The EY 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 banking and capital markets, wealth and asset management and insurance, as well as academics and policymakers 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 comprising 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 are also practical in their focus. The Journal will be published three times a year.

gfsi.ey.com
The Journal of Financial Perspectives

Part 1: Strategic

13 Driving a strong risk culture: a Swiss Re perspective
David Cole, Group Chief Financial Officer, Swiss Re

25 Future directions for foreign banks in China
Jack Chan, Managing Partner, Financial Services – Greater China, EY LLP
Geoffrey Choi, Assurance Leader, Financial Services – Greater China, EY LLP
Kelvin Leung, Banking & Capital Markets Leader, Financial Services – Greater China, EY LLP
Shelley Chia, Partner, Financial Services – Greater China, EY LLP

45 An E.U. financial transaction tax and the unintended consequences for risk management
Serge Wibaut, Professor of Finance, Université catholique de Louvain

55 Regulatory experience in the U.S. and its lessons for the European Union
Anthony Saunders, J.M. Schiff Professor of Finance, Stern School of Business, New York University

63 Firm structure in banking and finance: is broader better?
Markus Schmid, Professor of Corporate Finance, Swiss Institute of Banking and Finance, University of St. Gallen
Ingo Walter, Seymour Milstein Professor of Finance, Corporate Governance and Ethics, Stern School of Business, New York University

75 Six years after the crisis
Roy C. Smith, Kenneth Langone Professor of Finance, Stern School of Business, New York University

87 Transforming banks, redefining banking: global banking outlook 2014–15
Bill Schlich, Global Banking & Capital Markets Leader, EY LLP
Ian Baggs, Deputy Global Banking & Capital Markets Leader, EY LLP
Steven Lewis, Director - Global Lead Analyst, Banking & Capital Markets, EY LLP

Part 2: Tactical

109 Evaluating different approaches to quantitative easing: lessons for the future of central banking
Bluford H. Putnam, Chief Economist and Managing Director, CME Group

121 New results on the correlation problem in operational risk
Vivien Brunel, Head of Risk and Capital Modeling, Société Générale

129 Financial perspective: the unintended consequences of regulatory oversight and control – lessons from the banking and the asset/alternative funds industries
Frederick C. Militello, Jr., Adjunct Professor of Finance; Global and Investment Banking, Stern School of Business New York University and CEO and Senior Thought Leader, Future Change Management, LLC.

139 Risk management insights from Markowitz optimization for constructing portfolios with commodity futures
D. Sykes Wilford, The W. Frank Hipp Distinguished Professor of Business and Finance, The Citadel

149 Japanese patent index and stock performance
Takao Kobayashi, Professor of Finance, Aoyama Gakuin University
Yasuhiro Iwanaqa, Aoyama Gakuin University & Sumitomo-Mitsui Trust Bank, Ltd.
Hideaki Kudoh, Nomura Asset Management Co., Ltd.

149 An intraday event study methodology for determining loss causation
Alex Rinaudo, Senior Vice President, Compass Lexecon and Atanu Saha, Executive Vice President, Compass Lexecon

149 Are negative P/E ratio firms different from positive P/E firms?
The case of interlisted vs. non-interlisted firms in Canada
George Athanassakos, Ben Graham Chair in Value Investing, Ivey Business School, Western University, London, Ontario, Canada

149 Trust in banks after the financial crisis
Santiago Carbo-Valverde, Professor of Economics and Finance, Bangor Business School
Executive summaries

Driving a strong risk culture: a Swiss Re perspective
by David Cole, Group Chief Financial Officer, Swiss Re

Intelligent protection is driven through culture, not processes. Swiss Re applies the principle of intelligent protection to enable risk taking while ensuring appropriate oversight. Apart from helping to manage upside and downside risks to the benefit of the company’s bottom line, this also ensures consistency, develops market confidence and ultimately promotes a competitive advantage. The behaviors that result in intelligent protection are in turn determined and guided by a strong risk culture. The risk manager has a key part to play in enabling and shaping risk culture – within Risk Management, but much more importantly, across the entire organization. As the ability to demonstrate a strong risk culture is becoming the foundation of market confidence, the risk manager’s contribution creates essential value for the company. This article explores Swiss Re’s focus on risk culture, its drivers and Swiss Re’s ongoing effort to strengthen it. It also provides examples of how risk managers can support their partners across business units to further embed a strong risk culture throughout the firm.

Future directions for foreign banks in China
by Jack Chan, Managing Partner, Financial Services – Greater China, EY LLP, Geoffrey Choi, Assurance Leader, Financial Services – Greater China, EY LLP, Kelvin Leung, Banking & Capital Markets Leader, Financial Services – Greater China, EY LLP and Shelley Chia, Partner, Financial Services – Greater China, EY LLP

This article highlights the challenges and opportunities that foreign banks face in China. Based on interviews with 38 foreign bank Chief Executive Officers (CEOs) and senior executives between July and September 2013 – a period when several significant initiatives were underway, which will have a direct impact on the course that different foreign banks will chart over the next three to five years, such as the ongoing internationalization of the Renminbi, interest rate liberalization, and the form and scope of the recently announced Shanghai Free Trade Zone – this article provides insights on the future directions the foreign banks may pursue. We find that foreign banks will seek to grow fee income and develop their advisory role outside of China. The latter change will be assisted by Chinese corporates as they expand internationally, through acquisitions and greenfield investments. Such developments will foster foreign bank relationships, both offshore and onshore. Foreign banks will continue to evolve and choose different paths of expansion, based on their strengths and market opportunities. While some are building a retail presence and expanding their networks, others will focus on narrow market niches. Some of the larger foreign banks have divested their holdings in the large commercial banks, but simultaneously have moved into the securities and trust businesses. As the financial reforms take shape, new opportunities will emerge through the synergies gained by these diversification strategies.
An E.U. financial transaction tax and the unintended consequences for risk management
by Serge Wibaut, Professor of Finance, Université Catholique de Louvain
In the aftermath of the subprime crisis in the U.S. and the sovereign debt crisis in Europe, the opportunity for establishing a financial transaction tax (FTT) has become a topic of debate in the European Union. In this article, we survey the literature dealing with the possible theoretical and empirical implications of such a tax on market volatility. We then turn to the possible – and unintended – consequences of an FTT on savers and investors. We conclude that these consequences might outweigh the benefits of the FTT. More specifically, we find that an FTT is unlikely to meet its stated volatility control and revenue-raising objectives, i.e., an FTT is unlikely to decrease volatility, and indeed, volatility might increase as markets became less liquid. It might raise very little revenue and could work to create more risk and deter long term investment. And then there are serious unintended side effects to consider. Most importantly, for the financial security and safety of the whole financial system, an FTT might heavily penalize pension funds, as well as the banks in their liquidity management and risk management activities, to the detriment of a well-functioning financial system.

Regulatory experience in the U.S. and its lessons for the European Union
by Anthony Saunders, J.M. Schiff Professor of Finance, Stern School of Business, New York University
This paper looks at the challenges facing the Eurozone countries in establishing a banking union. Using the historical regulatory experience of the U.S. as a laboratory for comparative analysis, this article suggests that due to the diverse set of economies with competing local supervisors and diverse bankruptcy laws achieving the three “legs” of a European bank union, namely supervision, deposit insurance and restructuring/resolution, is by no means an easy task. Even today, the U.S., despite having introduced its first national banking regulation in 1863, still falls short of what might be viewed as a full banking union, despite the existence of a common currency from at least 1913 with the foundation of the Federal Reserve. Whether the Europeans can tackle this challenge with such a diverse group of countries remains to be seen.
Firm structure in banking and finance: is broader better?
by Markus Schmid, Professor of Corporate Finance, Swiss Institute of Banking and Finance, University of St. Gallen and Ingo Walter, Seymour Milstein Professor of Finance, Corporate Governance and Ethics, Stern School of Business, New York University

A focal point in strategic and regulatory debates about the pros and cons of diversification among financial institutions continues to be the issue of economies of scope in financial intermediation. In this paper, we summarize the theoretical research on the value of diversification in financial services firms, and survey the empirical research done so far on the conglomerate discount (the suggestion that diversified business are priced less favorably by the markets due to the fact that management expertise is spread too thinly, or that due to soft factors many of the benefits of conglomeration don't materialize, etc.) in U.S. and international financial services businesses. We also review research on the internal capital market efficiency of universal banks and financial conglomerates. The paper provides new empirical evidence on the conglomerate discount in U.S. financial intermediaries and how that changes between non-crisis and crisis periods, showing a decline in the discount under turbulent conditions. This suggests that the markets value the benefits of diversification among financial services firms during turbulent times as it helps protect the firm from severe issues in one part of its business. At other times, conglomerations are viewed less favorably by the markets.

Six years after the crisis
by Roy C. Smith, Kenneth Langone Professor of Finance, Stern School of Business, New York University

Since the great financial crisis of 2008, governments have set new records in market intervention, re-regulation of banking and financial markets, and in prosecution of banks for mismanagement. The new regulations touch just about everything in finance and completely change the financial regulatory system. However, these regulations have been hastily prepared and contain flaws that suggest they may not be effective in addressing systemic risk in the future. They also impose enormous compliance costs on the banking industry, even though no more than about 40 banks worldwide are systemically important. The costs will be passed on to customers of banks and users of financial markets, which may detract significantly from future economic growth prospects. Banks will have to adapt their business models to the new circumstances of their industry, or surely be challenged by shareholders until they do. But the way the regulatory platform is structured now, they will be adapting toward a more conservative, risk-averse form of what they were before. That may mean that too-big-to-fail becomes too-cautious-to-fail. If we have reduced the banks' capability to finance basic economic growth, a role that is equally important to our societies as systemic risk reduction, the regulatory response to the crisis may prove to be very mistaken. Risk has to be financed somewhere, if we are to return to the growth rates that we need to maintain and improve living standards. Most likely, the regulatory and other measures taken in response to the crisis will have to be modified in the future, but probably not in the near term.
Transforming banks, redefining banking: global banking outlook 2014-15
by Bill Schlich, Global Banking & Capital Markets Leader, EY LLP, Ian Baggs, Deputy Global Banking & Capital Markets Leader, EY LLP and Steven Lewis, Director - Global Lead Analyst, Banking & Capital Markets, EY LLP

Six years on from the tumultuous events of 2008, many aspects of the banking industry are almost unrecognizable from the way they were in the period leading up to the Global Financial Crisis (GFC). Whether it is the fixed income businesses of investment banks, the (mis-) selling of products to retail consumers, the demand for financial services from customers in emerging markets or the scale and intrusive nature of regulation and supervision, the industry has changed and will continue to change for some years to come. Responses by banks have been mixed. Many have been in a holding pattern, waiting for the regulatory landscape to become clearer and the global economy to improve. Some that embarked on strategic reviews in the wake of the GFC have been forced to revisit them as markets failed to recover and regulators imposed new rules, both domestically and internationally. This paper argues that the time for waiting may be over. Reshaping large global and regional banks is a multi-year undertaking and the broad direction of regulatory travel is now clear. The pendulum will swing from review to action, although the initial focus is likely to be on “no regrets” investments and initiatives. As they transform, banks must also incorporate and maintain enough flexibility in their models to respond to new rules as they emerge and survive further setbacks in the global economic recovery.

Evaluating different approaches to quantitative easing: lessons for the future of central banking
by Bluford H. Putnam, Chief Economist and Managing Director, CME Group

More than half a decade into the massive experiment with central bank balance sheet expansion known as quantitative easing (QE), research is turning to the question of what was actually achieved, how high were the costs, and just how difficult the exit process may prove to be. This article explores these questions by examining the lessons from the differences in how the U.S. Federal Reserve (Fed) practiced QE compared to the European Central Bank (ECB). Our analysis suggests that the first round of emergency QE in the fourth quarter of 2008 helped to avoid allowing a financial disaster to spiral into a depression. Moreover, the quick purchase of distressed assets by the Fed helped U.S. banks reduce their capital requirements and recover faster than European banks that only received emergency liquidity loans. Later and successive rounds of QE by the Fed purchased only liquid U.S. Treasuries as well as MBS, while the ECB remained focused on lending to banks with less emphasis on asset purchases. Our analysis argues that from 2010 onward, the Fed did not understand that U.S. job losses were coming exclusively from the government sector. Fed QE did not create any new jobs at all, while the unintentional consequences appear to have vastly complicated the exit process. By contrast, the ECB, mainly using loans to the banking system, appears to have avoided many of the exit risks faced by the Fed. Our conclusions are that central bankers are likely to use QE aggressively in the face of any new financial crisis, yet they may decide that the costs are not remotely worth the risks of using QE once an economy has entered a recovery period, even a weak one.
New results on the correlation problem in operational risk
by Vivien Brunel, Head of Risk and Capital Modeling, Société Générale
Internal models of operational risk are all built based on the same guidelines provided by the regulators. However, we observe a broad range of practices among banks concerning modeling choices and calibration methods. This paper discusses the relative importance of the main drivers and modeling choices of the operational risk capital charge. Many studies in the literature have focused on the modeling of the tails in the severity distributions. Here, we use a class of analytical models for operational risk in order to assess the relative importance of all parameters of the model. In particular, we show that the bank’s capital charge is not very sensitive to the dispersion in correlations, the average level of correlations being a much more critical parameter of the operational risk capital charge. We show that the assumption of uniform correlations is robust, contrary to what is often advised by internal auditors or regulators.

Financial perspective: the unintended consequences of regulatory oversight and control – lessons from the banking and the asset/alternative funds industries
by Frederick C. Militello, Jr., Adjunct Professor of Finance; Global and Investment Banking, Leonard N. Stern School of Business New York University and CEO and Senior Thought Leader, Future Change Management, LLC.
Regulators are largely problem solvers. They seek intended consequences. However, observation indicates that regulatory intentions and outcomes frequently do not coincide. If you will, there is a gap between the two; namely that of unintended consequences. Such consequences are the result of financial managers, and their organizations, facing increased regulatory induced business dilemmas. Such dilemmas must be managed, they cannot be solved. Moreover, they take the form of difficult choices – the essence of strategy – affecting the efficiency, innovativeness and competitiveness of financial organizations. As such, regulation has the unintended consequences of inspiring strategic thinking and organizational differentiation; reducing concentration and systemic risk. However, there is another unintended consequence afoot; namely, the regulatory induced impact on strategy and new business models is giving rise to increased interconnectivity of financial sectors and organizations. Intended consequences in one financial sector are spilling over into others, sometimes intentionally, sometimes not. Financial managers are busy assessing the results of regulation on comparative organizational capabilities, resulting in greater partnering across organizational and financial market segments. Clearly, when it comes to unintended consequences, and their systemic implications, there is a paradox. Regulators are becoming more proactive regarding these unintended consequences; increasing their regulatory reach. However, the paradox of unintended consequences is not a problem to be solved, it is a dilemma that must be managed with interactivity of the public and private sectors.
Risk management insights from Markowitz optimization for constructing portfolios with commodity futures
by D. Sykes Wilford, The W. Frank Hipp Distinguished Professor of Business and Finance, The Citadel

Diversification, via commodity investment, is touted as essential to every investor. How many times must we hear buy gold or own an oil well to “learn” that commodities in our portfolios will save us during the next crisis? Starting with the assumption that this advice is correct, this paper analyzes with a naïve Markowitz approach the hidden risks of incorporating commodity exposures into a portfolio. Using recent history, the very simplistic methodology chosen highlights portfolio problems that may occur. Naïve long-only diversification (with commodities) does not add a great deal relative to the event risk that can (and often does) occur. The opposite is concluded for long/short approaches. In such a context, commodities can be key drivers in risk mitigation over time, and even during crisis conditions, but they are only truly useful once the risk of using them is fully considered. Herein lies the problem with most commodity diversification plans. Simply adding commodities to a portfolio may actually create unintended risks, not intended diversification. Given that world markets can, and do, periodically erupt into turbulence, this paper is a useful guide to the pitfalls that arise when applying simplistic approaches to portfolio diversification with commodities.

Japanese patent index and stock performance
by Takao Kobayashi, Professor of Finance, Aoyama Gakuin University, Yasuhiro Iwanaga, Aoyama Gakuin University & Sumitomo-Mitsui Trust Bank, Limited and Hideaki Kudoh, Nomura Asset Management Co., Ltd.

In the global economy, high technology serves as a source of competitive advantage for Japanese companies. In Japan, there is a patent value indicator which is unique among other patent value indicators developed in the U.S. The uniqueness lies in its focus on measuring the exclusivity and technological competitiveness of each patent using data based on the number of actions taken by third parties against the patent. The construction of such a patent value metric became possible thanks to the Japanese Government’s active disclosure of information on patent attacks. This paper is our first attempt to study the relationship between technological competence and firm performance using this technology indicator.
An intraday event study methodology for determining loss causation
by Alex Rinaudo, Senior Vice President, Compass Lexecon and Atanu Saha, Executive Vice President, Compass Lexecon

In this paper, we set out an intraday event study methodology relying on minute-by-minute data and formulate an analytical framework to determine the window of time, i.e., the event window, over which stock prices fully reflect relevant new information. Event study is the commonly used method in securities litigation to determine loss causation, materiality and damages. While the typical securities litigation is brought against the issuer, damages have also been claimed against other financial market participants, such as underwriters and analysts. In addition, securities litigation claims are increasingly brought by pensions, asset managers and other financial institutions. While the traditional daily-price-data-based event study approach assumes the event window to be a full day, this paper’s methodology allows the data to determine the length of the event window. This is particularly relevant because many prior studies have shown that stock prices reflect new information within a matter of minutes. Our event study model not only provides a more accurate measurement of the stock price impact of a relevant event, but also determines whether the impact is statistically significant.

Are negative P/E ratio firms different from positive P/E firms?
The case of interlisted vs. non-interlisted firms in Canada
by George Athanassakos, Ben Graham Chair in Value Investing, Ivey Business School, Western University, London, Ontario, Canada

The main purpose of this paper is to examine whether negative P/E stocks are different from positive P/E stocks, and whether negative P/E stocks outperform, on average, the universe of positive P/E stocks. Moreover, the paper examines (a) whether interlisted and non-interlisted firms behave similarly or there are distinct differences between them, and (b) whether differences in positive and negative P/E stocks appear only in one group of firms or differences are equally driven by both interlisted and non-interlisted firms. The key finding is that firms with negative multiples are indeed different than firms with positive multiples in that (a) a relatively small number of firms with negative multiples experience high forward stock returns even though the majority of them do not, resulting in a large difference between mean and median returns, and (b) the value, size, liquidity and business risk premiums behave differently for negative versus positive P/E firms. Moreover, the paper also shows that there are key differences between interlisted and non-interlisted firms both in the positive and negative P/E space, and that the majority of the premiums referred above are driven primarily by the larger group of non-interlisted stocks. As a result, not only must negative P/E firms be segregated from positive multiple firms, but also interlisted firms ought to be segregated from non-interlisted firms in related research as aggregation would undermine the clarity and generality of findings, affect the homogeneity of the sample, and dilute findings and tests of significance.
Trust in banks after the financial crisis

by Santiago Carbo-Valverde, Professor of Economics and Finance, Bangor Business School

In this article, we examine the status of trust in financial institutions during and after the financial crisis and analyze the possibilities that financial institutions have to rebuild and improve trust from their customers. There are three pillars involved in the evolution of trust which interact with each other: i) changes in customer’s perception about financial services provided by banks; ii) regulatory actions on the financial sector; and iii) the transformations in the way banks provide their services and establish the relationships with their customers. The evidence shown in this article suggests that even if the three pillars are essential, changes in the levels of trust are mostly in the hands of the banks themselves. In this sense, relying upon recent empirical results for Spain – where the banking crisis has been among the most severe and long-lasting, and where trust in financial institutions is quite low compared to international peers – we show that by changing specific attributes of the services that they provide, financial intermediaries can improve trust from their customers and even offset the negative effects of the financial crisis.
Part 1: Strategic

Driving a strong risk culture: a Swiss Re perspective

Future directions for foreign banks in China

An E.U. financial transaction tax and the unintended consequences for risk management

Regulatory experience in the U.S. and its lessons for the European Union

Firm structure in banking and finance: is broader better?

Six years after the crisis

Transforming banks, redefining banking: global banking outlook 2014-15
Driving a strong risk culture: a Swiss Re perspective

David Cole
Group Chief Financial Officer, Swiss Re

Abstract
Intelligent protection is driven by culture. Swiss Re applies the principle of intelligent protection to enable risk-taking while ensuring appropriate oversight. Apart from helping to manage upside and downside risks to the benefit of the company's bottom line, this also ensures consistency, develops market confidence, and ultimately promotes a competitive advantage. The behaviors that result in intelligent protection are in turn determined and guided by a strong risk culture. The risk manager has a key part to play in enabling and shaping risk culture – within Risk Management, but much more importantly, across the entire organization. As the ability to demonstrate a strong risk culture is becoming a foundation for market confidence, the risk manager's contribution creates essential value for the company. This article explores Swiss Re's focus on risk culture, its drivers, and Swiss Re's ongoing effort to strengthen it. It also provides examples of how risk managers can support their partners across business units to further embed a strong risk culture throughout the firm.

1 As of 1 May 2014, David Cole was Swiss Re's Group Chief Risk Officer since March 2011 and is exercising that function until 31 August 2014.
1. Introduction
Over the last 25 years, the evolution of risk management within the insurance sector has been a key contributor to the industry’s resilience to a variety of external shock events, including the 2008–09 global financial crisis. Fundamental changes in risk management – and the rigor with which it is approached – were triggered by the events of 9/11 and natural catastrophes, such as winter storm Lothar in 1999 and hurricane Katrina in 2005. This impacted risk modeling and assumptions in particular, with dependencies across business lines subsequently seen in a new light.

Re/insurers are exposed to complex risks that require sophisticated risk management techniques. This is not an industry where simplification will always improve understanding. In many ways, the growth in complex risk modeling has served the business remarkably well. Such modeling is based on attempting to “think the unthinkable.” It takes a scenario-type approach that, instead of judging potential future outcomes based purely on past experience, shifts the question to “if this future event occurred, what would the outcome be?”

One of the arguments emerging from the global financial crisis is that people have become too dependent on models, and that models failed. A response to this argument is to ask for more “common sense” in risk management. The adage that one should only buy, sell or hold risks that one can understand is back in vogue; after all, “models do not make decisions, people do.”

Models should provide outputs that help us better understand the risks associated with a decision, enabling good decisions to be made within the complex global financial system and economy. However, discipline is required to challenge the assumptions of a model and understand its limitations, and to ask what stress tests have been undertaken and how the core principles of the modeling approach are shaping the results.

It is risk culture that delivers this “common sense” and discipline. Focusing on the cultural aspects provides context for technical modeling and helps us recognize the importance of a Risk Management function that enables and empowers individuals to act in a decisive and preemptive way, often against prevailing wisdom. Risk culture thus enables risk management to deliver competitive and resilient value for companies.

Recent regulatory developments are also driving the focus on risk culture. Members of the Board of Directors and senior management have an explicit responsibility to establish and maintain the “right” culture. Because of the ongoing international regulatory discussions, we expect this trend to accelerate.

With this greater regulatory focus, the ability to demonstrate a strong risk culture will have an increasing influence on market confidence. For these reasons, Swiss Re has sought to articulate a risk management vision encompassing our desired risk culture and engages in a number of initiatives aimed at further strengthening it across our organization. These initiatives focus on reinforcing Swiss Re’s risk management framework with clear expectations on employee behavior, talent management, skills development and incentives.

This article looks at the evolution of risk management in insurance and examines why risk culture is currently in the spotlight. It explores the various elements that underpin an effective risk culture and sets out some of the steps Swiss Re is taking to continuously embed and enhance risk culture within our organization.

2. Evolution of risk management
Risk management as a discipline has evolved from a broadly intuitive approach to one based on a set of objective elements, namely risk and capital modeling, risk governance frameworks and, recently, also steadily increased risk transparency and disclosure.

These elements are the cornerstones of an integrated approach to assessing and controlling risks. They provide the framework and tools that support business decision-making. However, they must be implemented in practice by talented individuals, particularly the ones who can balance an independent role as risk managers with being trusted partners of the core business. To contribute effectively to sustainable business success, they need to be both enablers and controllers of a company’s risk taking.


16 The Journal of Financial Perspectives
2.1 Risk and capital modeling

Through advances in technology, large and sophisticated companies are using group-wide risk and capital models to quantify risk and allow enhanced allocation of the cost of risk capital with the aim of increased capital efficiency. These models are enormously sophisticated, are a considerable achievement and can provide competitive insight. However, models are tools to inform and support decisions made by management.

Models represent a powerful but simplified and imperfect reflection of reality. They need continuous improvement to incorporate lessons learned from extreme situations. For instance, the last financial crisis revealed that certain diversification effects did not work as expected, especially financial market exposures. On the other hand, other diversification effects for insurance risks continue to behave as many models predicted, though this must be periodically reviewed and challenged.

It is of critical importance to understand the limitations of risk models and to be aware of when to use them and when to rely more on judgment.

2.2 Governance

The role of the Chief Risk Officer (CRO) is a symbol for the development of a holistic risk management approach. First established in the re/insurance sector by Swiss Re in 1997, this position gives risk management a seat at the top table of the organization. It reflects the increasing importance of creating a holistic and independent overview of the risks faced and translates this into actionable material for decision-making by the senior executive team. As with many aspects of risk management, the positioning of the CRO role reflects the challenge of ensuring that the risk manager is tightly integrated into the business and yet can maintain independence in both thought and action.

Our belief – and experience – is that this combination is best served by the CRO being a direct part of the top executive team, rather than reporting to the Chief Financial Officer, and that they should not have profit and loss responsibilities. Additionally, at Swiss Re, the CRO has a reporting line into the Finance and Risk Committee of the Board of Directors. This ensures that the CRO is fully aware of the organization’s risk appetite and can set risk limits appropriately, while also maintaining independence of judgment.

Under the stewardship of the CRO, a key step toward good risk governance is to ensure that there is a clear and complete definition – and documentation – of the organization’s risk appetite, the audit and review of any models used, and their appropriate disclosure. Furthermore, to ensure independence, it is important to clarify roles and responsibilities and to distinguish among the risk taker (business unit), controller (risk manager) and owner (Board of Directors), for the entire business and on specific transactions.

Individual risk takers and risk managers must understand the risk appetite of the risk owner. In governance terms, risk appetite should be set by the Board of Directors, based on available capital and required liquidity. The risk appetite limits are then defined – to determine where capital is best invested to generate the required returns – with a risk attitude statement describing the basis of engagement in risk taking activities.

This provides a framework for decision-making, where risk takers and risk managers can assess forces driving a deal, including the relationship between risk taking and remuneration and incentives. In this respect, risk managers must ensure that performance measurement reflects the risk/reward balance of the respective business or transaction.

A key part of the role definition for risk managers is detailing how the day-to-day risk oversight will work and how they will ensure their independence. It is vital that risk managers have the courage to speak up and make up their own minds: should they escalate the issue or persist in asking difficult questions until an answer emerges?

2.3 Disclosure and transparency

The disclosure of an organization’s risks to key internal and external stakeholders is a key driver of transparency. Swiss Re strongly believes that, going forward, this transparency will be a major factor in underpinning and maintaining credibility. Transparency is a non-negotiable aspect of the entire risk management process. Indeed, it can be argued that it is the major driver of the successful implementation of quantitative risk management and governance. Providing information to
stakeholders inside and outside the firm allows for independent validation and peer review of results. Of course, it also exposes weaknesses and opens up the potential for challenge. However, this creates trust between stakeholders, as material issues can be discussed in a practical, solution-oriented fashion.

2.4 Increasing focus on risk culture
As risk management has evolved, attention has now moved to the importance and effectiveness of the risk culture within an organization. Risk culture underpins all aspects of risk management. For the sake of a definition that may help the reader to understand what we mean by “risk culture” as a concept: risk culture captures the expected behaviors needed to provide confidence that a company is operating in accordance with its stated risk tolerance.

An effective and strong risk culture can deliver a number of significant benefits:

- Confidence, both within an organization and among stakeholders, regarding risk exposure and the way it is handled to deliver sustainable, profitable business. An effective and strong risk culture enables intelligent risk taking that optimizes return or avoids exposure to excessive loss.
- The ability to respond quickly and dynamically. Through the awareness and deliberate acceptance of the existence of different risks, a company can effectively evaluate their current position and take advantage of opportunities.
- Enabling preemptive assessment and reaction to emerging risks and developments as a result of a forward-looking mindset. This can potentially reduce the regulatory burden, as proactive engagement in developing and emerging risk can result in effective preemptive action that increases regulatory confidence and minimizes any possible regulatory censure.
- Awareness of the limitations of risk models and when to use them and when to rely more on judgment. A strong risk culture encourages employees to speak up against agreed wisdom and provide alternative perspectives.
- Reinforced risk management, control and awareness, enabling considered and effective responses to risk, as the organization’s culture has acknowledged the risks for what they are, for better or for worse.

Regulatory requirements reinforce the increasing focus on risk culture – companies are required to show not only what they do but how they do it. The concept that “meeting all the rules is sufficient” is increasingly rejected. Although not new, this is most effectively illustrated by qualitative risk management requirements and risk governance. For example, under Pillar 2 in Solvency II it is not sufficient to merely report capital numbers in accordance with the requirements. Instead, companies are required to produce their own risk and solvency assessment and to disclose how they manage risks, and their appetite for risk, within the organization.
3. Key characteristics underpinning risk culture

In order to build awareness of risk culture, an initial step is to ensure that the governance and risk management framework recognizes and articulates a desired risk culture. Within a risk management framework, the following four elements can be seen as underpinning risk culture and help to assess its depth:

- **Controlled risk taking** - to ensure sustainable value creation, it is important that risk taking is done in a controlled fashion. This can be achieved through a risk control framework that sets a clearly defined risk mandate as part of overall risk management standards. Key elements of the risk control framework are risk exposure limits, restrictions on contract features or contract types, and related referral processes. This framework sets the requirements and boundaries within which all risk takers are expected to act and should be reinforced through links to performance reviews, incentives and remuneration.

- **Clear accountability** - most companies operate on the principle of delegated authority. Ultimate responsibility for the outcomes of decisions will always remain with the delegating authority. However, within the framework of delegated authorities, individuals need to be clearly accountable for the decisions they take. This accountability should be reinforced through incentives aligned with overall business objectives.

On this basis, delegated authorities should be documented. In addition, when delegating authority, performance targets must be communicated, along with escalation paths.

- **Independent risk controlling function** - conflicts of interest can be avoided through the creation of dedicated units that monitor risk origination activities and provide independent assurance that risks are being appropriately modeled and managed and that adequate control instruments are in place. In order to discharge these responsibilities, Risk Management should be given unrestricted access to all information deemed relevant. In addition, there should be clarity that assurance can be provided by external and internal audit functions and Risk Management on the processes.

- **Transparency** - risk transparency and responsiveness to change are key to risk control processes. There should be processes that facilitate risk management knowledge sharing at all levels. The central goal of risk transparency is to create a culture of mutual trust and to reduce the likelihood of surprises regarding the source and potential magnitude of losses. This goal is achieved through regular dialogue and by establishing timely and appropriate risk reports, which document risk landscapes and loss potentials.

---

*Figure 1: Swiss Re’s Risk Management vision*

*Source: Swiss Re*

*Figure 2: Swiss Re’s holistic vision for “risk culture”*

*The Swiss Re risk culture model joins the dots between relevant frameworks and enablers to drive a consistent set of risk behaviors and attitudes across the organization.*

*Source: Swiss Re*
Leveraging off these characteristics, Swiss Re has been able to extend its universe of insurable risks and to enhance its ability to manage existing risk exposures, particularly through the use of risk radars (see Box 1).

**Box 2: Identifying emerging risks – Swiss Re’s SONAR process**

The SONAR (systematic observation of notions associated with risk) process is the key pillar of Swiss Re’s forward-looking risk management. Swiss Re defines “emerging risks” as newly developing or changing risks that are difficult to quantify and could have a major impact on society and the insurance industry.

SONAR is Swiss Re’s tool for identifying, assessing and managing emerging risks. By means of a network of experts across the company and an interactive web platform, Swiss Re collects early signals of emerging risks. All signals are reviewed, assessed and prioritized by a dedicated emerging risk management team, which closely interacts with a number of topic experts in Swiss Re’s various business areas. The findings are regularly summarized, distributed to relevant stakeholders throughout the company and made available to all employees via Swiss Re’s intranet. In 2013, Swiss Re published its SONAR report for the first time to a broad external audience.

Source: “Swiss Re SONAR: Emerging risks insights”. Available at swissre.com

**Box 3: Country Risk Management Advisory Group for a more holistic view on risk**

To provide business units with a more integrated view on risk, Swiss Re currently pilots Country Risk Management Advisory Groups, which include one Risk Management representative from each department – Actuarial Management, Risk & Governance, Operational Risk, Financial Risk Management and Governmental Affairs & Sustainability. Each of the participants takes responsibility to pull together the relevant information, communicate to the team, and thus generate a holistic view on risk. Insights from this RM Advisory Group are consolidated and integrated risk reports are provided to business.

An initial step was to articulate a central vision for Swiss Re’s risk management and risk culture (Figure 1). “Forward looking,” “integrated” and “trusted” are the three risk management dimensions which form the essence of our vision. In order to achieve our desired risk culture and encourage a more adaptive and intuitive approach to risk, this program aims to enhance three areas:

- “Intelligent protection” is about enabling a more “intelligent” means of protecting the business, by understanding planned risks, supporting the taking of good risks, and the top-down management of the biggest threats, both known and emerging.
- “Confidence” aims to increase the level of trust in Swiss Re’s risk management capabilities, both externally and between different internal functions and teams.
- “Competitive advantage” is about shifting the risk culture from a strategy of blanket protection or zero tolerance to one that encourages educated and transparent risk taking. Combined with enhancements to risk management capabilities, this will ultimately enable better commercial decisions, without incurring unsustainable cost, and therefore improve competitive advantage.

Swiss Re believes that it is the cultural aspects of risk management that can help individuals act in a decisive and preemptive way (see Figure 2), often against the prevailing wisdom, and know how to act when faced with challenges that can be seen as inconsistent with the interests of the company and wider stakeholders.
The Journal of Financial Perspectives

Risk culture forms part of the overall culture of Swiss Re. How our employees communicate to clients, regulators and the market will drive the perception of Swiss Re. As our environment and our employees change, our risk culture will continuously evolve.

It is, therefore, necessary to look at the ways in which we can engage the whole organization and embed the desired risk culture into the employee lifecycle. Maintaining and developing our risk culture needs to be done across the company and cannot be the sole responsibility of the Risk Management department.

This means that we need to take a more a holistic view to driving risk culture across the organization.

4. Actions to embed risk culture at Swiss Re
Swiss Re’s ERM EP was set up to help drive the alignment of our risk culture and vision. It has utilized the core building blocks that are already in place at Swiss Re in terms of risk appetite, limits, governance, models and reporting. The focus of this enabling program is on seven areas where ERM can provide enhanced value to deliver the vision. All of them support aspects of our desired risk culture.

1. Greater enlightenment: improving forward-looking analysis through enhancement of the “top risk” identification process. This is to be achieved by bringing together the top-down and bottom-up processes to enable sharing of risk identification insights from different areas of expertise. This process allows different risks to be prioritized, monitored and mitigated. The resulting information is then fed back to the respective business units for further analysis, deeper modeling or product development (see Box 2 and 3).

2. Greater embeddedness: integrating risk management and capital management perspectives into business planning, with the aim of enhancing the common understanding of both risk managers and business partners of the range of potential outcomes and drivers over the planning horizon.

3. Greater empowerment: further enhancing the “target operating model” with mapping of all the key risk management activities (identification, assessment, monitoring, execution and control) to relevant owners in the business and Risk Management.

4. Greater effectiveness: implementation of a single risk register for all operational risks, alongside a top-down mapping of where these risks emerge in the process landscape. Use of principles-based standards to drive group-wide assessment against identified risks. Use of common risk rating definitions for all issues and improvement plans regardless of the source of identification, with the ratings


Box 4: Example of “Get FIT”

To identify changes needed, four behaviors were chosen which support our risk management vision:

- Collaborating with business partners
- Anticipating and preparing for future risks
- Expressing complex ideas concisely
- Speaking up

A survey was undertaken to gather the perception of how the Risk Management function demonstrated these four behaviors. The findings from the survey highlighted the different perceptions of our partners against Risk Managements’ own self-assessment. Risk managers perceived they were demonstrating the behaviors to a higher level than perceived by their business partners.

To address this, the Get FIT project has recruited a group of 65 risk change leaders, all volunteers from within Risk Management, to work on 10 initiatives aimed at addressing these findings. During their work in these initiatives, individuals are being coached on change management techniques. They are encouraged to give feedback and to personally explore incorporating the desired behaviors while developing solutions within their initiatives.

Swiss Re will repeat the internal survey toward the end of 2014 to assess the impact of these initiatives.

Risk culture forms part of the overall culture of Swiss Re. How our employees communicate to clients, regulators and the market will drive the perception of Swiss Re. As our environment and our employees change, our risk culture will continuously evolve.
considering both residual risk and gap versus tolerance level. This should foster a common universe so that all functions have a shared view on materiality and prioritization.

5. **Greater entrepreneurship:** positioning the Risk Management function to provide more value-adding advice to the business, both on request and proactively as a valued business partner. To successfully implement such a change in the relationship, the emphasis is on risk managers establishing themselves both as active listeners to their business partners and as credible and skilled contributors to the debate. This is also an area where the risk manager’s advisory or business enabling role moves to the forefront, compared to their traditional controller role.

6. **Greater efficiency:** examining ways to improve processes, systems, accessibility and communication with the aim of being fully integrated with applications that serve the wider business and ensure that the right information is provided by Risk Management and risk takers as efficiently as possible. A “get-it-right-first-time” attitude is an important corollary to the process and systems architecture.

7. **Greater engagement:** taking steps to emphasize that risk management is not just the responsibility of the Risk Management function, but something that involves the entire organization.

These initiatives and messages are supported and clearly communicated through sponsorship from the Board of Directors and Executive Committee. However, in order that they cascade to every level of the business, they are also being strongly driven by senior and middle management.

Swiss Re’s Get FIT project (reflecting the vision of “forward looking, integrated and trusted”) is a key pillar of our enabling program. The objective of the project is to create stronger engagement between Risk Management and their partners, risk takers and risk owners, across the organization. The project explores ways the Risk Management function can articulate and influence the alignment of our risk culture vision more broadly. This exercise has highlighted the challenges of promoting and stimulating changes in behavior.

