a vigorous attack on Glass-Steagall, reflected in a drumbeat of political initiatives to get the rules changed. These included “high-road” arguments that the structure of financial intermediaries should be driven by competitive and strategic consideration, not anachronistic legislation. They also included “low-road” initiatives such as political lobbying and audacious tactical initiatives. Bankers Trust's technically illegal, from a Glass-Steagall perspective, underwriting of commercial paper in 1985 led to litigation by the investment banks protecting their turf and forced the courts to decide the merits (Bankers Trust won). Citicorp's technically illegal merger, again from a Glass-Steagall perspective, with Travelers to form Citigroup in 1998 was an audacious bet that Glass-Steagall would soon be swept into the dustbin of history by creating facts on the ground. Such bold pre-emptive strikes by commercial banks were soon validated by passage of the Gramm-Leach-Bliley Financial Services Modernization Act of 1999 (GLB).

Passage of GLB by an overwhelming Congressional vote of 343 to 86 put the final nail into the coffin of functional separation in U.S. financial intermediation. At the retail level, it allowed commercial banks to gather assets into both bank deposits and securities accounts such as money market mutual funds, helping to stem the incursion by broker-dealers into their traditional client base and broadening their ability to respond to changes in client preferences. At the wholesale level, GLB allowed commercial banks to underwrite and trade in corporate debt, corporate equities, and municipal revenue bonds, and compete head-on with the broker-dealers. The Riegel-Neal Interstate Banking and Branching Efficiency Act of 1994 had previously repealed the McFadden Act, which limited interstate branching. So, by 2000 the stage was set for a return to full-blown universal banking in the U.S., with few regulatory constraints on scale and scope in financial intermediation.³

In sweeping aside functional separation in U.S. financial intermediation the Gramm-Leach-Bliley Act arguably had unintended consequences, as did the passage of Glass-Steagall 66 years earlier. From a public interest perspective, for good or for ill?

Within two years of its enactment, every major commercial bank that took full advantage of its new freedom of access to investment banking was involved in the most serious spate of corporate financial scandals of modern times – including the collapse of Enron and WorldCom – resulting in large losses for the banks themselves and their investor clients, major fines and legal settlements, and a general erosion of confidence in financial markets. Using their enormous balance sheets, the new financial conglomerates had become fee-chasing Goliaths that were vulnerable to rogue clients playing them off against each other and against the independent investment banks.

3 Among the remaining constraints, the 1999 deregulation did not remove the restrictions on banks under the Bank Holding Company Act of 1956 (BHC), which prevented financial institutions from owning non-financial corporations. It conversely prohibited corporations outside of the banking sector from entering deposit taking and commercial lending. This prompted many nonfinancial corporations, such as General Electric and BMW, to set up industrial loan corporations (ILCs), mainly chartered in Utah, which enabled them to take FDIC insured deposits and make commercial loans. It also allowed broker-dealers and investment banking units of financial conglomerates to set up FDIC-insured ILCs to offer their clients in the form of brokerage sweep accounts. Also left in place was a cap of 10% on total U.S. deposits booked by any single bank holding company, although the largest financial conglomerates soon lobbied for the cap to be lifted.

All of the new investment banking units of bank holding companies became embroiled in major regulatory violations and exploitation of conflicts of interest, including corrupted equity research, facilitating late trading and market timing by hedge funds against the interests of ordinary shareholders of in-house mutual funds, and acting simultaneously as principal and intermediary in corporate actions. The much-criticized Sarbanes-Oxley Act followed as one effort to try to redress abuses and restore market integrity.

At the same time, the power of the newcomers in competing for investment banking market share was impressive. Less than a decade after the 1999 deregulation, only five major independent investment banks remained – Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch, and Morgan Stanley. All the rest succumbed to the newly unchained financial conglomerates' use of their massive balance sheets to capture investment banking business, ranging from underwriting and market-making to merger advice and proprietary trading. The dominant U.S. financial conglomerates were joined by three foreign universal banks in the global wholesale league tables – Credit Suisse, Deutsche Bank, and UBS AG.

Together, the top-ten global wholesale firms led over 80% of investment banking transactions by value in 2006, the year before the start of the global financial crisis. As a consequence, all ten global wholesale banks were at the epicenter of the crisis. Besides encountering "pipeline" exposure to market, credit, and liquidity risk in pursuit of what proved to be toxic flow-business, many also took on "warehouse" exposure on their massive balance sheets or in off-balance sheet conduits set up to avoid regulatory capital requirements [Acharya and Richardson (2009)].

In the event, of the independent investment banks one (Bear Stearns) failed and was absorbed by JP Morgan Chase with the aid of large-scale taxpayer assumption of risk. A second (Merrill Lynch) was absorbed by Bank of America with the encouragement of the Treasury and Federal Reserve on highly favorable terms for Merrill shareholders and employees. A third (Lehman Brothers) was allowed to fail, triggering substantial turmoil in global financial markets. The other two (Goldman Sachs and Morgan Stanley) became bank holding companies in order to obtain access to Federal Reserve refinancing. So ended the era of globally-competitive, independent U.S. investment banks.

Among the foreign-based wholesale universals, UBS and Royal Bank of Scotland were rescued by their respective governments. Others like Credit Suisse, Barclays, and Deutsche Bank managed to survive on their own, but some would doubtless have failed in 2008 had they not by then become “systemic” institutions, benefiting from critical central bank funding and government debt guarantees that transferred to taxpayers key risks they had assumed in executing their wholesale banking strategies.

Emerging from the global financial crisis and the ensuing public-sector debt problems in the Eurozone, the dominance of large, complex, interconnected bank-based financial conglomerates active across the spectrum of banking and “shadow” financial intermediation – renamed “systemically important financial institutions” (SIFIs) – was greater than ever before. Events demonstrated that in such a financial architecture there is no substitute for government risk absorption *in extremis*. They also demonstrated that efforts to monitor and constrain systemic risk centered on such firms – and to resolve them when they fail – remain unproven. Except for the U.K., proposals to reconfigure the financial architecture toward smaller, more specialized, easier to regulate and resolve, and ultimately less systemic financial institutions have been uniformly turned back.

The SIFI issue is amplified in the context of today’s global financial markets, where different institutional structures compete with one another, as well as international financial firms competing in various local markets around the world. Passage of the 2010 Dodd-Frank Act in the U.S., the Basel III capital and liquidity rules, regulatory initiatives to “ring-fence” retail banking in the U.K., and regulatory initiatives in the E.U., Switzerland, and elsewhere have moved the debate along. It invariably comes back to the linkage between size and organizational structure of financial intermediaries – either as a source of systemic risk or as a key issue in crisis-resolution after systemic damage materializes. Again, implicit in all of these efforts is the assumption that the large financial conglomerate model will dominate the financial architecture going forward. [See Acharya et al. (2009 and 2011) for a detailed discussion of the issues in retrospect and prospect].

Value creation and value destruction in SIFIs

If SIFIs today dominate global financial intermediation and are indeed the source of serious systemic risk, then the lack of interest in systemic risk-reduction and regulatory simplification

by separating some of their key functions in the wake of the last crisis may not survive the next one. Given the severity and duration of the fallout from the 2007-09 financial turbulence, the perceived inequity of its impact, and the high-pressure lobbying by banks to undermine some of the key regulatory reforms that have been enacted, it seems unlikely that the political economy will tolerate another “made in finance” debacle anytime soon – especially if resolution of failing firms turns out to be impossible without once again calling on taxpayer support. Next time around, SIFIs could be targeted as a malignant institutional form that tends to “manufacture” systemic risk. Yet breakup of SIFIs *in extremis* is unlikely to be costless, imposing potential efficiency offsets to any improvement in safety and soundness that is achieved. Why? Because SIFIs are, at least in part, the product of market forces whose benefits would have to be sacrificed in any institutional restructuring that breaks them up.

Like good architecture, “form follows function” in optimum financial system design. Industrial economics suggests that the structural form of firms in particular lines of economic activity should follow the dictates of institutional comparative advantage. If there are significant economies of scale that can be exploited, it will be reflected in firm size. If there are significant economies of scope that can be exploited – either in costs or revenues – it will be reflected in the range of activities in which successful firms are engaged. If productive linkages can be built across geographies or client segments or business lines, this too will be reflected in the breadth and geographic scope of the most valuable and productive intermediaries.

In a pure market-driven context, optimum institutional structure is driven by the production functions of financial intermediaries on the one hand, and preference functions of end-users on the other. Distortions in markets for financial intermediation in the form of taxes, explicit or implicit subsidies, and regulatory constraints will alter structural optimization and create efficiency losses in the financial system. Financial institutions become larger or smaller, broader or narrower than what is optimal for the intermediation tasks to be performed. From a public interest perspective, these distortions are justifiable only if the associated social gains – including improved safety and soundness and reduced moral hazard – exceed the associated efficiency losses. This balance is not easy to achieve, yet it is at the core of the financial architecture debate.

The underlying drivers of the structure of financial intermediaries thus center on two questions. Is bigger better? Is broader better? Put simply, scale and scope. Better for whom? For shareholders? For clients? For regulators? For society? The empirical evidence so far allows no definitive conclusions on some of the key issues, with the datasets needed to find defensible evidence reflecting many moving parts and making *ceteris paribus* assumptions difficult.

Scale

Whether significant economies or diseconomies of scale exist in financial services has been at the heart of strategic and regulatory discussions about optimum firm size in the financial services industry for decades. Are larger firms associated with increased scale economies and hence profitability and shareholder value? Can increased average size of firms create a more efficient financial sector? Credible empirical work must isolate the impact of pure size-effects imbedded in the production function of the financial intermediary, and isolate them from other gains or losses related to firm size.

Cost estimation has consistently found that economies of scale are achieved with increases in size among small commercial banks (e.g., below U.S.\$100 million in asset size). A few studies have shown that scale economies may also exist in banks falling into the U.S.\$100 million to U.S.\$5 billion range [Cornett et al. (2003)]. There is limited evidence so far of scale economies in the case of banks larger than U.S.\$5 billion in asset footings. There is some scattered evidence of scale-related cost gains for banks up to U.S.\$25 billion in asset size [Houston and Ryngaert (1994)]. But there is no credible evidence so far of significant firm-wide scale economies among very large banks. Some studies have found the relationship between size and average cost to be U-shaped. This suggests that small banks can benefit significantly from economies of scale, but that large banks seem to suffer from diseconomies of scale, resulting in higher average costs as they increase further in size. Overall, the consensus seems to be that scale economies and diseconomies generally do not result in more than about 6% difference in unit costs interquartile range among U.S. commercial banks.