The first approach was to undertake a top-down method of messaging and senior management engagement. However, we found that this did not have the impact we wanted and, as a result, efforts have been refocused on enhancing risk culture by also initiating bottom-up projects and programs. This has created greater energy and more sustainable outcomes, both in terms of
Box 5: Swiss Re’s High Performance with High Integrity Initiative

Swiss Re’s business depends on our reputation, as our core business is the sale of promises. In exchange for value, we promise our clients to take on their risks – and to be there to pay them when certain events occur. Without integrity, the credibility and trust of our clients in the value of our promises could easily be jeopardized.

A culture of integrity clearly reduces the risk of misconduct and improves employee and business performance. A strong culture of integrity correlates positively with long-term total shareholder return, as an ethical company should be able to secure the trust of its customers and better attract and retain talent.

The challenge
While employee perspectives may be the most valuable potential source of information about integrity and risk, in reality, most employees do not talk about these types of issues. As a result, only a fraction of the potential information about key risks gets communicated and escalated. Statistics show that 50% of observed misconduct never gets reported by employees, and 60% of reported misconduct does not escalate beyond direct managers. The main reasons for this are fear of retaliation or a belief that nothing happens anyway. (Source: Corporate Executive Board - Compliance and Ethics Leadership Council).

The solution
It is neither possible nor healthy for diverse, complex, dynamic and global organizations to provide rules and policies for any situation that employees may face. Instead, such organizations have to rely on the creation of a culture and a set of values that guide employees to make the right judgment call. It is about creating a cohesive roadmap that works everywhere and all the time, rather than varying from region to region and time to time. It is about instilling global immovable values into local culture. It is about a mind-set and behavior resulting from the values.

Interactive dialogue sessions
Swiss Re launched its High Performance with High Integrity initiative in 2013. We started the initiative by holding interactive dialogue sessions with our senior leaders and their management teams, using real life case studies from across the company. Each session considered what actions should have been taken, the lessons learned, and how to ensure we consistently drive the right values and behaviors in Swiss Re, and within our own teams.

Since then we have completed dialogue sessions with more than 1,000 of our employees around the globe. We are continuing this in 2014. In addition to the dialogue sessions, we have introduced High Performance with High Integrity into our leadership and management programs. It will be included as a module within our High Growth Market Development Program this year. We also link it directly with our leadership imperatives.

As we continue to drive this cultural change across Swiss Re, we are starting to see the development of local initiatives to support our efforts. For example, the Americas team now has an “Integrity Guru” section in their monthly newsletters. One division had an award for their 2013 Integrity Champion. Another division has incorporated High Performance with High Integrity into their vision and mission statements.

Assessing and enhancing risk culture at Swiss Re
In addition to the work being done by the ERM EP, Swiss Re has been working with EY LLP to assess Swiss Re’s existing risk management framework and to identify efficient steps that could be taken to deliver our vision and enhance awareness.

The following dimensions were used to assess the current state and identify areas most likely to be effective in building on our existing framework.

EY assessed the elements of our framework and existing initiatives surrounding the various categories using the segmented circle illustrated in Figure 3. The assessment highlighted certain areas where focus on awareness and embedding could have potentially material benefits in enabling
Driving a strong risk culture: a Swiss Re perspective

For a company to grow and foster business, it is essential to take operational risk in all activities. Two years ago, in order to promote our risk culture, Swiss Re implemented a control-related behavior assessment for key risk takers and controllers. Despite initial uncertainty over how management would react, the assessment was well received and has clearly boosted ownership and accountability for managing operational risks within the company.

At Swiss Re, key risk takers are defined as those individuals whose compensation is most directly tied to value creation through risk taking activities. This includes members of the Executive Committee, Group Management Board and also the majority of the next management level, the Managing Directors.

Key risk controller positions include Property & Casualty/Life & Health Actuarial Management and the heads of Swiss Re’s Integrated Assurance functions (IAF), i.e. Operational Risk Management, Compliance and Group Internal Audit.

To further embed operational risk awareness, Swiss Re’s IAF have established a mix of measurable criteria as the basis for a periodic assessment of the behavior of key risk takers and controllers:

- Self-assessment quality
  This is a review of the self-identified issues, mitigating action plans and level of completion in undertaking individual risk and control self-assessments. This includes making issues transparent to stakeholders and escalating them to senior management as appropriate.

- Issue prioritization
  Ensuring issues are included in overall work priorities and are properly differentiated in a systemic and sustainable fashion.

- Issue deterioration/resolution
  Resolving issues timely and completely through transparent action plans. Ensuring systemic and sustainable resolution mechanisms are in place to avoid reoccurrence of the same or similar issues.

- Other qualitative factors
  Focusing on proactiveness and openly maintaining a strong risk culture within both their own department and throughout the group, evidenced by factors, such as: (1) driving initiatives for improving risk management practices and operating controls within their own department and across the group; (2) use of operational event (losses and incidents) analysis as a key means of identifying and applying learnings and continuous improvement; and (3) open and transparent sharing of information required to help partners achieve risk management and control objectives.

For a company to grow and foster business, it is essential to take operational risk in all activities. Two years ago, in order to promote our risk culture, Swiss Re implemented a control-related behavior assessment for key risk takers and controllers. Despite initial uncertainty over how management would react, the assessment was well received and has clearly boosted ownership and accountability for managing operational risks within the company.

At Swiss Re, key risk takers are defined as those individuals whose compensation is most directly tied to value creation through risk taking activities. This includes members of the Executive Committee, Group Management Board and also the majority of the next management level, the Managing Directors.

Key risk controller positions include Property & Casualty/Life & Health Actuarial Management and the heads of Swiss Re’s Integrated Assurance functions (IAF), i.e. Operational Risk Management, Compliance and Group Internal Audit.

To further embed operational risk awareness, Swiss Re’s IAF have established a mix of measurable criteria as the basis for a periodic assessment of the behavior of key risk takers and controllers:

- Self-assessment quality
  This is a review of the self-identified issues, mitigating action plans and level of completion in undertaking individual risk and control self-assessments. This includes making issues transparent to stakeholders and escalating them to senior management as appropriate.

- Issue prioritization
  Ensuring issues are included in overall work priorities and are properly differentiated in a systemic and sustainable fashion.

- Issue deterioration/resolution
  Resolving issues timely and completely through transparent action plans. Ensuring systemic and sustainable resolution mechanisms are in place to avoid reoccurrence of the same or similar issues.

- Other qualitative factors
  Focusing on proactiveness and openly maintaining a strong risk culture within both their own department and throughout the group, evidenced by factors, such as: (1) driving initiatives for improving risk management practices and operating controls within their own department and across the group; (2) use of operational event (losses and incidents) analysis as a key means of identifying and applying learnings and continuous improvement; and (3) open and transparent sharing of information required to help partners achieve risk management and control objectives.

Following, the assessment by EY, we are now exploring three core areas:

- Communication: seeing how we can join the dots more visibly within the organization so that it is clear how initiatives fit together. In order to make the change, everyone needs to be aware of what is expected of them and how they contribute concretely.

- Tangible insights: making views on risk appetite more tangible and relevant for decision-makers across the company. We have had great success with the approach that High Performance with High Integrity has used and hope to leverage this for wider discussion on other aspects of our risk culture.

- Performance pay: continue to evolve the performance assessment to highlight the importance of not just what employees do but how they do it, when considering their contribution and performance. A particular focus is on further developing our key risk takers assessments.
Key risk takers assessment (KRTA)
The current implementation of the KRTA has been developed from an operational risk perspective and, as the name suggests, is focused on our key risk takers. Our ultimate aim is to roll it out beyond key risk takers and enable the reporting of results at an individual, department, business line or company-wide level (Box 6).

Conclusion
With the increasing regulatory focus and need to address the negative sentiment toward insurers, both from investors and customers, a strong risk culture provides a critical component for sustainable economic success. Swiss Re has put risk culture at the heart of its drive to deliver its risk management vision and safeguard our competitive advantage. This exercise and its focus cannot be limited to Risk Management and must form part of a company-wide perspective. As this is not something where process and rules can, of themselves, deliver results, there is a conscious need to engage in continuous improvement. Swiss Re is committed to driving this continuous process for an ever-stronger risk culture. We also believe that the industry as a whole will benefit from sharing insights into the key drivers of risk culture and experiences gained across the financial services sector.
Future directions for foreign banks in China

Jack Chan
Managing Partner, Financial Services - Greater China, EY LLP

Geoffrey Choi
Assurance Leader, Financial Services - Greater China, EY LLP

Kelvin Leung
Banking & Capital Markets Leader, Financial Services - Greater China, EY LLP

Shelley Chia
Partner, Financial Services - Greater China, EY LLP

Abstract
This article highlights the challenges and opportunities that foreign banks face in China. Based on interviews with 38 foreign bank Chief Executive Officers (CEOs) and senior executives between July and September 2013 – a period when several significant initiatives were underway, which will have a direct impact on the course that different foreign banks will chart over the next three to five years, such as the ongoing internationalization of the Renminbi, interest rate liberalization, and the form and scope of the recently announced Shanghai Free Trade Zone – this article provides insights on the future directions the foreign banks may pursue. We find that foreign banks will seek to grow fee income and develop their advisory role outside of China. The latter change will be assisted by Chinese corporates as they expand internationally, through acquisitions and greenfield investments. Such developments will foster foreign bank relationships, both offshore and onshore. Foreign banks will continue to evolve and choose different paths of expansion, based on their strengths and market opportunities. While some are building a retail presence and expanding their networks, others will focus on narrow market niches. Some of the larger foreign banks have divested their holdings in the large commercial banks, but simultaneously have moved into the securities and trust businesses. As the financial reforms take shape, new opportunities will emerge through the synergies gained by these diversification strategies.
This paper examines the future directions the foreign banks may take in China. The findings are based on interviews with CEOs and senior bank executives of 38 foreign banking institutions in China. These interviews were taken at a time when there were significant expectations on the path to financial reform in mainland China. Several significant initiatives are underway that will have a direct impact on the course that different foreign banks will chart over the next three to five years, such as the ongoing internationalization of the Renminbi, interest rate liberalization, and the form and scope of the recently announced.

This article is organized as follows: In section 1, we look at the current situation facing foreign banks in China. In section 2, we look at the regulatory framework, followed by an analysis of how the Chinese market is opening up in section 3, and what the growth potential looks in section 4. In section 5, we look at product and segment developments and in section 6, we analyze how human resources have developed. Section 7 concludes the article.

1. Overview of the current situations

1.1 Foreign bank presence

As of the end of 2012, forty-two locally-incorporated foreign banks, 95 branches and 197 representative offices operated in China.1 Within this group of foreign banks, 37 locally-incorporated foreign banks and 25 branches were permitted to offer derivatives, 6 locally-incorporated foreign banks were authorized to issue RMB financial bonds, and 3 locally-incorporated banks were approved to issue proprietary credit cards. According to the China Banking Regulatory Commission (CBRC), foreign banks were present across 59 cities in 27 provinces.

Although these foreign banking institutions had experienced growth in the amount of assets they held, by 10.66% in 2012 to RMB2.3t (U.S.$376b), their market share based on total assets dropped to 1.82% in 2012. This level of market share was similar to that recorded eight years earlier in 2004. Deposits held by foreign banks increased by 7.74% to RMB1.43t (U.S.$234b) from 2011 to 2012 while loans totaled RMB1.04t (U.S.$17b), an increase of 6.4% from 2011 to 2012.

As a group, the foreign banks generated after tax profits of RMB16.34b (U.S.$2.67b) for 2012.

CBRC has also reported the following ratios for the locally-incorporated foreign banks at the end of 2012: liquidity ratio - 68.77%, nonperforming Loans (NPL) ratio - 0.52%, capital adequacy ratio (CAR) - 19.74%, and core CAR - 19.25%.

Foreign banks continue to be overshadowed by the large Chinese commercial banks, the joint-stock banks, city commercial banks, policy banks and rural commercial banks (Figure 1). Despite the steady asset growth of foreign banks since 2008, they have failed to gain market share. Setting aside the large Chinese commercial banks, other groups, such as the city commercial banks and rural commercial banks have gained much stronger market positions than foreign banks.

Furthermore, the large domestic banks have started to expand their overseas presence. At the end of 2012, the CBRC reported that 16 Chinese banks had more than 1,000 overseas establishments in 40 countries. It is expected that this footprint

---

1 Source: CBRC Annual Report.
will expand dramatically over the next five years as Chinese banks grow their international operations by establishing new subsidiaries and branches and through acquisitions.

A key question is, if the Chinese economy is now on the cusp of economic and financial reform, will the foreign banks be able to play a more important and diversified part in the Chinese financial services sector?

1.2 Some background statistics on the participants

The 38 banks interviewed employ 29,229 people and anticipate this number will expand to 35,274 by 2016, a 20.68% increase. However, three European banks anticipate that their number of employees will remain unchanged within the three-year period.

Some of these banks have an active retail banking presence, ranging from a customer base of less than 100,000 up to several hundreds of thousands. The number of multinational clients ranged from just 10 to 15, up to several hundred. For the state-owned enterprises (SOEs), most banks said they had relationships with between 20 and 35 SOEs while a few indicated up to 100. Only 10 banks were able to calculate the number of privately-owned enterprise (POE) clients, and this estimate ranged from 10 up to 800.

Twenty-three banks provided estimates of their cost to income ratio in 2013 and in 2016. This ranged from 90% to just 19% in 2013. All but three banks were above 75%. By 2016, none of these banks envisaged a cost to income ratio above 70%, although three banks were unable to provide forecasts.

Fourteen banks provided estimates on their loan to deposit ratios (LDR). Only one bank was above 75% while the lowest ratio was 40%. Three banks were unable to provide estimates for 2016. Four banks projected increases in their LDR by 2016.

The projected asset growth for the participants in 2013 was below that anticipated in 2016. Several participants attributed the slowdown to increased cost of funds and economic slowdown in 2013. Six banks anticipated no asset growth in 2013, while a further 10 banks expected growth of up to 10%. Thirteen banks anticipated an asset growth rate above 10%. Projecting forward to annual asset growth in 2016, one participant expects shrinkage, 10 banks fall below 10%, but 9 banks expect expansion above 20% with 2 banks expecting to double in size between 2015 and 2016.

The foreign banks' parent banks are strongly committed to their Chinese operations. Thirty-seven of the 38 banks stated that their parent’s support was resolute. There was only one dissenting bank, which stated that due to home country issues China was now a lower priority than in the past. Several participants said that the level of involvement between the Chinese operation and head office had increased. This is in part based on an expectation that the internationalization of the economy will present more opportunities for the foreign banks outside China and in particular, where those banks have a strong footprint, for example, a Spanish bank with strong representation in Latin America.

Twenty-four banks placed China in the “Top 3” most important target growth markets in the world. This group often includes countries such as India, Indonesia, Malaysia, Turkey, and parts of Africa and Latin America.

Going forward, the majority of foreign banks believed that their market share will stay the same nationally. Twenty-six banks believed it would remain static, while 9 banks forecasted an increase and 3 banks a decline. In contrast, 16 banks predict an increase in the secondary cities while 15 banks believe market share will remain the same. Seven banks predict a decline.

Therefore, despite cautious optimism expressed elsewhere in this report regarding new product offerings, market liberalization and internationalization, most foreign banks do not anticipate any significant change in market share based on total assets.

The foreign banks have a much larger market share in Shanghai and the creation of the Shanghai Free Trade Zone may allow further gains. However, market share based on total assets is probably not the best measure going forward.

The branch networks of the “Big Six” retail banks, Bank of East Asia, Citibank, DBS, Hang Seng Bank, HSBC and Standard Chartered Bank are concentrated on the eastern side of the country with a heavy emphasis on the Tier 1 cities of Guangzhou, Shanghai, Tianjin and Beijing.
1.3 The challenges facing the respondents in competing in the Chinese banking industry?
Participants were asked to score three different groups of challenges they faced in operating a foreign bank in China, namely regulatory, operational and market growth.

1.3.1 Regulatory challenges
The participants believe that they face a plethora of rules and regulations which expand in scope each year. One participant maintained that they are required to file 6,300 different reports annually, in contrast to their parent bank, which files 400 reports with its home regulator.

Foreign banks are frustrated with their access to the China bond market which, at around U.S.$4t, is the fourth largest in the world. Some of the larger foreign banks do have access to the market. Furthermore, the issuance of Renminbi Qualified Foreign Institutional Investors (RQFII) quota to foreign banks, allowing their clients to invest in Chinese bonds, may become an attractive new product offering for foreign banks.

The third most important regulatory challenge was capital and liquidity constraints. These issues are discussed elsewhere in the report. The capital issue has become more critical for foreign banks since the survey interviews took place. In October 2013, CBRC proposed that the minimum registered capital for newly incorporated foreign banks should be raised from RMB300m to RMB1b (U.S.$164m). Registered capital is the amount required for a foreign bank to be granted a new bank license.

1.3.2 Operational challenges
The most important operational challenge facing foreign banks is attracting and retaining well-qualified and skilled personnel. Although staff turnover appears to be slowing, many foreign banks continue to feel vulnerable to competitors poaching key personnel.

The legal environment is also a significant operational challenge, followed by the related issue of governance in client companies. The foreign banks are also conscious of the pressure on future profitability.

1.3.3 Market growth challenges
The top three market growth challenges facing the foreign banks were margin compression, attracting and retaining profitable retail customers, and domestic bank competition.

Margin compression will, in the short term, affect the appetite of foreign banks to expand their loan books and impact profitability. The locally-incorporated foreign banks with a retail focus continue to expand, but domestic banks are pursuing the same target customer base with improving service. China Merchants Bank, for example, is acknowledged by the foreign banks as a formidable competitor in retail banking.

1.3.4 How would you rate the health of credit and risk in the corporate and retail markets?
Domestic banks: when asked to record an opinion on the well-being of both corporate and consumer credit in domestic banks, 78% of the respondents expressed concern that corporate credit was deteriorating. This was in part a reflection of the slower economic growth rate, interest rate liberalization, and recognition that the nonperforming loan (NPL) ratio in domestic banks is rising. This view did not extend to consumer credit, where 17 of the 22, who expressed an opinion on consumer credit, felt that it remained stable.

While 80% contended that operational risk for domestic banks remains the same, this is not the case for market risk, which just over 50% believed was worsening.

Foreign banks: a much more stable view is articulated by the foreign banks when they report on credit and risk assessment of their own operations in China. They recorded their highest value deteriorating score for market risk at 24%, but most believed there has been minimal change for corporate and consumer credit and operational risk.

2. The regulatory framework
To get a better understanding of the views of the foreign banks toward the regulatory environment, two important documents were reviewed. One published annually by AmCham China, which reflects the views of the American banks, and the second is published by the European Chamber of Commerce (E.U. Chamber) in China, which itemizes a list of regulations that are of concern to its members.
Table 1, which is based on issues raised by the E.U. Chamber and required each respondent to rate the headlines on a scale of 1 to 10, in terms of importance to their bank, demonstrates that the foreign debt quota is the most important regulatory issue for foreign banks.\(^2\)

In early 2012, the long-term debt quota allocated to foreign banks was doubled to U.S.$24b, permitting more money to flow into China from the outside. While some foreign banks appeared to be content with their existing quotas, there were a number of participants who have not enjoyed increases and contended that the quotas hinder their ability to meet demand for loans and investments.

The second most important regulatory issue was the 75% loan-to-deposit ratio followed by the foreign guarantee quota and the desire to broaden the LDR by including other liquidity sources in the ratio, both coming third. Foreign banks contend that financing for foreign invested companies often requires a guarantee from a foreign bank or foreign parent company. These foreign guarantees are included in the foreign debt quota.

Other concerns, all scoring around 8 out of 10, were the 10% withholding tax, the 5% business tax, and the tax assessment on representative offices based on their costs. Interestingly, at the bottom of the list were the restrictions on branches and sub-branches. Foreign banks have often complained about the “one branch at a time” approval process. The positioning of this concern suggests that participants believe that the process is speeding up and becoming less burdensome. As foreign banks expand into secondary cities, their plans are experiencing a more ready approval by the regulators.

2.1 Critical regulatory issues for foreign banks

Thirty-seven participants answered this open-ended question, and responses revealed that the two most critical regulatory issues were the loan-to-deposit ratio and the net open position (NOP) rule. The LDR, which refers to the 75% ratio of loans to deposits, continues to pose major challenges for a number of foreign banks because of their limited branch networks and ability to obtain deposits. The NOP rule, which was introduced by the State Administration of Foreign Exchange (SAFE) in May 2013, links a bank’s onshore foreign exchange LDR with its NOP floor. The rule requires that once a foreign bank’s foreign exchange LDR ratio exceeds 100% (75% for domestic banks), a new NOP floor will be set for the start of the next month.

Other regulatory matters that were of concern to the respondents include:

- Liquidity
- The “three measures and one guideline” – this regulation came into existence three years ago and covered a new set of risk management rules. Although most foreign banks have adjusted to these rules, several smaller participants noted that it continued to be onerous.
- Branch network expansion – the CBRC’s “one branch at a time” policy is viewed as tedious and limits growth. Foreign banks can not submit simultaneous branch applications. The relatively lower score given to this factor in the previous question may suggest that some banks (the larger ones) are no longer finding this difficult.

<table>
<thead>
<tr>
<th>Issues raised by the E.U. Chamber</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remove foreign debt quota</td>
<td>8.5</td>
</tr>
<tr>
<td>Remove the 75% LDR</td>
<td>8.3</td>
</tr>
<tr>
<td>Remove foreign guarantee quota</td>
<td>8.2</td>
</tr>
<tr>
<td>Include other stable liquidity sources in the ratio</td>
<td>8.2</td>
</tr>
<tr>
<td>Coordination by regulators</td>
<td>8.1</td>
</tr>
<tr>
<td>Waive the 10% withholding tax on offshore funding</td>
<td>8.1</td>
</tr>
<tr>
<td>Waive the 5% business tax on onshore and offshore lending</td>
<td>8.0</td>
</tr>
<tr>
<td>Remove the cost based income taxation of rep office</td>
<td>7.9</td>
</tr>
<tr>
<td>Improve communication between regulators re-target and policies</td>
<td>7.7</td>
</tr>
<tr>
<td>Licence approval process</td>
<td>7.6</td>
</tr>
<tr>
<td>Remove the two times capital limitation in the CFETS interbank market (PBOC)</td>
<td>7.5</td>
</tr>
<tr>
<td>Access to bond underwriting market</td>
<td>7.3</td>
</tr>
<tr>
<td>Greater ownership and scope for foreign vs domestic</td>
<td>7.1</td>
</tr>
<tr>
<td>Easier branch/outlet expansion</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Table 1: Scoring of high priority regulations for foreign banks

Participants scored each factor on a 1 to 10 scale where 10 is maximum importance.

---

\(^2\) The National Development and Reform Commission administers the long-term foreign debt quota, while State Administration of Foreign Exchange administers the short-term debt quota (covering periods less than one year).
Future directions for foreign banks in China

- Foreign debt quota
- Capital inflow regulations
- Regulatory overlap and insufficient clarity - this also included what some foreign banks believe are at times contradictory policies between the People’s Bank of China (PBOC) and SAFE, different interpretations between provinces and the need to sometimes deal with three to five regulators. One bank cited an example where National Development Reform Commission (NDRC) approved of their decision to bring CNH (offshore RMB) deposits on-shore while PBOC disapproved.
- Basel III challenges
- Capital changes on derivatives
- RMB loan quota
- Employee initiated financial crime
- New regulations - when new regulations are first announced, the guidance on implementation is often weak.
- Regulation of wealth management products
- Brokerage, full access still restricted

2.2 How will interest rate liberalization affect your bank in China?
The move to liberalize interest rates and make them market accountable has been viewed as a basic requirement in the rebalancing of the Chinese economy. To date, interest rate liberalization has been asymmetric with action on lending rates but no changes to deposit rates. The general consensus among participants was that margins will be squeezed in the short term as lending rates decline. This will lead to reduced bank profitability.

The foreign banks believe that if deposit rates are deregulated and a market rate-driven environment emerges, this will benefit them. These foreign banks feel that since they are better skilled at managing risk and adapting to change, such a move would allow them to better differentiate their products on both price and risk. They also felt that mid-sized domestic banks would be particularly hit by such a move.

While a number of foreign banks were uncertain about when deregulation of deposit rates would occur, quite a few believe that it will happen soon, perhaps during 2014. A couple of banks linked the timing of a deposit rate deregulation to the introduction of a national bank deposit insurance scheme.

Some participants offered a different interpretation. Given that the domestic banks can now determine their own lending rates, they may choose to lower rates and ease the burden on the SOEs and local government financing vehicles (LGFVs). This may benefit the large domestic banks and place them in a stronger market position relative to all the other types of banks in the marketplace.

2.2.1 The launch of the loan prime rate
In October 2013, interest rate liberalization took another step forward when the interbank funding center launched a new loan prime rate (LPR).³

The funding center, which is under the control of the PBOC, will strike an LPR calculated on a weighted average of the costs to the best customers of the top nine banks (namely Industrial and Commercial Bank of China, China Construction Bank, Agricultural Bank of China, Bank of China, Bank of Communications, China Citic Bank, Shanghai Pudong Development Bank, Industrial Bank and China Merchants Bank). Commentators believe that the creation of the LPR means that the PBOC has transitioned away from benchmark rates toward the new bank or market “influenced” rates. The first LPR rate was 5.71%, below the PBOC benchmark rate of 6.0%.

2.2.2 Future liberalization of deposit rates
On the deposit front, it is argued that the Memorandum of Understanding signed between the PBOC and the U.S. Federal Deposit Insurance Corporation (FDIC) on 24 October 2013, also suggests that China’s desire to relax restrictions on deposit rates and implement a deposit insurance scheme is well underway. This will be further discussed in the following section.

2.3 Deposit insurance
The PBOC has indicated that it plans to introduce a deposit insurance scheme. In general, foreign banks are supportive of this initiative and see it as an integral part of interest rate liberalization and a further step toward a practice followed in other countries. The respondents noted that while it will add costs, the heaviest burden will probably be borne by the big domestic banks.

³ This topic is discussed again below as it relates to the impact of margin compression on foreign bank performance.
Foreign banks with a retail presence welcomed the move. One bank suggested that the insured amount would be around RMB 500,000 (U.S.$82,000). This is below the U.S. where the FDIC insures deposits up to $250,000. In Canada, it is C$100,000 and in countries such as Germany, France, Netherlands and Spain, it is €100,000.

A number of respondents suggested that margin compression will inevitably lead to a crisis for a number of mid-sized domestic banks; hence the launch of a deposit insurance system seems both timely and prudent. An article in the *China Daily* on 3 September 2013, quoted an interview with Shang Fulin, CBRC Chairman, published in *Qiushi Journal*, in which he said that: “risk control remained a monumental task given the slowing economy, rising bad assets, industrial overcapacity and shadow banking.”

Box 1 – BIS’ portrayal of the Chinese banking system

At the end of 2012, 511 commercial banks were registered in China, with total bank assets (including off-balance sheet assets) amounting to RMB129t (approximately U.S.$20t), circa 240% of GDP. The financial system is dominated by the five largest commercial banks, which hold about 60% of total banking assets. One Chinese commercial bank – Bank of China – is classified as a global systemically important bank (G-SIB). The CBRC is in the process of developing a framework for domestic systemically important banks (D-SIBs).

Under the new Basel III standards, the weighted average total capital ratio of Chinese banks stood at 13% in 2012 while the Tier 1 ratio and the CET1 ratio were both 10%. The ratio of nonperforming loans (NPLs) of commercial banks over total loans amounted to 0.95% at the end of 2012.

The average provisioning ratio (the ratio of loan loss provisions to total loans) for commercial banks was 2.82% (the CBRC requires all banks to meet a minimum provisioning ratio of 2.5% by 2016); and the provisioning coverage ratio (ratio of loan loss provisions to NPLs was 296%.

While the Chinese banking system is growing rapidly, the core of its banking business remains traditional, concentrated on credit products and services. This is reflected in a high proportion of loans relative to total assets and a relatively high share of RWA for credit risk as percentage of total RWA. Bank credit has increased from RMB16t in 2003 to RMB67t in 2012, equivalent to more than 120% of GDP. Overseas assets and assets denominated in foreign currency are relatively small as the focus of Chinese banking, thus far, is primarily domestic.

The CBRC has been cautious in allowing banks to engage in complex financial activities. Some examples include correlation trading (which has not been permitted so far), and a closely monitored opening into securitization products and complex OTC derivatives. Reflective of this is the relatively small proportion of trading in financial activities, with an average market risk RWA of less than 1% of total RWA. The securitization markets are small and at a pilot stage. At the end of 2012, the volume of outstanding asset-backed securities was less than RMB20b.

In recent years, a market for wealth management products (WMPs) has developed. These products offer retail customers an alternative to traditional bank deposits and have seen strong growth, in part due to the existing caps on bank deposit rates. Banks offer WMPs directly but also indirectly through trust companies. In spring 2013, the CBRC issued new regulations for banks with regard to the prudential treatment of WMPs. The Assessment Team discussed the nature and structure of WMPs and the possible economic similarity to securitizations. According to the CBRC, there is no tranching of credit risk and banks do not provide liquidity facilities to WMPs, and therefore WMPs are not classified as securitizations. The CBRC applies the standard credit risk rules if the WMPs are held on a bank’s balance sheet. As a result, WMPs were not considered by the team when assessing the CBRC’s implementation of the securitization framework.

Source: Basel Committee on Banking Supervision, Regulatory Consistency Assessment Program, Assessment of Basel III Regulations – China, September 2013. A pdf copy of this report is available on the BIS website, www.bis.org
Relaxation of deposit rates is seen as a critical step in financial reforms. Increased competition may lead to bank failures, and although the Chinese public implicitly believes that troubled banks are, and will continue to be, supported by the government, this may no longer be the case.

A number of respondents argued that a number of mid-sized banks will find it difficult to compete effectively when reforms take hold, and this may result in their acquisition by the larger banks.

The deposit insurance scheme could play an important role if such a scenario was to unfold. Clearly in the absence of such a scheme, the clean-up process could be much more challenging.

2.4 How far have the Chinese banks progressed on Basel III? What is its likely impact on the local banking market?
Foreign banks believe that Basel III will be a challenge to the big Chinese banks. The participants’ views varied on how much progress the domestic banks have made toward the implementation of Basel III. One view was that Basel III will require a higher capital ratio which will slow the growth of the domestic banks. An Asian bank suggested that it will reduce the level of available credit in the real economy. As a result, this may improve the competitiveness of the foreign banks.

A North American bank noted that the proposed steps toward financial reform will take place ahead of Basel III. Most participants agreed that the Chinese banks will require more capital. The Bank for International Settlements (BIS) published an assessment of China’s regulatory consistency with Basel III in September 2013. A summary of the BIS portrayal of the Chinese banking system is presented in Box 1.

The government’s cautious approach to the banking market has been demonstrated by its refusal to permit any new domestic entrants. The last one to gain approval was China Minsheng Bank in 1996. This bank ranked number 9 in China and 54 globally in the 2012 Banker rankings based on Tier 1 capital. The Banker records 96 domestic Chinese banks in its 2012 Top 1,000 Global Banks.

3. Opening up the market
3.1 RMB internationalization
The respondents indicated that the increased internationalization of the renminbi would provide a plethora of new opportunities for foreign banks across a range of areas. These will include trade settlement, offshore lending, offshore deposits, offshore and onshore investment products, and offshore raising of capital.

Industry experts have suggested that the path toward RMB internationalization will involve three steps:
- Step 1 for trade finance
- Step 2 for investment
- Step 3 – the emergence of the renminbi as a reserve currency

There has been significant progress made on the trade front, and as the financial reform process unfolds, developments will focus on investment flows. The foreign banks are optimistic that, as the market opens up, they will be able to provide increasingly innovative and sophisticated services, and leverage their international networks.

A number of foreign banks discussed the prospects of the development of RMB offshore centers in their home markets, such as Frankfurt, Hong Kong, London, New York, Paris, Singapore, Sydney, Taipei, Tokyo and Toronto.

The respondents believed that the scope of new opportunities will be based on the opening up of the capital account. Although they expressed uncertainty about the pace of financial reform and the ability of the new leadership to implement far reaching changes, early indications, such as the launch of the Shanghai Free Trade Zone and recent moves to facilitate the flow of capital into and out of China are positive developments. Nevertheless, further developments regarding capital account liberalization, such as expansion of the corporate bond and derivatives market, will be needed. The respondents believe they can play a vital role in this process and hope that the process will gather momentum. As one participant commented, there is no reverse gear for RMB internationalization. One interesting observation made by a small foreign bank questioned whether the new opportunities associated with RMB liberalization will be restricted by the different regulators to the large global banks or whether all foreign banks will be able to share in the benefit.
3.2 The recently announced Shanghai Free Trade Zone

In September 2013, the government announced plans for the China (Shanghai) Pilot Free Trade Zone (FTZ). Initially, 11 financial institutions, including two foreign banks (DBS and Citibank), were given permission to register in the FTZ. At the time of writing this article, two additional foreign banks (HSBC and Bank of East Asia) have been added.

While the interviews were conducted prior to the official announcement, in anticipation of the FTZ's creation, participants were asked if whether it would allow foreign banks to introduce new services and expand more rapidly. Despite the limited availability of information at that time, most respondents were optimistic about future opportunities. Three respondents did not think that foreign banks would benefit while a further nine banks remained neutral. However, five banks agreed it would foster expansion and a further five banks strongly agreed.

Foreign banks believe that the success of the FTZ will depend on the financial reform process and much rests on the speed of transition toward RMB convertibility. Several banks made reference to the often cited “Shanghai 2020” deadline. They believe that the reform process will need to be comprehensive and expedient if the goal of creating an international financial center is to be fulfilled by 2020. The general opinion can be summarized as cautious optimism. However, several banks believed that reforms would be made expeditiously and solid opportunities would emerge.

From the foreign banks' perspective, a number of significant reforms are anticipated and these include:

- Opening-up of the financial sector with a relaxation of investment controls and innovation in trade supervision systems.
- Reforms that will advance RMB internationalization.
- Faster approval for foreign bank branch approval. Foreign banks located in the FTZ will no longer be required to operate as representative offices for two years before applying for branch status.

Further clarifications were provided in a CBRC notice on banking regulation in the FTZ. It noted that qualified foreign banks will be able to set up subsidiaries, branches specialized institutions, and Sino-foreign joint venture banks in the FTZ. It commented that sub-branches of foreign banks in the FTZ may be upgraded to branches and research will look at shortening the period it takes for representative offices to become branches, and for foreign banks to conduct RMB business.

The CBRC notice said that banking institutions in the FTZ will be supported to develop cross-border financing services including, but not limited to, commodity trade financing, supply chain trade financing, offshore shipping financing, modern service sector financial support, domestic loans under overseas guarantee, commercial paper, etc.

There will be support to undertake cross-border financial services including, but not limited to, cross-border M&A loans, project loans, overseas loans under domestic guarantee, cross-border asset management and wealth management, real estate trust, and investment funds, etc. It further suggested that a “green and fast access channel” for the banking sector will be established in the FTZ, aided by a time limit mechanism to improve access efficiency.

Of particular attraction to the foreign banks is the reference to “optimize the calculation methods and regulatory requirements for certain indicators such as the LDR and liquidity.”
Finally, a further interpretation document was issued by the Shanghai Municipal Government on 15 October 2013. The document posed the following question, “Has the Shanghai FTZ rolled out any new measures in terms of financial innovation?” The answer given was as follows, “The Shanghai FTZ has put forward the following four measures of financial reform:

- **RMB convertibility under capital account:** Under a controllable risk level, the Shanghai FTZ will implement a trial program of RMB convertibility under the capital account, and innovate business and management modes through separate accounting.

- **Interest rate liberalization:** The Shanghai FTZ will cultivate an independent pricing mechanism in line with the development of the real economy, and gradually promote the reform of interest rate liberalization.

- **RMB cross-border usage:** Enterprises in the Shanghai FTZ may initiate cross-border RMB businesses based on their actual situation, thereby facilitating the cross-border usage of RMB.

- **Foreign exchange management:** The Shanghai FTZ will set up a foreign exchange management system to facilitate trade investment.”

### 4. Performance and growth

When commenting on their own bank’s financial performance over the next three years, 25 banks believe that it will increase slightly while seven banks optimistically believe it will improve significantly. At the opposite end of the scale, one participant predicted a significant decrease. A similar pattern was predicted for foreign banks as a whole. Twenty-one banks predicted a slight increase and four banks said they expected foreign banks as a group will see significant increases. Nine banks, however, believe the status-quo will remain and three forecast slight decreases (Figure 2).

These findings suggest the foreign banks expect to make steady, consistent progress toward improved profitability. However, the optimistic predictions elsewhere in this report surrounding regulatory reforms and the potential for improved growth are not matched by significantly improved profitability.

#### 4.1 Corporate banking profitability

Figure 3 displays profitability (in terms of capital allocated) to nine different lines of corporate or wholesale banking. 31 banks provided responses to the corporate banking question and 18 of them said that this line of business was profitable for them, while eight banks stated that it was very profitable.

In trade finance, 18 banks indicated being either marginally profitable or profitable, while 11 banks said this product line was very profitable.
4.2 Retail banking profitability estimates
Just seven banks answered this question. Success was mixed. Under the general heading of “retail banking” – lending, deposit-taking and transactional banking – three banks responded that they were profitable, two were marginally profitable, and two acknowledged that they were making losses.

Although several participants offer credit cards (either in their own name or as a joint card with a Chinese bank), only one responded and said they were marginally profitable. Foreign banks have acknowledged that they do not have the critical mass to make credit cards profitable. Personal loans have limited demand and one participant commented that the “Three measures, one guideline” regulation which requires the lender to identify a loan’s purpose and usage, makes this particularly challenging.

Secured lending, primarily in mortgages, was recorded as profitable by four banks. Only two banks commented on private banking, with one saying that it was profitable while the other said it was only marginally profitable. A number of foreign banks are active in this area but business is directed to, and handled by, their parent bank outside of China.

4.3 Expected return on equity (RoE) for 2014 and projected for three to five years
The participants provided predictions for RoE in 2014 and in three to five years’ time. While these are only rough estimates, they do provide insight into the foreign bankers’ assessment of the market’s potential.

Thirty-four banks offered opinions on RoE in 2014, and they were as follows: six banks expected an RoE of less than 10%, 10 banks between 10% to 12%, eight banks between 13% and 15%, six banks between 16% and 18%, and four banks between 19% and 21%. Within the same group of 34 banks, projecting forward three to five years, two Asian banks, eight European banks, and five North American banks predicted an RoE of 16% or higher.