A basic problem is that most of the available empirical studies focus entirely on firm-wide scale economies, even though the important scale issues are encountered at the level of individual

businesses. Economies of scale appear to drive operating economies and competitive performance in areas such as global custody, transactions processing, and institutional asset management, for example. Economies of scale may be far less important in other areas such as private banking and M&A advisory services. Unfortunately, information on the business-line cost functions needed to identify and calibrate economies of scale are generally proprietary and rarely available to researchers. Nevertheless, findings of relatively small-scale economies in firm-wide cost functions, in the presence of significant observed scale economies in specific activity lines, suggest that important unobserved diseconomies of scale exist in other parts of the firm-wide cost functions.

Moreover, the presence of significant economies of scale at the business-line level does not necessarily provide a compelling reason for market dominance by large financial conglomerates, since these benefits could also be harvested by firms that specialize in scale-driven activities. Examples include financial specialists in scale-sensitive activities like custody, clearance and settlement, and transactions processing. Specialized scale-driven firms such as Bank of New York Mellon, State Street, Euroclear, and BlackRock are among the most successful in the industry.

Operating efficiency

Adjusting for scale, there is ample evidence that significant performance differences exist among financial intermediaries – as reflected for example in cost-to-income ratios among banks, insurance companies, and investment firms of comparable size [for a survey, see Walter (2006)]. Operating efficiency differentials in production functions center on efficiency and effectiveness in the use of labor and capital, sourcing and application of available technology, and acquisition of inputs and organizational design, alongside the design of compensation and incentive systems. This suggests that the way financial intermediaries are run may be more important than raw size or the selection of businesses that they pursue [Berger and Humphrey (1992)].

This is good news for smaller financial intermediaries, suggesting significant managerial degrees of freedom in the drive for efficiency – although if very large institutions are systematically better managed than smaller ones, there may indeed be a link between firm size and operating efficiency. It is also possible that

very large organizations may be more capable of the massive and “lumpy” capital outlays required to install and maintain cutting-edge information-technology and transactions-processing infrastructures. If extremely high recurring technology spend-levels result in greater operating efficiency, then large financial services firms will tend to benefit in cost-competition with smaller ones.

Scope

Adjusting for scale, the range of activities engaged in by financial intermediaries can have both cost and revenue benefits.

Cost reductions may be achieved by selling a broader rather than narrower range of products – i.e., the joint production of two or more products or services may be accomplished more cheaply than producing them separately. Cost economies of scope can be harvested through the sharing of IT platforms, branch networks and other overheads, information and monitoring costs, and the like. Like economies of scale, cost-related scope economies should be directly observable in production functions of financial services suppliers and in aggregate performance measures. Most empirical studies have failed to find significant cost-economies of scope in the banking, insurance, or securities industries [De Long (2001a)]. They do suggest that some cost-diseconomies of scope are encountered when firms in the financial services sector add new product ranges to their portfolios – as the product range widens, unit-costs seem to go up, although not dramatically so. Like scale economies and diseconomies, significant cost-scope economies or diseconomies would be reflected in valuations of financial intermediaries, and therefore affect optimum institutional structure.

Revenue economies of scope in financial intermediation arise when the all-in cost to the buyer of multiple financial services from a single supplier is less than the cost of purchasing them from separate suppliers. This includes search, monitoring, and contracting costs. Additionally, financial intermediaries that are diversified into several types of activities or several geographic areas tend to have more contact points with clients, supporting cross-selling. On the other hand, revenue-diseconomies of scope could arise from management complexities and conflicts of interest associated with greater breadth. Because of the proprietary nature of in-company data required to confirm or refute the existence of revenue economies of scale at the

individual product level, few empirical studies of revenue-scope economies in financial services are available so far [De Long (2001b), Houston and Ryngaert (2001)].

In terms of revealed behavior, persistent management focus on cross-selling suggests that revenue economies of scope in financial intermediation may indeed exist at both the wholesale and retail levels. But they are likely to be very specific to the types of services cross-sold and the types of clients served [Mitchell and Onvural (1996)]. As always, the devil is in the detail – especially in the design of incentives and organizational structures to ensure that cross-selling actually occurs.

Universal banks and financial conglomerates like to argue that any imposition of size and scope constraints would inhibit their ability to provide financial intermediation to non-financial corporate and institutional investor clients that are themselves large and diverse. On the other hand, it seems likely that the broader the activity-range of financial firms in the presence of imperfect information, the greater the probability that the firm will encounter potential conflicts of interest, the higher will be the potential agency costs facing clients, and the more difficult and costly will be the internal and external safeguards necessary to prevent conflict exploitation. The competitive consequences associated with conflict exploitation may counteract the realization of revenue economies of scope among financial services firms.

Market power

As in other industries, financial intermediaries will seek to dominate markets in order to extract economic rents – sometimes referred to as economies of “size” as opposed to classic economies of “scale.” Exploitable market power allows intermediaries to extract rents from consumers or users of financial services and redistribute them to shareholders and employees, cross-subsidize other areas of activity, invest in wasteful projects, and reduce pressures for cost-containment. It may also affect safety and soundness, since excess margins allow financial intermediaries to be better capitalized and more resistant to shocks – an argument that may apply to the comparatively good performance of the big Canadian and Australian banks in the 2007-09 financial crisis, for example. In effect, end-users of financial services may pay a monopoly-surcharge, part of which takes the form of an insurance premium that supports improved firm robustness.