A similar question asked the respondents to comment on their Return on Assets (RoA) for 2014 and again in three to five years. Only 16 banks responded to this question and three of them suggested that this metric is not meaningful and was, therefore, not measured by their bank.

Seven banks said that in three to five years, RoA would be between 0.8% and 1.1%, five banks between 1.1% and 1.4%, and three banks between 1.4% and 1.7%.
Future directions for foreign banks in China

4.4 Future growth opportunities

Figure 4 presents the views of the respondents on the growth potential of six different categories over the next three years. The banks believe that corporate lending, trade finance, small and medium enterprises (SME) and commercial lending, cash management and treasury, and corporate advisory services will all increase either slightly or significantly over this period.

Corporate advisory services are also expected to grow, with 10 banks expecting significant and 13 expecting moderate growth. These growth predictions are premised on the expectations that Chinese companies will continue to invest outside of the country. The “other” category refers to cross-border transactions and consumer loans.

4.5 Revenue growth for individual participants

Figure 5 presents the respondents’ estimates of projected revenue growth for 2013 and 2016. There are two noticeable positive outliers, with one bank expecting 100% growth in 2013 and in 2016, while another predicting 80% growth in 2013 and 100% in 2016. Otherwise, most banks clustered around the 10% to 25% growth mark for 2013 and 2016. The one negative outlier expects 0% growth in 2013 and a contraction of as much as 8% by 2016.

4.6 Top areas for continued revenue growth

To provide further insights into their revenue projections, the respondents were asked to identify their top three revenue growth areas during the previous 12 months. It was assumed that these areas of growth would continue into the future.

The two most frequently cited areas of recent revenue growth were corporate lending and trade finance. Within the group of 38 banks, around 50% mentioned both of these product areas.

Other frequently mentioned sources of revenue were cash management and treasury, foreign exchange and FX derivatives, and for three banks, SME lending.

4.7 Current and future levels of margin compression

When asked whether they had experienced margin compression, 89% of the respondents (33 banks) stated that they had indeed experienced interest rate margin compression in 2013, and that they also expect it to continue into 2014.

This will put pressure on profitability for all banks, both foreign and domestic. In response, the foreign banks will use three different strategies to alleviate the problem. They plan to increase fee income, rebalance their portfolios, and reprice their loan book (Figure 6).

Each option presents real challenges. They plan to increase fee income by providing more services to Chinese companies as they expand abroad. Many participants said rebalancing their loan portfolio would be very difficult and time-consuming. An Asian bank suggested that it was already transitioning to better quality credits and expanding transaction services. A North American bank said it was attempting to broaden client relationships and reduce its focus on lending.

In July 2013, the PBOC relaxed its control over lending rates by allowing lenders to price their loans based on prime rates. The LPR is based on the lending rates offered by nine commercial banks to their best corporate clients. The banks’ influence on the rate is weighted based on their market share of outstanding loans in the preceding quarter.

At the end of June 2013, the five largest domestic lenders represented 44% of the industry total. It is accepted that the new LPR is a better reflection of credit supply and demand. It will lead

---

4 In addition to the ones mentioned above, the following areas were also mentioned in no order of priority: acquisition activity (this area was often mentioned in the context of the markets where the banks have a strong presence, such as the U.S., Germany, Latin America and Russia), bond trading, equipment financing, energy financing, syndications, interest rate swaps, corporate deposits, expanding network in Western China, debt capital markets, transaction banking and mortgages and insurance.
to a growth in swap products as corporations use derivatives to hedge interest rate exposure. This will represent increased opportunities for the foreign banks.

5. Product and segment developments

5.1 Products in demand

The respondents were provided with separate lists of different retail and corporate banking services, and were asked to rank the three services from each group that they believe will increase in importance for foreign banks by 2016.

5.1.1 Retail products

Twelve banks responded to questions about retail products and put investment products and mobile banking well ahead of private banking and credit cards (Figure 7).

5.1.2 Corporate products

Thirty-seven banks responded to the corporate services questions, and placed debt capital markets well ahead of securitization and interest rate swaps. Structured products, cross currency swaps, and online transaction systems followed (Figure 8).

The participants’ perspectives on the securitization market has been influenced by the PBOC’s announcement in October 2013 that locally incorporated foreign banks would be permitted to join a pilot plan that allows banks to package loans into tradable securities referred to as collateralized loan obligations (CLOs). The pilot scheme was previously limited to just domestic banks and was placed on hold during the global financial crisis. It is suggested that the scheme will be limited to relatively small deals in the U.S.$150m to U.S.$1.5b range.

5.1.3 Wealth management

Only 13 banks said they were targeting the wealth management segment. The respondents agreed that the products offered were generic and it was difficult to differentiate their offerings from those of domestic banks. Some mentioned serving Chinese clients with a link to their home market.
Figure 9: Priority scores of foreign banks
One bank said that they provided low risk quality products, while another said that the market’s future potential was phenomenal. Several Asian banks mentioned that it was a challenge to compete with the large, long-established European banks. Another Asian bank said that their involvement in the wealth management market was being supported through a joint venture with a trust company.

It goes without saying that definitions of wealth management clients differ between banks.

5.1.4 Mutual funds
According to the Asset Management Association of China, the Chinese mutual fund industry totaled U.S.$650b (4% of deposits in the banking system) as of May 2013. Several foreign banks have been licensed by the China Securities Regulatory Commission (CSRC) to sell mutual funds, including HSBC, Citibank, JPMorgan Chase, Standard Chartered Bank, DBS, Hang Seng Bank, Bank of East Asia and Nanyang Commercial Bank.

Twelve banks shared their opinions on the difficulties of distributing mutual funds. Three quarters agreed that it would present challenges. However, it was considered a useful addition to the foreign banks’ product line. Some said they would offer just three to five fund company products, while others said they would offer a tailor-made fund advisory service.

5.1.5 Foreign banks’ interest in other parts of the financial sector
Foreign banks have also been active in recent years in forming joint ventures with securities companies, such as Deutsche Bank, Goldman Sachs, Morgan Stanley, Citigroup, JP Morgan and UBS.

Cinda Asset Management, for example, which has been given the go-ahead to list in Hong Kong, has two foreign bank shareholders, UBS and Standard Chartered Bank. Cinda is one of four asset management companies set up in 1999 to absorb problem assets held by China’s four largest banks.

5.1.6 Foreign banks’ use of social media
The banks were asked to comment on their use of social media to promote foreign bank services. Five banks (all with retail operations) within a group of 33 respondents indicated that they use Sina Weibo. Sina Weibo is a micro-blogging site with similarities to both Twitter and Facebook. It is only active in the Mainland and some other Chinese-speaking markets, and had over 500 million registered accounts at the end of 2012. By June 2013, it reached 54 million active daily users. In January 2011, a rival service called Weixin or WeChat was launched by the Chinese internet company Tencent. WeChat already has over 100 million users outside of China.

Citibank has the largest number of followers. It uses Sina Weibo to promote credit card usage at Dining City in Shanghai, Beijing, and Guangzhou.

Both DahSing Bank and DBS feature investment products on their pages. DBS appears to target an internationally-oriented customer by promoting rates for both U.S. and Australian dollar deposits.

Finally, Standard Chartered Bank highlights Priority Banking and its team of investment advisors. All of these examples and statistics were obtained from Sina Weibo pages and were not discussed separately in the survey interviews (Table 2).

5.1.7 Customer segments: future opportunities for growth
Figure 9 displays the levels of interest of the foreign banks in six different market segments: SOEs, POEs, SMEs, financial institutions, global corporates, and high networth individuals. Respondents were asked to score their level of interest in each segment on a scale of 1 to 10 where 10 represents maximum interest.

<table>
<thead>
<tr>
<th></th>
<th>Followers</th>
<th>Weibo posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citibank</td>
<td>150,864</td>
<td>334</td>
</tr>
<tr>
<td>Dah Sing Bank</td>
<td>2,388</td>
<td>256</td>
</tr>
<tr>
<td>DBS</td>
<td>109,280</td>
<td>5,089</td>
</tr>
<tr>
<td>Standard Chartered</td>
<td>41,695</td>
<td>1,623</td>
</tr>
</tbody>
</table>

Table 2: Foreign banks using Sina Weibo
Source: Sina Weibo, 21 November 2013

5 Cinda Asset Management completed its IPO in December 2013.
6 The four asset management companies and the banks whose bad assets they manage are: Cinda (China Construction Bank), Great Wall (Agricultural Bank of China), Orient (Bank of China) and Huaxiang (Industrial and Commercial Bank of China).
Future directions for foreign banks in China

SOEs: out of a total of 34 banks, 23 attributed an interest score of eight or higher in SOEs. Nine banks recorded the maximum score of 10.

POEs: although POEs are a growing part of the Chinese economy, they are currently less attractive to the foreign banks than the SOEs – 15 attributed a score above eight.

SMEs: the pattern of interest changes with SMEs. Only four participants assigned a score of 10, and overall only seven banks scored eight or higher. Ten banks out of a total of 23 banks showed minimal interest.

Financial institutions: with the exception of just one bank, the foreign banks interact closely with other financial institutions in the Chinese market. 18 banks assigned a score of eight or higher in terms of their interest in this segment.

Global corporates: 18 respondents scored eight or higher and 29 of the 35 respondents scored six or higher. As expected, global corporates that are active in China remain a key target segment.

High networth individuals: only 19 banks responded to the importance of high net worth individuals. Just eight banks recorded a score of eight or higher, suggesting that this segment is of interest to just a small group of foreign banks.

5.2 Level of interest in investing in domestic banks and other parts of the financial sector

76% of the respondents believe that foreign bank interest in investing in Chinese banks has declined. One of the most frequently cited reasons for this change is Basel III, and the need to hold extra capital against minority stakes in other financial institutions.

Foreign banks that helped many of the larger Chinese banks go public have in recent years sold their stakes. For example, according to Dealogic, U.S. banks acquired a stake of at least U.S.$14.8b in Chinese banks between 2002 and 2010, and then sold U.S.$37.3b worth of these positions from 2009 to 2013.

Currently, foreign bank ownership is limited to 20%, and as one European banker noted, even if this was permitted to rise to 49%, foreign banks would still be limited in their management and control.

Some participants suggested that the pending process of financial liberalization may create instability among mid-sized city commercial banks. A subsequent period of consolidation might then provide new opportunities for foreign banks if the CBRC permit takeovers of troubled banks.
In addition to investing in banks, participants were asked to suggest other areas of the financial sector that might offer future potential for foreign banks. The banks suggested the following: securities companies, trust companies, leasing companies, factoring companies, consumer finance companies, asset management companies, auto finance companies and life insurance companies.

6. Human resources developments

6.1 Levels of demand for different foreign bank positions

The respondents were provided with a list of 17 different job functions and asked to rank the top five in order of importance (Figure 10). At the top of the list was corporate relationship bankers, which was selected as the most important job function needed by 20 banks. The second position in highest demand was legal and compliance personnel, ranked by 26 banks. In third position was risk management personnel, ranked by 16 banks.

The numbers can be further refined to display the needs of particular types of foreign banks. For example, relationship bankers in retail banking were selected among the top three positions by six banks within a group of eight banks.

Three banks ranked branch personnel in the Top 3. Staff functions that recorded very low levels of interest were accounting, finance, new product development and taxation.

6.2 Skill shortages

Figure 11 presents the results of the respondents’ rankings of seven types of skills required in banking. This evaluation supplements the results presented in Figure 10. It also highlights some of the cultural challenges faced by the foreign banks.

As Figure 11 shows, foreign bank employees are viewed on a relative basis to lack leadership skills and international experience. Six banks highlighted leadership and four banks highlighted international experience as very serious human resource issues.

At the opposite end of the scale, three factors were viewed as not of great concern (based on an average score of less than 3 for 38 banks). They were, in order of decreasing concern, product knowledge, team orientation and client relationships.

A number of participants noted that because of the limited products permitted in China, product knowledge was not a significant concern.
6.3 Salary increases expected
A number of banks observed that salaries may not grow as rapidly in 2013 as in the past. One banker said that they were “not as crazy” this year while another said they anticipate only a modest increase during 2013. The most common estimated increase was 8%, made by 10 banks (Figures 12 and 13).

When asked to comment on the foreign banks as a whole, 13 of the 32 respondents also opted for 8%.

There were nine banks that expected salary increases of more than 8%, including two European banks and one North American bank expecting increases of the magnitude of 15%, 18% and 20%, respectively. One of these banks commented that they had not received a salary increase for three years and this was a catch-up move.

Overall, 36 banks envisaged salary increases in 2013 and only two banks said they would remain unchanged.

Participants confirmed that there had been no remuneration changes in 2013 in relation to cash bonuses, equity options and bonus deferrals.

A Hong Kong foreign bank said that they now abide by CBRC rules on cash bonuses and deferrals. Another Asian bank indicated that they had begun to structure packages around staff retention strategies.

6.4 Foreign bank versus domestic bank salaries
The bankers were asked to compare levels of salaries for foreign banks versus their domestic counterparts across three different levels: senior management, middle management and support staff.

As Figure 14 presents, the belief remains that foreign banks, as a rule, continue to offer higher salaries. However, this is only part of the remuneration package and when other components, such as housing and car allowances and performance bonuses are included, the domestic bankers’ package becomes more attractive.

One participant summarized the position by saying that the elite domestic bankers earn a lot more while the mid-level bankers are more or less the same.

6.5 Staff retention
The average turnover rate for the 38 respondents was 14.2% in 2012 and they anticipated that this would drop to 12.9% in 2013.

One bank mentioned that it had lost 40% of its staff in 2012, including many front office positions. Another small foreign bank said it lost 30% in 2012, while a bank that lost 20% advised that if you record more than 30% turnover, a call could be received from the CBRC.

Figure 15 shows that 18 banks predicted 15% or higher turnover in 2013 versus 19 banks in 2012.
However, a review of individual bank percentages shows that five banks expected higher turnover in 2013 versus 15 banks which anticipated a decline. This suggests that retention is improving and this may be a reflection of the overall economic situation.

7. Conclusion
The future of foreign banks in China will much depend on the scope and implementation of the financial reforms. Interest rate liberalization, RMB internationalization and the FTZ offer a real prospect for accelerating the breadth of opportunities for the foreign banks in China. The overall sense is that foreign banks will have an important role to play in the opening-up of markets and a rebalancing of the Chinese economy for the years to come.
An E.U. financial transaction tax and the unintended consequences for risk management

Serge Wibaut
Professor of Finance, Université catholique de Louvain

Abstract
In the aftermath of the subprime crisis in the U.S. and the sovereign debt crisis in Europe, the opportunity for establishing a financial transaction tax (FTT) has become a topic of debate in the European Union. In this article, we survey the literature dealing with the possible theoretical and empirical implications of such a tax on market volatility. We then turn to the possible – and unintended – consequences of an FTT on savers and investors. We conclude that these consequences might outweigh the benefits of the FTT. More specifically, we find that an FTT is unlikely to meet its stated volatility control and revenue raising objectives, i.e., an FTT is unlikely to decrease volatility, and indeed, volatility might increase as markets became less liquid. It might raise very little revenue and could work to create more risk and deter long term investment.

And then there are serious unintended side effects to consider. Most importantly, for the financial security and safety of the whole financial system, an FTT might heavily penalize pension funds, as well as the banks in their liquidity management and risk management activities, to the detriment of a well-functioning financial system.

1 The author was formerly CEO of Axa Bank (Belgium) and Chief Investment Officer for Axa Northern and Central Europe. The author wishes to thank the anonymous referees for useful comments. All remaining errors remain his. This research was partly funded by a research grant from the CME Group Foundation.
1. Introduction

The merits of a financial transaction tax (FTT) have remained in the news as the E.U. debates proposals for such a tax. Interest during the last few years in an FTT for Europe appears partly motivated by the European sovereign debt crisis and the fiscal austerity that followed. Many political leaders, who were both upset by the market volatility and in need of finding new sources of revenue, found “speculators” in the financial markets to be suitable villains to blame.

This review of the issues and challenges associated with an FTT suggests among other things that:

- An FTT is not likely to decrease volatility, and indeed, volatility might increase as markets became less liquid.
- An FTT might raise very little revenue.
- An FTT could work to create more risk and deter long term investment.
- An FTT might heavily penalize pension funds, as well as the banks in their liquidity management and risk management activities, to the detriment of a well-functioning financial system.

The idea of this kind of tax had already been floated by John Maynard Keynes during the “Great Recession” [see Keynes (1936)] and was resurrected by James Tobin in his 1972 Janeway Lectures at Princeton. The future Nobel prize-winner economist suggested then it might be wise to introduce a tax on currency transactions in order to reduce the volatility that was surging after the demise of the Bretton Woods. Besides “throwing sand in the wheels of speculation” [Eichengreen et al. (1995)], such a tax, claimed Tobin, would increase social welfare as “vast resources of intelligence and enterprise are wasted in financial speculation, essentially playing zero-sum games” [Tobin (1991)].

Although economists’ DNA is programed to be hostile to transaction taxes (since these lead to inefficiencies in the allocation of resources), many economists agreed with Tobin. For instance, Summers and Summers (1990) claim that a financial transactions tax could eliminate “wasteful trading “ and “excessive financial engineering” and Greenwald and Stiglitz (1986) show that an FTT could be welfare-improving (in the Pareto sense) if markets are incomplete or if the information is imperfect.

Most proponents of FTT also agree that the tax should apply not only to spot transactions but also to derivative contracts, since these are close substitutes to spot positions taken in the underlying asset.

Not everyone agreed, however, with Keynes and Tobin. For instance, Friedman (1953) objected to the Keynes argument by saying that speculation could not be destabilizing in general since, if it were, speculators would, on average, lose money. Indeed, he claims: “People who argue that speculation is generally destabilizing seldom realize that this is equivalent to saying that speculators lose money, since speculation can be destabilizing in general only if speculators on the average sell when the currency is low in price and buy when it is high.”

Other critics pointed out that although it might be deemed fair to tax “excessive” speculation, it is however practically impossible to distinguish speculators from “ordinary” portfolio managers and even relatively active commercial risk managers. As such, taxation would be blindly applied to all and hence lead to inefficiencies in the allocation of resources as well as to the potential for considerable unfairness against different types of hedgers, risk managers, and portfolio managers.

In this paper, and after examining the basic concepts and tenants in the European proposal as it stood in the summer of 2013, we briefly review the theoretical and empirical (respectively in section 3 and 4) literature surrounding the FTT, focusing essentially on the capacity of taxation to curb volatility and deter speculators from engaging in “wild” trading.

In section 5 we examine what could be the consequences of an FTT, expected and unintended, on some financial practices, such as asset management and risk management.

2. The scope of the European debate

The FTT debate continues to evolve. Over 2012, 2013, and into 2014, 11 member countries of the E.U. have engaged in a vigorous discussion aimed at applying a consistent and uniform FTT across France, Germany, Belgium, Spain, Italy, Portugal, Estonia, Austria, Greece, Slovakia and Slovenia. Other countries are either opposed to the measure (most notably the U.K. and Sweden) or have uncertainties about their position (Luxemburg and the Netherland, for instance).
While the breadth of the debate may have narrowed a little as political compromises have been discussed, the original scope of the debate [European Union (2013)] was quite large “covering transactions relating to all types of financial instruments as they are often close substitutes for each other.”

Moreover, part of the desire was to apply an FTT to all transactions as long as at least one party to the transaction was established in the territory of a participating member state, irrespective of whether it is acting on its own account or on behalf of another entity. Under such a broad international reach, any of the following transactions would be subject to the tax: (a) the purchase and/or sale of a financial instrument before netting and settlement; (b) the transfer between entities of a group of the right to dispose of a financial instrument as owner or any operation implying the transfer of the risk associated with the financial instrument; (c) the origination (or modification) of derivative products; (d) repos, reverse repos and securities lending operations; and (e) any exchange of financial instruments.

The initial E.U. Directive would not apply to certain activities, mainly: (a) primary market transactions (underwriting), (b) transactions with central banks of member states and the ECB, (c) transactions with the European Union, the European Investment Banks, the European Financial Stability Facility and the European Financial Mechanism, as well as certain other international organizations. Retail activities, such as mortgages, loans and consumer credits would remain outside the scope of the FTT. Ironically, when one refers to the original idea aired by James Tobin, the E.U. version of an FTT would not apply to currency transactions on the spot market (but derivatives on currencies would be subject to taxation).

While determining whether a financial institution is considered to be established in the territory of a participating member state required that a number of conditions were met, it essentially boiled down to whether the institution had been authorized by the authorities of a member state to act as such in respect to transactions covered by the Directive. This included the existence of a registered office or a branch. It also included transaction where the institution is a party, in its own name or in the name of a third party, to a financial transaction with a counterparty established in a member state (residency principle) or in one of the financial instruments covered by the Directive (issuance principle).

The initial E.U. Directive defined financial institutions as investment firms, regulated markets, credit institutions, insurance and reinsurance undertakings, UCITS, pension funds, alternative investment funds and securitization special purpose vehicles. It did, however, exclude from its scope central counterparties, central securities depositaries and the member states themselves.

During the discussions, it was also suggested that member states should be allowed to set their own tax rates, but that it could not be lower than 0.1% for all transaction except derivative contracts, which would be taxed at a minimum rate of 0.01%. And, adding a bit complexity for less liquid securities, the tax rate would be the fair (market) value of the transaction, or the notional amount if the transaction is a derivative contract.

Several European countries have (or have had) some experience with different forms of FTT. For instance, the U.K., Belgium and Switzerland impose a stamp duty on equity transactions. Sweden had imposed a fully-fledged FTT (with a very high rate) in 1983 but abandoned the scheme in 1991 after most of the financial transactions had fled the country.

At any given time, the actual proposals would embrace the narrower constructs to achieve a compromise agreeable to all 11 countries participating in the discussions. But the political winds can, and do, shift, as so does the economic environment. And, as long as political populism perceives financial institutions as easy targets on which to blame market volatility, the likelihood of an E.U. FTT being adopted cannot be dismissed.

3. The theoretical impact of a FTT on volatility
The FTT has been thoroughly studied over the years. These studies have adopted different approaches, such as the tax being viewed as a surrogate taxation of capital income [Grandcolas (1986)], as a means of improving the efficiency of the financial sector [Summers and Summers (1990)] or to eliminate the distortions against fixed capital formation [Tornell (1990)]. More recent studies [Lo (2009)] have focused on potential reforms following the financial crisis of 2007-08.
We shall narrow our focus here to only those studies that look at the impact of an FTT on volatility and liquidity and later examine the possible implications for risk management practices.

The classical financial literature can hardly explain excess liquidity or volatility since it assumes agents’ rationality and efficient markets, which together lead to stable equilibrium. Consequently, and not surprisingly, to examine the impact of an FTT, most authors start with the assumption that some agents are irrational. These “noise” traders can be assimilated to “chartists” or plain retail investors who act irrationally or use irrelevant information to build their portfolios. Rational, or fundamentalist, traders try to arbitrage against such agents but cannot do so in a very efficient way as they are uncertain of the noisy traders’ reactions.

In most of these models, the volatility of the market is driven by the ratio of noisy traders to rational agents and the degree to which the former are irrational.\(^2\) In this context, some economists have tried to study how a transaction tax would affect the volatility of financial markets, and their findings, as we shall see in the brief review of the literature that follows, are rather inconclusive.

Hau (1998) developed one of the first such models, in which he considered a market where agents formed heterogeneous expectations. The introduction of an extra trader increases the depth and liquidity of the market but also adds noise and trading risk. In the absence of a tax, the equilibrium in this market is characterized by excessive market entry and “highly” volatile prices. The introduction of a tax would result in a reduction in both trading activity and price volatility.

Haberer (2004) looks at how a transaction tax would fare in a market where efficiency is measured as the market’s capacity to absorb new information. In an inefficient market (i.e., one in which some traders form their expectations using “wrong” techniques), Haberer shows that at low levels of trading, increased liquidity reduces volatility, but that as trading picks up the uncertainty due to increasing speculation causes more volatility. Introducing a tax in such a market would increase volatility if trading was thin but would curb it once volumes increase, as it would reduce the incentive for speculative trading.

Shi and Xu (2009) look at a model where a transaction tax could change the proportion of noisy-to-rational traders on a market characterized by entry costs. In their model, the rational agents’ expectation formation depends on the number of noisy traders. They show that in such a context three equilibria may arise. If the number of noisy and informed (rational) traders were the same, then the introduction of a tax would have no impact on volatility. On the other hand, if the ratio of noisy to rational agents is not equal to one, then the increased entry cost (due to the tax) will discourage rational agents from trading, which would cause volatility to increase as the proportion of noisy traders would increase. The third equilibrium would occur if the entry cost would be so high that it would discourage noisy traders from entering the market. In that case, levying a tax would have no effect on volatility.

Some economists also looked as to how the microstructure of the market could influence the effectiveness of an FTT. For instance, Pellizzari and Westerhoff (2009) show that in markets characterized by continuous double auction (which is the most common method used by modern markets to fix transaction prices) levying a tax would have positive, though small, impact on volatility, since it would reduce liquidity and hence increase the impact of an order on price variation. However, in a market operated by dealers and in which they are the providers of liquidity, a tax could decrease volatility by discouraging noisy trading.

The impact of an FTT has also been studied in the context of so-called zero-intelligence (ZI) models. This approach models individual agents who behave without strategy within a specific market mechanism so that the impact of the market characteristics might be isolated. By assuming that individuals behave randomly, the market mechanism can be observed independently of agent’s specific strategies. As such, it is then possible to distinguish the driving forces of market mechanism from the agents’ strategies. Using the ZI technique, Farmer et al. (2005) show that a given market may exhibit larger volatility if the ratio of market orders to limit orders increases. They explain this result by the fact that traders passing orders at market prices demand more liquidity than limit-order agents. If the ratio of

---

market to limit orders increases the liquidity on the market will decrease, bringing about an increase in volatility.

Building on findings of Farmer et al., Ehrenstein et al. (2005) evaluate, within a ZI model, the impact of a Tobin tax on volatility and conclude that it results in a decrease in volatility as long as the tax rate is not “too” high and does not significantly affect liquidity.

Their results are, however, contested by Mannaro et al. (2008), who find that if the tax is levied on a single market at a rate of less than 0.5%, volatility should increase if some noise traders are present in the market. They also conclude that if the traders can switch to another market, the taxed market would become more volatile than the untaxed one as the tax will reduce liquidity.

While necessarily incomplete, this brief sampling of the research on the theoretical impact of an FTT on market volatility demonstrates the lack of agreement within the economics profession. While a majority of models tend to suggest that such a tax would reduce market volatility, many others contend the contrary. Since the consensus among economist today is to say that many different heterogeneous agents coexist in the markets, or that many agents change their strategies through time (passing from rational behavior to noisy trading and vice-versa), it is to be feared that many of these models also coexist at the same time in financial markets. Consequently, we can only say that the economic literature is inconclusive as to the impact of an FTT on market volatility. Moreover, the studies that do suggest that an FTT might reduce volatility typically rely on a very unappealing set of critical assumptions, such as irrational traders willing to lose money consistently or noisy traders with no strategies and random processes, etc.

4. Empirical evidence

Empirical studies about the FTT have mainly three concerns. First, does such a tax affect volatility? Second, does it affect trading volume? And thirdly, does it have an impact on the price of securities themselves?

These studies, however, face at least three hurdles. First, while there are cases of FTT which have been observed in the past, it might be difficult to infer robust results from these very different cases, often with missing or incomplete data. Since such a tax would amount to an increase in transaction costs, however, one can look at the impact of transaction costs on volatility and liquidity to surmise how an FTT would fare in a financial market.

Second, it might be quite difficult to split transaction volumes into their “fundamental” and “noisy” components. As a consequence, it is practically impossible to determine how each of these components is affected by the tax.

The third problem concerns the ways prices are affected by the tax, since these prices will also be influenced by the way traders will react to the tax. Will they trade in assets that are close substitutes but that are not covered by the tax? How will they react to the drop in liquidity that should follow the FTT? etc..

It needs to be realized, therefore, that empirical studies will only shed partial and incomplete light on the problem. Most of the literature devoted to the analysis of transaction costs on trading report a positive correlation between the costs and volatility. Those studies mostly concern transaction costs on equity markets. Mulherin (1990) examines the evolution of transaction costs on the NYSE between 1897 and 1987 and concludes that if the imposition of a transaction tax does result in reduce trading, it will not mitigate volatility.

Jones and Sequin (1997) corroborate his findings, as do Atkins and Dyl (1997), who find that a tax would prevent stock prices from adjusting to new information instead of reducing short-term speculation. Liu and Zhu (2009) reach the same conclusion for the Japanese equity markets.

Aliber et al. (2003) examine how foreign exchange markets are affected by transaction costs for four currencies (D-Mark, GB pound, Yen and Swiss Franc) over the period 1977 to 1997. They found that, although transaction costs are quite small, an increase of 2bp in these costs leads to an increase of 0.5% in volatility.

We now turn our attention to articles studying equity returns in those countries that have imposed equity taxes. Umlauf (1993) studied the Swedish markets during the period 1980–87, which is before and after the imposition of a brokerage tax, and finds no sign of a decrease in volatility after the introduction of the tax.

Saporta and Kan (1997) examine the impact of the U.K.’s stamp
duty on volatility of stock prices and find no correlation there either. Hu (1998) examines the effects of transaction taxes on volatility in Japan, Hong-Kong, South Korea and Taiwan between 1975 and 1994 and does not find a statistically significant impact.

Looking at the impact of the tax on trading volumes, Umlauf (1993) reports that when the rate of the FTT in Sweden was increased from 1% to 2% in 1986, 60% of all the trading volume of the most actively traded Swedish stocks migrated abroad (mostly to London). Hu (1998), on the other hand, finds that in the 14 cases of tax changes in the Asian markets he examined, turnover was hardly affected by the change.

Finally, and with regards to the impact of the taxes on prices, Umlauf (1993) reports a drop of 2.2% on the day a 1% tax was introduced and a 0.8% drop when the tax was raised to 2%. Hu (1998) reports a drop of 0.6% in Hong-Kong and of 1.6% in Taiwan on the announcement date of the tax. Saporta and Kan (1997) find that on the day the U.K. stamp duty was increased from 1% to 2%, the market dropped by 3.3%. There is no mention of whether those price falls were compensated once the tax had come into effect and became part of business as usual.

From this brief review of the empirical evidence of the implications of an FTT, we find that the results concerning trading volumes and price levels are mixed and probably incomplete. A more important message from the empirical research is how little volatility seems to be affected by the introduction of a transaction tax. While typically left beyond the scope of these research studies, one may contemplate that market participants are essentially attempting to manage the natural and manmade risks that arise from the interplay of the fundamentals of supply and demand as they react to changes in government policies and regulations. Hence, the view might be taken that the source of market volatility is only partly related to the mechanics of the marketplace.

5. (Unintended) consequences of an FTT
An FTT will have many consequences in terms of public finances, of income distribution (who will eventually pay the tax) and on the cost of capital. We shall focus here on the probable implications on some lines of financial activities, namely UCITS (and pension funds), risk management and liquidity management.

Before commenting on the impact of the FTT on these lines of activities, however, one should bear in mind that the nominal rate of the tax may not give an accurate indication of how much tax will eventually be paid. Indeed, as was noted in a publication from Clifford Chance (2011), the financial system is so complex nowadays that cascade effects will multiply the impact of the FTT. As will be demonstrated below, the sale of a security puts into motion a whole series of events involving numerous participants in the financial markets, the result of which might be that the tax may have to be paid at different stages of the process even if only one “effective” transaction was initiated.

Imagine that a pension fund wishes to sell a derivative through a regulated exchange. It will address itself through a broker, who in turn will execute the order and settle the transaction with the clearinghouse. Note that the central counterparty (CCP) is exempted from paying the tax. The buyer counterparty will in turn use the services of a broker, an exchange and a clearinghouse. All in all, the FTT may have to be paid as many as five times. This cascade effect may very well induce transactions to move away from regulated markets and back to the OTC markets, thereby increasing credit and settlement risk.

The FTT may, therefore, very well disincentive the use of CCPs, especially for derivative clearing, and as a consequence decrease the protection offered by regulated markets.

Turning now to the impact of the FTT on asset management, the cost of the tax would depend on the turnover ratio of the UCIT (i.e., how fast the assets under management are replaced on average with other holdings in a given year). A recent study by EFAMA (2013) tends to show that the average turnover of long-term investment UCITS (i.e., excluding money-market funds (MMFs)) is 0.9 and about 6.5 for MMFs.

MMFs would thus be severely hit by the FTT, since their clients use those funds on a continuous basis to invest and disinvest their cash. It appears that the average holding period of short-term MMFs is less than 60 days. This short holding period implies a high turnover of the assets, as a result of which the tax will have to be paid several times a year.

MMFs will be practically be doomed if the FTT were to be introduced, since the FTT would on average amount to 65bp per
year. In the current low-rate environment, one should expect this industry to disappear if the FTT came into effect.

Also, note that many equity funds that are based on quantitative algorithms have a much higher ratio than 0.9, as they try to profit from small inefficiencies in the markets. One can only suppose that such funds will struggle to survive an FTT.

Many fixed-income funds make extensive use of derivative instruments either to manage their duration or to hedge credit exposures. Under an FTT scheme, these practices will be penalized and the return to the client will decrease.

One should also take into account the indirect effects of FTT, essentially an increase in spreads (see section 3 above). On top of this increase, one should also assume that stock and bond lending would probably become unprofitable, resulting in lower returns for the final investor. Indeed, levying the FTT on security lending will start a cascade effect, as the lending is most often followed by an onward sale of the borrowed securities. Once the borrower sells to another investor it will result in a double taxation, both on the sale and repurchase of the asset and then again when the borrower repurchases the securities in the market to meet their obligations toward the lender.

In view of these different impacts, one can expect that for a long-term investor, such as a pension fund, the expected costs – indirect and indirect – of the FTT would be in the range of around 25bp, which once capitalized over 30 years would amount to a final loss of about 7.7%.

Liquidity management will also become more expensive if FTT was levied. Indeed banks’ liquidity management relies heavily on repo transactions. A repo is a secured lending/deposit operation that allows the borrower to benefit from a low interest rate because of the collateral provided to the lender. The European Directive, in its proposed form, classifies a repo as a standard sale and repurchase of a bond, meaning that the tax would be levied twice.

This feature of FTT is completely at odds with the current – and recent – banking regulation that requires banks to hold more liquid assets or assets of such quality that they can be eligible for a repo operation. The adverse effect on liquidity will impact not only banks but also corporates that rely on banks for their own liquidity. The tax could, therefore, prevent corporates from efficiently accessing financial markets and funding, and the problem would probably be even more acute for SME’s.

Turning now to the impact of the FTT on risk management in banks, it should be noted that the tax may dissuade banks from certain hedging activities as these may become too expensive. As such, the FTT will increase systemic risk in the financial sector. Let us look at a few examples of hedging activities that may become too expensive to carry out.

Certainly, all short-term activities will become heavily penalized since they have to be renewed frequently. Examples of such operations are FX swaps hedging money market operations whose maturity ranges from a few weeks to three months, leading to paying the tax at least four times a year and probably much more than that.

The same also holds true for forward rate agreements (FRAs), whose short leg usually has a maturity of a few weeks. Once again, the FTT will have to be paid several times a year. In the present extremely low interest rate environment, the cost of these hedges may become relatively quite expensive.

Even in the realm of risk management, hedging of plain vanilla instruments, such as a bond or a bond portfolio, may become very expensive. As we have seen, once a transaction is carried out on a regulated exchange or a CCP platform the number of operations increases with the number of intermediaries. So that, for instance, if a financial institution wants to buy an option to hedge itself against adverse movements in interest rates, it will have to follow the whole chain discussed above. If, on top of this, the cost of the option is to be paid with a repo operation, a risk manager may think twice before starting the operation or chose a less costly (and less efficient) hedging strategy.

All in all, if the tax applies to all types of operations and/or is high, banks will think twice before initiating some hedging activities. This could result in less hedging or in a less efficient hedging strategy. Taxing hedging operations goes clearly against the spirit of the current banking regulatory environment that aims to decrease the risk of financial institutions.
Finally, let us remark that the Basel regulations force banks to carry more capital if their operations are riskier. Banks will then have the choice to either take on more risk, because hedging becomes more expensive, and thus hold more capital or hedge at a higher cost with the same level of capital, and hence decrease their return on equity. Caught between a rock and a hard place, banks may very well choose to pass the entire burden of the tax to their clients, which is very far from the FTT’s aim.

6. Conclusions

We have seen that the theoretical literature dealing with the Tobin tax does not yield a clear-cut view on whether levying such a tax will decrease volatility and speculative activities. Researchers that are willing to make critical assumptions about the irrationality of certain market participants are able to find the potential for some reduction in volatility. By contrast, many models founded on certain realistic conditions tend to suggest that the FTT could very well increase volatility.

Empirically, it appears that few attempts to impose a Tobin-like tax were successful. At best, a duty-stamp has been levied on equity trading for some time in many countries and does not seem to have stemmed speculation.

More confusingly, the tax could create more risk and deter long-term investment. We have mentioned that the tax, at least in its proposed form, would heavily penalize pension funds, a quite unacceptable result in western societies that have to cope with aging populations.

Banks will also be penalized in their liquidity management, as well as in their risk management activities. This puts them at odds with the current banking regulation that forces banks to hold a larger cushion of liquid assets and favors better risk management practices and a lower risk profile.

We may conclude that while many may have some sympathy for the Tobin tax, particularly governments in need of new revenue source, one should not ignore the side effects of an FTT. Governments should recognize that the direct, and especially the unintended indirect costs, may very well exceed any real benefits. Risk management practices may be altered in a way that is detrimental to the overall safety and security of the financial system. Moreover, any benefits are highly likely to be overestimated once market participants are faced with the reality of the tax and begin to alter their behavior in response to the new cost structure. The real question, therefore, may become whether an FTT is the best instrument to fight destabilizing and “excessive” speculation. And, that is assuming we can in fact define and measure “excessive” speculation in the first place.
References


Regulatory experience in the U.S. and its lessons for the European Union

Anthony Saunders
J.M. Schiff Professor of Finance, Stern School of Business, New York University

Abstract
This paper looks at the challenges facing the Eurozone countries in establishing a banking union. Using the historical regulatory experience of the U.S. as a laboratory for comparative analysis, this article suggests that due to the diverse set of economies with competing local supervisors and diverse bankruptcy laws achieving the three “legs” of a European bank union, namely supervision, deposit insurance and restructuring/resolution, is by no means an easy task. Even today, the U.S., despite having introduced its first national banking regulation in 1863, still falls short of what might be viewed as a full banking union, despite the existence of a common currency from at least 1913 with the foundation of the Federal Reserve. Whether the Europeans can tackle this challenge with such a diverse group of countries remains to be seen.