Still, despite very substantial consolidation in recent years in the most concentrated segment of the financial services industry, wholesale banking and capital markets, there is little evidence of market power based on standard metrics like the Herfindahl-Hirschman index.⁴ This is reflected in the meager risk-adjusted total returns to investors who own shares in the principal players in the industry and the low valuation metrics exhibited by many of these firms. Post-crisis consolidation in retail financial intermediation, on the other hand, may have reached unacceptable concentration in some national markets. E.U. competition authorities have required substantial post-consolidation divestitures, although rising banking concentration levels in the U.S. are not yet considered problematic.

Proprietary information

One argument supporting large, diverse financial intermediaries is that internal information flows are substantially better and involve lower costs than external information flows available to more narrowly-focused firms. Consequently, a firm that is present in a broad range of financial markets, functions, and geographies can find client-driven trading, financing, and investment opportunities that smaller and narrower firms cannot. Moreover, some areas of financial intermediation, like wholesale banking and asset management, have become the repository of dedicated expertise – embodied human capital in specialist businesses that are conducted by specialists to meet specialist client requirements.

It seems unclear whether size or breadth has much to do with the proprietary-information attribute of financial intermediation. Indeed, there has been plentiful anecdotal evidence of high-performance talent deserting SIFIs to join hedge funds and investment banking boutiques which depend on external information and derive competitive advantage from the ability to interpret and act on it.

Diversification and financial stability

Greater diversification of earnings attributable to multiple products, client segments, and geographies may be associated with more stable, safer, and ultimately more valuable financial institutions. The lower the correlations across the cash flows from a firm's various financial intermediation activities, the greater the

benefits of diversification. This should produce higher credit quality and higher debt ratings (lower bankruptcy risk), and therefore, lower cost of capital than faced by narrower, more focused firms. Likewise, greater earnings stability should bolster share prices.

Empirical evidence does indeed suggest relatively low correlations among cash flows from various businesses conducted by financial conglomerates, based in part on studies of earnings stability in hypothetical mergers of different types of financial intermediaries over time [see simulations in Saunders and Walter (1994)]. If true, this would be an important argument against breaking up financial conglomerates and universal banks.

Conglomerate discount

It is often argued that the shares of multi-product firms and business conglomerates tend to trade at prices lower than shares of more narrowly focused firms (all else being equal). There are two basic reasons why this “conglomerate discount” is alleged to exist.

First, it is argued that, on the whole, conglomerates tend to use capital inefficiently. This may be attributable to managerial discretion to engage in value-reducing projects, cross-subsidization of marginal or loss-making projects that drain resources from healthy businesses, misalignments in incentives between central and divisional managers, and the like. The bulk of value erosion in conglomerates is usually attributed to over-investment in marginally profitable activities and cross-subsidization [Berger and Ofek (1995)].

Second, a source of a possible conglomerate discount is that investors in shares of conglomerates find it difficult to “take a view” and add pure sectoral exposures to their asset portfolios. Investors may want to avoid such stocks in their effort to construct efficient asset-allocation profiles. This is especially true of performance-driven managers of institutional equity portfolios who are under pressure to outperform cohorts or equity indexes. Why would a fund manager want to invest in yet another (closed-end) fund in the form of a conglomerate – one that may be active in retail banking, wholesale commercial banking, middle-market lending private banking, corporate finance, trading, investment banking, asset management insurance, and perhaps other businesses as well?

Both the capital-misallocation effect and the portfolio-selection effect may weaken investor demand for shares of universal banks

⁴ The HHI for the composite of wholesale banking activities has been less than 1,000 (sum of the squares of percentage market shares $0 < \text{HHI} < 10,000$), and the industry remains subject to ruthless competition in most of the constituent businesses.

and financial conglomerates, lower their equity prices, and produce a higher cost of capital than if the conglomerate discount were absent. Recent empirical studies using large data panels have attempted to ascertain whether or not functional diversification is value-enhancing or value-destroying in the financial services sector. They conclude that there is a value-reduction effect of about 20% [Laeven and Levine (2007), Schmid and Walter (2009)].

The impact of business scope on the balance between greater stability (low earnings correlations among the businesses) and value destruction imbedded in highly diversified structures (associated with the conglomerate discount) raises some intriguing implications. Shareholders of universal banks and financial conglomerates may in effect be paying for diversification-driven systemic risk reduction through share-price erosion and relieving taxpayers of some of their exposure to systemic risk. This would be in line with an objective to pass some of the costs of potential bailouts back to firm shareholders, who might then pressure their boards to redirect strategy toward greater focus and less systemic firm characteristics, while at the same time lifting some of the conglomerate discount.

Monetary policy and surveillance

Two important issues often neglected in the universal banking debate are the effects on monetary policy and bank surveillance. Here it seems that both are somewhat open issues.

First, with respect to monetary policy, if there are just a few large banks in the financial system (such as in Canada and Australia), open-market operations can be quickly transmitted to the economy via the lending decision of just a few intermediaries. That is, the transmission time would be fairly short. In contrast, if there are large numbers of widely dispersed small and big banks in the system, the effect of central bank open-market operation might first affect large banks and then gradually feed down to smaller banks and small firms and consumers. This may be one reason why the lag between a monetary policy action and its effect on the real sector has been long and relatively unpredictable in the U.S. A second potential monetary policy advantage of a universal banking system is the possibility of the central bank or appropriate regulatory agency imposing moral suasion. It would seem far easier for regulators to apply pressure to undertake certain actions (e.g., increase corporate lending) in a system with 8 banks rather than 8,000.

With respect to surveillance, there are clear costs and benefits of a universal banking structure. On the one hand, regulators are faced with auditing a highly complex organization with systemic links both domestically and internationally. This may be especially so given likely differences in financial sophistication of a relatively underpaid bank examiner up against a relatively overpaid financial engineer. On the other hand, the cost and expense of auditing on an annual basis 8 rather than 8,000 banks may go some way to offset these costs, at least in an operational if not a systemic risk sense.

Benchmarking the role of SIFIs

The previous section of this paper suggests that the growing dominance of SIFIs in the global financial architecture reflects a complex balance of positive and negative economic drivers – scale, scope, market power, diversification, and conglomeration. The facts on the ground suggest that the positives must outweigh the negatives. Otherwise, SIFIs would play a much more limited role in financial intermediation than they actually do. It follows that efforts to break them up or to “ring-fence” certain activities, in the interest of reduced systemic risk and improved regulatory feasibility, will result in significant offsets in terms of erosion of financial system performance.

On the other hand, the fact that these institutions are systemic may itself explain their dominant role in the financial architecture. The argument is well-known. Without state assurances, uninsured depositors and other liability holders demand a risk premium. When a financial institution is not permitted to fail, the risk premium is no longer necessary. So, given the unacceptable consequences of institutional collapse, debt and equity investors in systemically sensitive financial services firms that surpass certain thresholds of size, complexity, and interconnectedness expect to be bailed-out by taxpayers. This results in materially lower cost of capital in comparison with non-systemic financial intermediaries, tilts the competitive playing field in their favor, and in turn allows them to capture progressively greater market share.⁵

⁵ This argument is similar to that associated with Fannie Mae and Freddie Mac in the U.S., two government sponsored enterprises (GSEs) that gradually encroached on private sector mortgage finance to achieve dominant positions in the market, based on an implied government debt guarantee and hence a AAA debt rating. In the event, the implied guarantee of the two toxic GSE-SIFIs became a reality during the 2007-09 financial crisis, with no structural solution yet at hand.

Moreover, financial intermediaries that expect to be covered by bailouts have an incentive to increase their risk in order to enjoy higher equity returns (moral hazard). In effect, they receive an un-priced subsidy from the taxpayer, which arguably distorts the underlying scale, scope, and other economic drivers of firm structure that drive their structure. Second-order distortions turn up in employee compensation levels far in excess of other sectors of the economy, out of proportion to value added to the real economy by the financial sector, and encouraging financial activity that is wealth-redistributing rather than wealth-creating. SIFIs' competitive dominance, this argument suggests, is largely a product of their own systemic importance, with the taxpayer held hostage in the process.⁶

In the U.S., bailout policy toward SIFIs became explicit as early as 1984, when the Comptroller of the Currency testified to Congress that 11 banks were so important that they would not be permitted to fail. In other countries the same policy existed, and covered even more of the local financial systems than in the U.S. There were numerous examples of government support of banks in France, Switzerland, Norway, Sweden, Finland and Japan during the 1990s. The financial turbulence of 2007-09 further hardened the SIFI bailout presumption in virtually all of the affected countries, and, along with crisis-driven consolidation, further reinforced their competitive dominance.

The key question is how SIFIs contribute to, or detract from, the performance of the financial system, viewed from the perspective of the public interest. We suggest there are four benchmarks against which the financial system should be calibrated:

▶ **Static efficiency** – the metrics include the weighted mean spread between what ultimate savers (predominantly households) receive and what ultimate users of capital (households, nonfinancial businesses, and governments) have to pay. That spread is some composite of operating costs, regulatory costs, and intermediation losses. Fairness and transparency are additional aspects of static efficiency.

- ▶ **Dynamic efficiency** – here the metrics are less transparent, and include product and process innovation and technology change in financial intermediation and the role of the financial system in promoting economic growth by continually allocating and denying capital to competing uses in the production function that drives the real economy.
- ▶ **Stability** – the financial system itself should be sufficiently robust to withstand shocks that will inevitably emanate from the real sector from time to time. Furthermore, the financial sector itself should not generate shocks that spill-over into the real sector, generating serious negative externalities.
- ▶ **Competitiveness and robustness** – in a macro sense, the financial sector is an industry like any other, and generates income, employment, and international trade in services. Countries compete vigorously to maintain financial systems that add value in this regard.

These four benchmarks may well be in conflict with one another, and involve trade-offs that are often hard to identify and measure. Nevertheless, options for financial reform should be consistently calibrated against these benchmarks.