1 For a Spanish language publication of this paper, see “La Experiencia Regulatoria en Estados Unidos y sus Lecciones para una unión bancaria Europea,” Papeles de Economía Española, 137, 19–27. I would also like to thank FUNCAS for their financial support and Santiago Carbo-Valverde and Francisco Rodriguez Fernandez for comments.
1. Introduction
In this paper, I will discuss the progress and potential barriers to banking union in Europe using the historical regulatory experience of the U.S. as a laboratory for comparative analysis. I will begin my analysis by highlighting the fact that unlike what many are lead to believe, the U.S., despite having introduced its first national banking regulation in 1863, the National Banking Act of 1863, still falls short of what might be viewed as a full banking union, despite the existence of a common currency from at least 1913 with the foundation of the Federal Reserve. This paper is organized as follows. In Section 2, I will analyze the development of U.S. banking "union" in the three areas deemed to be of importance for a European banking union, namely (i) a common supervisor, (ii) a common private deposit insurance system and (iii) a common failure/bailout/restructuring system organization or approach. To my mind, it is these three "legs" of policy that define the nature and degree of a banking union in any common currency area whether U.S. dollar or euro. In Section 3, I will discuss the developments so far in the Euro region and the prospects for unification related to the three "legs" of policy discussed above. Section 4 concludes.

2. The evolution of banking union in the U.S.
2.1 States supervision
It is often forgotten that in the U.S., considerable bank regulatory power exists at the state level. In the U.S., both prior to and following the civil war, considerable powers in bank licensing were in the hands of individual state regulatory bodies. Even today, so-called "states rights" play a central role in the regulation and licensing of banks. Prior to 1863 and the passage of the National Banking Act, U.S. banks were state chartered, small localized institutions and poorly if at all supervised. Moreover, they had the power to issue their own dollar currency. That is, dollar notes issued were linked to the bank that issued them. Notionally at least, these notes were backed by gold. As the number of state-chartered individual banks grew and transportation improved, it was not uncommon for dollar bills issued by a bank, say in Iowa, to trade at a discount in New York, reflecting the transactions cost of actually redeeming the gold backing such dollar bills. Some estimates put the number of U.S. banks operating in the U.S. in the mid-1860s at the tens of thousands. To provide some order to the process of licensing and supervision in the U.S. the National Bank Act was passed in 1863, establishing the Office of the Comptroller of the Currency (OCC). The Office of the comptroller of the Currency (OCC) established a National Charter for banks in Competition with the State Charters offered by state regulators. The primary function of the OCC is to charter so-called national banks as well as to close them. In addition, the OCC examines national banks and has the power to approve or disapprove their merger applications. However, instead of seeking a national charter, banks can, even today, be chartered by any of the 50 individual state bank regulatory agencies. The choice of being a nationally chartered or state chartered bank lies at the foundation of the "dual" banking system in the U.S. While most large banks, such as Bank of America, choose national charters, this is not always the case. For example, Morgan Guaranty, the money center bank subsidiary of J.P. Morgan Chase was until recently chartered as a state bank under New York state law. In September 2009, 1,492 banks were nationally chartered and 5,419 were state chartered, with approximately 69% and 31% of total commercial bank assets, respectively.

Thus, it is very clear that the U.S. does not have, even today, a single bank regulator. Moreover, the regulatory structure is even more complicated because the Federal Reserve, in addition to its monetary policy powers, can and does regulate bank holding companies and in addition provides important services, such as wire transfer and lender of last resort facilities to national banks and certain state banks that wish to become members of the Federal Reserve System.

For example, Citigroup is a holding company of Citibank, its main subsidiary. It should be noted that Citigroup has hundreds of subsidiaries other than Citibank, including insurance, investment banking and finance subsidiaries. However, Citigroup (the holding company) is viewed as the major “source of strength” to Citibank and is thus regulated by the Federal Reserve. At the same time, Citibank’s major regulators are the OCC and the deposit insurance provider, the Federal Deposit Insurance Corporation (FDIC).

As is clear, the supervisory, examination and chartering organization of U.S. banking is highly complex, and in the eyes of many, inefficient. For example, on any day, three different regulatory agencies may be visiting Citigroup, i.e., the Federal...
Reserve, the OCC and the FDIC. Although these agencies and their examiners are meant to coordinate and share examination and supervision information, considerable concern has been raised about their ability, and even willingness, to do so because of so-called “turf” battles.

An additional problem is the issue of a consistent examination methodology across regulators, especially national bank and state bank regulators, given that under the U.S. system it is the banks that choose which charter to adopt (note that adopting a state charter inhibits a bank from branching outside its home state). While both national and state examiners are meant to apply the same rating system, the so-called CAMELS system, which rates the bank from 1 (good) to 5 (bad) based on six factors: its capital (C), its assets (A), its management (M), its earnings (E), its liquidity (L) and its sensitivity to market risk (S), some studies have found that state regulators are far more generous in their rating examinations than national (OCC) regulators. This may have important implications for Eurozone supervision issues as to whether the ECB will focus only on large banks and leave smaller banks to national regulatory bodies or seeks to regulate all Eurozone banks irrespective of size.

In summary, the supervisory system for banks in the U.S. is highly fractured with multiple regulators and examination standards that are inconsistently applied based on the nature of the bank’s charter. An immediate lesson for any impending Eurozone banking union is the need or feasibility of having a single regulator who in turn can apply the same examination standards across all banks.

The supervisory structure of the US banking system is summarized in Figure 1.

2.2 Deposit insurance
While some individual U.S. states had tried and ultimately failed to set up operational private deposit insurance funds prior to 1933, the first and existing national deposit insurance fund (FDIC) was established in 1933 in the wake of the great depression when thousands of banks failed. The original purpose of FDIC was to protect “small depositors” from bank failures by levying insurance premiums on banks to create a centralized privately (bank) paid for deposit insurance fund. The original level of individual depositor insurance coverage at commercial banks was U.S.$2,500, which was increased (six times since 1934), to U.S.$100,000 in 1980, and to U.S.$250,000 in October 2008. Between 1945 and 1980, commercial bank deposit insurance clearly worked; there were no runs or panics, and the number of individual bank failures was very small. Beginning in 1980, however, bank failures accelerated, with more than 1,039 failures in the decade ending in 1990, peaking at 221 in 1988. This number of failures was actually larger than that of the entire 1933–79 period. Moreover, the costs of each of these failures to FDIC were often larger than the total costs of the mainly small bank failures in the 1933 to 1979 period. As the number and costs of these closures mounted in the 1980s, the FDIC fund, built up from premiums paid by banks (and the reinvestment income from those premiums), was rapidly drained. Any insurance fund becomes insolvent if the premiums collected and the reserves built up from investing premiums are insufficient to offset the cost of failure claims. FDIC’s resources were virtually depleted by early 1991, when it was given permission to borrow U.S.$30b from the U.S. Treasury. After 1991, there was a dramatic turnaround in the fund’s finances and a drop in failures — partially due to record profit levels in banks. Specifically, as of March 2008, the FDIC’s Deposit Insurance Fund (DIF) had reserves of U.S.$52.8b. The reserves to insured deposits ratio

---

Figure 1: U.S. bank supervisory structure

---

was 1.19%. In 2007 there were three DI failures; in 2005 and 2006 there were no DI failures.

However, the recent financial market crisis hit the banking industry very badly. In 2008, 26 DIs failed (at a cost to the FDIC of U.S.$20b) and in 2009, 140 additional failures occurred (at a cost of over U.S.$36b). By September 2009, the FDIC’s DIF reserves had fallen to ~U.S.$8.2b, less than zero for only the second time since its founding in 1933. However, unlike the near bankruptcy of the FDIC in 1991, the low balance in the FDIC's DI insurance fund did not result in the insurer's failure. Rather, the FDIC and the federal government took several steps to ensure the fund would have sufficient resources to deal with any and all depository institution failures. To address the falling balance in the fund, the FDIC levied one special assessment in early 2009 and a second in the fall of 2009, in addition to raising the rates banks were charged for deposit insurance. Further, the agency took the unprecedented step of requiring banks to prepay U.S.$45b of insurance premiums by the end of 2009. The premiums covered the fourth quarter of 2009 and all of 2010 through 2012. Finally, the FDIC was also given approval to tap U.S.$500b in additional funding from the Treasury Department through the end of 2010.

What are the major lessons to be learned from the experience of the FDIC? The first is that a privately funded deposit insurance fund can be organized to successfully protect small depositors against idiosyncratic bank failures, for example, due to inefficient management or portfolio diversification policies of a given bank. Second, it is extremely unlikely that a privately (bank) funded deposit insurance scheme can survive a period of systemic or contagious bank failures as occurred during the 1980s in the U.S. during a real estate crisis or in the recent 2008-09 crisis. In both cases, government intervention was required either through lines of credit to the fund being drawn upon or outright bailouts by the U.S. Treasury of large failing insolvent banks, as reflected in the U.S.$100b plus bailout of the 10 largest U.S. banks in the initial stage of the 2008 TARP. Third, setting the level of deposit insurance premiums to be paid by banks is extremely difficult. Indeed, in the last few years U.S. bank deposit insurance premiums paid into the FDIC fund have dramatically increased, yet they were still insufficient to prevent FDIC insolvency. This is despite them being made more risk-sensitive (risk-based) and forward looking. Some have even argued that high deposit insurance premiums, by draining bank profits, actually weaken banks and the banking system and could lead to more, not less, bank failures. Fourth, at what level should insured deposits become uninsured deposits?

Currently in the U.S., deposits up to U.S.$250,000 are insured and those above U.S.$250,000 are technically “uninsured.” However, while uninsured depositors may take a loss or haircut in a world where failures are idiosyncratic and not too large, the problem remains as to their treatment during a major crisis, when the government intervenes and bails out a bank, believing it to be “too big to fail” (TBTF), as it did through the TARP program. In such cases, all depositors are fully protected and have little or no incentive to monitor the performance of a bank. Thus, a key issue to be faced by any organized deposit insurance system is how to discipline depositors in the largest systemically important TBTF banks.

2.3 Bailout/resolution/closure policy

The approaches of the U.S. to resolving failed or failing banks have been extremely varied. What is often forgotten is that after the 10,000 or so bank failures and the Great Depression, the U.S. actually established a Reconstruction Finance Corporation (RFC) in 1934 to resolve or restructure banking institutions. Between 1934 and 1957 (when it ceased operations), its chosen methods of restructuring were either to make direct loans to these banks or buy their preferred stock. One clear question facing a European bank entity, such as the ERM, should it directly intervene in buying equity of a bank to help it recapitalize, is what type of equity should it buy? Preferred stock provides no direct corporate governance control over a bank's performance and pays a fixed return, so there is no real sharing of any upside gain from improved bank performance. On the other hand, preferred stock has greater protection against downside risk than common stock. Clearly, one issue for any Eurozone bank bailout policy is the mix of loans, preferred and common stock to be used in any particular bank intervention.\(^3\) As of 2013, virtually all the preferred stock issued by the banks to the U.S. government has been redeemed at a profit to the U.S. Treasury.

---

\(^3\) It is interesting to note that the U.K. Government took a common stock stake in a failing RBS (Royal Bank of Scotland).
After the savings bank crisis in the U.S. in the 1980's, which rendered 100s of banks insolvent, as a result of a property crash, a special purpose asset resolution agency was established to sell-off the assets of the failed banks. Specifically, the Resolution Trust Corporation (RTC) was charged with the task of resolving the assets of over 700 failed savings and loans between 1989 and 1995 (the end of its mandate). In performance, it resembled a mega “bad bank,” especially created to liquidate the assets of the failed banks (some U.S.$394b). In recent years, the good-bad bank structure has been used at the private individual bank level with private financing or limited government funds; in the RTC case it was mainly government funded with a second organization, the Resolution Funding Corporation, issuing long-term debt to fund the bad asset resolution. However, at times, the RTC entered into investment equity partnerships with private firms to resolve certain assets. On the one hand, the RTC has been viewed as a success as a resolution agency since such private deals allowed it to share in increased asset values as the economy recovered. On the other hand, it has been accused of giving deals to its private equity partnerships that were deemed by some to be too generous. As noted earlier the RTC closed in 1995.

With respect to the 2008–09 crisis, no special corporation was created to liquidate assets, as happened with the RTC and savings and loans, but rather the FDIC itself acted as the failure resolution agency by entering into deals to share the loss risk with either private buyers or more commonly with banks willing to acquire a failed bank’s assets. Under such deals, a loss-sharing agreement would be created between the bank and the FDIC. For example, in selling, say U.S.$100m of assets of a failed bank to a sound bank, the FDIC might agree to bear the first 80% of any asset value loss with any remaining loss in value, if it occurred, being borne by the asset purchasing bank. By 2009, as the crisis had begun to wind down, it was estimated that the FDIC had agreed to assume a loss-share on some U.S.$80b of assets with a potential loss share of U.S.$14b.4 However, as with the RTCs equity partnership arrangements, the FDIC has benefitted from recently rising house and land values that have reduced its loss exposure. It should be noted, however, that at the time, in 2009, the potential loss exposure from the asset loss-sharing program exceeded the funds in the FDICs reserves, implying that there was also a contingent risk-sharing between the FDIC and the U.S. government as a back-up supplier of funds to the FDIC.

3. Implications for European banking union

3.1. Supervision

All three “legs” of any potential banking union among Eurozone, and possibly non-Eurozone, countries are in their very early stages. The most advanced is the issue of common bank supervision.5 While it is generally agreed that the ECB will be the common supervisor, it still remains unclear whether it will regulate only the largest bank or all banks. If it chooses to supervise only the largest banks it will slip into a “dual banking system” similar to the U.S., with smaller banks remaining the supervisory province of local country regulators with diverse standards and policies. Of course, a natural question is where should the line be drawn between large and small or ECB versus national supervisors?

Either choice involves costs and problems. If the ECB attempts to supervise even the smallest banks it will either have to rely on the national supervisors to produce “soft information,” such as the bank’s management quality, which is often central to survival of small banks, or else assemble a very large multinational, linguistically skilled and mobile examination body very quickly to carry out on-site, as well as off-site, supervision. Moreover, if it relies on local supervisors applying the same methodology to evaluate the safety and soundness of a bank – such as the CAMELS system in the U.S. – it will require considerable training of local supervisors so as to avoid the problem faced in the U.S. of uneven rankings among national and state bank examiners.

What is clear from the above is that implementation is likely to be slow and extremely difficult. It seems likely and logical that in the initial stages the ECB will only regulate the largest banks, leaving smaller banks to national supervisors. To my mind, this makes sense if the large bank group contains systemically important banks, whose failure would impose a contagion effect on other Eurozone banks. The non-systemically important banks could initially remain under the supervisory control of the local regulator. While some have argued for a definite timetable for


implementation, it is difficult to believe that such a timetable could be adhered to in practice.

A further issue that seems to have been overlooked is regulatory equality between ECB supervision of the 17 Eurozone countries and supervision in the 10 non-Eurozone countries. This may well lead to another dual supervisory system, between Eurozone and non-Eurozone countries. It is not farfetched to imagine a scenario of a so-called “race to the bottom,” in which a non-Eurozone country puts in place a less stringent supervisory system in an effort to induce banks to move or charter a new bank in their country. In the discussion so far, it seems the only barrier to such a risk is the inducement to establish a common college of supervisors to enter into voluntary cooperation with the ECB.

3.2 Privately funded euro deposit insurance system
Currently, a widely diverse system of deposit insurance mechanisms exists among the European Union countries. What are the issues and barriers inhibiting the creation of a unified privately funded (by banks) deposit insurance system? The three main issues are (i) deposit coverage, (ii) insurance premium pricing and (iii) government backing.

As discussed earlier, the maximum coverage level for so-called insured depositors in the U.S. is U.S.$250,000, which was raised from U.S.$100,000 during the recent financial crisis. Determining what is the appropriate upper limit of coverage for a small relatively unsophisticated saver is going to be far from easy at an economic level, simply because of the differences in income, wealth and financial sophistication across the Euro countries. In the end, coverage is likely to be as much a political as an economic decision.

The second issue is premium pricing. As noted in the section 2 discussion regarding the FDIC, no privately funded deposit insurance scheme can withstand a systemic shock unless very large special assessments for such risks are built into the premiums. At a minimum, the premiums should create a fund of sufficient size to meet a “normal” level of idiosyncratic bank failures. Since 1994, this has been done in the U.S. by setting premiums as a function of bank risk, initially measured by a bank’s CAMELS supervisory score, and its capital ratio. Over time, this has evolved into a multi-dimensional credit-scoring model containing a large number of bank risk characteristics. It seems only reasonable that in its earliest stages, any Euro-union (or Eurozone) deposit insurance system should be built on a very simple set of risk characteristics, perhaps mimicking the initial supervisory score and capital ratio approach used in the U.S.. A further issue will be whether the deposit insurance system would cover just the Eurozone banks or be extended to all banks in the European Union.

The third issue is the deposit insurance fund’s solvency in a crisis. As discussed earlier, the FDIC has occasionally, in periods of systemic risk, had to draw on its line of credit with the U.S. government. In advocating a Eurozone privately funded deposit insurance system, the relevant governments will have to stand ready to provide support funds to keep the insurance system solvent. The problem here is that there will be 17 (and possible 27) governments involved rather than one, and a crucial issue will be the sharing rule adopted by governments in providing, or standing ready to provide, such funds. Again, such a sharing rule may well end up being a political rather than economic decision.

3.3 Restructuring bailout and resolution mechanisms
With respect to a common policy on restructuring, there seems to be a view developing that such funds might be provided by the existing European Stability Mechanism (ESM) to recapitalize insolvent banks in a similar manner to how the U.S. government used its TARP program to recapitalize some 600 U.S. banks. Currently, there seems to be some opposition to this (e.g., by Germany) based on the belief that ESM funds should only be used to help sovereign countries and not privately held banks. This seems to be shortsighted; for once funds are lent by the ESM to a particular country it is not obvious how that country can be prevented from recapitalizing its own banks. Indeed, it might be argued that a direct ESM bank capital investment mechanism might be preferable to an indirect (via sovereign country) mechanism.

If the ESM was empowered to recapitalize individual Eurozone country banks it will be faced with similar dilemmas as the U.S. government during the recent crisis, namely (i) which banks are worth recapitalizing rather than allowing them to close, (ii) should the funds made available to any bank be limited, e.g., to a certain percent of a bank’s risk-assets, and (iii) what should be the form of recapitalization, i.e., common stock, preferred stock
or subordinated debt, each has a different trade-off in terms of governance power versus loss risk.

The question of who should close a bank and using what criteria is also a thorny issue. It is being proposed that the ECB will have both licensing and closure powers. However, unlike the U.S., which has common bankruptcy rules regarding claimholder preference (it is viewed as being debtor friendly), the ECB will face 17 different countries with 17 different bankruptcy laws, varying from the creditor to the debtor friendly. Any closure of a bank by the ECB may be faced by legal challenges by either creditors or debtors until the European Union develops a common bankruptcy law, defining the rights and priorities of claimholders should the ECB need to close a bank.

Finally, if banks are being closed during a systemic crisis should a Eurozone government-backed resolution fund be established? The RFC, RTC and FDIC loss sharing — all provide possible models for such a Euro-fund, with the question of which one to pick largely being political.

Interestingly, despite these apparent difficulties, European Union finance ministers in June 2013 agreed on a novel “bail in” rescue plan for troubled Euro-banks. The novelty of the proposal is that the shareholders and creditors will take a first loss before ESM funds, or other special state funds, are used. Basically, this initial proposal is that the shareholders and creditors take the first 8% loss, i.e., 8% of liabilities are wiped out, closing the negative net worth gap between liabilities and assets. Only after that will ESM or other state money be used. Further, depositors are fully protected up to €100,000.

This proposal will have to go through numerous stages of legislative approval, including the European parliament and local country legislatures, but it does seem to be a rational alternative to a bail-out policy that merely increases instability and systemic risk in the banking system.

In addition, in November 2013 the ECB produced a more formal proposal for a separate “Single Bank Resolution Fund.” Behind the fund is a single resolution mechanism that complements the single supervisory mechanism. The fund would be permanent and established from ex-ante risk-based contributions of Eurozone banks. Operationally, the ECB as supervisor would carry out three levels of tests on banks: (i) a solvency test, (ii) an asset quality test and (iii) a stress test. If they fail these tests, they will be passed to the separate resolution authority in charge of the fund to carry out the appropriate restructuring.

4. Conclusion
The regulatory system has taken 150 years to develop in the U.S. Even today, it is far from unified with four supervisory agencies overseeing banking organizations, i.e., 50 state regulatory bodies, the OCC, the FDIC and the Federal Reserve, with a wide variety of individual and often-overlapping powers, such as in examination and supervision. The Eurozone is at the very start of what will be a long and difficult process; and whether unification is achievable across such a diverse set of economies with competing local supervisors and diverse bankruptcy laws remains to be seen. Nevertheless, the U.S. is a useful laboratory to examine the benefits and costs of different approaches to the three “legs” of European bank union, i.e., supervision, deposit insurance and restructuring/resolution.

---

Firm structure in banking and finance: is broader better?

Markus Schmid
Professor of Corporate Finance, Swiss Institute of Banking and Finance, University of St. Gallen

Ingo Walter
Seymour Milstein Professor of Finance, Corporate Governance and Ethics, Stern School of Business, New York University

Abstract
A focal point in strategic and regulatory debates about the pros and cons of diversification among financial institutions continues to be the issue of economies of scope in financial intermediation. In this paper, we summarize the theoretical research on the value of diversification in financial services firms, and survey the empirical research so far on the conglomerate discount in U.S. and international financial services businesses. We also review research on the internal capital market efficiency of universal banks and financial conglomerates. The paper provides new empirical evidence on the conglomerate discount in U.S. financial intermediaries and how that changes between non-crisis and crisis periods, showing a decline in the discount under turbulent conditions.
1. Introduction
As the structural transformation of national financial systems and the global financial architecture continues, three questions are often raised during the debates: is it better for financial services firms to be (1) bigger, (2) geographically expansive and (3) functionally diverse? These questions have implications for the strategic positioning of banks and financial firms in optimizing revenues, risks and costs. They are also of interest to regulators and the general public in terms of the trade-offs that exist between efficiency, innovation and stability in setting the rules under which financial services firms are required to operate.

These debates have informed a rich array of both conceptual and empirical studies that address the underlying issues and attempt to confirm or refute the central hypotheses in financial intermediation. This paper addresses the third of the aforementioned issues – whether functional diversity in financial intermediation creates or destroys value among financial intermediaries. This literature is less well developed than work on the size and scale of financial firms, but no less important in driving the strategic and regulatory debates.

In the U.S., the Gramm-Leach-Bliley Act of 1999 (GLB) removed many of the restrictions on financial services firms. Well before the GLB legislation, U.S. banks were already shifting away from lending activities toward a broader array of financial services, in part through separately capitalized subsidiaries of bank holding companies (BHCs). But the Act opened the way for full financial integration by explicitly allowing U.S. BHCs to engage in brokerage, advisory and underwriting activities, among others. In Europe, the scope for functional diversification of financial intermediaries had been deregulated earlier and more completely. The E.U. Second Banking Coordination Directive of 1989 allowed European banks, as universal credit institutions, to diversify directly (without BHCs) into various activity areas such as commercial banking, investment banking and insurance. In terms of strategy, financial services firms can change their degree of diversification either by divesting or acquiring assets, or by redirecting their activities into new business segments. And a firm’s portfolio of activities can shift over time due to divergent growth rates in the existing business segments.

Various arguments have been made in favor of diversification as a value-enhancing strategy for financial services firms. The most important are so-called “synergies” resulting from cost and revenue economies of scope. Other potential benefits are lower tax burdens as a result of tax-efficient intra-firm transactions and more efficient internal, as compared to external, capital markets resulting from a better coordination across highly specialized activity lines, better monitoring and control of capital expenditures, and sharing of managerial best-practices. Reduced bankruptcy risk due to less than perfectly correlated revenue streams across functional segments may likewise result in improved debt ratings, higher debt capacity, lower taxes and lower weighted average cost of capital (WACC) of diversified financial firms resulting in higher share prices as compared to more specialized financial firms.

Separately, too-big-to-fail backstops provided by the public at zero or below-market cost through the central bank or public guarantee agency may support the creditworthiness of the banking unit of a financial conglomerate, and by extension the entire financial firm. Depending on the backstop beneficiaries, extending from retail depositors to subordinated debtholders and even shareholders, explicit or implicit public guarantees can benefit a financial firm’s WACC and mask the impact of scale and scope economies and diseconomies, creating serious competitive distortions in the process.

Arguments against diversification in financial services firms include a cross-subsidization among business lines, which may result in inefficient capital allocations and reduced performance incentives in the profitable business segments. Diversification may also lead to overinvestment in low-NPV projects due to excess free cash flow and unused borrowing capacity, as well as non-materiality of individual capital allocation errors in relation to the firms’ overall market value. In financial conglomerates, conflicts of interest among clients and activity-areas may create incremental exposure to reputation risk resulting in higher debt costs and lower share prices.

Whether the benefits of diversification and financial conglomerate outweigh specialization is a key issue in defining the strategies of financial intermediaries and the evolving

---

1 For a detailed review of arguments for and against diversification in financial services firms, see Walter (2004), Chapter 3.
architecture of national and global financial systems. There already exists a large body of literature on the valuation effects of corporate diversification in non-financial firms. However, the literature on the valuation effects of diversification in financial intermediaries has been much more limited, but has grown substantially in recent years.

In section 2, we summarize the findings of the existing literature on the valuation effects of diversification in financial intermediaries. In section 3, we extend the empirical analysis in Schmid and Walter (2009) by another eight years to investigate the value of diversification in a consistent empirical setup based on a data panel covering nearly 30 years. The empirical findings confirm a conglomerate discount in financial services firms for the sample period 1985–2012. However, when compared to previous studies by Laeven and Levine (2007) and in particular Schmid and Walter (2009), the average discount, while still statistically and economically significant, is somewhat reduced. We show that this reduction in the average discount appears to result from increased diversification benefits in the crisis years 2006–08 when the diversification discount turned insignificant. Hence, as could be expected, the risk-reducing effect of diversification seems to be particularly valuable during a crisis.

2. Theoretical research
The focus of this article is on empirical research. Nevertheless, we begin by briefly summarizing the findings of the theoretical literature on the value implications of functional diversification in financial intermediaries.

Boot and Schmeits (2000) show that, in a competitive environment with modest returns to capital, diversification is valuable as it helps to protect these returns and reduces the risk-taking incentives resulting from limited liability. They show that the introduction of internal cost-of-capital allocation schemes creates internal market discipline that complements the weak external market of conglomerate firms. They conclude that, under certain circumstances diversification may very well create value and argue that this is particularly likely in the banking industry.

Wagner (2010) develops a model that shows that even though diversification reduces financial institutions’ individual probability of failure, by exposing them to the same risks, and hence making them more similar to each other, it increases the risk of systemic crises. When such crises induce costs exceeding those emanating from individual bank failures, efficient diversification encompasses a trade-off between lower overall probabilities of banking failures and a higher probability of systemic failures. These arguments can be extended beyond diversification and also apply to interbank liquidity insurance and bank mergers. The two main conclusions are that full diversification is not optimal and that capital requirements for banks with more diversified portfolios should be higher rather than lower.

Ibragimov et al. (2011) develop a model in which negative externalities arise because diversification decisions at financial intermediaries may be optimal for the individual firms by reducing the risk of individual failure, but at the same time they may be suboptimal for society as a whole due to a higher risk of systemic failure. The results depend on the distributions of risks that the financial intermediaries assume. When these risks are thin-tailed, diversification is optimal for both the individual firms and society. When these risks are moderately heavy-tailed, diversification is optimal for the individual firms but suboptimal for society.

Gatzert and Schmeiser (2011) provide a conceptual framework in which the conglomerate discount of financial intermediaries is quantified by employing an option-based approach, with potential diversification benefits being calculated using the tail value-at-risk and shortfall probability. They show that under competitive conditions, the potential benefits from diversification are small.

Freixas et al. (2007) investigate in a theoretical model framework the risk-taking incentives of financial conglomerates that combine banking and non-banking units. They show that the conglomerate’s liability structure determines the level of market discipline faced by the firm, which in turn determines the risk-taking incentives of the conglomerate. Their paper shows that conglomerates extend the reach of the deposit insurance safety

---

2 See Martin and Sayrak (2003) and Erdorf et al. (2013) for literature reviews.

3 Acharya and Yorulmazer (2007, 2008) show that banks may choose to increase the likelihood of joint failures to induce regulators to bail them out.
net to their non-banking divisions, resulting in higher risk-taking, which in turn may outweigh any diversification benefits. Their model also makes predictions about the relative valuation of financial conglomerates versus stand-alone firms — because financial conglomerates have less access to the deposit insurance fund, they should trade at a discount relative to stand-alone financial institutions.

2.1 Empirical research on U.S. data
The U.S. has been the focus of most empirical studies because of its large market size, the number and variety of financial institutions, a single currency and a single accounting and regulatory regime.

One way to investigate the value of functional diversification of banks is to perform an event study analysis of merger announcements distinguishing between diversifying and focusing mergers. In fact, earlier empirical research was mainly interested in analyzing the potential cost savings and revenue gains from bank mergers. For example, Houston and Ryngaert (1994) show that the overall gains from 153 bank mergers announced between 1985 and 1991 are positive but statistically insignificant. Moreover, they find that greater operations overlap is associated with higher merger gains.

Houston et al. (1999) find that in-market (i.e., concentrating) mergers tend to create value upon announcement. Specifically, they find that target firms gain and acquirers do not lose in in-market mergers. They argue that such in-market merger transactions are expected to cut costs faster and more dramatically than market-extending mergers.

DeLong (2001) finds that the markets react positively to bank mergers that focus activities (and geography) while diversifying mergers do not create value. In contrast to most other studies, DeLong (2001) uses cluster analysis of stock returns to determine the activity focus of banks. Bidders in diversifying mergers destroy value on a statistically significant scale, bidders in non-diversifying mergers do not destroy any value, and targets earn approximately the same returns in both groups.

Stiroh (2004a) analyzes the potential diversification benefits resulting from banks shifting from interest income to fee income, trading revenue, and other types of non-interest income. The study finds little evidence of diversification benefits resulting from a shift toward non-interest income. Most importantly, a greater reliance on non-interest income, and in particular on trading revenue, is associated with higher risk and lower risk-adjusted profits.

DeYoung and Roland (2001) investigate the relationship between bank profitability, volatility, and the share of different revenue sources resulting from different activities of large commercial banks. Consistent with Stiroh (2004a), the authors find that a larger fraction of fee-based revenue increases the volatility of both bank revenues and bank earnings.

Stiroh (2004b) investigates the relationship between activity diversification and risk-adjusted performance of small community banks in the U.S. The study distinguishes between diversification of broad activity classes, such as shifting between interest-generating activities, fiduciary operations and, trading revenue and diversification within these broad activity classes. The results show that diversification across broad activity classes is not associated with benefits, but that diversification within the lending and non-interest activities is associated with a higher risk-adjusted performance.

A further stream of research investigates the stock market reaction to announcements of regulatory changes affecting banks’ possibilities to diversify. This work suggests that the market reacted favorably to the announcement of the passage of the U.S. GLB legislation in 1999, which opened the way for integrated multi-functional financial intermediaries. For example, Lown et al. (2000) find positive returns associated with both commercial and investment banks’ stocks upon announcement by President Clinton on 22 October 1999 that the passage of GLB was imminent. Consistent with this work, Yu (2003) reports positive returns to the announcement of GLB for large securities firms, large insurance companies, and bank holding companies already active in some securities businesses i.e., firms with so-called “Section 20 subsidiaries” allowing limited investment banking activities. The market seems to have expected gains from functional diversification of financial services firms, or at least increased opportunities to diversify and possibly benefit

4 In contrast, they show that in financial holding company conglomerates capital arbitrage may increase market discipline and raise welfare.
from cross-product synergies, extensions of “too-big-to-fail” guarantees, or other facets of diversification.

Stiroh and Rumble (2006) use a sample of roughly 1,800 U.S. FHCs from 1997 to 2002 to investigate whether U.S. financial holding companies benefit from diversification into non-interest income activities. The results based on a panel regression framework with firm fixed effects – which allows for quantifying the impact of changes in diversification on risk-adjusted profits – suggest no value impact of diversification, but instead show a negative impact of the ratio of non-interest income to net operation income on risk-adjusted profits of FHCs.

Schmid and Walter (2009) investigate the valuation effects of diversification across the entire range of financial intermediation functions – commercial banking, investment banking, insurance, asset management, and financial infrastructure services (clearance, settlement, payments, custody, etc.) – in a sample of roughly 4,000 U.S. bank-years on 660 banks from 1985 to 2004. The results from fixed effects regressions show a substantial and persistent conglomerate discount of between 9% and 15% among financial intermediaries. Financial intermediaries choose to diversify for unknown reasons. Hence, diversification is an endogenous choice variable [Campa and Kedia (2002)]. In addressing these issues econometrically, Schmid and Walter (2009) confirm the existence of a conglomerate discount and suggest that it is diversification that causes the discount, not the fact that troubled firms diversify into other more promising activity areas. The study also investigates whether the conglomerate discount depends on the firms’ main activity-area or on the specific financial activity-areas that are combined within the conglomerate. It finds a significant conglomerate discount in all three main activity-areas (credit intermediation, securities and insurance) but no conglomerate discount associated with investment banking. Moreover, there is little difference in the conglomerate discount across different combinations of financial activity-areas with two notable exceptions: combinations of commercial banking and insurance segments and combinations between commercial banking and investment banking segments are associated with a significant premium.

2.2 Empirical research on international data
Laeven and Levine (2007) use a sample of 836 banks from 43 countries over the time period 1998-2002 to investigate whether banks that engage in multiple activities are traded at lower prices than specialized banks. As a measure of bank value, the authors use a modified version of the Lang and Stulz (1994) “chop-shop” Tobin’s Q (market value to book value ratio), which controls for general valuation differences across financial services industries. They differentiate focused banks by their interest income versus non-interest income and by loans versus other earning assets. While this is arguably a crude measure, it may also help to overcome the problem of using self-reported segment data. The results show a statistically significant and economically large conglomerate discount which survives controlling for the endogenous self-selection of banks and a battery of other robustness tests.

Baele et al. (2007) analyze whether there is a diversification discount in a sample of 255 banks from 17 European countries over the time period 1989-2004. Bank value is measured by an adjusted Tobin’s Q ratio, and the authors find a positive relationship between bank value and the degree of functional diversification, arguing that the stock market seems to anticipate higher bank profits resulting from diversification. The authors also investigate the relationship between bank diversification and both idiosyncratic and systematic bank risk. Their results suggest a nonlinear relationship between diversification and idiosyncratic risk. Most of the banks in the sample are able to reduce idiosyncratic risk by diversifying. Consistent with Stiroh (2006), the study finds that a higher share of non-interest income in total income is associated with higher systematic risk. This has conflicting implications for different stakeholders – large shareholders or bank managers are mainly interested in idiosyncratic risks, while diversified investors mainly care about systematic risk.

Van Lelyveld and Knot (2009) investigate whether diversification activities combining banking and insurance create or destroy shareholder value in a sample 135 financial firms in the European Union. Specifically, they compare the excess value measure proposed by Berger and Ofek (1995) across 45 focused banks, 45 focused insurance companies, and 45 diversified financial intermediaries active in both commercial banking and insurance provision. However, due to some filters in the data management process and data availability only 25 firms (with 132 observations over the time period 1990-2005) end up in their multivariate tests. Their results suggest no consistently
significant conglomerate discount resulting from combinations of banking and insurance activities, which is consistent with the results in Schmid and Walter (2009). However, they show that there is quite some variation in cross-section analysis and that a diversification discount may apply depending on firm size, complexity and risk.

Elsas et al. (2010) investigate the relationship between revenue diversification and bank value in a sample of 3,348 bank-year observations on 380 listed banks from nine countries during 1996-2008. The nine countries covered by the study are Australia, Canada, France, Germany, Italy, Spain, Switzerland, the U.K. and the U.S. and the sample covers only large banks with total assets exceeding U.S.$1b in at least one sample year. The authors find a positive relationship between revenue diversification and bank profitability resulting from higher margins from non-interest activities and lower cost-to-income ratios. This higher profitability also seems to translate into higher market values of diversified banks.

Mercieca et al. (2007) investigate whether diversification of small European banks into non-interest income activities improves performance. Their sample includes 755 small banks, i.e., banks with an asset size below €450m as of year-end 2000, from the 15 European Union member countries before the 2004 E.U. enlargement over the period 1997-2003. As the sample includes both listed and unlisted banks, bank performance is measured based on accounting and not market measures (ROA, ROE). The authors find no evidence of diversification benefits within and across business lines and a negative valuation effect associated with higher shares of non-interest income in total income. They conclude that — in an increasingly competitive environment due, for example, to the Financial Sector Action Plan, the Takeover Directive of the E.U. and Basel II — small European banks diversify into activity areas where they lack expertise and experience, resulting in value losses.

Berger et al. (2010) use a sample of 464 bank-year observations on 88 Chinese banks over the period 1996-2006 to investigate whether there is a diversification discount in Chinese financial firms. The results suggest that four different dimensions of diversification — loans, deposits, assets and geography — are all associated with reduced profits, higher costs and a significant valuation discount. Moreover, they find that foreign ownership and/or conglomerate affiliation is associated with a smaller diversification discount in Chinese banks. The authors argue that these findings are consistent with better monitoring and access to more and better networks, and partnerships help Chinese banks to overcome inefficiencies associated with diversification of bank activities.

2.3 Bank diversification and internal capital markets

Internal capital markets within financial conglomerates can either serve shareholders’ interests by mitigating external capital constraints if run efficiently, or destroy company value if resulting in an inefficient cross-subsidization among business lines and as a consequence reduced performance incentives in the profitable business segments. Available evidence on non-financial conglomerates suggests that internal capital markets result in inefficient cross-subsidization and, on average, destroy shareholder value [e.g., Scharfstein (1998), Rajan (2000)].

Houston et al. (1997) investigate the existence of internal capital markets in bank-holding companies by analyzing the cash flow sensitivity of loan growth, and the correlation between different subsidiaries’ loan growth, within a bank holding company. They find that the loan growth of subsidiaries is more sensitive to changes in the holding company’s cash flows and capital position than the subsidiary’s own cash flow and capital position. They also find that loan growth is negatively correlated across the subsidiaries within a bank holding company. In short, their findings confirm the existence of internal capital markets in bank holding companies.

Campello (2002) investigates whether internal capital markets help banks alleviate the impact of external financial constraints on investments, and whether they promote investment efficiency. He compares the responses of small subsidiary and independent banks to monetary policy changes. The results show that internal capital markets in financial conglomerates relax credit constraints

---

5 This finding is also consistent with Cubo-Ottone and Murgia (2000), who find positive and significant merger announcement returns associated with European banks diversifying into the insurance industry.

6 The reason for focusing on small rather than large banks is that large banks can offset Fed policies on the margin because of their better access to non-reservable forms of deposits at relatively low costs (e.g., see Kashyap and Stein (2000)).
faced by stand-alone banks and lessen the impact of a tight monetary policy on bank lending activity. The study also finds that frictions between conglomerate headquarters and external capital markets result in a less efficient investment allocation process in small bank holding companies but not in large ones.

Klein and Saidenberg (2010) find evidence consistent with a conglomerate discount in bank holding companies. Specifically, bank holding companies with more subsidiaries are found to be less profitable and show lower Tobin's Q ratios than bank holding companies with fewer subsidiaries. The authors also attempt to separate the effects of organizational structure from those of diversification, and find that neglecting the effect of organizational structure results in an inflated conglomerate discount. Finally, they show that the benefits of an internal capital market can be realized best within firms with fewer independently chartered sub-units due to significant costs associated with managing complex, multi-unit organizations.

3. New empirical analysis

In this section, we extend the sample used in our earlier work – Schmid and Walter (2009) – by eight additional years (2005–12), resulting in a 28-year sample period extending from 1985 to 2012. The sample consists of all financial firms from 1985 to 2012, with data reported on both the Compustat segment and industrial annual data files and total assets of at least U.S.$100m, and we apply the same filters as in our earlier study. The dataset covers the broadly-defined U.S. financial services sector, including commercial banks and bank holding companies, insurance companies, asset managers and broker-dealers, but excludes investment companies as well as foreign firms listed as American Depository Receipts (ADRs).

To examine whether diversification increases or decreases firm value, we use the excess value measure proposed by Berger and Ofek (1995) and used in many other studies, including ours. This excess value measure compares a firm’s value to its imputed value if its segments were operating as standalone entities. Each segment of a diversified firm is valued based on the median sales (assets) multipliers for single-segment firms in that industry. The imputed firm value for each segment is then calculated by

<table>
<thead>
<tr>
<th>Year</th>
<th>Observations</th>
<th>Focused (%)</th>
<th>Diversified (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>47</td>
<td>46.81</td>
<td>53.19</td>
</tr>
<tr>
<td>1986</td>
<td>141</td>
<td>53.90</td>
<td>46.10</td>
</tr>
<tr>
<td>1987</td>
<td>152</td>
<td>52.63</td>
<td>47.37</td>
</tr>
<tr>
<td>1988</td>
<td>150</td>
<td>51.33</td>
<td>48.67</td>
</tr>
<tr>
<td>1989</td>
<td>144</td>
<td>52.08</td>
<td>47.92</td>
</tr>
<tr>
<td>1990</td>
<td>130</td>
<td>54.62</td>
<td>45.38</td>
</tr>
<tr>
<td>1991</td>
<td>151</td>
<td>61.59</td>
<td>38.41</td>
</tr>
<tr>
<td>1992</td>
<td>178</td>
<td>63.48</td>
<td>36.52</td>
</tr>
<tr>
<td>1993</td>
<td>239</td>
<td>67.36</td>
<td>32.64</td>
</tr>
<tr>
<td>1994</td>
<td>270</td>
<td>67.04</td>
<td>32.96</td>
</tr>
<tr>
<td>1995</td>
<td>293</td>
<td>69.97</td>
<td>30.03</td>
</tr>
<tr>
<td>1996</td>
<td>317</td>
<td>72.56</td>
<td>27.44</td>
</tr>
<tr>
<td>1997</td>
<td>307</td>
<td>69.38</td>
<td>30.62</td>
</tr>
<tr>
<td>1998</td>
<td>263</td>
<td>54.75</td>
<td>45.25</td>
</tr>
<tr>
<td>1999</td>
<td>223</td>
<td>54.26</td>
<td>45.74</td>
</tr>
<tr>
<td>2000</td>
<td>209</td>
<td>53.59</td>
<td>46.41</td>
</tr>
<tr>
<td>2001</td>
<td>207</td>
<td>53.62</td>
<td>46.38</td>
</tr>
<tr>
<td>2002</td>
<td>200</td>
<td>52.00</td>
<td>48.00</td>
</tr>
<tr>
<td>2003</td>
<td>218</td>
<td>52.75</td>
<td>47.25</td>
</tr>
<tr>
<td>2004</td>
<td>221</td>
<td>51.13</td>
<td>48.87</td>
</tr>
<tr>
<td>2005</td>
<td>229</td>
<td>41.48</td>
<td>58.52</td>
</tr>
<tr>
<td>2006</td>
<td>241</td>
<td>44.40</td>
<td>55.60</td>
</tr>
<tr>
<td>2007</td>
<td>226</td>
<td>49.56</td>
<td>50.44</td>
</tr>
<tr>
<td>2008</td>
<td>200</td>
<td>52.50</td>
<td>47.50</td>
</tr>
<tr>
<td>2009</td>
<td>210</td>
<td>52.86</td>
<td>47.14</td>
</tr>
<tr>
<td>2010</td>
<td>240</td>
<td>56.67</td>
<td>43.33</td>
</tr>
<tr>
<td>2011</td>
<td>252</td>
<td>54.76</td>
<td>45.24</td>
</tr>
<tr>
<td>2012</td>
<td>262</td>
<td>54.58</td>
<td>45.42</td>
</tr>
<tr>
<td>Mean</td>
<td>211.43</td>
<td>55.77</td>
<td>44.23</td>
</tr>
</tbody>
</table>

Table 1: Sample overview by calendar year

The table reports the number of total observations, the percentage of focused firms, and the percentage of diversified firms for each sample calendar year. The sample includes all financial firms (NAICS 520000–529999, excluding 525900–525990) covered by the Compustat Annual and Compustat Segments databases. All firm-years for which more than one segment (with differing NAICS codes) is reported are classified as diversified.

7 Compustat defines sales for financial companies as total current operating revenue plus net pretax profit or loss on securities sold or redeemed minus non-recurring income.
multiplying the segment's sales (assets) by the median ratio of the market value to sales (assets) for single-segment firms in the same industry. The industry median ratios are based on the narrowest North American Industry Classification System (NAICS) grouping that includes at least five single-segment firms with complete data and total assets of at least U.S.$100m. The imputed value of the complete firm is calculated as the sum of the imputed segment values. The imputed value provides an estimate of the firm if all of its segments were operated as stand-alone entities, assuming that in this case the segment's valuation would correspond to that of the median focused firm in the same industry. Excess value is then calculated as the log of the ratio of a firm's value to its imputed value. A negative excess value indicates that a firm trades at a discount and a positive excess value implies that the firm trades at a premium as compared to the median focused firm in the same industries.

We use alternative measures of diversification. The first is a dummy variable, which is equal to one if a firm reports more than one segment in Compustat, Number of segments is the number of reported segments, Herfindahl (sales) and Herfindahl (assets) are a sales- and asset-based Herfindahl–Hirschman index computed as the sum of the squares of each segment's sales (assets) as a proportion of total sales (assets) for the firm, Ln(total assets) is the natural logarithm of total assets, Leverage is book leverage, and ROA the return on assets. All regressions include firm fixed effects. The standard errors (in parentheses) are based on the cluster-robust variant of the Huber–White sandwich estimator, which clusters at the firm level. a, b and c denote statistical significance at the 1%, 5% and 10% levels.
From 1998 to 2012, there is no clear trend in diversification with the percentage of diversified years fluctuating between 43.3% (2010) and 58.5% (2005).

Table 2 reports the results from panel regressions of the excess value measure based on sales (Columns 1–3) and assets (Columns 4–6) on alternative measures of diversification and controls. All regressions include firm fixed effects to control for unobserved firm characteristics that are constant over time. To account for within-correlation between observations on the same firm, the standard errors are clustered at the firm level. Consistent with our earlier findings in Schmid and Walter (2009), the results indicate a significant diversification discount in all six regression specifications. Based on the dummy variables in Columns 1 and 4, diversified firms trade at a discount of between 7.8% and 9.0% as compared to a portfolio of focused firms. This diversification discount is still highly significant but somewhat smaller than the discount reported for the years 1985–2004.

In Table 3, we re-estimate the regression specification in Column 1 of Table 2 for each sample year from 2005 to 2012 separately to investigate whether the value (or cost) of diversification changed over the more recent sample years. The results show that the diversification discount is statistically and economically significant in the years 2009–12 but insignificant in the years 2005–08. While the coefficient on the Diversified dummy is still sizable in 2005, indicating a diversification discount of 11.0%, the estimate of the discount is much smaller and insignificant at the onset and throughout the recent financial crisis.

This finding suggests that, as expected, the risk-reducing effect of diversification is particularly valuable during the crisis years and outweighs the negative effects of diversification. Our results are consistent with Kuppuswamy and Villalonga (2012), showing that the value of diversified non-financial firms relative to focused firms increased in the recent financial crisis due to a debt coinsurance feature of conglomerates and more efficient internal capital markets.9

8 In 1998, SFAS 131 superseded SFAS 14 in the regulation of segment reporting and SIC was replaced by NAICS.14 One reason for this regulatory change was a perceived underreporting of segments. Consistently, Berger and Hann (2003) show that the introduction of SFAS 131 has resulted in firms reporting a greater number of segments.

9 It is possible that subsidies in the form of implied public guarantees are responsible for part of the reduction in the discount between the pre-crisis and crisis periods, although we control for firm size in the regressions. However, a linear size control may not be sufficient to pick up this effect.
To summarize, the new empirical results reported here confirm the finding of a conglomerate discount in financial services firms over an extended time period covering almost 30 years. However, the average discount reported is somewhat smaller than in previous studies [e.g., Laeven and Levine (2007), Schmid and Walter (2009)]. This reduction in the average discount seems to emanate from increased diversification benefits in the crisis years when the diversification discount turned insignificant. Hence, similar to effective risk management in financial firms, corporate diversification may provide benefits in a crisis environment but at a cost in a good market environment [e.g., Aebi et al. (2012), Ellul and Yerramilli (2013)].

4. Conclusion
Scope and scale lie at the heart of the debate on the static and dynamic efficiency of the financial architecture. Management of banking and financial services firms are mandated to take both issues into consideration, when considering strategic positioning of their enterprises in the competitive playing field of today’s financial markets. Much of the research on scale economies in banking suggests that scale effects on firm-wide cost functions dissipate at surprisingly small-size levels. However, scale benefits differ widely by function – global custody is heavily scale-driven while in corporate advisory services scale economies are virtually immaterial, for example. Much less research has focused on cost and revenue economies of scope, in part because the natural source of datasets (with large numbers and types of financial firms within a single market with one currency and one tax and regulatory environment), did not exist for a country, such as the U.S. because it had a history of separation between banking and non-banking financial intermediation, from 1933 to 1999. The fact that empirical studies were hindered by the need to adjust for the many other factors besides scope that affect firm value metrics also played a part in the lack of literature.

In this paper, we provide an up-to-date survey of scope-related theoretical and empirical research, which offer a disparate array of results distinguished by the absence of consistently definitive conclusions. The exception is a predominant finding of a conglomerate discount in the financial services sector in most empirical studies – comparable to the findings in the nonfinancial sector – as well as the finding of greater stability associated with a broader scope of activities. This suggests that scope-effects generally favor debt-holders (and taxpayers if systemic backstops exist) relative to shareholders. This seems corroborated in our finding of a lower conglomerate discount during the recent period of financial turbulence relative to pre-crisis conditions.

References
Houston, J., and M. Rynhaaert, 1994, “The overall gains from large bank mergers,” Journal of Banking and Finance 18, 1155-1176
Stiroh, K. J., 2004a, “Diversification in banking: is noninterest income the answer?” Journal of Money, Credit, and Banking 36, 853-882
Stiroh, K. J., 2006, “A portfolio view of banking with interest and noninterest activities,” Journal of Money, Credit, and Banking 38, 1351-1361
Walter, I., 2004, Mergers and acquisitions in banking and finance, Oxford University Press, New York
Six years after the crisis

Roy C. Smith
Kenneth Langone Professor of Finance, Stern School of Business, New York University

Abstract
Since the great financial crisis of 2008, governments have set new records in market intervention, re-regulation of banking and financial markets, and in prosecution of banks for mismanagement. The new regulations touch just about everything in finance and completely change the financial regulatory system. However, these regulations have been hastily prepared and contain flaws that suggest they may not be effective in addressing systemic risk in the future. They also impose enormous compliance costs on the banking industry, even though no more than about 40 banks worldwide are systemically important. The costs will be passed on to customers of banks and users of financial markets, which may detract significantly from future economic growth prospects. Banks will have to adapt their business models to the new circumstances of their industry, or surely be challenged by shareholders until they do. But the way the regulatory platform is structured now, they will be adapting toward a more conservative, risk-averse form of what they were before. That may mean that too-big-to-fail becomes too-cautious-to-fail. If we have reduced the banks’ capability to finance basic economic growth, a role that is equally important to our societies as systemic risk reduction, the regulatory response to the crisis may prove to be very mistaken. Risk has to be financed somewhere, if we are to return to the growth rates that we need to maintain and improve living standards. Most likely, the regulatory and other measures taken in response to the crisis will have to be modified in the future, but probably not in the near term.
1. Introduction
The great, global financial crisis began in 2007 with the decline in U.S. home prices, peaked in the fourth quarter of 2008 after the failure of Lehman Brothers, the bailout of AIG and numerous large banks in the U.S. and Europe, and plunged the U.S. and much of the rest of the world into a steep recession and an extended period of low to no-growth in the world economy. The crisis lingered through 2009 and then spread to the sovereign debt markets of the Eurozone in 2010, culminating in a prolonged period of distress for the euro and its member countries.

As global financial crises go, this one was surely the worst since the 1930s, and arguably it was worse than even that one. A major difference between the two was in the actions taken by the U.S. Federal Reserve, which moved quickly and boldly to restore liquidity to shattered financial markets that had experienced “systemic” collapse.

The mortgage crisis will be remembered for its avalanche-like impact on financial markets, the rapid transmission of this impact to the global economy, and for the resulting mind-set of affected U.S. and European taxpayers (and politicians) who believed that they had been unfairly called upon to bail out rich bankers and other capitalists whose greedy and reckless behavior was the cause of the crisis.

2. A pair of bubbles
The crisis will also be remembered as one of a pair of similar financial “bubbles” that inflated and burst during the first decade of the twenty-first century, bringing to bear the idea of uncontrollable financial cycles of great destructive capacity being a consequence of the increasingly global commitment to free market economics. The earlier crisis (sometimes called the Tech Wreck) began in 2000 with the melt down of the NASDAQ stock market, extended through the bankruptcies of Enron and WorldCom in 2001-02 and ended after three consecutive years of declining stock prices and record-level bankruptcies in the U.S..

The effort to boost growth after the Tech Wreck by loosening the money supply and dropping interest rates to (then) record-low levels is now seen as having contributed to rapidly rising home prices and the formation of the second bubble – the mortgage-backed securities (MBS) bubble. From 2000 through 2007, the MBS market doubled in size to U.S.$5.5t – these high-rated, high-yielding securities had become the ideal investment for pension funds and other institutional investors in response to low interest rates and stock prices.

Demand for MBS quickly exceeded the supply of extant mortgages, though these were bundled into securitized packages as rapidly as banks could process them. Before long, a new supply was found in the subprime (lesser credit quality) sector, and these too were securitized at a record pace. When home prices leveled off, and began to decline for the first time in many years, some investors began to fear that the value of the collateral of subprime mortgages was threatened and withdrew from the sector. Still, how bad could it be? Subprimes represented perhaps as much as 20% of all MBS, and if they were devalued by 10% or 20% the resulting price-adjustment to the overall portfolio would be tolerably modest.

One of the unsolved mysteries of the MBS financial crisis is why the markets declined as much as they did, and spread so widely to other credit markets, in light of the initially relatively modest increase in the expected default risk from declining home prices. (Home prices fell more rapidly once the crisis had impacted the real economy). The answer seems to lie in the realm of psychological factors that dominate the field of behavioral finance, but it may simply be a case of following a crowd into something new only to depart in a rush once it became clear that the fashionable new thing was not only no longer fashionable, but actually dangerous.

3. Market size and market risk
After three decades of deregulation, globalization and market integration, and of enormously enhanced technological capacity, the global value of all marketable securities in 2007 was U.S.$206t (or 3.6 times global GDP), up from U.S.$56t in 1990 (2.6% times global GDP). Global markets had become enormous, interconnected and efficient. Publicly transparent market prices now rationally determined investment flows into assets that would optimize growth and productivity – something liberal

---

economists have dreamed of seeing on a global scale since the days of Adam Smith.

But there is a flaw in the market system. It is subject to volatility, which, in view of the increasing quantity of funds in the market, can unleash huge amounts into particular investment bubbles that can be reversed suddenly. When the bubbles are big enough, the impact of a sudden rush for the exits can also be huge.

In May 2008, the Bank of England published a Financial Stability Report showing that market losses since early 2007 on a portfolio of U.S.$900b par value of U.S. subprime securities were more than twice the amount estimated on the basis of projected delinquency rates (then estimated at around 19%; the actual default experience proved to be much lower). The study occurred after the rescue of Bear Stearns in March 2008 that was thought to have calmed the markets. Nevertheless, the difference between market values (58% of par) and historical cost adjusted for changes in expected default rates (81%), the Bank of England report explained, was largely due to “increased uncertainty about eventual credit losses, greater investor aversion to such uncertainty, [and] because investors require bigger discounts to invest in illiquid markets.”3

In other words, the fast-dropping market, fueled by anxious sellers and margin lenders, far overshot its estimated equilibrium level, as markets often do. Unfortunately, the overshooting extended further as the market digested the federal takeover of Fannie Mae and Freddie Mac in early September 2008, and the extreme market mayhem that followed the bankruptcy of Lehman Brothers a week later. As a result, mark-to-market write-offs by those holding the securities (including many highly leveraged large banks) were well beyond anything that might have been expected. Transcripts of Federal Reserve meetings in September 2008, published in February 2014, reveal that few, if any, of those present at the meetings anticipated the intensity of the market run that occurred after Lehman’s failure.4

Citigroup, which booked losses of U.S.$143b during the 2007–09 crisis, wrote off more than all of its shareholder capital, which had to be replaced with government assistance.5 Several other banks had to do the same. (Basel III would not have made much difference). Altogether, banks around the world, according to Bloomberg data, wrote off U.S.$1.8t of capital in the 2007–09 period. The top ten global capital market banks accounted for about a third of this amount.

But, would we have had a crisis at all if the price of subprime MBS had simply moved to their risk-adjusted equilibrium level i.e., 81%? Or, would the crisis have escalated into what it became if the Federal Reserve had treated Lehman as it had Bear Stearns? Would we have had to create Basel III, or Dodd Frank or the European Banking Authority if prices had stabilized at or close to their equilibrium levels?

But the markets did overshoot and, because so many banks were involved in assembling and warehousing MBS, the instant mark-to-market write-offs were enormous and the crisis ensued.

4. Market intervention
Governments intervened extensively to offset the effects of the crisis. The Federal Reserve recognized the liquidity collapse in the markets and acted quickly once it did. Its unprecedented direct intervention in various financial markets over an approximate six-month period after September 2008 involved the deployment of nearly U.S.$8t.6 These bold market-stabilizing actions were effective and necessary to prevent financial market liquidity from drying up altogether.

4.1 Massive and addictive
The U.S. Treasury also intervened in 2008 with U.S.$350b of Troubled Assets Relief Program (TARP) investments in securities of banks and automobile companies, and in 2009, President Obama signed into law the U.S.$787b American Recovery and Reinvestment Act, a Keynesian stimulus program. Subsequently, the Federal Reserve, in its various “quantitative easing” programs, invested a further U.S.$2t to U.S.$3t in securities purchases to help stimulate economic growth and increase

Six years after the crisis

employment. After the first U.S.$8t of emergency intervention by the Federal Reserve, it is not yet clear whether the other actions taken by the U.S. Government were effective or necessary in economic terms, but there seems to be little doubt that they were necessary in political terms.

In Europe, similar stabilization and stimulus efforts were undertaken, though on a smaller scale than in the U.S. In 2010, a U.S.$1t European Financial Stability Facility was created to address distressed sovereign credits in the Eurozone, and the European Central Bank invested about U.S.$1.5t to U.S.$2.5t in various market support operations. The Bank of England also intervened actively in British markets, boosting its balance sheet by U.S.$500b to U.S.$1t.

Altogether, approximately U.S.$14t to U.S.$15t was expended by U.S. and European governments to shore up their weakened economic and financial systems. Additional large-scale stimulus and quantitative easing efforts also took place in China and Japan.

Such extraordinary and unprecedented levels of government intervention, however, have also had considerable price-distorting effects on the markets for stocks and bonds. Markets are not “free” with this level of government intervention, and without free markets, economic output may be weakened (or distorted) over the longer term. What is still unclear, six years after the crisis, is whether (and when) the governments will discontinue intervention altogether. The public has accumulated a taste for this type of intervention and presumably it will expect governments in the future to intervene whenever economic conditions appear to need a boost. This is a discouraging thought to advocates of free market discipline.

4.2 Re-regulation

Within two years after the peak of the crisis in 2008, the U.S. and European governments moved to create new regulations and institutions to control their respective financial systems. These included a third version of the Basel Accord,7 (called Basel III, to be adopted by 28 countries), the creation of a Financial Stability Board by the Group of 20 Industrial Countries, a new European Banking Authority for the E.U., and the massive (2,300 pages) Dodd-Frank Wall Street Reform and Consumer Protection Act. Also, the Swiss government installed extra heavy capital adequacy standards for their largest banks (the Swiss Finish) and the U.K. introduced a “ring-fencing” system for separating banking from investment banking activities and imposed a special levy on bank held assets above U.S.$35b. In 2013, the E.U. also approved a securities transaction tax, restricted bonus payments by banks, awarded regulatory authority over the top 130 European banks to the European Central Bank, and moved to create a new European Banking Union separate from its newly created European Banking Authority.

These various regulatory initiatives have been occurring in different jurisdictions, and at different speeds, with minimal effort to coordinate or balance them out. But they all apply fully or in part to all large banks operating globally, thereby extending the scope, complexity and cost of compliance by the world’s largest and most important lending and capital market banks. The new rules affect capital adequacy ratios, leverage, liquidity reserves, trading activities, bank executive compensation, reporting, and consumer financial protection. They also provide for risk mitigation and “orderly liquidation” procedures, and “living wills” for bank holding companies.8

4.3 Regulating in the dark

One observer of this frantic process of creating new regulations, Roberta Romano of Yale University Law School, has noted that it was done hurriedly, without the benefit of good underlying research, and therefore was very likely to be seriously defective. As such, she said, it was equivalent to “regulating in the dark.”9

The Dodd Frank Act requires 398 new regulations to be written by the eight agencies that are subordinate to the new Financial Stability Oversight Council. Only about half of the new rules were in place on 31 December 2013.10 One of these, the Volcker Rule, to limit “proprietary trading of banks’ own money” has an

---

7 The Basel Committee was formed in 1987 by the G-10 group of industrial countries (plus Switzerland and Luxembourg, in conjunction with the Bank for International Settlements, located in Basel) to provide a robust set of risk-adjusted minimum bank capital adequacy standards to protect the global banking system from a domino-like collapse if a major, undercapitalized bank should fail.

8 In 2013 the U.S., the U.K., Japan and the European Central Bank were all engaged in forms of “quantitative easing” programs to assist economic growth through means of monetary policy.


addendum of nearly 1,000 pages describing how the rule should be applied. Rules on trading and clearing of financial derivatives are also very lengthy, complex, internationally dissimilar, and require a vast amount of reporting of data to regulators, the value and usefulness of which is unclear. Except for capital adequacy and leverage ratios, which are more severe for the largest banks, the new rules and regulations apply to all banks in the U.S. (7,000) and Europe (9,000).

Few, if any, of the new regulations have been subjected to any kind of cost-benefit analysis. A 2012 study by Standard & Poor’s estimated that compliance costs for the top eight U.S. banks could reach U.S.$34b annually (more than one-quarter’s earnings), not including initial set up and programming costs. It is obvious that the cost of installing new compliance systems and preparing the reports required will be very considerable and that these costs will be passed on to users of financial services in the future, increasing the cost of capital for all American and European enterprises.

The new regulations have many other flaws. One is the assumption that systemic risk will be abated by increased capital adequacy requirements. The regulators have not actually defined what systemic risk is, other than to assert that banks with assets in excess of U.S.$50b (Dodd-Frank) have it, as do those on a list of 29 “systemically important banking institutions” designated by the Financial Stability Board of the G-20. Whatever risks are involved, more capital reserves is likely to be safer than less, but future concentrations of systemic risk may be greater than any feasible amount of capital reserves.

4.4 Systemic risk is market risk
At New York University’s Stern School, the Volatility Lab led by Nobel laureate Robert Engle has calculated something called “S-Risk,” which is a measure of how much replacement capital banks would have to raise if a stock market decline of 40% should occur over a six-month period. S-Risk assumes the stock market is a proxy for broad-gauge market risk, which in the crisis proved to be the source of most of the asset write-offs and funding troubles. S-Risk, therefore, is a function of the volatility (or beta) of stock prices, which have had seven spikes above 25%, four above 40%, and one of 80% in the last 10 years. The betas of bank stocks have been at historically high levels since the crisis, with the average beta for the top ten global capital market banks at 31 December 2013 being 2.01, or twice the volatility of the stock market as a whole, a very high level of volatility for a group of large, bureaucratic, heavily regulated financial institutions.

Another way of looking at market risk is in the extent of “roll-over” exposure that a bank has. A bank that invested only its capital and customer deposits in loans and other assets would have no exposure to the risks of reissuing debt in the markets to replace maturing obligations, because it would not have issued such obligations. But, once a bank starts to expand its lending and trading activities by financing them in the markets it is exposed to the risk that investors will not renew their holdings, forcing the bank into a funding crisis. Thus, the greater the extent to which a bank depends on non-deposit funding of its liabilities, the greater is its exposure to market risk. In 2008, for example, the three largest U.S. capital market banks relied on non-deposit funding for 51.2%, and the five largest European capital market banks for 76.9%, of their total liabilities. Equity capital provided very little funding of these banks’ total assets in 2008 – 7.8%

<table>
<thead>
<tr>
<th>Three largest U.S. banks</th>
<th>2008</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basel leverage ratio</td>
<td>7.50%</td>
<td></td>
</tr>
<tr>
<td>Shareholder equity to total assets</td>
<td>7.40%</td>
<td>10.60%</td>
</tr>
<tr>
<td>Deposits to total liabilities</td>
<td>48.80%</td>
<td>56.80%</td>
</tr>
<tr>
<td>Tier 1 equity to risk-weighted assets</td>
<td>10.60%</td>
<td>12.75%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Five largest European banks</th>
<th>2008</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basel leverage ratio</td>
<td>3.70%</td>
<td></td>
</tr>
<tr>
<td>Shareholder equity to total assets</td>
<td>2.40%</td>
<td>4.50%</td>
</tr>
<tr>
<td>Deposits to total liabilities</td>
<td>23.10%</td>
<td>38.20%</td>
</tr>
<tr>
<td>Tier 1 equity to risk-weighted assets</td>
<td>10.80%</td>
<td>15.40%</td>
</tr>
</tbody>
</table>

Table 1: Certain balance sheet ratios for capital market banks (as of 31 December)

12 Compliance Week, 13 August 2012.
13 Dodd-Frank Wall Street Reform and Consumer Protection Act (2010), Title I, Subtitle C.
14 See: www.mfmgroup.org/documents/ENGLE.pptx.
15 The next two largest U.S. capital market banks, Goldman Sachs and Morgan Stanley, are former investment banks with very little deposits and heavy reliance on market funding.
Six years after the crisis

of the U.S. banks’ but only 2% of the European’s, demonstrating the extent to which their balance sheets and their exposure to market risk were increased by leverage. Nevertheless, because of their ability to game the Basel I risk weightings, their Tier 1 capital ratios appeared to be adequate as demonstrated in Table 1, which also shows how much the banks have reduced leverage since 2008.

Basel I, focused mainly on risk-weighted assets, is a measure of exposure to credit risk, not market risk. Basel III, which essentially doubled the amount of risk-adjusted capital banks must retain, does the same, though the importance of market risk was acknowledged in the form of new leverage and liquidity rules. But setting these rules has been somewhat arbitrary and they vary considerably between countries.  

The trouble with market risk, however, is that you never know how much of it you may have to face as volatilities and correlations among asset classes shift, sometimes very suddenly.

The chances are that there will be a lot of market risk in the future. With global market capitalization of tradable securities increased to approximately U.S.$230t at the end of 2013, financial market activity today is vast, global, interconnected and subject to many different volatility factors that can suddenly send trillions into or out of particular investments with hugely powerful pricing effects that are then subject to instant mark-to-market accounting rules.

4.5 No orderly liquidation

The Dodd Frank law provides for a complex process for the government to intervene in banks to “mitigate” risks or to orchestrate an “orderly liquidation.”17 This process, however, can only begin after the Financial Stability Oversight Council, America’s über-regulator, has voted first to designate a firm as a “grave threat to the financial stability of the U.S.” The designation may be appealed in the courts by the bank to which it was applied, but even so, most experts believe in a crisis the process could begin within a few days of designation. Mitigation could mean many things, but investors in the bank’s debt and equity securities will assume it to mean loss-sharing by them though “bail-ins” or “haircuts” as have been described by the Federal Deposit Insurance Corporation, the agency appointed to manage mitigations and liquidations.18 However, as soon as any grave threat designation is proposed for a vote, well before any mitigation can occur, we can assume investors will head for the exits and refuse to roll over maturing obligations of the designated bank. Very likely, contagion would subject many of the banks’ peers to the same problem. Regulators would thus be creating the institutional run on the bank they are trying to prevent from failing by starting the not-so-orderly intervention process in the first place. Anticipating such a result should dissuade most regulators from ever making the designation, which, de facto, neutralizes this power to intervene.

In the E.U., the European Central Bank has been given the authority to regulate the largest European banks, and to apply similar orderly liquidation procedures in doing so. How it might do so is not yet clear, but doing so at the ECB level raises political issues that Dodd Frank does not face. Can, for example, the ECB liquidate a Spanish bank without the consent of Spanish authorities?

In any event, mitigating actions that would make a difference to a bank’s solvency (such as shutting down a failing investment banking unit) are large-scale matters that can take a long time to implement. If a regulator waits until “grave threat” conditions become clear and indisputable, it may be too late to mitigate anything. For mitigation to become a useful tool of regulators, they would have to be able to foresee the need for it (which has not happened often in the past) and to persuade the bank to act appropriately even if the bank opposed the finding, which for most large banks is quite likely.

Thus, the real power of the regulators is not much changed by the orderly liquidation mechanisms. To implement them in a crisis is likely to be dangerous, and to impose them before a crisis begins may be very difficult to do. More likely, the Federal Reserve or the European Central Bank will continue to keep systemically important banks afloat indefinitely by advancing roll-over funds through lender-of-last-resort facilities or market support

---

17 Dodd Frank (2010), Title II.
18 Statement Of Martin J. Gruenberg, Chairman, Federal Deposit Insurance Corporation, on mitigating systemic risk through Wall Street Reform before the committee on banking, housing and urban affairs of the U.S. Senate, 11 July 2013.
programs. This is the situation that existed before the crisis, though perhaps future availability of the lender-of-last-resort funds will come with requirements for risk-reducing actions. Unless, of course, the regulators decide to let the troubled bank fail, in which case the liquidation procedures, including loss sharing by debt investors, would be applied.

4.6 No bailouts
The great dilemma of bank regulators has always been whether or not to use government funds to prevent the bankruptcy of a bank unable to roll over its maturing liabilities. Bailouts are only undertaken to save the financial system from collapse, but the government assistance directly benefits managers and investors in the bank’s bonds and other non-deposit securities. Despite stated public policy not to do so, regulators are likely to use the powers they have to assist systemically important banks rather than risk a collapse. Of course, there is a precedent for government assistance to banks. All banks benefit from deposit insurance arrangements in which ordinary depositors are protected from losses up to insured amounts. And, central banks have long accepted the duty to supply liquidity to banks facing a sudden, extended withdrawal of deposits, or “run”.

In 1984, Continental Illinois Bank, then the seventh largest U.S. bank, faced a run by institutional depositors and investors after experiencing difficulties in rolling over maturing instruments in the markets. To insure the bank’s solvency, the U.S. Government guaranteed all creditors (not just insured depositors). Continental had been reckless in its lending activities, and weak in controlling them – like many of its competitors at the time – so bailing out bad managers was controversial. The bailout, however, sent a signal to the market that confined the problem to Continental Illinois and averted a system-wide crisis.

The Federal Reserve, then chaired by Paul Volcker, acted firmly in the Continental Illinois case to dispel the moral hazard such interventions invoke. He immediately replaced Continental’s board of directors and senior management, its shareholders were wiped out, and the bank’s bad assets were quickly transferred to a “bad bank,” so that the “good bank” could be recapitalized and returned to duty. It was later privatized and the government recovered most of what it expended in the rescue.

This and other actions by the Federal Reserve stabilized the banking system. In Europe, governments did similar things in the late 1980s and 1990s to strengthen their banks. The French government used U.S.$30b of taxpayer funds in 1993 to bail out Credit Lyonnais, the largest French bank. Many other European and Asian governments did the same at that time, but little public complaint was heard. As disagreeable as these bailouts may have been, there was general agreement that they were preferable to letting the financial system fail.

This time, public attitudes about bailouts are different. During the crisis, hundreds of banks (with the exception in the U.S. of Lehman Brothers, an investment bank, and Washington Mutual, a savings bank) were protected from failure, but in the regulatory aftermath, reflecting public opinion, efforts were made to deny the continued existence of any sort of default protection beyond deposit insurance. In the future, investors in bank bonds or other liabilities are told to expect to participate in any losses that might be sustained by the bank in exchange for government assistance.

4.7 Ending too-big-to-fail
Indeed, the preamble to The Dodd Frank Act notes that “ending too-big-to-fail” and “protecting American taxpayers from bailouts,” are among the stated purposes of the law. The law prevents the use of any taxpayer funds (such as a future TARP) to assist banks and limits the scope of the Federal Reserve’s intervention authority. The Federal Reserve still is able to intervene in secondary markets to boost prices and to make loans of last resort to banks, and the ECB retains similar powers, so quiet, low-visibility assistance is still possible.

Dodd Frank requires that systemically important banks and non-banks be subject to more intense, on-the-scene regulation by the Federal Reserve. The law permits the Fed to restrain

21 Ibid, pp. 185–191.
22 Preamble to the Dodd Frank Act: “To promote the financial stability of the United States by improving accountability and transparency in the financial system, to end “too big to fail”, to protect the American taxpayer by ending bailouts, to protect consumers from abusive financial services practices, and for other purposes...”
Six years after the crisis

activities it deems to be risky, and, conceivably, to insist on a change of chief executive. Prior to the crisis, the Fed was not an aggressive regulator — generally it permitted large banks to do whatever they wanted — so it will be important to see how strict it will be in the future. The task of regulating large banks is much more extensive than before, but the people doing the regulating (and the budgets they have to work with) are much the same as before the crisis.

However, Dodd Frank and public sentiment are unambiguous in wanting to see the end of too-big-to-fail. They want to see the full load of protection against systemic risk carried by the banks themselves. This has forced banks to bulk-up to convince creditors that they are completely safe on their own, despite continuing to be large, cumbersome and not very transparent institutions.

Few banks have found this easy to do. Moody’s ratings for the top capital market bank holding companies now range from A3 to Baa3. In announcing downgrades of JP Morgan, Goldman Sachs and Morgan Stanley in 2013, Moody’s noted “rather than relying on public funds to bail out one of these institutions, we expect that bank holding company creditors will be bailed-in and thereby shoulder much of the burden to help recapitalize a failing bank.”23

Such comparatively low ratings are hardly strong endorsements for the world’s most active, systemically important capital market banks. JP Morgan’s senior debt rating, which was downgraded to A3 in 2013, was Aa2 in 2007.

5. Regulation by prosecution

The crisis has left large segments of the general population, politicians and the media continuing to blame the banks for the Great Recession and high unemployment that resulted from it. Investigations of banks and their various activities in creating mortgages, distributing mortgage backed securities, and processing foreclosures, as well as for trading abuses, tax fraud, money laundering and other matters has revealed market practices that were shoddy at best. A deluge of lawsuits resulted, many of them from private parties affected by these practices, but there has also been an unprecedented wave of prosecutions of banks by government officials in the U.S. and in Europe. Altogether, these various lawsuits generated out-of-court settlements of approximately U.S.$120b to U.S.$130b through 2013.24 There continue to be on-going, high-profile investigations of alleged market rigging of the interbank, foreign exchange and commodities markets.

Prosecutions of senior officials of banks deemed to be responsible for the inappropriate actions, however, have not accompanied these settlements. To be sure, bank chief executives and other senior officials have been investigated thoroughly, but without producing evidence sufficient to bring charges. Recklessness or stupidity by bankers is not a crime, so charges are brought against the banks-as-corporations instead, which invariably their boards of directors prefer to settle (rather than face the uncertainties of a jury trial) with the judgments and considerable costs paid by shareholders.

This produces a perverse result. Shareholders are punished for actions of management that were not chargeable, but for which investors have already paid in write-offs and much reduced share prices and dividends. In addition to depleting bank capital at a time when regulators want it to be increased, the uncertainty of the prosecutions increases S-Risk and substantially increases institutional caution and the cost of internal surveillance in banks. These may prove to be high costs to the financial system, to be endured over a long time, just to extract a symbolic victory for politically minded regulators and prosecutors.

5.1 Implications for the financial services industry

Six years after the crisis began, the financial services industry is still adapting to the extraordinary increase in regulatory, reporting and compliance costs. The largest banks and non-banks, especially those designated to meet higher Basel capital requirements, carry the heaviest burden. There are not very many of these “systemically important financial institutions” (SIFIs) – 30 or 40 in the world – but they are especially large, several with assets of more than U.S.$2t, and significantly exposed to market risk because of their extensive investment banking activities.