The more complex the industry, the greater the challenge to sensible regulation. This is nowhere as striking as in massive, complex, global financial services conglomerates that may be too hard to manage, too hard to oversee and govern, and almost certainly too hard to monitor and regulate. Consequently, the organizational structure of financial firms invariably enters into the regulatory debate which, in the real world, confronts heavily entrenched and politically connected players, and runs up against the personal financial interests of some of the brightest minds and biggest egos in business.⁷

There are two ways to address the central issue of internalizing systemic risk. One relies on market discipline and the other relies on functional separation. We believe the first is clearly preferable to the second, but depends more heavily on overcoming political economy, regulatory arbitrage, and corporate governance

⁶ Determining precisely when a financial institution becomes systemic raises some difficult issues. Citicorp was already the largest bank holding company in the U.S. before it merged with Travelers in 1998. So the bailout benefits of becoming still larger or more complex may be marginal if a firm already enjoys SIFI status.

⁷ Practitioners [e.g., Ackermann (2009)] argue that forcing structural change on systemic financial firms would be bad for global allocation of capital and risk, and new mechanisms would have to be found to preserve as much efficiency and innovation as possible in cross-border financial flows.

frictions that threaten its ability to produce a balance across the four benchmarks that is both durable and in the public interest.⁸

Under the market-discipline approach, if explicit or implicit subsidies of financial intermediaries, such as too-big-to-fail support, can be eliminated or properly priced, so that systemic externalities are successfully internalized, then the managements and boards of SIFIs will presumably rethink their strategies and their firm structures and react accordingly, probably becoming less systemic and certainly easier to monitor and regulate. Benefits (net of offsets) go to society by reducing its exposure to systemic risk. Benefits also go to shareholders by reducing or eliminating any conglomerate discount and allowing them to choose in which functions of financial intermediation to invest their capital – as opposed to holding shares in multi-function financial conglomerates.

Specifically, under Basel III, designated SIFIs will be subject to a capital surcharge for being systemic, and a further surcharge reflecting the potential systemic risk imbedded in their business models. Under this rule, Basel III capital ratios for Citigroup, HSBC, JP Morgan Chase, and Deutsche bank will be 9.5%, while BNP Paribas and Barclays will be 9%. Eight banks, including Bank of America, Goldman Sachs, Morgan Stanley, and Royal Bank of Scotland will be at 8.5%, and fourteen large banks that are more retail-oriented will be at 8%. These BIS Basel III capital levels do not preclude national regulators adding further surcharges. Switzerland announced a proposed “Swiss finish” with core capital ratios up to 16% for its two global banks. It would be hard to imagine bank managers and boards not rethinking their basic business models were such capital levels to be widely adopted. If they do not, investors may do it for them in the form of metrics that assign low values to current businesses and future prospects.

Indeed, there are grounds for optimism. For example, UBS in late 2012 announced a strategic redirection for its investment banking activities, to exit fixed income and commodities trading and focus on activities that support its valuable private banking franchise – with a loss of some 10,000 jobs. At the new capital levels, the risk-adjusted returns of these highly competitive activities make no business sense. Similar rethinking has been underway at other major wholesale banks such as Barclays, Citigroup, Royal Bank of Scotland, and others.

If the market-discipline solution prevails and systemic risk externalities are substantially internalized, then the issue of optimum organizational structure will resolve itself, and the associated incremental costs (offsets) will be disseminated among the end-users of financial intermediation as well as shareholders in an efficient general equilibrium solution. But if full application of the market-discipline solution to internalizing systemic risk is not possible – due to political blockage, regulatory arbitrage, or unresponsive corporate governance that leaves a residue of systemic risk – then resorting to mandated structural change involving activity constraints on SIFIs may be a second-best alternative to achieving better balance across the financial architecture benchmarks. Some of the separated financial firms will remain systemic, but they will be narrower and more specialized, and therefore more amenable to competent functional oversight by specialized regulatory agencies.

Two very different approaches to mopping-up residual systemic risk in the ongoing debate are the “Volcker Rule” under the Dodd-Frank Financial Reform and Consumer Protection Act of 2010 in the U.S. on the one hand, and the “ring-fence” proposal of the 2011 U.K. Independent Commission on Banking (Vickers Commission) recommendations as well as the E.U.’s 2012 Liikanen Report on the other.

Under the Volcker Rule, financial institutions with access to government lender of last resort (Federal Reserve) facilities and performing a vital public utility role in the payments system and as a transmission belt for monetary policy, should not be engaged in activities that generate systemic risk. Following vigorous debate, these prohibited activities now include proprietary trading in financial instruments and their derivatives (except U.S. government and municipal securities) and principal investing in hedge funds and private equity funds in excess of a cumulative total of 3% of core capital. Passage of the Volcker Rule quickly led to shut-downs and divestitures of proprietary trading desks, although the distinctions with respect to market-making and client-driven trading remained subject to debate. Nevertheless, the necessary risk-taking no longer allowed in bank-related financial firms equally quickly migrated to independent hedge funds and private equity firms without major disruptions in the market [Acharya et al. (2011)].

⁸ This includes periodic threats of geographic migration to more permissive regulatory environments that allegedly would accept a lender of last resort role.

The U.K. “ring fence” proposal basically limits taxpayer guarantees to retail banking and certain other commercial banking services from the bulk of wholesale banking conducted unguaranteed outside the “fence” but within the same institution – with separate governance structures and differential capital buffers. It remains unclear whether either of these models neutralizes whatever systemic risk remains after firms have adjusted to capital buffers and other efforts at systemic risk internalization [Vickers (2011)].