24 As of 31 December 2013, largest U.S. banks had charged approximately U.S.$115b against legal reserves. Large European banks were involved in settling charges of LIBOR market rigging and other offenses, involving charges of approximately U.S.$5b – U.S.$10b.
Most SIFIs depend on capital market activity for trading opportunities and profits, but the sheer volume of new regulations, and the continued uncertainty over the lengthy time it has taken for them to be finalized, has considerably dampened market activity, innovation and initiative-taking. These banks have been downgraded by bond rating agencies, and their share prices are still well below pre-crisis levels (several still trade below book value). In addition, they have been required to reduce leverage substantially and to increase capital reserves, which have driven their returns on equity to levels that are very low by historical standards. In each year since 2008, the average of the top ten capital market banks’ net return on equity (or Economic Valued Added, returns on equity after the cost of equity capital) has been negative. For 2013, the EVA for these banks ranged from 1.23% to -16.99% and averaged -8.4%, with nine of the ten negative. Prolonged periods of net negative returns suggest a collapse of economic viability in the banks’ basic business models (Table 2).

By contrast, some of the systemically important banks with significantly less capital market activity have fared better since the crisis. Wells Fargo (the U.S. bank with the highest market value), HSBC, Royal Bank of Canada, Toronto Dominion and State Street all reported positive net returns on equity capital in 2013. Several European banks in this category, however, have struggled to regain their financial equilibrium and continue to report negative net returns.

Non-systemic banks (U.S. banks with assets of less than U.S.$50b, and their European equivalents) are less affected by the new regulations than those that are systemic, but still must comply with almost all of the new regulations.

Non-systemic non-banks (insurance companies, other asset managers and institutional investors) are unaffected by most of the new regulations, but still are experiencing somewhat higher regulatory and compliance costs.

This hierarchy of regulatory costs creates competitive differentials. Large capital market banks are likely to be tied up in ways that their smaller competitors are not. Smaller but traditional banks may become bolder in seeking to take lending

<table>
<thead>
<tr>
<th>Ranked by market capitalization</th>
<th>Market cap (U.S.$b)</th>
<th>Market value to book value</th>
<th>RoE (%)</th>
<th>Beta</th>
<th>Cost of capital</th>
<th>EVA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP Morgan Chase</td>
<td>221</td>
<td>1.09</td>
<td>9.00</td>
<td>1.68</td>
<td>11.33</td>
<td>-2.33</td>
</tr>
<tr>
<td>Citigroup</td>
<td>159</td>
<td>0.80</td>
<td>7.00</td>
<td>2.90</td>
<td>19.44</td>
<td>-12.44</td>
</tr>
<tr>
<td>Bank of America</td>
<td>166</td>
<td>0.83</td>
<td>5.00</td>
<td>2.77</td>
<td>18.65</td>
<td>-13.65</td>
</tr>
<tr>
<td>BNP Paribas</td>
<td>98</td>
<td>0.91</td>
<td>6.10</td>
<td>1.78</td>
<td>12.67</td>
<td>-6.57</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>79</td>
<td>1.14</td>
<td>11.00</td>
<td>1.30</td>
<td>9.77</td>
<td>1.23</td>
</tr>
<tr>
<td>UBS</td>
<td>72</td>
<td>1.35</td>
<td>6.70</td>
<td>1.87</td>
<td>13.21</td>
<td>-6.51</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>61</td>
<td>1.04</td>
<td>5.00</td>
<td>1.41</td>
<td>10.44</td>
<td>-5.44</td>
</tr>
<tr>
<td>Barclays Bank</td>
<td>56</td>
<td>0.88</td>
<td>1.00</td>
<td>2.66</td>
<td>17.99</td>
<td>-16.99</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>49</td>
<td>0.65</td>
<td>2.00</td>
<td>2.28</td>
<td>15.69</td>
<td>-13.69</td>
</tr>
<tr>
<td>Credit Suisse</td>
<td>49</td>
<td>1.01</td>
<td>7.50</td>
<td>1.60</td>
<td>11.58</td>
<td>-4.08</td>
</tr>
<tr>
<td>Average</td>
<td>101</td>
<td>0.97</td>
<td>6.03</td>
<td>2.03</td>
<td>14.08</td>
<td>-8.05</td>
</tr>
<tr>
<td>Best 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Global banks’ economic value added (31 December 2013)

EVA = RoE - cost of equity capital

25 The top ten banks are selected on the basis of their market share of new financings and mergers at the beginning of calendar years. Returns of equity are those reported by the banks after all write-offs and special charges. Cost of equity capital is determined by the Capital Asset Pricing Model, with 10-year U.S. Treasury securities as the risk free rate.
Six years after the crisis

business away from their larger, more tightly regulated brethren. Traders and corporate finance bankers in systemic firms may prefer to move to non-banks (e.g., hedge and private equity funds) to avoid regulatory burdens, increased cautiousness and compensation restraints.

Indeed, the market share concentration of the top banks has already been substantially weakened since the crisis. In 2013, the top ten capital market banks accounted for 70% of global loans and securities new issues and merger and acquisition advisories (the top five accounted for 45%); in 2006 the top ten banks handled 94%, and the top five, 57% of such transactions.26

These numbers suggest some significant changes in the global capital market industry. Regulatory changes have been especially hard on some European universal banks, which have been forced to retreat from the more competitive areas of the markets, and their future role as industry leaders has to be questioned.

5.2 Risk shifts to the shadows
Willingly, or by competitive erosion, the systemically important banks can be expected to change their business models by transferring risk to the “shadow banking” system. The banks can do this by becoming distributors, rather than retainers of risk. By reducing inventories of tradable assets, which must be marked to market, they will reduce their risk-weighted assets against which capital must be retained, while also reducing non-deposit funding for the assets.

This is already happening. Banks in the U.S. and Europe have reduced their holdings of risk-weighted assets, shed inventories of corporate securities, and raised lending standards across the board. Various reports indicate a reduction in bank lending over the past few years, together with an increase in corporate lending by private equity funds and the issuance of collateralized loan obligations (securitized bank loans). The Bank for International Settlements recently reported that 50% of trading in corporate securities was, for the first time, occurring in the shadow system that largely bypasses the banks. That is probably a good thing, because the shadow banking arena includes few, if any, firms that might threaten the financial system should they fail. Indeed, hedge funds (included among shadow banks) fail periodically without a systemic affect.

Some worry that risk can become invisible in the shadows and hard to regulate. But Dodd Frank provides the authority to the Financial Stability Oversight Council to designate any non-bank as being systemic, and thus subject to much tighter regulation. So far the Council has only designated three large financial firms as systemic (Prudential Insurance, GE Capital Corp., and AIG), though it is now studying large asset management firms, such as BlackRock and Fidelity, to see whether it should deem them to be systemic as well.

Shadow banks, however, do not act as intermediaries assisting clients. They act strictly as investors, not as market makers providing liquidity to enable financial transactions. So that role, prominently played by capital market banks over the past decade, may come to be in short supply. Shadow banks can relieve the risk held by systemically important banks, but they cannot replace the liquidity supplied by them.

6. Is all the new regulation worth its cost?
Dodd Frank touches just about everything in finance and completely changes the financial regulatory system. Even so, only two of the law’s 16 titles, or sections, relate to systemic risk, which was meant to be its main purpose. The statute identifies about 40 banks as systemic – not very many considering the scope and enormous expense of the law, which must be borne by the industry and its clients and customers. European reforms promise to be equally, if not more expensive to the banking system than the American ones.

The new wave of regulation will have powerful effects on the industry. It takes a lot to get huge, powerful banking institutions to change their ways, but the regulatory moves will be enough to force several large banks to adapt their business models to the new circumstances of their industry, or surely be challenged by shareholders until they do. They will reduce leverage, risk-weighted assets and rollover dependencies, which will reduce their exposure to market risk. That is certainly desirable in the effort to reduce the likelihood of another finance-driven global economic crisis.

26 Rankings compiled by the author, based on Thomson Reuters Securities Data files, 2006 and 2013.
But the way the regulatory platform is structured now, they will be adapting toward a more conservative, risk-averse form of what they were before, one in which banks become “too-safe-to-fail.” That may mean too-cautious-to-fail, which may not be what we want either if it comes at the cost of reducing the banks’ capability to finance basic economic growth.

What has been, and continues to be, the greatest cost of the financial crisis is lost global economic output, employment and social services. This shows up in prolonged, unusually low GDP growth rates in the U.S. and Europe.

In 2013, U.S. GDP growth was 1.9%, after a disappointing fourth quarter, down from 2.8% in 2012. GDP growth in Europe for 2013 was negligible with continuing fears of deflation and stagnation. The average U.S. growth rate in the 20th century was 3.5% – for the 13 years since 2000, however, it has been only about 2.0%. When the growth rate falls below a healthy, sustainable average, the broader economic support mechanisms of the society are endangered.

It is never clear how much economic growth and productivity depends upon the availability of finance for business and investment, but finance is certainly one of the factors to which growth is attributable. Maintaining adequate sustainable growth rates is equally important to our societies as managing systemic risk.

Regulators, regulating in the dark, certainly need to be careful of installing growth-killing measures across the Euro-American industrial base, as they appear to have done and will need to reverse in due course. At the same time, however, banking regulators will have to recreate themselves in a more viable and pragmatic form. This means regulating ordinary things more effectively – not allowing giant mergers that threaten the loss of control over risk and management, preventing excessive leverage obtained by gaming the Basel system and neglecting large amounts of off-balance sheet obligations. Regulators need to be able to do their jobs better than they did prior to the crisis but they are certainly entitled to insist that banks maintain minimum standards of solvency and stability in exchange for the various benefits they receive from the system.

6.1 Waiting for the banks

UBS was the first capital market bank to feel the full weight of the financial crisis. Its “warehouse” of mortgage-backed securities forced write-offs sufficient to bring about a government bailout and a complete change of management. The Swiss government then imposed draconian, “never-again” capital requirements that forced the bank to examine its aggressive global banking business model that had achieved, after several large acquisitions, a prestigious sixth ranked market share based on aggregated capital market origins and advisory assignments.

The UBS share price dropped nearly 80% from the beginning of the crisis in early 2007 until late 2009, about the same as Deutsche Bank and Barclays, but much more than UBS’ rival Credit Suisse’s.

Under new leadership since 2011, UBS made a major strategic change. It would sharply reduce its investment banking activities and their associated risk-weighted assets, concentrate on strengthening its core wealth management businesses and promise to pay out at least 50% of profits once it had shored up its balance sheet to achieve a fully applied Basel III Tier 1 ratio of 13%.

It was betting that its regulators, wealth management clients and the stock market would appreciate this fundamental change once it saw that it was starting to work.

In 2013, UBS’ investment banking global market share fell to ninth from sixth in 2008. Its total originations (book running assignments for global new issues of syndicated loans, bonds, and equities and merger and acquisition advisory assignments) were U.S.$515b, less than what they had been in 2008.

Risk-weighted assets (RWA) at UBS were reduced to U.S.$250b, or 22% of its U.S.$1.1t balance sheet. Nonetheless, its investment banking division, with RWA reduced to U.S.$69b (28% of total RWA), reported pre-tax operating profits of U.S.$2.8b (after excluding a $1.9b charge for litigation expenses related to prior activities), representing a return on attributed equity of 30.6%.

UBS’ fully applied Basel III Tier 1 ratio of 12.8% was the highest among all capital market banks (and just below its 13% target
for increased dividend payout) and its 4.7% leverage ratio well exceeded the 3.7% average for its European peers.

The strategic change (together with its associated austerity), does seem to be paying off. In 2013, with profits before litigation charges up by 44% to U.S.$4.7b, the bank increased common stock dividends by 67% while investors await achieving the trigger point of a dividend payout of 50% of earning. UBS’ stock price rose by 35% during the year, more than any of its European peers.

At March 31, 2014, UBS’ shares traded at 1.39 times book value, compared to its European peer group average of 0.88 times. By continuing to deemphasize its investment banking business – to a point where it was no longer dangerous and could trade at 2-3 times book value (Lazard Frères, with very little market risk exposure, traded at 9.5 times book value) – and investing further in its asset management business should enable the bank to be valued more as a leading private banking asset manager with multiples of book value in the 2 times range, similar to those of BlackRock (1.9x) and Julius Baer (2.0x), than as a capital intensive, risk-laden investment bank. At 2 times book value, UBS’s stock price would rise by a further third.

Only two of UBS’ capital market peers have undertaken major strategic changes – both Citigroup and Morgan Stanley were changed by the former’s sale of its Smith Barney brokerage business to the latter. For Citigroup the move was a reluctant but necessary effort to shrink the bank. For Morgan Stanley, the conversion into the world’s largest retail brokerage firm has been a difficult and lengthy transition – still, Morgan Stanley was in fourth place in 2013. Citigroup, which was in sixth place in 2013 had been in first place in 2003.

The stock market does not appear to value this market share very much (JP Morgan, the leader, trades at 1.14x book value, Citigroup at 0.77x and Morgan Stanley at 0.99x in March 2014).

Six years after the crisis that dumped an ocean of regulatory change and reputation damage onto the capital markets industry, the principle banks involved are little changed. Only UBS has announced a major move – to distance itself from the capital markets business, to save its crown jewel and restore shareholder value, and the stock market has begun to buy into it. There is plenty of room and incentive, and still time, for others to make similar strategic changes.
Abstract
Six years on from the tumultuous events of 2008, many aspects of the banking industry are almost unrecognizable from the way they were in the period leading up to the Global Financial Crisis (GFC). Whether it is the fixed income businesses of investment banks, the (mis-) selling of products to retail consumers, the demand for financial services from customers in emerging markets or the scale and intrusive nature of regulation and supervision, the industry has changed and will continue to change for some years to come. Responses by banks have been mixed. Many have been in a holding pattern, waiting for the regulatory landscape to become clearer and the global economy to improve. Some that embarked on strategic reviews in the wake of the GFC have been forced to revisit them as markets failed to recover and regulators imposed new rules, both domestically and internationally. This paper argues that the time for waiting may be over. Reshaping large global and regional banks is a multi-year undertaking and the broad direction of regulatory travel is now clear. The pendulum will swing from review to action, although the initial focus is likely to be on “no regrets” investments and initiatives. As they transform, banks must also incorporate and maintain enough flexibility in their models to respond to new rules as they emerge and survive further setbacks in the global economic recovery.
1. Introduction

“The more things change, the more they stay the same.” That may be true in some aspects of life, but not in the banking industry. At a global level, the scale of change in the industry since 2008 has been unprecedented. The industry is anything but the same.

Banks have been trying to restructure themselves and rehabilitate the industry in the face of significant headwinds. The regulatory rulebook is being completely rewritten, the industry has suffered major reputational damage and major economies have struggled to emerge from recession. Results have reflected these challenges, with return on equity (RoE) at many banks failing to cover the cost of equity.

Complying with new regulations and operating within regulatory constraints has been the first priority for banks and will remain so for some time to come. However, banks must also consider how to meet the expectations of customers, shareholders and markets within those constraints.

On the regulatory front, some rules remain unwritten, especially in relation to structural reform. There is also the potentially counterproductive march toward regulatory nationalism. Although the mist has not cleared completely, there is more clarity in key areas and this certainty should prompt more banks to embark on much needed strategic transformation programs of both business and operating models.

Banks will need to incorporate enough flexibility in their models to respond to new rules as they emerge, essentially developing a strategy around remaining uncertainties. The initial focus is likely to be on “no regrets” investments and initiatives. That is understandable, but is only a start.

Regulation will remain the primary driver of reform for the foreseeable future, but it is not the only one. We see five unstoppable forces that we believe will drive banks to change over the next few years, namely regulation, customer, technology, competition and society. Overlaying all of those forces will be a fluctuating and unpredictable economy.

Banks’ responses to those forces will be severely constrained by regulation in all its forms but, with “wait and see” no longer a viable strategy, we believe that more banks will be transitioning from planning to execution, and those responses will coalesce around four themes, namely business models, customer relationships, organizations and infrastructure.

How effectively the banking institutions respond to these challenges will determine their prospects going forward. In this article, we will look at these challenges in more detail and discuss how these institutions can, and should, respond to them.

2. The banking industry of today is almost unrecognizable

The global financial crisis (GFC) of 2008-09 unleashed not only a major economic slowdown, but also an unprecedented regulatory response to how the risks within the banking system should be managed.

Despite the severity and breadth of the reforms, it has taken many years for these changes to be implemented at the national level. For example, in the immediate aftermath of the crisis, in 2009, the G20 economies agreed to the comprehensive regulatory reform of the banking industry that resulted in the Basel III accord. Yet, it has taken several years for politicians and regulators to translate those global standards into national rules and implement other aspects of regulatory reform. Many remain unwritten. In the European Union (E.U.), the Capital Requirements Directive (CRD) IV was only finalized in 2013 and many rules under the Dodd-Frank Act (2010) have yet to be finalized.

Combined with a sluggish economic recovery, these delays have served as a block on efforts by banks to begin the arduous process of structural reform. Rather than pushing ahead with transformational change programs, many have opted to wait until more of the regulatory agenda becomes clear. The focus has been on tactical steps to “fix what’s broken.”

Most of the final rules have multi-year lead-in times before they take effect. Markets have been much less patient. As we have already seen with Basel III capital standards, banks have been pressured to provide commentary on compliance, and some have been forced to raise additional equity capital. There is an expectation that market pressures will also extend to liquidity and leverage ratio requirements as those are finalized.

One of the biggest challenges so far has been dealing with
the chasm that has appeared between global design and local execution. In most cases, the problem has not been regulatory arbitrage; rather, local interpretation and implementation of standards have begun to erode what should have been a globally consistent framework. Potentially even more significant are locally driven initiatives, such as the Foreign Banking Organization (FBO) and Volcker rules in the U.S. and ring-fencing proposals in the E.U., which will require major structural reform.

Global banks must deal with additional rules at home, different interpretations of the same rules across jurisdictions and the threat of overreach by home and host authorities incorporating extraterritorial powers into regulations. National supervisors will often have justifiable reasons for the rules they introduce, but an inconsistent global framework risks damaging the more effective aspects of the international capital markets.

2.1 Challenges reflected in the numbers
The economic and regulatory headwinds faced by the banking industry have been reflected in the numbers. In addition to decidedly mixed revenues, banks across regions have also been hit by a series of compliance and regulatory breaches, and by claims related to the GFC, with fines and settlements already running into the billions of dollars and no end in sight.

More recent GDP growth and falling credit losses have enabled some banks to release provisions against loan defaults. Given the flattening revenues and rising costs across many parts of the industry, these releases have been a key driver of profitability — which clearly is not sustainable. There have also been distinct regional differences in performance.

The U.S. banks, having restructured their balance sheets in the immediate wake of the GFC, have been able to take advantage of the domestic recovery and opportunities abroad. Notwithstanding recent concerns in some Asia-Pacific economies, banks in the developed and emerging markets of the region have benefited from solid economic growth to deliver some strong performances.

Indeed, there has been a positive story across most emerging markets over the last few years, reflected in the solid capital base and RoE levels that most banks have been able to maintain. Strong GDP growth has fueled higher wages and increased levels of disposable income, driving consumer spending and demand for financial products from business and retail customers. Government-sponsored investment programs have strengthened cross-sector infrastructure, paving the way for further growth.

However, the emerging markets are not without their challenges. Rising costs are pressing margins in a number of industries. There is a lack of depth in domestic capital markets that will be a drag on growth, particularly with the U.S. quantitative easing (QE) coming to an end. As a result, banks will be forced to seek opportunities for faster growth and funding beyond their traditional bases. We believe significant, further infrastructure investment is needed to reduce reliance on commodities and broaden the bases of these economies. Banks and governments will benefit from collaborating to ensure the financial services sector can promote, rather than hamper, future growth.

Compared to peers elsewhere, European banks have fared less well, as the Eurozone sovereign debt crisis effectively halted recovery in the region. Most banks in developed markets have undertaken some level of restructuring, and risk-weighted assets (RWAs) have fallen. They have exited non-core business lines and sold assets. Some have also acknowledged their lack of market leadership and wound down sub-scale business lines. Unfortunately, all this is still a work-in-progress for many European institutions and they remain weighed down by non-core and underperforming assets.

As noted above, banks have also struggled to deliver top-line revenue growth, as market volatility has deterred activity and customers across many developed markets have continued to pay down debt. The low interest rate environment has continued to squeeze net interest margins, with the Americas banks in our sample seeing a decline of more than 15% in the period 2009–13.

The long-awaited recovery of the housing market in the U.S. has spurred growth in the mortgage market, particularly for refinancing. This initial rush has slowed significantly, however, and there are concerns that the threat of higher interest rates, once QE ends, will discourage further growth. The European market has also shown some signs of life, supported by policy initiatives, such as the “Help to Buy” scheme in the U.K. However, many are wary of encouraging too much growth in this sector at a time when interest rates are at an all-time low.
Revenue growth has also been challenged within the investment banking divisions of major global banks. While most products continue to see soft demand, resulting in overcapacity across the industry, others are picking up. Advisory and origination have shown signs of improvement in 2013, with market performance, low rates and investor appetite supporting growth in equity and debt capital markets. However, trading revenues have been lackluster and profitability-squeezed. Capital concerns and the increased price transparency of electronic platforms are driving a continued decline in fixed income, currencies and commodities (FICC) businesses, with revenues down 26% between 2009 and 2012. Poor performances in the second half of 2013 will also have undone gains in the first six months. A partial recovery in equities will not have been enough to compensate, given the size disparity of these business lines.

Although compensation costs have declined as a proportion of revenue for the top 50 global banks (from 33% in 2009 to just under 30% in 1H13), non-staff costs have continued to rise faster than revenues, particularly for European banks. At a total cost level, only the Asian banks in our group have been able to grow revenues faster than costs.

The precise implications may be unknown, but we have already seen what the expectations about the termination of QE can do to markets, yields and prices. The International Monetary Fund’s October 2013 Global Financial Stability Report noted that market losses on bond portfolios could reach U.S.$2.3t as a result of tapering. Its eventual termination, likely toward the end of 2014, will exacerbate some of the challenges facing banks and make the case for structural reform even stronger.

2.2 Banks across many developed markets have struggled to recover

At the bank and the industry level, much more change is required to deliver strategic restructuring and to deal with overcapacity, both in terms of numbers of banks, and the products and services they provide to customers. In emerging markets, banks must consider how to equip themselves for the next phase of growth in the face of rapidly evolving customer requirements and uncertain economic growth.

Responses in the industry have been mixed. We have noted that some banks already have acknowledged the need for change, with transformational reform programs under way. Others are just recognizing the scale of what is required.

Before we consider how banks should be responding to an environment in flux, it is important to understand how those changes are manifesting themselves. What will be the key forces that drive change over the next couple of years and beyond? Why does reform need to cut across structure and strategy of both banks and the industry? Essentially, why do we need to redefine banking?

Economic shifts will continue to have an influence as the global recovery enters its next phase and trade routes between markets, both developed and emerging, are redefined. Beyond that, there are multiple change drivers, but we have separated them into five core areas: regulation (global and local), customers (demands and expectations), technology and innovation (enabler, differentiator and disruptor), competitors (old and new) and society (politicians, citizens, activists). For some, the challenges ahead will be immense. Even for those banks that have begun the process of strategic restructuring, the evolutionary nature of these issues will force many to revisit their plans.

2.2.1 Regulation

Regulation has been the primary driver of structural change within banking since the full implications of the GFC became clear, and many believe that it is likely to continue for some time.
However, the scope and scale of regulatory change affecting banks, particularly those that operate cross-border, are such that a mere tactical response is no longer a viable option.

Initially, incremental steps to comply had seemed to be the only sensible response due to the lingering uncertainty and the new rules being developed across multiple jurisdictions. That approach must now give way to a strategic restructuring of business models and operating models.

We have seen already the impact that increased capital charges are having on some business lines, such as FICC. The Institute of International Finance (IIF)/EY 2013 risk management survey, *Remaking financial services: risk management five years after the crisis*, illustrates this, with the majority of respondents (69%) expecting capital requirements for their trading books to at least double under Basel III (Figure 1).

The proposals from the Basel Committee’s second consultative document on a fundamental review of banks’ trading books, if adopted, will have further implications in this area. Banks may be required to calculate risks and allocate capital according to a standardized methodology, instead of using an in-house approach. As well as increasing costs, this could also affect the liquidity of some markets as banks choose to hold fewer illiquid assets on their trading books.

In addition, as a result of over-the-counter (OTC) derivatives reform, most of the banks surveyed also expected capital charges on derivatives to rise by at least 100% as a result of counterparty risk charges alone. With a corresponding impact on RoE (several percentage points), and new margin rules to contend with, delivering acceptable returns will be increasingly elusive for those banks that do not enjoy the scale benefits of market leadership.

Final rules for liquidity under the net stable funding ratio (NSFR) have yet to be established by the Basel Committee, as have rules for leverage ratios by a number of jurisdictions, including the U.S. These are likely to drive additional changes to business and funding models. For global banks, the exercise will be further complicated by the increasing preference for subsidiaries over branches by national regulators and potential differences in adoption of liquidity and leverage ratios.

**Balkanization**

Structural reform, in the shape of ring-fencing and product restrictions, will require banks operating in the U.S., U.K., France, Germany and elsewhere to either cease certain activities or ring-fence them to protect the deposit-taking entity. Across a number of emerging markets, regulators are also implementing much more stringent regimes for foreign banks around capital, liquidity, funding sources and locally housed infrastructure. Combined, these issues will demand a structural redesign of the organization, its systems and processes, and its approach to capital management, liquidity management and governance.

Once national regulators agree upon and finalize recovery and resolution plans (RRP), making banks resolvable is likely to require some major restructuring. But it is not yet clear exactly how far regulators will expect banks to go and whether there will be cross-border acceptance of national resolution regimes. Host jurisdictions may be reluctant to rely on a single point of entry approach and may insist that local subsidiaries be sufficiently capitalized to withstand shocks without relying on capital or loss-absorbing debt from the parent institution in the home country. This may reduce the risk that banks are “too big to fail,” but it will also reduce the efficiency of the global banking system and its ability to support the global economy.

Many of these regulatory initiatives illustrate the seemingly unstoppable march toward balkanization, with global standards giving way to national rules that contradict the stated aim of international cooperation. Variations may be justifiable at a national level, but they will force banks to rethink their global strategies as the playing field becomes much less level.

The benefits of a global banking system are also being threatened by the preference of some governments and regulators for what some term overreach, as evidenced by the U.S. Foreign Accounts Tax Compliance Act (FATCA) and the Eurozone’s attempt to introduce a financial transaction tax (FTT). The largest 130 banks in the Eurozone will face an additional challenge in 2014 as they undergo the European Central Bank’s Asset Quality Review. This is expected to be a much more rigorous assessment of the banks’ balance sheets than the previous exercise by the European Banking Authority. Asset sales and additional capital raising to strengthen balance sheets are expected, both in anticipation of and in response to the results.
In the U.S., the new FBO rules will have a significant impact on the capital, liquidity and regulatory reporting requirements of those banks affected, as well as their organizational and legal entity structures. The compliance and broader business implications will inevitably require some institutions affected by the rules to reconsider their operations in the U.S.

Beyond prudential regulation, the broader issue of business conduct is becoming a considerable force for change as many banks continue to reel from the effects of multiple failures and breaches. As more institutions have been implicated in issues ranging from anti-money laundering to mis-selling and rate manipulation, a further wave of regulation is likely to deal with conduct risk. Remediation will be time-consuming and expensive as banks examine their own processes, procedures and broader behavioral issues in parts of their organizations.

2.2.2 Customer
Customer requirements are evolving rapidly. Banks are under pressure from retail and business customers, across both developed and emerging markets, to offer the levels of service and flexibility experienced in other sectors.

In rapid-growth markets, there is also demand for new types of products for the first time. Retail customers are looking for both loan and investment products as income levels rise. For business customers, particularly those looking to expand abroad, demand is growing for hedging products and access to international capital markets. For many domestic banks in these markets, providing these products profitably will be a challenge that requires an innovative response.

There is an increasing expectation across all markets that customers can transition seamlessly between channels and be able to use their channel of choice for different activities, whether they be undertaking simple transactions or seeking advice. More customers are looking to use smartphones for a range of banking and payment activities, including contactless payments and payments direct from the customer’s account to a merchant’s account without the use of debit or credit cards. In an era when most customers are multi-banked and switching between providers is increasingly easy, banks must address a number of key issues to protect both the relationship and the stability of their retail funding base.

Table 1: Most important key features or benefits sought from primary financial service providers

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keeps your personal information safe</td>
<td>35%</td>
</tr>
<tr>
<td>Protects your financial information</td>
<td>35%</td>
</tr>
<tr>
<td>Provides easy access to branches and ATMs</td>
<td>31%</td>
</tr>
<tr>
<td>Is transparent about what they charge for and makes it clear to you</td>
<td>29%</td>
</tr>
<tr>
<td>how to avoid paying fees</td>
<td></td>
</tr>
<tr>
<td>Offers excellent online banking features</td>
<td>26%</td>
</tr>
<tr>
<td>Reaches out to you as soon as possible if they believe a problem may</td>
<td>24%</td>
</tr>
<tr>
<td>exist with your account</td>
<td></td>
</tr>
<tr>
<td>Has an excellent reputation</td>
<td>24%</td>
</tr>
<tr>
<td>Offers low cost banking options</td>
<td>24%</td>
</tr>
<tr>
<td>Works with you when you need help or encounter a problem</td>
<td>20%</td>
</tr>
<tr>
<td>Handles your request quickly</td>
<td>19%</td>
</tr>
</tbody>
</table>

Table 1: Most important key features or benefits sought from primary financial service providers
Percentage shown is the percentage of total respondents that selected the benefit as one of their top five most important benefits
Source: EY global survey of retail banking customers 2014

Transparency on rates and fees is essential, as highlighted in EY’s 2014 global survey on consumer banking, as are excellent online banking features (Table 1). However, the top two features that customers are looking for relate to the security of personal and financial information.

Under attack
This prioritization reflects concern by customers, regulators and the banks themselves. The nature, speed and sophistication of cyber crime are evolving rapidly. Hackers operating for financial gain may be the more immediate concern of customers, but “denial of service” attacks that disable systems and delete or steal data are potentially much more damaging. Banks recognize the seriousness of these threats. They also recognize that it is impossible to protect themselves against all attacks and so they must focus on securing the data and systems they regard as their “crown jewels.” However, budgetary concerns and a lack of skilled staff are preventing many from delivering those fixes. This is dangerous not only for the individual bank but also for the broader financial system, as attackers hunt for the weakest link.

Suitability
Underpinning a number of the responses outlined in Table 1 is the need for banks to rebuild trust with their customers. Consumer protection agencies across multiple jurisdictions are introducing
new rules on suitability, but banks must also demonstrate that they are delivering value and prioritizing long-term customer satisfaction over short-term product sales.

There is also an expectation of change among business and corporate customers. For many, banks are too focused on pushing potentially unwanted products onto them rather than working with them to provide more appropriate solutions. Outdated systems and processes was also raised as a key concern in recent survey of multinational corporations.¹ Banks also need to consider how to serve these customers more consistently across geographies (Figure 2).

Banks recognize that they may have been less focused on their customers as a result of multiple business and regulatory challenges. As the EY/Financial Times survey on structural reform shows, changing customer expectations was ranked as the third most important issue as banks consider structural changes to their businesses (Figure 3).

2.2.3 Technology and innovation
Technology will be a crucial driver of change from multiple angles. We have already noted above the importance retail customers place on security, and the threat of data theft and other cyber-crime looms large across the industry. Yet, data is a much broader issue for banks. The amount of data that banks are required to collect, analyze, store and mine is growing exponentially. Even before banks can think about using the data to provide a more effective service to customers, they must overcome the significant burden of regulatory compliance. Tracking and reporting requirements have become much more onerous and legacy systems are struggling to cope.

For example, the Basel Committee’s principles of risk data aggregation must be implemented by 2016 and will require major investments in data, systems and reporting capabilities. According to the IIF/EY risk survey, nearly 40% of banks expect spending to increase by more than 30% in the next two years — and that estimate is just looking at information technology (IT) to support the bank’s risk architecture.

As customers become increasingly omni-channel and make greater use of social networks in their interactions with banks, the sources of data are also becoming more complex. With increased connectivity required between channels, and the growth of mobile and online banking, banks are finding that current systems are ill-suited to a more digital business environment. Some banks are starting to lead with a digital strategy, recognizing the brand benefits of differentiation and the potential savings from driving more activity to lower-cost channels, but these initiatives need to be part of the broader business transformation.

Technology is also enabling new ways for customers to buy banking products and services from new, non-traditional sources. The “disruptive technology” moniker is usually associated with new ways for customers and merchants to exchange payment for goods. As new payment and mobile wallet technologies become

¹ EY, 2013, “Successful corporate banking: focus on fundamentals.” EY LLP.
more widely used, banks will be in danger of losing transaction fees and prominence in the mind of the customer.

The term also applies to peer-to-peer (P2P) lending, both to retail and business customers. However, it could be applied equally to developments in the capital markets. We continue to see the further “electronification” of sales and trading, and technology is being used to disrupt other aspects of investment banking. The OTC derivatives reforms have proved to be a catalyst for new entrants looking to take advantage of the requirement for these products to be traded on exchange and cleared centrally.

2.2.4 Competition

Competition has always been a driver of change in any industry. Factor in the issues raised above and it becomes clear that the competitive landscape is likely to look considerably different in years to come.

Much of the focus has been on new, non-traditional players, but incumbent banks will also engender a new-look environment. Particularly where domestic markets are still over-banked, the combination of competitive and regulatory pressures must lead to further reshaping if recent revenue trends are to be avoided in the future (Figure 4). For example, the better capitalized banks in the Eurozone peripheral economies are attracting deposits from their struggling rivals; at the same time, they are also able to borrow more cheaply from the wholesale markets. Stronger banks will exploit this advantage to lure customers with better products and rates. International business customers may also be tempted to reconfigure their group of primary banks, as some on the list are no longer able to offer the same product range or geographic footprint. Domestic banks in emerging markets are also likely to face more severe competition in areas, such as corporate banking and wealth management, from regional banks looking to broaden their reach and from global banks in search of growth opportunities.

Capital and regulatory reporting requirements are squeezing the profitability of FICC sales and trading, and it is likely that only market leaders will have the scale to thrive and the funds to invest in the necessary technology. Competition may be intense in the short-to-medium term, but beyond that, only a few banks are likely to survive in a global, multidisciplinary market.

We are also likely to see the emergence of more new entrants, leveraging technology to focus on a particular niche, especially where regulators have imposed restrictions on banks or where the move to use technology is relatively recent, such as in corporate bond trading. Regulation has served as a catalyst for new trading exchanges and swap execution facilities, as well as more non-bank broker dealers. These “shadow banks” may lose some of that advantage over time as they are subjected to more scrutiny. However, the concern for banks must be that, as they are distracted by the regulation and remediation, shadow banking entities will take market share and the move away from the old status quo will be permanent.

Regulation is already driving some change in the payments arena, as banks in Europe, the U.S. and elsewhere are facing restrictions on card transaction fees. Historically, banks would have been reluctant to put this revenue at risk. Now, however, technological innovation may be required to prevent further losses from new competitors.

The humble current (or checking) account is also becoming an increasingly competitive battleground. Again, technology is playing a part as start-ups use new platforms and outsourcing

![Figure 4: Total revenues from top 50 global banks, 2008-13](source: Bank reports, EY analysis)
arrangements to offer a fresh take on banking. It is likely that more retailers will expand their product ranges to include current accounts, in the hope of becoming a customer’s primary financial services provider. Technology and telecommunications firms may also expand the scope of their services to include more financial products.

P2P lenders may only represent a small fraction of the loan market, but the concept is attracting significant attention from institutional investors, which are looking for opportunities to boost yield. As more asset managers establish loan funds, they may use P2P lenders as intermediaries to source borrowers instead of using banks. Recognizing the potential growth opportunities and the type of borrowers that may be attracted to these sites, banks are beginning to work with P2P lenders on larger loan requests.

The broad arena of mobile money is also likely to see much more intense competition as technology and telecommunications firms move further into mobile payments and mobile banking. The loss of fees will be one concern for banks, but more troubling will be the potential weakening of their relationship with both retail customers and merchants, as digital wallets diminish brand awareness and customer connection.

2.2.5 Society
The final driver encapsulates why the other four will lead to significant change within the industry. Indeed, it is society – whether in the form of politicians, shareholders, customers or the media – that has helped to create the environment for the other forces of change to develop as they have. The cumulative impact on the industry’s public reputation of compliance breaches, product mis-selling, computer failures, data leaks and benchmark rate manipulation is impossible to measure precisely, but it has been severe. Coming after the effects of the GFC, these failings have left confidence in banks at what may be an all-time low.

Customers expect change and investors are looking for credible and sustainable business models. Investors are also looking for the balance of rewards to shift back from managers and employees toward shareholders. Scrutiny from politicians and the media will continue to be relentless. Activism, by both consumers and shareholders is on the rise in both developed and emerging markets.

The pendulum of responsibility has swung far away from caveat emptor to a much more extreme form of consumer protection. Irrespective of whether that swing may be too far, the current environment is probably here to stay for years to come.

As the IIF/EY 2013 risk management survey highlighted, banks recognize the need for a behavioral shift in attitude and approach. Conduct has become a major concern, and board directors and chief risk officers are increasingly focused on operational and reputational risk. Banks’ risk appetites may be more conservative than they were precrisis, but embedding that into the culture is an ongoing process.

The industry needs to change the way it interacts with customers, sells to customers and manages itself. The result needs to be banks that are more transparent and bankers who are more clearly customer-oriented.

3. Banks transformed, an industry reshaped
While demand for banking services is expected to grow over the coming years, the organizations that provide them and the manner in which they are delivered is changing. With the myriad challenges and issues forcing the need for change, what response do we expect to see from banks over the next few years and beyond?

When EY and the Financial Times spoke to senior banking executives across Europe to conduct a survey on structural reform earlier this year, we asked banks to categorize themselves now and in the future. For their current business model, 68% described it as universal banking. For their future model, 68% also believe they are likely to use a universal model. However, how universal banking is defined is likely to change, and not just in Europe. Large banks will be universal in their domestic markets, and perhaps a few other core jurisdictions, but many will need to be much more selective about the services they offer elsewhere. Across emerging markets, banks may strive to provide a universal banking service to their customers but will sometimes lack the resources, technology and distribution capabilities to do so single-handedly. Banks will be transformed as a result and the industry will be redefined.

In the wake of the GFC, the immediate priority for banks has been to comply with the wave of new regulatory requirements
on capital, liquidity and leverage and for recovery and resolution plans (RRPs). Considerable progress has been made but work is not yet complete. The challenge of compliance has already forced banks to rethink their business models as they learned to operate within new constraints. Future RoE targets continue to be revised down. Some banks will face further hurdles as rules on liquidity and leverage are finalized and additional capital is required.

As the regulatory constraints have become clearer, and banks are generally much stronger, the focus is now on having the right business and operating models to serve customers, deliver profitable growth and provide an acceptable return to investors. Some aspects of structural reform are being driven by regulation, such as resolution planning and ring-fencing, but change must go to the core of the business model and responses must reflect all of the drivers outlined above. There are fundamental, strategic questions that must be answered that will shape the business and determine the extent to which the future bank will indeed be universal.

3.1 Rush to execution?
Most banks have been through that strategic review process since the GFC, some more than once, so they have a reasonably clear understanding of their future business models. Yet many have been reluctant to commit to, and operationalize the new strategy, for fear of missing a product recovery or underestimating the impact of a business line exit on other businesses. Regulatory and economic uncertainty has also been a major factor in this “wait and see” approach.

However, as many would agree, the time for waiting may be over. Reshaping large global and regional banks is a multiyear undertaking, and the broad direction of regulatory travel is clear. The pendulum will swing from review to action, although the initial focus is likely to be on “no regrets” investments and initiatives. Banks must incorporate and maintain enough flexibility in their models to respond to new rules as they emerge.

Regulatory constraints will continue to drive significant activity as banks strive to comply with new rules and meet new requirements in areas, such as capital, liquidity and leverage. Beyond that initial focus, strategic transformation will coalesce around four core areas: business models, customer relationships, organizations and infrastructure.

3.1.1 Business models
As the final strategic review process takes place, there are a number of factors to consider. Regulation and economic events
continue to drive a bigger wedge between weaker and stronger banks, with the latter group able to borrow more cheaply and attract more deposits. For example, while loan-deposit ratios (LDRs) are declining across Europe, some institutions, particularly those in the Eurozone’s peripheral economies, still face significant challenges.

There has already been some consolidation in the European market, from 8167 in Q1 2011 to 7761 in Q3 2013, with markets such as Spain witnessing significant restructurings, with more than 50 institutions reduced to 12 in three years. We expect to see more consolidation in the medium term in those markets that remain overbanked, such as Germany and Italy. Although the LDR is considerably below 100% in the U.S., further consolidation is also likely here. Many of the small and mid-size banks will have neither the scale nor the niche attractiveness to win and retain enough business. Hit by the increasing costs of complying and competing and by falling net interest margins, they will struggle to survive without radically reshaping their business models.

Banks in emerging markets are also suffering the effects of a low interest rate environment and will be exploring new revenue opportunities to secure new funding capital and maintain returns. Notwithstanding recent policy rate rises to combat inflation, competition is intensifying and banks remain under political pressure to reduce spreads on credit products to retail customers and small businesses. At home, they will focus on product innovation and broadening channel reach to attract new customers, though they must resist the temptation to win volume at the expense of appropriate risk-based pricing.

These factors highlight some of the changes we expect to see in the banking landscape. Small and mid-sized banks operating within a single market may continue, but to remain viable, their models will evolve to incorporate alliances and partnerships with others (Figure 5). Stronger national banks will also collaborate with others, and some will look to further targeted expansion abroad.

Looking at banks in Asia, Africa and Latin America, we find that many are developing regional expansion strategies to exploit the growing demand for financial services in frontier markets and leverage their domestic expertise. They will continue their selective expansion abroad to support their corporate customers and serve key capital and trade flows. Given margin pressure and rising cost inflation, these emerging markets banks will also need to focus on scale and cost efficiencies to maintain viability of the business model.

Regulation as inspiration
Estimates vary on the cumulative impact of regulation on RoE, but it is impossible to ignore. As Figure 6 shows, more and more institutions are, and will be, reviewing aspects of their business models. Banks will continue to exit business lines and geographies as they streamline operations and strengthen balance sheets. As an example, the reforms mandating that OTC derivatives are traded on exchange and cleared centrally will have a significant impact on investment banking revenues, as will new margin requirements. Dominant players are hoping to benefit from scale and recover some lost margin from an increase in volume; for others, recovery options and long-term business viability will be limited.

These changes threaten the viability of global wholesale banking. As a result, we will see fewer genuinely global organizations and the emergence of more strong regional institutions in Africa, Asia and Latin America. These banks will benefit as markets return to stronger growth. Some selectively international banks in markets such as Australia, Canada and Japan, where limited domestic growth opportunities are forcing them to look elsewhere, will also

Figure 6: Effect of combined liquidity and capital changes under Basel III on business models
Source: IIF/EY, 2013, “Remaking financial services: risk management five years after the crisis.”
benefit as they continue their expansion into new markets, albeit on a highly selective basis.

The retrenchment across markets and products risks are having a significant impact on customers, as banks that retrench to become solid national banks are less able to fulfil their roles as financial intermediaries and support their customers across different countries and product areas. Banks in emerging markets are approaching the same problem from the opposite side, as some struggle to meet customer demand for new products and services, such as wealth management and investment banking. Alliances with others, whether in the form of informal collaboration or more formal partnerships will become increasingly important.

Universal 2.0
It is in the rapid-growth markets where we are already seeing examples of how these challenges can be overcome, such as with Japanese banks taking strategic stakes in Vietnamese banks to combine strengths. We are also finding that banks in Indonesia, Malaysia, Mexico and Nigeria are considering partnerships with genuinely global or strong regional banks to strengthen their product propositions.

Going forward, we expect different forms of partnership to emerge across all markets as thoughts turn to future growth opportunities. Institutions will look to harness their strengths and exploit their competitive advantages while recognizing that they may not be able to satisfy every aspect of customer demand single-handedly. More banks will question whether all products need to be manufactured in-house, and whether they need to both originate and distribute to customers. Some will hone their position as focused specialists, targeting sub-sectors of the industry, such as wealth management, asset servicing, M&A advisory or credit cards. Collaborating with other institutions rather than retaining sub-par business units or committing to major organic expansion programs will provide banks across markets with more agility to respond to new revenue prospects as well as potential economic shocks.

There is potential to extend the open-architecture philosophy beyond wealth management into more vanilla savings and investment products, as well as some credit products. This may be particularly appealing to those local or national banks that are smaller and lack scale. Some new entrants into banking, including retailers, have adopted this approach and focus on the front-end relationship with the customer, contracting with existing banks to provide other services and functions.

In addition to potential white-labeling and third-party sourcing options, there are already examples of banks partnering with telecommunications firms to use new technology to provide customers with offers, discounts and promotions, such as in Spain and the U.K. Although banks need to address any data-sharing concerns, there are opportunities here to develop much needed new revenue streams. We expect to see more banks looking at opportunities to collaborate with each other, and with organizations from other sectors, to provide new products and services to customers.

Some non-traditional new entrants, the innovative disruptors, will be a source of competition as well as potential collaboration. Yet many of them, whether they are shadow banks or technology firms, do not wish to develop the infrastructure needed to run a bank, nor do they wish to become full-scale banks with all of the associated regulatory requirements. Banks may be able to benefit by providing middle and back-office support to these innovative disruptors, and they may also be able to partner with them and benefit from their innovation.

For banks that need to consider these alternative solutions, perhaps the key dilemma will be determining which parts of the value chain they need to own, and which elements of the overall customer relationship they must retain, if they can no longer own the value chain end-to-end. This will vary by institution depending on its competitive advantages, but we think this collaborative approach is a strategy that will suit many banks. Even the genuinely global banks will choose collaboration over organic growth or full-scale acquisition in some markets. The future business model may be universal banking, but it will not be solo banking.

As banks transition to new models and adapt to what will be a lower return environment for many, compensation models will also need to be redesigned. The E.U. bonus cap makes that both inevitable and problematic for banks within the union. Elsewhere, investor demands and supervisory requirements will also push banks to act but they will need to ensure that variable bonuses are not replaced by high-fixed salaries.
3.1.2 Customer relationships
The worldwide focus on greater consumer protection, through legislation and regulation, will change the way banks interact with customers, sell products, reward staff and deal with complaints. As banks across multiple jurisdictions focus on restoring credibility and regaining trust, they must also turn their attention to implementing these new regulations and ensuring that adoption of the broader regulatory agenda supports efforts to restore customer relationships.

Which customers?
One of the broader dilemmas facing banks will be which of its customers to keep relationships with once the new strategy has been implemented. Many banks that have tried to be all things to all customers will find that promise difficult to keep. More are investing in data analytics to develop a clearer picture of the full, risk-adjusted value of a customer’s overall relationship with the bank. Although no business likes to turn away customers, we may see some institutions take that step and focus on a smaller group of core customer segments where they have genuine competitive advantage. As an alternative, we may also see much greater segmentation across channels, with lower-value customers more reliant on a technology-enabled, self-service model with limited interaction with a relationship manager.

Solutions, not products
That focus is likely to result in more sustainable relationships, and the message it sends to staff will also support the board sponsored agenda that many banks have to realign culture and behavior. There is recognition that too many sales staff and relationship managers were incented inappropriately in the past. With an increased focus on articulating and embedding a more customer-focused culture, we will also see banks ensure that internal target-setting and reward strategies are aligned to restoring trust and developing long-term relationships with customers.

From a product and service perspective, the priority should be on enabling more personalization and being much more transparent with rates and fees. Bank customers are willing to help banks provide them a more tailored solution.

Growth from digital
Across different sectors such as media, aviation and telecommunications, customers have much more freedom to create a more personalized solution that best suits their needs. Organizations in these industries are leading with a “digital first” strategy, where solutions are developed with mobile and online devices in mind rather than being retrofitted to support those channels. They have a better record than most banks of using customer data effectively. Customers are also experiencing better complaints handling from other sectors; improving effectiveness in this area will boost retention.

As their customers become accustomed to this flexibility and service quality, and start to experience it from new, non-traditional financial services providers, we anticipate that more incumbent banks will follow these examples. They certainly should. Pre-defined bundling is not necessarily bad, but we expect more banks to respond to customer expectations of greater choice.

Restructuring for business customers
All customers expect increasingly seamless service throughout the organization, and the speedy and effective resolution of problems. At a time when customers have more choice than ever before, structural reform that complements the new customer strategy is crucial to minimizing attrition.

Aspects of the regulatory agenda, designed to reform bank structures, will have an inevitable impact on business customers, particularly those that require products and services from both sides of a potential ring fence. The decisions banks make on booking models, legal entity strategy and location strategy in the coming months and years will also affect customers. As banks consider whether to establish an operating entity that serves multiple divisions, use shared service centers or use industry-wide utilities, it will be critical to understand what impact those decisions would have on different customer groups.

Customers as collaborators
We find that more businesses are expecting to refinance loan or debt obligations in the next year. At a time when interest rates and bond yields are well below historical levels, we are seeing more examples of asset managers establishing loan funds or, in the case of a Canadian pension fund, lending directly to companies. Indeed, for low-risk, longer-term projects, a pension fund or insurance company may be a better fit than a bank with a constrained balance sheet.
Banks have the customer base and the skills in loan origination and credit risk management. What many lack is the funds to lend or the capacity on the balance sheet to hold those loans to maturity. We expect to see more examples of banks working with asset managers to harness different skill sets and resources. Given the overdependence on bank lending in Europe and most emerging markets, banks should also be looking at how to help smaller companies secure funding from debt and equity capital markets.

3.1.3 Organizations
With significant changes to business models expected over the next few years, existing operating models will also need restructuring. We anticipate a noticeable shift of emphasis as banks transition from planning and tactical fixes to execution of broader, integrated, structural reforms.

Regulation is a significant driver of these reforms in two ways. First, banks will be required to implement certain changes to ensure resolution plans can be executed effectively and to comply with specific rules on ring fencing. For those banks affected by the proposed rules on FBOs in the U.S., structural changes will be inevitable to comply and the cost implications may necessitate strategic changes as well. Global booking models are also under threat. Second, the broader regulatory agenda, as it relates to capital, liquidity and OTC derivatives, makes structural reform an economic imperative. There has been a permanent step-change in the cost of doing business. Combined, there are multiple incentives to minimize duplication, improve capital and liquidity efficiency, and streamline legacy – and often overly bureaucratic – structures. Unfortunately, regulators are more concerned with stability than economic efficiency and some duplication will be inevitable. Partly as a reaction to this, the future operating models for many banks are likely to be simpler, reflecting their withdrawal from different markets and business lines.

Streamlining structures
Key considerations relate to which entities should be deposit-taking, which should be restricted to capital markets and investment banking activities, whether the ultimate parent should be a bank or a holding company and whether to create separate operating entities to manage key operations and technology functions. Banks must also adapt to restrictions on the activities possible within foreign branches, as more regulators favor locally capitalized subsidiaries.

We have seen a number of banks embark on a review of their legal entity structures already. More are expected to follow and rationalize their portfolios, disbanding unnecessary entities to improve capital efficiency. As banks embark on this exercise, they should also incorporate a review of internal structures and hierarchies. A simpler bank should be able to remove duplicate management and de-layer the organization.

These reviews should also incorporate the future-state structure of the business across jurisdictions. The increasing rise of regulatory nationalism, with rules above and beyond global standards, will have a major impact on structural reform. Local regulatory requirements in individual jurisdictions have become clearer, particularly around capital levels and the shift from branch to subsidiary structures. As a result, there will be a spotlight on the organization’s geographic footprint to determine which locations should serve as booking centers for different products. Trapped capital and liquidity are increasing risks for global banks, as are proposed tax changes, such as Europe’s FTT, and we are likely to see further rationalization of where certain activities are undertaken to mitigate this issue.

Transforming cultures
Banks across jurisdictions have been hit by wave upon wave of compliance failures and conduct issues. Immediate responses necessarily have been tactical as banks try to understand the scale of the problem and implement a remediation program. There are broader issues of culture and behavior that also need to be addressed, and a number of institutions have undertaken a more strategic assessment of how their organizations need to change. Much of this focuses on risk culture, but it also incorporates broader behavioral issues throughout the business.

There is a clear requirement for the bank’s risk appetite to be reflected in its culture, but organizations must also strike a balance between managing risks and delivering (sustainable) growth. Linked to that is the question of whether the organization needs to have a single culture across every part of the business or whether, instead, there are common core values but cultural variations among departments that reasonably reflect their different activities.
The major challenge for banks that have been through this process is ensuring that changes have been embedded throughout the organization and that new standards are being adhered to. Unfortunately, many lack the systems and processes required to help monitor conduct and behavior and to track how well changes are being embedded. The broader risk governance agenda will continue to be a key priority for banks, but we expect to see much more focus on this area to strengthen internal transparency and thereby reduce operational risk, particularly as it relates to conduct.

Connecting to shared services
Compliance with ring-fencing and resolution planning will also require affected institutions to consider which group functions and services should be shared across different businesses and subsidiaries, and whether some should be housed in separate operating entities. Regulators are keen to ensure that any “run the bank” operations, such as payment processing and account systems, do not suffer from the collapse of a single entity within the group. Establishing these entities will be expensive and time-consuming, diverting funds and time away from growth-focused initiatives. As banks invest in creating these structures, it is possible that some will consider whether they might be shared by other banks. Other large competitors may be reluctant to use another bank’s operating company, but it may be a significant advantage for smaller banks due to cost savings.

We also expect to see more banks introduce initiatives to transform core functions, such as finance, that may suffer from duplication and inefficiency as a result of legacy structures. Streamlining processes and controls across the bank and increasing levels of automation will form part of this and should result in significant reductions in back-office headcount. This transformation is also likely to include a revised location strategy, as banks will be exploring offshoring and near-shoring options to benefit from lower-cost centers. However, it is vital to reengineer the function first to avoid exporting inefficient processes to the new location. It is also important to consider how skills will be transferred between the head office and new centers, to ensure head office staff understand core processes and functions.

Emerging markets not exempt
Banks across a number of emerging markets have embarked upon significant expansion programs in recent years, with further plans for major growth in the future. Branch networks have been extended, product ranges enhanced and staff numbers increased. Many of these institutions are recognizing the need to reform structures and key functions to manage a business that is now several orders of magnitude bigger.

Risk management will require particular focus as branch networks test operational risk capabilities, rapid growth in staff numbers affects the organization’s culture and increased demand for credit products stretches the capacity of the credit risk function. Strengthening systems, processes and oversight will be a crucial area of focus to minimize the risk of failure and provide comfort to management and investors. “Run-the-bank” technology will be a crucial investment area in the short to medium term to deliver these upgrades, as well as to ensure that the bank has the management information available to comply with local and international regulatory standards. Other corporate functions and operating procedures will also need attention, ranging from payment processing to human resources, as these banks adapt to operating on a different scale. As more of these banks look to partnership or collaboration opportunities with foreign firms, demonstration of robust governance processes will be a prerequisite for any deal.

3.1.4 Infrastructure
The pressure on bank infrastructure is immense. Antiquated platform and multiple legacy systems were struggling to cope before the onslaught of new demands. The spend required to deliver new systems will be significant, and with the vast majority of bank IT budgets already committed to maintenance and regulatory initiatives, there is limited scope for investment, especially in Europe.

Pre-GFC profitability levels had also masked the inefficiencies inherent in many bank organizations, including duplicate activities across divisions and poorly designed processes. Taking these factors together, banks will need to redesign their infrastructures to reduce complexity and meet new requirements more efficiently. And to manage costs, more organizations will explore options to outsource, relocate and share systems and processes with other organizations.

Silos dismantled
Different divisions within banks have historically been guilty of
recreating rather than sharing systems, particularly if existing ones did not quite meet the requirements of a particular division or business line. Organizational silos and systems inherited through mergers and acquisitions have exacerbated the situation, leaving banks with complex and costly systems architectures and an army of staff to manage them.

We are already seeing a few banks develop shared systems across core functions, such as finance and risk to save money. Standardization of tools, platforms and processes will need to become much more common if costs are to be brought under control. We should also see further consolidation of processing centers, data centers and server facilities, with banks exploring outsourcing options and greater use of cloud technology.

The redesigned infrastructure should extend to a much tighter strategy for the development and support of systems and applications. Eliminating unused or underused applications will provide some cost savings, but banks should also review their broader approach to new development. Going forward, more banks are likely to configure third-party solutions instead of bespoke development in-house. We should also see consolidation of suppliers around a core group of vendors that provide both development and maintenance support.

Dismantling silos and harmonizing systems will provide banks with more reliable, consistent data across the business and enable them to harness “golden sources” of data that are easier to protect from cyber attacks. This will support compliance with regulatory requirements, ranging from risk data aggregation to counterparty exposure. It will enable banks to benefit from organization-wide capital and liquidity management. It should also allow the organization to have a single view of customer data across channels and use more sophisticated data analytics to strengthen customer relationships and aid business development.

In other industries, particularly within the manufacturing sectors, much smaller profit margins forced the adoption of “lean manufacturing” several years ago. Banks must consider how they can reengineer the business to industrialize more processes and make further use of straight-through processing, both to cut costs and to reduce the opportunity for errors.

As banks review their options, the question is whether they develop solutions in splendid isolation or consider industry-wide options for sharing infrastructure. To that end, we expect more banks will overcome the aversion to sharing, both internally and externally, typified by the “not invented here” syndrome.

Sharing is good
Although some IT systems are a source of competitive advantage for banks, many that store data for regulatory monitoring and compliance are not. Some industry-wide solutions in areas such as payments processing are well established, but we are starting to see examples of banks looking to share more back-office costs, particularly within corporate and investment banking. One group is looking at potentially outsourcing customer compliance checks (e.g., know your customer, anti-money laundering) to a new, third-party, industry utility. Another is exploring options to create a centralized library that provides a common source of client reference data, including legal entity identifiers and background checks.

We also expect to see further progress being made within retail banking, such as potentially storing all customer account details in a central utility. Not only would this reduce costs, it would satisfy regulatory requirements for increased portability of accounts. In addition to industry-wide solutions, we expect more banks to take advantage of third parties that offer other services. There is likely to be more white-labelling and outsourcing of financial infrastructure, such as payments, custody and settlements. Options to create industry utilities should also be explored by banks in emerging markets that are looking to develop new systems. In the medium term, we expect to see much less duplication of ATMs and point-of-sale machines in emerging markets.

We have seen examples of banks in Germany and Italy leveraging new technology, such as cloud computing and storage, to obviate the need for investment in physical servers and storage capabilities. Data security concerns are preventing some from adopting this innovation, but many are likely to follow as fears are allayed, possibly using a combination of public and private clouds, depending on the sensitivity of the data.

These decisions will not be taken lightly. Eliminating duplication, both within banks and across the industry, may make sense.
economically but the reaction of regulators to these options will be crucial. Concerns around continuity of service will need to be satisfied, as well as those related to data. But cost pressures should drive further progress in this direction. Concentration risks must also be addressed.

Some of these options would incur significant upfront development costs, but they would be cheaper than the cumulative cost of multiple institutions running similar systems. Indeed, third-party providers and industry utilities may also be more secure than stand-alone systems at individual banks, as these services will be their primary function. The initial focus is likely to be on assessing which systems do not confer any proprietary advantage and which processes could be extracted and transferred to a third party.

Digital first
Given the changes in customer behavior, and the emergence of new technology, we believe banks will follow other sectors such as media, down the path of “digital first” solution development strategies. Smartphone ownership continues to grow rapidly, and both retail and business customers are increasingly expressing a preference to manage certain transactions digitally instead of in-person or via a branch. Instead of retrofitting existing solutions onto a digital platform, we expect more banks to lead with their digital solutions and make them available in branches. As customer channel preferences continue to evolve, and more customers choose to visit branches for advice on complex solutions rather than to manage simple transactions, there is also opportunity for a new branch strategy.

Banks have significant resources tied up in their physical branches and could apply lessons from the retail sector where multiple brands have “concessions” within a single department store. Sharing physical branches may not work in busy urban centers where the customer numbers justify stand-alone branches, but it could be an option in less populous locations. This is likely to form part of a broader review of, and likely consolidation of, banks’ real estate footprints, particularly where they still have multiple offices in single cities.

4. Conclusion
There are multiple, unavoidable forces driving change in the banking industry, regulation being the strongest. New regulation is designed to make the industry more stable, make it less likely to require future taxpayer bailouts and ensure consumers are adequately protected.

Many of the reforms, if not all, are needed and yet expensive to implement. The impact on the cost of doing business is even more severe and regulatory constraints have rendered many current business models unprofitable. Banks need to change.

There is recognition that change is required, but banks are serving multiple masters. For regulators, stability is the primary concern. Investors may accept a lower risk-adjusted return from safer banks, but that return still needs to exceed the cost of equity. Customers expect to be treated fairly but also expect to benefit from new technology, and they are being tempted away from traditional banks by alternative, more tech-savvy providers. Banks need to innovate.

New entrants are using disruptive technologies to shake up the status quo. Many are just start-ups and their current share of the market may be small, but they are bringing to banking the sort of change we have seen in other industries. Some of these new ideas will work, some will not. Banks do not necessarily need to imitate in an attempt to compensate for not being a first mover. They can benefit from “second mover” advantage by developing an understanding of how customers react to new propositions, adapting as needed to suit their own base. We have already seen a few banks partner with innovators, such as P2P lenders. Acquiring innovative start-ups is also an option, but banks must be careful not to quash the creativity they are buying by “corporatizing” it.

Responses necessarily cut across business lines and extend throughout the industry. They need to restore credibility and trust. They must fix broken processes and replace antiquated systems. There is also a need to redesign compensation models as banks adapt to a lower return environment. Collectively, they should result in an industry that can generate sustainable growth and support the next phase of economic recovery in both developed and emerging markets.
Once plans are implemented, we expect banks to develop mechanisms and tools to monitor progress, particularly on the less tangible aspects of reform, such as culture and behavior. Institutions will be keen to strengthen their reputations and recover profitability.

The business of banking will continue, but these responses should change the way it is conducted. Irrespective of which model banks adopt, we will see more alliances and collaborations across business lines, markets and functions as universal banking is redefined.
Part 2: Tactical

Evaluating different approaches to quantitative easing: lessons for the future of central banking

New results on the correlation problem in operational risk

Financial perspective: the unintended consequences of regulatory oversight and control – lessons from the banking and the asset/alternative funds industries

Risk management insights from Markowitz optimization for constructing portfolios with commodity futures

Japanese patent index and stock performance
Evaluating different approaches to quantitative easing: lessons for the future of central banking

Bluford H. Putnam
Chief Economist and Managing Director, CME Group

Abstract
The era of quantitative easing (QE) is going to be intensely studied for lessons to guide to the future conduct of central banking. There were major differences in the motivations and objectives of central banks for their QE programs, which led to very different approaches to their asset expansion practices. Our research compares and contrasts the activities of the Federal Reserve (Fed) and the European Central Bank (ECB).

Quick and aggressive Fed and ECB actions after the disruptive and badly handled bankruptcy of Lehman Brothers and bailout of AIG more than likely prevented another Great Depression. Fed QE programs adopted from 2011, even as the U.S. economy was already recovering, may not have helped job creation at all. This means that it is essential for studies of the efficacy of QE to divide the analysis between the initial crisis response which worked well and the later asset expansion programs which do not appear to have helped. The ECB took a different approach. The ECB’s focus on liquidity loans calmed financial markets, but did not assist banks in shedding distressed assets and may have hindered economic growth compared to the Fed’s approach to purchase assets, reducing bank balance and capital requirements and leading to a faster economic recovery.

Exit strategies are also in the spotlight. The Fed’s exit from QE is likely to be highly complex, involving delays in returning to a more traditional short-term interest rate policy, dramatically diminished contributions to the U.S. Treasury from central bank net earnings and the potential for loss of some of the Fed’s independence over time as the U.S. Congress increases its oversight concerning the size of the Fed’s balance sheet and potentially large unrealized portfolio losses. By contrast, the ECB’s approach to QE mostly through loans to the banking system is allowing for the easiest and most natural exit path among the major central banks.

1 All examples in this report are hypothetical interpretations of situations and are used for explanation purposes only. The views in this report reflect solely those of the authors and not necessarily those of CME Group or its affiliated institutions. This report and the information herein should not be considered investment advice or the results of actual market experience.
1. Introduction
After the financial panic of September 2008 and the recession that followed, several of the key central banks of the mature, industrial world expanded their balance sheets aggressively, in what has become known as the era of quantitative easing (QE). Yet, there were major differences in the motivations and objectives of central banks for their QE programs, which led to very different approaches to their asset expansion practices. In turn, the different implementation methods resulted in very different outcomes, posing strikingly different challenges for their respective economies, with critical implications for future central bank policies.

Our research compares and contrasts the activities of the Federal Reserve (Fed) and the European Central Bank (ECB). We come to relatively obvious as well as some potentially controversial conclusions, including the following:

- Quick and aggressive Fed and ECB actions after the bankruptcy of Lehman Brothers and badly managed bailout of AIG more than likely prevented another Great Depression.
- Later Fed QE programs adopted from 2011, even as the U.S. economy was already recovering, may not have helped job creation at all.
- The ECB's focus on liquidity loans calmed financial markets, but did not assist banks in shedding distressed assets and may have hindered economic growth compared to the Fed's approach to purchase assets, reducing bank balance and capital requirements and leading to a faster economic recovery.
- The Fed's exit from QE is likely to be highly complex, involving delays in returning to a more traditional short-term interest rate policy, dramatically diminished contributions to the U.S. Treasury from central bank net earnings and the potential for loss of some of the Fed's independence over time as the U.S. Congress increases its oversight concerning the size of the Fed's balance sheet and potentially large unrealized portfolio losses.
- The ECB's approach to QE mostly through loans to the banking system is allowing for the easiest and most natural exit path among the major central banks.

Central bank motivations are critical to appreciate. As the financial panic unfolded in September 2008, both the Fed and the ECB focused on the immediate challenge of stabilizing markets. Later, as the economic recovery was taking hold in the U.S., the Fed shifted its focus to perceptions of fragility in labor markets and set its objective as lowering the unemployment rate at a more rapid pace than otherwise. The ECB, confronted by the sovereign debt crisis, viewed rebuilding the credibility of the single currency as its prime objective.

With different motivations came different methods and approaches. In the initial financial panic of the fourth quarter of 2008, the Fed bought distressed exposures from failing financial institutions. After the economic recovery began in late 2009, and the Fed shifted to worrying about labor markets, it began to accumulate U.S. Treasury securities, extended their maturities and purchased mortgage-backed securities. By contrast, the ECB's focus on rebuilding the credibility of the euro in the face of the sovereign debt crisis meant assuring the liquidity of the financial system. Thus, the ECB chose mainly to expand its balance sheet with emergency liquidity, low interest term loans to banks, with some, but much less, of an emphasis on purchasing the debt of weaker nations.

Progress toward achieving objectives has been different as well. The quick and aggressive actions by the Fed, the ECB and other major central banks in September 2008 and the following few months probably saved the world from the financial panic spiraling downward into another Great Depression. We will argue, however, that the Fed's QE programs from 2011 onward did little to nothing to assist the U.S. labor market and help in the creation of jobs, and it may well have been a hindrance due to the negative signals it sent about economic confidence. The ECB has enjoyed considerable success in stabilizing the euro and the E.U. banking system – its primary objectives – yet there has been a lack of meaningful economic growth, although the picture finally appears to be brightening in 2014.

As we look to future, the unintended consequences of QE are becoming more apparent. The Fed now has a balance sheet equal to about a quarter of one year's nominal GDP and faces severe challenges in how it manages the exit from QE. Specifically, the Fed is highly vulnerable to very large unrealized losses in its vast U.S. Treasury and MBS (mortgage-backed securities) portfolios should any inflation pressures emerge down the road. Together, these factors may well lead to heightened scrutiny by the U.S. Congress over Fed activities and may even result in some loss.
of independence. By contrast, the ECB is already seeing its balance sheet shrink naturally as banks are now paying back their emergency term loans relatively aggressively. Thus, the ECB does not face the prospect of portfolio losses, nor does it have the challenges the Fed has in reducing its balance sheet. Yet, the modest strength of the euro and remaining challenges in the banking system raise the possibility of unwanted deflationary pressures and some further policy adjustments by the ECB.

We will adopt the case study approach and examine the Fed and the ECB in turn. We will describe their motivations and methods in detail, and then focus our attention on the results, unintended consequences and implications for the future of their economies. Our concluding section will summarize the lessons learned so far from the QE era and what it means for the future of central banking.

2. U.S. Federal Reserve
2.1 Motivation and Implementation
During the financial crisis period, the U.S. Fed was led by Professor Benjamin Bernanke, a scholar of the Great Depression of the 1930s. Professor Bernanke was also a critic in the late 1990s of the Bank of Japan's unwillingness to take extraordinary measures, including asset purchases, to counteract the malaise faced by the country's economy after the equity and real estate boom of the 1980s had collapsed. When the financial panic hit after the bankruptcy of Lehman Brothers and the highly disruptive manner in which AIG was rescued, Fed Chair Bernanke turned to the playbook he had suggested for the Bank of Japan a decade earlier, and aggressively purchased assets.

Phase one in the U.S., or QE-1 (17 September 2008 through 31 December 2008), was all about crisis management. The Fed lowered its target federal funds rate to near zero, and it purchased approximately U.S.$1t of distressed securities and derivative exposures between mid-September and December 2008. We would argue that this part of the QE legacy was critical to the U.S. avoiding a depression, and it was in line with the spirit of the Fed's original charter to be a lender of last resort to the banking system. The Fed used an interpretation of its emergency powers to justify buying the various securities and exposures in QE-1, but it clearly did not want to be a long-term holder of distressed debt and derivative exposures. During the next phase (31 December 2008 through 29 December 2010) of the US QE era, the distressed assets and exposures were sold at a profit as markets calmed and the economic recovery commenced. The money received from the sale of the distressed assets and exposures was largely used to purchase MBS.

From 2011 onward, various programs of QE were introduced to expand holdings of U.S. Treasuries and MBS, as well as lengthen maturities. These next rounds of QE were not motivated by the immediate concerns of financial panic, as was the case in the last quarter of 2008. Equity markets had started a powerful recovery rally in March 2009. Economic growth had commenced in late 2009. Private sector employment growth had begun in early 2010. A second round of asset purchases, QE-2 (29 December 2010 through 6 July 2011), was initiated because the Fed felt the economic recovery was exceedingly fragile and net job growth was painfully slow. QE-2 focused on the purchase of U.S. Treasuries, adding about U.S.$600b to its portfolio, of which

---


only U.S.$40b were long-term securities of 10-year maturities or more. That is, QE-2 mostly comprised the purchase of short- and medium-term Treasury securities.

Still unhappy with the perceived fragility of the economy, pace of recovery, and labor market conditions, the Fed shifted gears and adopted a Maturity Extension Program (MEP) 4 (6 July 2011 through 26 December 2012) to actively try to lower longer-term U.S. Treasury yields. The MEP was essentially a modern version of the Fed’s 1961 attempts to lower long-term bond yields known then as “Operation Twist.” 5 In this phase, the Fed sold some of its holdings of short- and medium-term U.S. Treasury securities and used the proceeds to purchase long-term U.S. Treasury securities, dramatically lengthening the maturity profile of its Treasury holdings.

In the latter part of 2012, the Fed shifted gears again, terminating the maturity extension program and announcing a new, larger and open-ended asset purchase program, QE-3 (26 December 2012 through 25 December 2013). 6 This phase initially involved purchasing U.S.$45b of U.S. Treasuries and U.S.$40b of MBS each month, adding about U.S.$1t to the Fed’s balance sheet in 2013, an amount equivalent to about 6% of nominal GDP. During May 2013, Chairman Bernanke commenced a public discussion of when and how the Fed might taper its asset purchases. The official decision to start reducing the monthly asset purchases was made at the December 2013 Federal Open Market Committee (FOMC) meeting 7 and was re-affirmed with further cuts in asset purchases at the January 2014 meeting, when the Fed Board’s gavel was handed over from Ben Bernanke to the new Fed Chair, Janet Yellen.

2.2 Market impact

The most notable impact on long-term bond yields and probably directly attributable to the Fed’s actions occurred with the announcement and implementation of the MEP in 2012 and ended abruptly when Chairman Bernanke initiated his discussion of tapering the QE program. The Fed’s MEP and QE-3 programs temporarily succeeded in lowering 10-year Treasury yields in 2012 below the core inflation rate. Once the “taper talk” began, 10-year Treasury yields immediately rose back above the core inflation rate.

That is, the Fed’s asset purchases in QE-2 of U.S.$600b did not appear to make a noticeable difference in bond yields. This was probably because the purchases of long-term bond yields were a relatively small part of the mix. During this period, the Fed purchased mostly short- and medium-term U.S. Treasuries. Since the short-end of the yield curve was already anchored near zero by the target federal funds rate, the impact of these asset purchases on rates further down the yield curve was small to non-existent. Also of note, during the debate, announcement, and implementation of QE-2, bond market volatility was most probably dominated by the worsening sovereign debt crisis in Europe.

By contrast, when the Fed shifted to the MEP, it was selling the short end and buying the long end of the Treasury yield curve, and this made a material difference in bond yields.
Our estimates, as well as others,8 put the impact of the Fed’s MEP at approximately 100 basis points. This was re-affirmed by the market impact when “taper talk” commenced. At the end of April 2013, the 10-year Treasury was yielding 1.67%. By early July, this yield was more than 2.75%. And for the rest of 2013, the U.S. 10-year Treasury yield traded roughly in the 2.6% to 3.0% range.

The impact on U.S. equities from the various QE programs is highly debatable. As already noted, there is almost no question that QE-1 helped to stabilize the financial system and alleviate market fears of a meltdown. U.S. equities began their recovery rally in March 2009 as panic faded, and economic growth resumed in late 2009.

It is the impact on equities of QE-2, the MEP and QE-3 that is controversial and in question. Indeed, we would argue that studies of U.S. QE that combine and do not separate the impacts of QE-1 in the last quarter of 2008 amidst severe financial panic from the later QE programs that were conducted in the context of a recovering and modestly growing economy may miss critical lessons from the QE experiment.9

For his part, Chairman Bernanke in his speech at the Jackson Hole conference of August 2012 cited the theory of portfolio balance and asset substitution as the channel by which asset purchases by the Fed could raise stock prices.10 On the surface, this seems reasonable and is well-supported with economic theory, but there were other things happening, however, that were also likely to have contributed to equity rally. Corporate profits were recovering at a very sharp pace between 2010 and 2012. Based just on the 2010–12 experience, it seems likely that the lower bond yields and reduced bond market volatility attributable to the MEP did aid equity prices, but the credit should be shared with the powerful surge in earnings growth.

---


That is, to simplify:

Equation 1:  
\[
\text{Equity index growth} = \frac{\text{Earning growth expectations}}{(1 + \text{Treasury bond yield} + \text{market volatility (risk) premium})}
\]

That is, QE was operating mostly in a positive manner on stock prices with its influence on the denominator, lowering the hurdle rate with reduced bond yields and lower volatility. Earnings growth from a recovering economy was a positive factor working through the numerator.

What makes the hindsight analysis of the equity market reaction to QE-2, MEP and QE-3 even more interesting, though, is what happened once Chairman Bernanke commenced his “taper talk” in May 2013. The Treasury bond yield rose as did market volatility, suggesting lower equity prices, subject to earnings growth expectations. Earnings growth was certainly positive in 2013, but the growth rate was decelerating a little from the blistering pace in the first few recovery years. Equity prices, however, appeared to totally ignore the coming end of QE and the rise in Treasury yields. The S&P500 Index went from new record high to new record high over the whole second half of 2013. The performance of U.S. equities in the second half of 2013, in the face of higher bond yields, modestly increased volatility and slightly slowing observed trailing earnings growth, strongly suggests that there was a very powerful signaling effect in play. Namely, a decision by the Fed to exit its QE program would imply that the Fed had finally come to conclusion that the U.S. economy was strong enough to grow on its own without emergency and extraordinary assistance from the Fed. If this interpretation represents the main influence of QE, then the announcement of the various versions of QE-2, MEP and QE-3 may also be interpreted as having had a depressing influence on economic confidence, and the economy might well have been better off if there had never been the extended and expanded QE programs once the economic recovery had commenced. For sure, even though a private sector jobs recovery was well under way, major U.S. corporations went on a cash hoarding spree during the first few years of economic recovery. We would argue that one of the primary incentives to hoard cash came from the negative signaling by the Fed that it thought the economy was so fragile it could easily collapse back into recession and needed an extraordinary, never-before-tried, emergency life support.