The Liikanen report took a different ring-fencing approach. It envisages a strict ring-fencing of proprietary and third-party trading activities both in terms of risk management and capital allocation, essentially through a holding company structure which would control separately capitalized subsidiaries with credible firewalls between them. The trading division would have to hold its own capital and stands or falls by its own activities and cannot contaminate the firm’s retail and commercial banking units. The objective is to protect taxpayers from future bailouts without protecting stockholders or subordinated debt-holders of the trading unit of the holding company from loss, thereby promoting market discipline. The Liikanen Report also advocates a better alignment of executive pay with long-term corporate and taxpayer interests by raising the “debt” component of bankers’ bonuses by linking them to long-term profitability.

Observers who have thrown in the towel on a pure market-discipline solution – the core supporters of Volcker, Vickers, and Liikanen – point to the continued presence of specialists in the financial system. Despite government supported and implicitly subsidized financial conglomerates, this suggests that a modern version of functional separation would not be ruinous when benchmarked against the four aforementioned criteria. It suggests that a new financial architecture would emerge, populated by relatively transparent financial firms that lend themselves to comparatively straightforward oversight by functional regulators in tandem with a systemic risk regulator empowered to deal with banks as well as non-bank financial firms. SIFIs tend to argue that separation limits synergies that are essential to their business models, while the other side argues that a host of nonbank financial firms will step up to conduct those activities that contribute static or dynamic financial efficiency gains without risking systemic consequences, as long as they are appropriately regulated along functional lines with systemic risk oversight.

Conclusions

Regulators face the daunting task of designing an “optimum” regulatory and supervisory structure that provides the desired degree of stability at lowest possible cost to efficiency, innovation, and competitiveness - and to do so in a highly politicized environment in a way that effectively aligns such policies among regulatory authorities functionally and internationally and avoids “fault lines” across regulatory regimes that can be gamed and arbitrated. There are no easy answers. There are only “better” and “worse” solutions as perceived by the public, to whom the regulators are ultimately accountable.⁹

On the surface, functional separation in the U.S. from 1933 to 1999 may have in fact done little harm and a lot of good. Growth was respectable, financial efficiency and innovation eventually attracted imitators around the world, American firms dominated global financial markets, and financial crises were largely avoided. But that scenario hinges on the plausibility of the counterfactual that these gains would not have been realized if pre-1933 universal banking had persisted in the U.S. The rapid emergence of financial conglomerates thereafter coincided with plenty of problems – a spate of corporate scandals in which financial conglomerates were the leading facilitators, financial innovation aimed at thwarting regulation and redistributing wealth, market bubbles in tech stocks and real estate, slower growth, a financial crisis of massive size and long-lasting effects, and unprecedented taxpayer assumption of risk and losses. Circumstantial evidence? Possibly so.

Certainly, the emergence of the financial system from the latest crisis has provided a window of opportunity to once again recalibrate the rules in the direction of the public interest. To be sure, the 2010-11 Financial Crisis Inquiry Commission (Angelides Commission) was a pale imitation of the Pecora Hearings leading up to the bold reforms of the Banking Act of 1933. The findings included a clear rejection of forced structural change in SIFIs. The U.K.’s Vickers Commission and E.U.’s Liikanen likewise shied away from full-blown

⁹ Perhaps the most respected and experienced observer of the US financial system, former Federal Reserve Chairman Paul A. Volcker, noted in a speech in April 2008 that “... today’s financial crisis is the culmination, as I count them, of at least five serious breakdowns of systemic significance in the past 25 years – on the average one every five years. Warning enough that something rather basic is amiss... Simply stated, the bright new financial system – for all its talented participants, for all its rich rewards – has failed the test of the market place... [A] demonstrably fragile financial system that has produced unimaginable wealth for some, while repeatedly risking a cascading breakdown of the system as a whole, needs repair and reform.” Paul A. Volcker, remarks at a meeting of the Economic Club of New York, 28 April 2008, transcript at http://econclubny.org/files/Transcript_Volcker_April_2008.pdf.

SIFI reforms in favor of ring-fencing. Regulatory initiatives at the E.U. level uniformly accept SIFIs as a fact of life in the financial architecture. All accept the *raison d'être* and institutional dominance of large, complex, interconnected financial conglomerates, that Basel III and other efforts to internalize systemic risk will eventually succeed, and that the economic and social costs of forced structural change exceed the benefits. The good news is that all are converging on some form of separating the "casino" business from the core intermediation functions that form the financial air supply of the modern economy, either (a) by letting market forces push them into structurally simpler nonbank intermediaries that are either non-systemic or easier to oversee or regulate and resolve, or (b) by locating them in separately capitalized affiliates of financial holding companies, with credible firewalls capable of preventing contagion.

Arguably the most defensible approach to improving the financial architecture by addressing systemic risk, given the facts on the ground – and assuming it can be carried out in a disciplined, consistent, internationally coordinated and sustained manner with a firm eye to the public interest – will have to be heavily reliant on market forces in response to higher regulatory costs.

Forced to pay a significant price for generating systemic risk, SIFIs will have to draw their own strategic conclusions in the context of the microeconomics and industrial organization of global wholesale financial intermediation. But this assumes market discipline works effectively. Those who have become cynical about the political economy of regulation and "regulatory capture" have continued to advocate specific activity carve-outs as a second-best alternative to breaking the systemic risk grip of universal banks and financial conglomerates as a dominant form of organization in the financial architecture.

References

- Aachrya, V., and M. Richardson (eds), 2009**, *Restoring financial stability*, New York: John Wiley & Sons
- Aachrya, V., T. Cooley, M. Richardson, and I. Walter (eds), 2011**. *Regulating Wall Street*. New York: John Wiley & Sons
- Ackermann, J., 2009**, "Smaller banks will not make us safer," *Financial Times*, 30 July, 2009
- Berger, A. N., and D. B. Humphrey, 1992**, "Measurement and efficiency issues in commercial banking," in Griliches, Z. (ed.), *Output measurement in the service sector*, Chicago, IL: University of Chicago Press
- Berger, P. G., and E. Ofek, 1995**, "Diversification's effect on firm value," *Journal of Financial Economics* 37, 39-65
- Cornett, M. M., and H. Tehranian, 1992**. "Changes in corporate performance associated with bank acquisitions," *Journal of Financial Economics*, 31, 211-234
- Committee on Capital Markets Regulation, 2006**, Interim report of the Committee on Capital Markets Regulation (The Paulson Report). Washington, D.C.: US Government Printing Office
- DeLong, G., 2001a**, "Stockholder gains from focusing versus diversifying bank mergers," *Journal of Financial Economics*, 59, 221-252
- DeLong, G., 2001b**, "Focusing versus diversifying bank mergers: analysis of market reaction and long-term performance," Working Paper, CUNY
- Gande, A., M. Puri, A. Saunders, and I. Walter, 1997**, "Bank underwriting of debt securities: modern evidence," *Review of Financial Studies*, 10(4), 1175-1202
- Gande, A., M. Puri, and A. Saunders, 1999**, "Bank entry, competition and the market for corporate securities underwriting," *Journal of Financial Economics*, 54(2), 165-195
- Houston, J., and M. Ryngaert, 1994**, "The overall gains from large bank mergers," *Journal of Banking and Finance*, 18, 1155-1176
- Houston, J., C. James, C., M. Ryngaert, 2001**, "Where do merger gains come from? Bank mergers from the perspective of insiders and outsiders," *Journal of Financial Economics*, 60, 285-331
- Johnson, S., 2009**, "The quiet coup," *Atlantic Monthly*, May
- Kane, E. J., 1987**, "Competitive financial reregulation: an international perspective," in Portes, R., and A. Swoboda (eds.), *Threats to international financial stability*, Cambridge: Cambridge University Press
- Kane, E. J., 2001**, "Relevance and need for international regulatory standards," *Brookings-Wharton papers on financial services*, 87-115
- Laeven, L., and R. Levine, 2007**, "Is there a diversification discount in financial conglomerates?" *Journal of Financial Economics*, 85, 331-367
- Liikanen Report, 2012**, "Report of the European Commission's High-level Expert Group on Bank Structural Reform," Brussels, European Commission, October
- McKinsey & Co., 2008**, "Sustaining New York's and the US' global financial services leadership," Report commissioned by Mayor Michael Bloomberg and Senator Charles Schumer. Mayor's Office of the City of New York
- Mitchell, K., and N. Onvural, 1996**, "Economies of scale and scope at large commercial banks," *Journal of Money, Credit and Banking*, 28, 178-199
- Puri, M., 1994**, "The long-term default performance of bank underwritten security issues," *Journal of Banking and Finance*, 18(2), 397-418
- Puri, M., 1996**, "Commercial banks in investment banking: conflict of interest or certification role?" *Journal of Financial Economics*, 40(3), 373-401
- Puri, M., 1999**, "Commercial banks as underwriters: implications for the going public process," *Journal of Financial Economics*, 54(2), 133-163
- Saunders, A., and I. Walter, 1994**, *Universal banking in the United States*, New York: Oxford University Press
- Schmid, M. M., and I. Walter, 2009**, "Do financial conglomerates create or destroy economic value?" *Journal of Financial Intermediation*, 18(2), 193-216
- US Department of the Treasury, 2009**, "Bank regulatory reform. Rebuilding financial supervision and regulation," Washington, D.C. US Government Printing Office
- Vickers Commission, 2011**, Final report of the independent commission on banking: Recommendations, HM Government, London
- Walter, I., 1985, (ed.)**, *Deregulating Wall Street – commercial bank penetration of the corporate securities market*, John Wiley & Sons, New York
- Walter, I., 2004**, *Mergers and acquisitions in banking and finance*, Oxford University Press, New York

Ernst & Young

Assurance | Tax | Transactions | Advisory

About Ernst & Young

Ernst & Young is a global leader in assurance, tax, transaction and advisory services. Worldwide, our 167,000 people are united by our shared values and an unwavering commitment to quality. We make a difference by helping our people, our clients and our wider communities achieve their potential.

Ernst & Young refers to the global organization of member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit www.ey.com.

© 2013 EYGM Limited.

All Rights Reserved.

EYG No. CQ0069

The articles, information and reports (the articles) contained within *The Journal* are generic and represent the views and opinions of their authors. The articles produced by authors external to Ernst & Young do not necessarily represent the views or opinions of EYGM Limited nor any other member of the global Ernst & Young organization. The articles produced by Ernst & Young contain general commentary and do not contain tailored specific advice and should not be regarded as comprehensive or sufficient for making decisions, nor should be used in place of professional advice. Accordingly, neither EYGM Limited nor any other member of the global Ernst & Young organization accepts responsibility for loss arising from any action taken or not taken by those receiving *The Journal*.