The parallel perspective that QE-2, MEP and QE-3 did not create jobs is given more credence by a look into the details of the labor market data. Under Chairman Bernanke, the Fed appeared to be mostly focused on lowering the unemployment rate and seeing improvements in monthly net new job creation – the payroll employment data.

While the overall picture is one of a more sluggish jobs recovery than in previous cycles, going well back in time, that perspective changes when one divides the employment data into private sector jobs and government jobs. Indeed, private sector jobs in 2010–13 recovered at more or less the same pace as in 2002–06.

What was strikingly different in the post-financial crisis period was the behavior of government jobs, driven mostly by state and local governments and to a lesser extent by the U.S. postal service. From the end of 2001 through mid-2007, government jobs grew by 852 million. From April 2009 through March 2013, 817 million government jobs were lost. That is, the activity in the government job market was a huge setback for job creation in the post-crisis recovery.
During the housing boom period, property taxes were rising and supporting increased expenditures by local governments in the U.S. Financial recessions are, however, quite different from more typical cyclical recessions due to the role of debt and significant deleveraging that occurs after the crisis.11 In this case, the high levels of debt were concentrated in the home mortgage market, and the recession saw housing prices plummet. This meant that state and local government revenue expectations had to be reset at much lower levels. Hence, there was substantial retrenchment in government jobs, taking several years to run its course. We would argue that the Fed asset purchases did absolutely nothing to help state and local governments, and that private sector job growth was relatively robust, supported by strong corporate profit growth. There was no need for QE-2, MEP or QE-3, and other than temporarily reducing long-term bond yields in 2012 and early 2013, there was little achieved — although the price to be paid in the future may involve some rather serious unintended consequences.12

2.3 Unintended consequences
Regardless of one's opinions of the success or failure of the Fed's QE programs in meeting their objectives, the Fed's balance sheet and financial situation is profoundly different in the aftermath of the QE experiment. As of the end of 2013, the Fed's total assets stood near U.S.$4t, compared with U.S.$850b at the end of August 2008 before the financial panic began. This expansion took the size of the balance sheet from 6% to almost 25% as a share of U.S. nominal GDP.

Within that almost U.S.$4t are U.S.$580b U.S. Treasury securities of 10 years or longer remaining maturity as well as U.S.$1.5t in long-term MBS. The interest rate risk embedded in this portfolio is exceptionally large. By illustration, a 10-year Treasury zero-coupon bond equivalent would lose 17.5% of its value if market yields went from 3% to 5%. The Fed stands to have hundreds of billions in unrealized losses should inflation pressures rise and bond yields track upward as they often do.

There is the likelihood that the Fed will face growing scrutiny from the U.S. Congress over the size of its balance sheet, role in the economy and its plans to exit QE. And, part of the interest from the U.S. Congress will be due to the impact on the U.S. budget. Once the Fed began the massive expansion of its balance sheet, its earnings also went up sharply. The Fed contributes most of its earnings to the U.S. Treasury, and this helps in deficit reduction. Prior to the financial panic of 2008, the Fed often contributed around U.S.$20b or a little more to the U.S. Treasury each year. In the last few years, those numbers have been much higher — U.S.$76.9b in 2011, U.S.$88.9b in 2012, and U.S.$79.6b in 2013, for example.

The Fed may choose not to consider the impact on the U.S. budget deficit of any decision to start increasing the target federal funds rate from its near-zero level held from late 2008 into 2014. Nevertheless, a decision to raise the target federal funds rate potentially will have a profound impact on Fed earnings. On the expense side of the ledger, the key item for the Fed is interest expense. The Fed pays zero interest on the U.S.$1t or so cash outstanding, but it does pay 25 basis points on the required and excess reserves of the banking system (i.e., the federal funds deposits held by the banking system on account at the Fed to satisfy reserve requirements or in excess of those requirements). Under a longer-term scenario where the

11 Reinhart, C. M., and K. Rogoff, 2009, This time is different: eight centuries of financial folly, Princeton University Press.
core inflation rate rises to 2.5%, if economic growth was positive, the Fed might choose to raise its federal funds target to roughly equal the core inflation rate. Under this circumstance, the Fed would probably have to raise the rate paid on reserves to enforce the new federal funds target rate. Thus, the Fed might see a U.S.$60bn or more increase in its interest expense on its U.S.$3tn of required and excess reserves. At the same time, the Fed would be getting the old, low coupons from its holdings of longer-dated Treasury securities and MBS, with some small increases in rates received on the short-term bills it might be rolling over. So even before considering the potential unrealized losses on the portfolio, the carrying costs would be zooming upward.

Fed accounting does, however, consider unrealized losses, which are taken as a charge against earnings for the purpose of calculating the contribution to be made to the U.S. Treasury. Put another way, if inflation pressures emerge, if the Fed responds by raising the target federal funds rate and if bond yields rise with inflation pressures, then the Fed’s contribution to the U.S. Treasury is likely to be sharply diminished, and even a zero contribution is conceivable for several years.

Symbolically, and possibly of considerable interest to the U.S. Congress, there is also the probability, under certain scenarios for core inflation and bond yields, that the Fed’s unrealized losses on its portfolio could exceed its U.S.$62bn paid-in and equity capital by a substantial amount, making the Fed technically insolvent. However, as Professor Bernanke noted in his research on Japan in the 1990s, central banks can create money and manage monetary policy regardless of whether they are technically insolvent or not. Unlike the economics, the politics of a technically insolvent central bank might not be nearly so benign.

The guidance from the Bernanke Fed was that it planned to hold its long-term securities to maturity, so any losses would remain unrealized. There is new leadership at the Fed, with Chair Yellen having taken the gavel on 1 February 2014, so plans may change. And, the Fed has considerable discretion as to how fast it might want to shrink its balance sheet even without selling its longer-term Treasury securities and MBS. As already noted, when the Fed starts to raise its target federal funds rate, it may need to pay higher rates on excess reserves to achieve its new target rate in the marketplace. But, the Fed can also increase the speed it drains excess reserves from the system. The faster it drains reserves, the less its interest expense. The Fed can drain reserves rapidly if it does not replace or roll over short-term maturing bills and notes, if it uses reverse repurchase agreements, and if it chooses not to reinvest MBS interest and Treasury coupons received. All of these complicated questions surrounding the QE exit strategy will come into sharp focus as soon as the Fed even contemplates raising its target federal funds rate.

3. European Central Bank
The ECB faced immediate liquidity issues in its financial system during the 2008 crisis, but then also had to deal with the additional pressures on an undercapitalized banking system during the sovereign debt crisis in 2011 and 2012. The ECB chose to focus most of its balance sheet expansion on liquidity loans to the financial system to preserve the banks’ ability to survive the sovereign debt crisis with less of an emphasis on outright sovereign debt purchases, although some debt was bought. The largest motivating factor for the ECB was probably the “do anything it takes” attitude to preserve the euro as the single currency for the subset of nations in the currency zone of the European Union (E.U.). After all, the ECB was explicitly created to be the central bank for the euro.
3.1 Providing banking sector liquidity
The ECB’s choice to lend to banks rather than to buy distressed assets from banks as the Fed did was classic central banking. Serving as a lender of last resort to secure the safety and integrity of the whole banking system is one of the primary purposes of central banks.

Notably, the U.S. Fed had been created in 1913, following the Panic of 1907, because at the time the U.S. had no central bank and no method for providing the liquidity to banks to stop a run on a few banks from cascading into a run on the whole system. The first big crisis after the Fed was created came with the stock market crash of 1929, and the Fed totally failed in its mission, did not serve as a lender of last resort and allowed a stock market crash to bring down the banking system and the economy, with the Great Depression as its result. Unlike the ECB, though, as the Financial Panic of 2008 emerged, Fed Chairman Bernanke went down the path of buying distressed assets instead of making emergency liquidity loans.

There are similarities and differences in the two approaches. Buying assets immediately shrinks the size of the balance sheet of financial institutions, providing them cash as well as reducing their capital needs. Lending to banks solves the immediate liquidity challenges, but does not relieve banks of their distressed assets or reduce their balance sheets and capital needs. That is, both methods address the immediate liquidity crisis, but have different impacts on balance sheet sizes and capital needs.

We would argue that taking a U.S.$1t of distressed assets and exposures out of the system, as the Fed did in its QE-1 program, turned out to help the U.S. banking system rebound faster since it also reduced the size of bank balance sheets. The ECB’s method of providing low-interest term liquidity loans did not get the distressed assets off the balance sheets and left European banks in a very difficult capitalization posture. This had economic consequences, because the longer banks remain undercapitalized, the less new lending they can do. And, economies simply do not grow in a healthy manner without a well-functioning financial system.

The implications for exit strategies differ, too. As banks regain financial health, even if it takes longer, they can repay their loans, thus shrinking the balance sheet of the central bank in a quite natural manner and with no issues about potential portfolio losses at the central bank. This is what started happening with the ECB in 2013. As banks repaid their emergency liquidity loans, the ECB’s balance sheet started contracting rather dramatically. Of
course, part of the motivation for repaying the loans is the bank stress tests the ECB will conduct in the second half of 2014. Banks perceive that they will perform better on the stress tests if they can show they do not need the emergency liquidity loans. Thus, these tests may work to depress bank lending in the second half of 2014, at least until the ECB announces which banks have passed the stress tests.

3.2 Sovereign debt crisis
The emergence of the sovereign debt crisis in Europe as a follow-on impact of the Financial Panic of 2008 has its roots in two historical developments. First, banks in Europe tend to hold large portfolios of sovereign debt and are also lenders to municipalities for their public development projects. This is not so much the case in the U.S., where banks are not a primary source of funds for the U.S. Federal Government and issuance in the municipal debt market is used to finance local public projects. The role of European banks as long-term holders of sovereign debt complicated the crisis, by tightly linking the need for government fiscal reform with E.U.-wide banking reform.

And second, when the single currency system was created, to entice more members into the Eurozone system, it was agreed that for regulatory and credit risk purposes the debt of any sovereign nation in the E.U. would be treated the same as any other country in the E.U. That is, there would be no capital risk haircut when amassing a portfolio of higher yielding sovereign debt from E.U. nations with weaker economies. We would argue that if the euro was born with a birth defect, it was not the lack of common fiscal policy as typically cited; instead, it was the lack of an E.U.-wide bank capital policy that evaluated risk appropriately and selectively by country.

As the sovereign debt crisis worsened and its scope spread well beyond Greece, the ECB lent money to the central banks of nations in distress as well as bought some of the weaker sovereign debt securities to calm the markets. Since this was essentially a government fiscal crisis, there were limits to what the central bank could do. The task fell to the E.U. finance ministers to hammer out programs in order to bail out countries. And, the individual countries had to take very tough fiscal austerity measures to get their financial houses in order. In the midst of this severe fiscal deleveraging by the weaker countries in the Eurozone, the ECB was mainly focused on avoiding any implosion of the banking system and stabilizing the euro.

We would also note that what sovereign debt securities the ECB did buy were largely of the distressed variety with low prices and high yields. Consequently, as the crisis faded into the background with yields falling and prices rising, and the ECB made a nice profit. We will return to this topic in the concluding section, but it makes a big difference whether a central bank is buying the distressed assets that are part of the immediate problem or whether the central bank is buying the most liquid, low-rate, flight-to-quality assets in the financial system, as the Fed was doing in its QE-2, MEP and QE-3 programs.

3.3 “Do whatever it takes (to preserve the euro)”
Shortly after becoming the president of the ECB, Mario Draghi famously told the markets that the ECB would do whatever it takes to preserve the euro. At this point, the worst was already over for the sovereign debt crisis, but markets were still unsettled. The media and financial analysts were regularly speculating on whether Greece and possibly other countries would drop out of the Eurozone. The Eurozone, however, has no protocol for kicking a country out of the club. And, arguably, Germany’s economy had been one of the largest beneficiaries of the Eurozone, with its implicitly fixed exchange rate removing currency risks from German exports to other E.U. countries. In any case, the commitment to do whatever it takes was credible, and so the ECB did not actually have to do much, and the euro bounced upward and stabilized relatively quickly after President Draghi’s remarks.

4. Lessons learnt (or perhaps not)?
Stick to using QE only for crisis management: given the general consensus that the first round of QE as the Financial Panic unfolded in September 2008 may well have saved the world from another Great Depression compared with the distinct controversies and unintended consequences of later QE efforts by the Fed after economic growth has already resumed, one lesson might be for central banks to only use QE for crisis management.

Have an exit plan: one key lesson for central bankers from the QE era is that large experiments with new methods of conducting monetary policy are likely to have equally large unintended consequences. This does not mean central banks should not be willing to act quickly and aggressively in a financial crisis,
but absent a crisis, and certain during a period of modest, if not robust economic growth, one would hope the lesson has been learned to have an exit plan at the time the new strategy is implemented. The Fed went into uncharted waters when it adopted asset purchase programs even as the economy was already in recovery mode (i.e., QE-2, MEP and QE-3). The Fed faces extremely complex exit issues that may well cause it to delay going back to a more tradition short-term interest rate policy as it faces more intense oversight from the U.S. Congress and sharply diminished net earnings and lower contributions to the U.S. Treasury. By contrast, the ECB’s use of term liquidity loans has a natural exit strategy as banks pay back the loans.

Central bank profits and losses: another interesting observation has to do with nature of the assets being bought in any QE program and the potential for central bank portfolio losses. In QE-2, MEP and QE-3, the Fed bought U.S. Treasuries and MBS when core inflation was close to 1%. If the U.S. economy were to achieve to Fed’s long-term objective of a 2% core inflation rate, and its shorter-term objective of 2.5% core inflation, this would imply rising bond yields and unrealized losses on the Fed’s portfolio. That is, according to the Fed’s business plan, economic success would mean huge portfolio losses on its bond purchases. By contrast, since the ECB primarily used term liquidity loans, and as these are paid back, the ECB faces no portfolio losses. And, when the ECB did buy assets, it tended to buy the weaker sovereign debt securities in the midst of the crisis, so as the crisis faded, the ECB holdings produced a tidy profit. The question facing central bankers is whether any QE program that will generate losses if economic success is achieved is worth the risk to its credibility and political independence.

QE and its role in helping banks recover: the Fed’s choice in QE-1, during the height of the crisis in late 2008, was to buy distressed assets from the financial system and take them out of play. Arguably, the Fed’s initial QE program allowed banks to reduce their balance sheets, lower their capital requirements and recover much faster – leading to faster economic improvement. The ECB’s choice to focus on term liquidity loans meant that it did not provide any assistance to help banks reduce their balance sheets and meet their capital requirements. As a result, the European banking system was much slower to recover and considerably less able to assist in any economic recovery.

The future of central banking: our best assessment is that as more and more research is done to evaluate QE and as the unintended consequences continue to play out, that there may be a diminishing appetite for the Fed’s experiment with QE during times of economic growth, even if the growth is modest. By contrast, the successes of both the ECB and the Fed in containing the Financial Panic in 2008, and later the ECB’s success in helping stabilize the euro during the European sovereign debt crisis, more than justifies future crisis management. If anything, both the Fed and the ECB can look back and take great credit for not allowing the Financial Panic to turn into a prolonged and deep depression.
New results on the correlation problem in operational risk

Vivien Brunel
Head of Risk and Capital Modeling, Société Générale

Abstract
Internal models of operational risk are all built based on the same guidelines provided by the regulators. However, we observe a broad range of practices among banks concerning modeling choices and calibration methods. It is thus relevant to discuss the relative importance of the main drivers and modeling choices of the operational risk capital charge. Many studies in the literature have focused on the modeling of the tails in the severity distributions. In this paper, we use a class of analytical models for operational risk in order to assess the relative importance of all parameters of the model. In particular, we show that the bank's capital charge is not very sensitive to the dispersion in correlations, the average level of correlations being a much more critical parameter of the operational risk capital charge. We show that the assumption of uniform correlations is robust, contrary to what is often advised by internal auditors or regulators.

1 I would like to thank Pavel Shevchenko for the very stimulating correspondence we had together. This article reflects the author’s opinions and not necessarily those of his employers.
New results on the correlation problem in operational risk

1. Introduction
The current regulatory framework allows banks to compute their capital charges for operational risk under an internal model, which is often based on the loss distribution approach (LDA). In this approach, loss distributions are calibrated at the cell level (a cell is the elementary risk unit per business line and type of risk) and the bank's capital charge is estimated by aggregating cell loss distributions under some dependence assumption [Chernobai et al. (2007)].

The Basel Committee [Bank for international Settlements (2011)] provides some guidelines about how banks should appropriately reflect the risk profile in their internal models. However, banks benefit from some flexibility in their modeling choices that may lead to some discrepancies in capital charges for similar risk profiles. The broad range of practices observed among banks results, in particular, from different distributional or dependence assumptions in the models.

Many studies have focused on the modeling of the tails in the severity distributions [Dutta and Perry (2007), Moscadelli (2004)], but the bulk of the correlation problem is still unresolved and controversial. There is a strong debate about the choice of the copula function for losses across cells, since the scarcity of the data makes it quite difficult to solve for this issue. The regulators advise banks to determine sound correlations and to retain conservative assumptions. Some institutions have selected the simplest option and use equal correlations between cell losses. This assumption is, of course, questionable and may embed some model risk, but regulators, as well as practitioners, have great difficulties in asserting arguments about realistic and conservative correlation levels. Some authors believe that correlations between cell losses are as low as 4% [Frachot et al. (2004)].

Most of the knowledge we have about operational risk quantification comes from complex models and heavy Monte Carlo simulations, and, as far as we know, there is no analytical model that takes into account risk and correlation dispersion among cells. This article fills this gap. Under the asymptotic single risk factor (ASRF) assumption, we obtain new results about the bank's capital charge sensitivity to the critical parameters of the model. In particular, we show that the capital charge is not that sensitive to correlation dispersion, and the constant correlation assumption is robust.

This new result is obtained with few specifications, and we conjecture that it remains valid, at least qualitatively, for real bank portfolios that have a finite number of cells. We believe that our approach is also relevant for pioneering a new way to compute capital charges and challenge internal model assumptions as exemplified in this paper.

This paper is organized as followed. In section 2, we provide some real data evidence about cell loss distributions and correlations. We will also solve the ASRF model with lognormal losses at the cell level, even when individual cells have various risk profiles. In section 3, we solve the case of non-equal correlations between cells and provide some key results about the capital charge sensitivity to the main critical parameters of the model. Section 4 concludes the paper.

2. Some empirical facts about cell loss distributions and correlations
In the LDA framework, the aggregate operational loss for cell number $i$ is equal to the sum of individual losses:

$$L_i = \sum_{n=1}^{N_i} X_i^n$$

where $L_i$ is the aggregate loss of cell number $i$, $N_i$ is the number of events over 1 year, and $(X_i^n)_{n=1}^{\infty}$ is the sequence of the individual loss severities for cell number $i$. The aggregate loss process is a compound Poisson process, and accordingly, the model is based on the following assumptions:

- The number of events and severities are independent
- Severities are independent and identically distributed random variables

2.1 Cell loss distribution parameters
Concerning loss distributions, there exists a number of studies that look at individual loss distributions [see for instance Dutta and Perry (2004) and Moscadelli (2004)], but there are very few empirical studies about aggregate cell losses. We aim to fill this gap here.

We have conducted our study based on the SAS OpRisk Global database, which as of November 2013 included 6,402 events that have occurred in financial firms since 2002, the date from which financial institutions started collecting and reporting
their operational losses systematically. We have calibrated the frequency of events and lognormal severity distributions for each of the 21 cells that have more than 30 losses. Direct calibration of the aggregate loss distribution from real data is, of course, impossible because there is only one observation per year. However, it is possible to assess the compliance with the lognormal distribution of the aggregate loss distribution obtained through the LDA.

Let us consider that the loss distribution for cell number i is lognormal with parameters $\mu_i$ and $\sigma_i$; the ratio between the expected value and any quantile depends only on the parameter $\sigma_i$:

$$\frac{\text{Expected value } (i)}{\text{VaR}_{q}(i)} = e^{\frac{q}{\sigma_i}}.$$  \hspace{1cm} (2.1)

$$\sigma_i = F_{q} \cdot \sqrt{F_{q} + 2 \ln \frac{\text{Expected value } (i)}{\text{VaR}_{q}(i)}}.$$  \hspace{1cm} (2.2)

where $\text{VaR}_{q}$ is the q-percentage of the lognormal distribution and $F_{q} = N\left(1 - q \right)$. Inverting equation (2.1) leads to two different solutions; we have chosen the one with a minus sign in front of the square root in equation (2.2) because we require the parameters $\sigma_i$ to decrease with the ratio of expected value to quantile for all cells. We observe that broader distribution assumptions for cell losses in the model can naturally be taken into account by choosing the plus sign solution in equation (2.2).

The LDA leads to the following ratios for each cell in the tail of the loss distribution ($q > 95\%$). For several values of the confidence level, Table 1 provides the observed average value and standard deviation of parameters $\sigma_i$ over all cells, implied from equation (2.2).

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Average</th>
<th>St.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>98%</td>
<td>41%</td>
</tr>
<tr>
<td>97.5%</td>
<td>99%</td>
<td>39%</td>
</tr>
<tr>
<td>99%</td>
<td>107%</td>
<td>44%</td>
</tr>
<tr>
<td>99.5%</td>
<td>112%</td>
<td>46%</td>
</tr>
<tr>
<td>99.9%</td>
<td>124%</td>
<td>48%</td>
</tr>
<tr>
<td>All</td>
<td>107%</td>
<td>42%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Average</th>
<th>St.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>98%</td>
<td>41%</td>
</tr>
<tr>
<td>97.5%</td>
<td>99%</td>
<td>39%</td>
</tr>
<tr>
<td>99%</td>
<td>107%</td>
<td>44%</td>
</tr>
<tr>
<td>99.5%</td>
<td>112%</td>
<td>46%</td>
</tr>
<tr>
<td>99.9%</td>
<td>124%</td>
<td>48%</td>
</tr>
<tr>
<td>All</td>
<td>107%</td>
<td>42%</td>
</tr>
</tbody>
</table>

Table 1: Parameter $\sigma$ implied value from real data

The range of values of parameters $\sigma_i$ is rather stable when the confidence level changes: the average value over all cells and confidence intervals is equal to 107%, and the observed standard deviation is equal to 42%. To assess the robustness of these estimates, we have computed the median of observed values for the parameters $\sigma_i$, which is equal to 108.5%, and is very close to the average value, and the med-med estimator (median value of the spread with the median) that is equal to 31%, which is lower than the measured standard deviation.

2.2 Cell loss correlations

For most of the studies [see for instance Aue and Kalkbrener (2007), Frachot et al. (2004)], cell loss correlations are generated by the dependence of the number of events between cells rather than the dependence of severities. Under the assumption of lognormal severity distributions $X_i \sim LN(m_i, s_i)$, Frachot et al. (2004) show that the loss correlation between cell 1 and cell 2 is:

$$\text{corr}(L_1, L_2) = \text{corr}(N_1, N_2) \cdot e^{\frac{r_1 - r_2}{\sigma_1 + \sigma_2}}.$$  \hspace{1cm} (3)

The correlation of the number of events is linked to the loss frequencies of cells 1 and 2. Bivariate Poisson variables are obtained by considering three independent Poisson variables $Z$, $Y_1$, and $Y_2$ with parameters $r$, $\lambda_1$ and $\lambda_2$, respectively; the variables $N_i = Z + Y_i$ are also Poisson with intensities $\lambda_i$, and their correlation is equal to:

$$\text{corr}(N_1, N_2) = \frac{r}{\sqrt{\lambda_1 \lambda_2}} \leq R = \min \left( \frac{\lambda_1 \lambda_2}{r}, \frac{1}{\lambda_1} \frac{1}{\lambda_2} \right).$$  \hspace{1cm} (4)

The upper bound R for the correlation comes from the inequalities $\lambda_i \geq r$ and $\lambda_i \geq r$. Aue and Kalkbrener (2007) observed that frequency correlations were about 10%, and higher correlations were specific to some couples of cells only. Frachot et al. (2004) claimed that loss correlations were as low as 4%; we recover this result when we take $\text{corr}(N_1, N_2) = 38.5\%$ (which is the expected value estimated by Brunel (2014)) and $s_1 = s_2 = 1.5$ (which is the lowest value observed by Frachot et al. (2004) for these parameters) in equation (3). From SAS OpRisk data, we observe that the parameters $s_i$ have an average value equal to 2.03, a standard deviation equal to 0.42 and are ranged between
1.34 and 2.90. Correlation upper bounds can be computed with equations (3) and (4) from this data. We find an expected value equal to 1.33% and a standard deviation equal to 1.61%. Correlation upper bounds all range between 0% and 4%, except a few of them; the highest correlation has an upper bound at 11.27% and is found between “Execution, delivery and process management” and “Internal fraud” cells of the retail brokerage business line. All these studies confirm that we expect low levels of correlation between cells.

2.3 Correlation parameters in the Gaussian copula model
In the Gaussian copula framework with lognormal marginal cell losses, the correlation parameter $\rho_{i,j}$ between two cells is related to the cell loss correlation:

$$\text{corr}(L_i,L_j) = \frac{e^{\mu_{i,j} - 1}}{\sqrt{(e^{\sigma_i} - 1)(e^{\sigma_j} - 1)}}$$  \hspace{1cm} (5)

This formula with parameters $\sigma_i = \sigma_j = 107\%$ and a conservative assumption for loss correlations $\text{corr}(L_i,L_j) = 4\%$, leads to a correlation parameter of the Gaussian copula equal to $\rho_{i,j} = 7.2\%$. External data support the assumption of very low correlation parameters in the copula framework, much lower than 10%.

3. Critical parameter analysis of the capital charge
3.1 A simplified LDA model
In the rest of this article, we build a simple portfolio model for operational risk. We assess that the bank’s operational risks is a portfolio of $N$ operational risks at cell level. We make the following four assumptions:

- Lognormal distributions: the loss for cell number $i$ is a lognormal random variable $L_i$ with parameters $\mu_i$ and $\sigma_i$. As shown in section 2, we assume that the $\sigma_i$ parameters have an expected value equal to $\sigma = 107\%$ and a variance $\nu = 18\%$ (except in section 3.2).
- Gaussian copula: pair-wise correlations $\rho_{i,j}$ may be different from each other. For numerical estimations, we assume that the average correlation is equal to 10% (this is a conservative assumption as seen in section 2.3).
- One factor model: cell losses are sensitive to the same systemic factor called $F$. This factor is assumed to be a standard normal random variable. The specific part of the risk is embedded in another independent normal random variable called $\epsilon_i$ (i = 1, ..., N). Systemic and specific factors are all assumed to be independent from each other.
- We assume that the parameters are not dependent on the number of cells $N$.

In this framework, the annual loss for a cell can be written as the exponential function of a normal random variable, which is a linear combination of the systemic and specific factors. We get for cell number $i$ (i = 1, ..., N):

$$L_i = e^{\mu_i + \beta_i F + \sqrt{\sigma_i^2 + \nu^2} \epsilon_i}$$  \hspace{1cm} (6)

The parameters $\beta_i$ are linked to the pair-wise correlations of the Gaussian copula: $\rho_i = \beta_i \beta_j$. Because cells may have very different risk characteristics, and because correlations may be very different for different pairs of cells, we assume that the parameters $\mu_i$, $\sigma_i$ and $\beta_i$ are the observations of i.i.d. random variables called $M$, $\Sigma$ and $B$, respectively. In the limit $N \to \infty$, the bank’s loss is equal to $N(L(F)$ and is a function of the common factor $F$, as in Vasicek’s model for granular homogeneous loan loss distributions [Vasicek (2002)]:

$$L(F) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} L_i = E[e^{\mu_i + \beta_i F + \sqrt{\sigma_i^2 + \nu^2} \epsilon_i} | F] = E[e^{\mu_i} e^{\beta_i F + \sqrt{\sigma_i^2 + \nu^2} \epsilon_i} | F] = E[e^{\theta X}]$$  \hspace{1cm} (7)

We are now going to study the following cases for the random...
variables $\Sigma$ and $B$:

- Case 1: the parameters $\sigma_i$ are normally distributed, i.e., $\Sigma \sim N(\sigma, \nu)$, and correlations are constant, i.e., $B = \sqrt{\rho}$.
- Case 2: the parameters $\sigma_i$ are constant, i.e., $\Sigma = \sigma$, and correlations are normally distributed, i.e., $B \sim N(\sqrt{\rho}, w)$.
- Case 3: the parameters $\sigma_i$ are constant, i.e., $\Sigma = \alpha$, and the variable $B$ is uniformly distributed between $\sqrt{\rho} - \sqrt{3w}$ and $\sqrt{\rho} + \sqrt{3w}$ (bounds are chosen so that the expected value and the variance are equal to $\sqrt{\rho}$ and $w$, respectively).

3.2 Heterogeneous risk and identical correlations

The parameter $\sigma$ drives the average risk of each cell and is, of course, a critical parameter of the model. We will show here that the bank’s capital charge is also very sensitive to the dispersion of the risks at cell level (measured by the parameter $\nu$) and to the correlation parameter $\rho$.

We assume that we are in case 1 and, for numerical computations, that $\sigma = 107\%$. In the limit $N \to \infty$, the bank’s loss can be computed exactly from equation (7) as a Gaussian integral [Brunel (2014)]. In Figure 1, we plot the bank’s capital charge as a function of the cell risk dispersion parameter $\nu$ (the value of the correlation parameter $\rho$ is equal to 10%). We see that the parameter $\nu$ is a critical parameter of the model because the bank’s capital charge is very sensitive to cell risk dispersion.

In Figure 2, we plot the capital charge impact as a function of the correlation parameter $\rho$ compared with the reference capital charge, which is computed with $\nu = 42\%$ and $\rho = 10\%$.

Surprisingly, we observe that the capital charge is almost an affine function of the correlation parameter, which could not be inferred a priori from the closed-form formulas (see Brunel, 2014). The bank’s capital charge is a function $K(\sigma, \nu, \rho)$ of the parameters of the model, and we have the approximate formula:

$$ K(107\%, \nu, 10\%) \approx K(107\%, \nu, 10\%) \cdot S(\nu) \cdot (\rho - 10\%) $$

For $\nu = 42\%$, the shape coefficient $S(\nu)$ is equal to 6.25, and this linear formula is very accurate ($R^2 = 99.97\%$). We can observe that linearity is maintained in a wide range of values for the parameter $\nu$ and remains very accurate ($R^2 \approx 99.5\%$) in the range of values we have studied for the parameter $\nu$; we plot the function $S(\nu)$ in Figure 3.

The shape function $S(\nu)$ is increasing, meaning that when cell risk dispersion is getting higher, the correlation parameter is getting more and more critical.
New results on the correlation problem in operational risk

3.3 Uncertain correlations
Correlations are not identical to each other, as illustrated in section 2, but estimating them from real data is a challenge from a statistical viewpoint. Data is scarce and limited only to one observation per year for the aggregate loss. Estimation of the number of events correlation is no longer robust for the same reason and severity correlations are only observable for cells that exhibit a sufficient number of events per year. Assuming identical correlations among cells is a current practice even if, in reality, correlations are unknown parameters. In what follows, we remain in the limit $N \to \infty$ and correlation uncertainty is included in the model by assuming that the random variable $B$ has an expected value equal to $\beta = \sqrt{\rho}$ and a variance equal to $w$. For the sake of clarity we assume that individual risks are all equal among cells, i.e., $\Sigma$ is a constant equal to $\sigma$.

As pair-wise correlations are equal to $\rho_{ij} = \beta_{ij}$, there is a direct link between the variance of $\rho_{ij}$ and the variance of the sensitivity parameters $\beta_{ij}$. Because of the independence of the $\beta_{ij}$, we can write:

$$\text{var} (\rho_{ij}) = E (\beta_{ij}^2) - E (\beta_i) E (\beta_j) = w (w + 2\beta^2)$$

This leads, by solving the second order equation in $w$, to:

$$w = \sqrt{\beta^4 + \text{var} (\rho_{ij})} - \beta^2$$  (8)

For $\beta^2 = \rho = 10\%$ and $\sqrt{\text{var} (\rho_{ij})} = 3\%$ (which is a conservative value compared to what is measured from observed data; see section 2.2), we have $w = 0.44\%$, i.e., the standard deviation of the parameter $\beta$ is equal to 6.6%.

The ratio of the capital charge including model risk ($w > 0$) to the capital charge without model risk ($w = 0$) measures the increase in capital due to dispersion or uncertainty on correlations. We plot this quantity in Figure 4 as a function of the mean deviation of the correlation parameter ($\sqrt{w}$) in cases 2 and 3.

The curves corresponding to cases 2 and 3 are close together: the shape of the distribution function for the random variable $B$ is not a critical choice. Moreover, we show that the impact of the mean deviation of the correlation parameter is lower than 2% for $\sqrt{w} = 6.6\%$ in both cases.

Even with a much more conservative choice for the individual cell risk parameter $\sigma = 200\%$, the impact of correlation dispersion on the bank's capital charge would be about $+5\%$. Our conclusion is that correlation dispersion (measured by parameter $w$) is, by far, not as critical as the other parameters of operational risk models (average cell risk $\sigma$, cell risk dispersion $\nu$ and average correlation parameter $\sqrt{\rho}$).

4. Discussion and conclusion
This article explores the relative importance of parameters and assumptions in operational risk capital models. In particular, we obtain some new results on the criticality of the risk parameters that are driving the capital charge and on the correlations.

Our approach is based on a class of simplified analytical models that incorporate dispersion in individual cell risks and correlation levels. The main finding of this paper is that uniform correlation is a robust assumption for capital charge modeling. This result is important because it means that model risk associated with the value of correlations is not a major issue for capital measurement: the assumptions related to the calibration of individual cell risk or to the average level of correlations is much more critical. Moreover, regulators often challenge internal models on their accuracy and require refined and complex models. This requirement is relevant when modeling individual
cells loss distributions. This paper shows that this is not the case for correlations. Differentiated pair-wise correlations have a less significant impact on the capital charge than uniform correlations, and generate more model risk in the calibration.

Concerning operational risk, there are still other unanswered questions that have as yet not been considered in the academic literature and for which no theoretical basis has been established. Among these issues, we can mention the number of cells (business lines and types of risks) problem. Because the calibration of the loss distribution is done independently across cells and cells risks are aggregated in a second step, the global calibration of the model is not based on a portfolio approach. As a result, the capital charge depends on the number of cells retained in the model whereas the global risk of the bank is independent of any risk classification. This is an open field of research that could have concrete practical implications on the design of internal models for operational risk.

These kinds of theoretical studies are necessary for banks when they are negotiating with the regulators to obtain approval for their internal models. Indeed, in many circumstances, data is too scarce to provide by themselves a formal proof of the model assumptions or methodological choices. Theoretical arguments associated with observed data (external or internal) are a powerful way to assess the robustness of the models.

References
Financial perspective: the unintended consequences of regulatory oversight and control – lessons from the banking and the asset/alternative funds industries

Frederick C. Militello, Jr.
Adjunct Professor of Finance; Global and Investment Banking, Leonard N. Stern School of Business New York University and CEO and Senior Thought Leader, Future Change Management, LLC.

Abstract
Regulators are largely problem solvers. They seek intended consequences. However, observation indicates that regulatory intentions and outcomes frequently do not coincide. If you will, there is a gap between the two; namely that of unintended consequences. Such consequences are the result of financial managers, and their organizations, facing increased regulatory induced business dilemmas. Such dilemmas must be managed, they cannot be solved. Moreover, they take the form of difficult choices – the essence of strategy – affecting the efficiency, innovativeness and competitiveness of financial organizations. As such, regulation has the unintended consequences of inspiring strategic thinking and organizational differentiation; reducing concentration and systemic risk. However, there is another unintended consequence afoot; namely, the regulatory induced impact on strategy and new business models is giving rise to increased interconnectivity of financial sectors and organizations. Intended consequences in one financial sector are spilling over into others, sometimes intentionally, sometimes not. Financial managers are busy assessing the results of regulation on comparative organizational capabilities, resulting in greater partnering across organizational and financial market segments. Clearly, when it comes to unintended consequences, and their systemic implications, there is a paradox. Regulators are becoming more proactive regarding these unintended consequences; increasing their regulatory reach. However, the paradox of unintended consequences is not a problem to be solved, it is a dilemma that must be managed with interactivity of the public and private sectors.
Introduction
This paper is about regulatory oversight and control, the managerial choices regulatory control and oversight engender and, most importantly, their unintended consequences. The paper utilizes examples from both the banking and the alternative/asset fund industries; however, the lessons learned are readily applicable to the broader financial services industry operating both in and out of the shadows.

It is the recognition of this paper that many of the unintended consequences noted herewithin may manifest themselves with or without regulation and regulatory pressures; however, the urgency and intensity of managerial choices – and their unintended consequences – increases in an environment of regulatory oversight and control. Moreover, it is the observation of this paper that many of the managerial choices today being made – and their unintended consequences – would not surface at all if it were not for the existence of certain regulatory initiatives.

This paper is meant to be forward-looking. It explores such questions as:

▶ What are intended and unintended consequences?
▶ What role do regulators and financial managers respectively play in their emergence?
▶ What are the regulatory drivers of unintended consequences?
▶ What are the managerial dilemmas or choices that give rise to unintended consequences?
▶ What unintended consequences can be observed in the asset/alternative fund industry; and, how do they carry over from/to the banking and broader financial services industry?
▶ What opportunities do unintended consequences offer both regulators and financial managers to work together for the improvement in the stability of the financial system and improved business performance?

Consequences — meaning and scope
Before delving into examples of the unintended consequences of today's regulatory environment, let us explore the meaning, business context and implications of consequences, both those that are “intended” and those that are “unintended.”

Consequences\(^1\), in and of themselves, are not categorically good or bad. Rather, they are a direct result of the problems we need to solve and the dilemmas, or extreme choices, we need to manage as a result of an increasingly complex and interconnected world.

Certainly, over the years there has been much debate about the meaning of unintended consequences – the subject of this paper. Economists seek to capture its intent in what has come to be called the “Law of unintended consequences.” However, while this “Law” makes it abundantly clear that such consequences are neither good nor bad there is a lack of definitional clarity and an assumed bias, by those that seek to apply it, toward explaining the perversity of outcomes.\(^2\) This paper seeks to realign this bias with a more balanced view of the economic and financial outcomes caused by the public and private sectors relative to the realities of today's regulatory initiatives, and offers a clearer description through definition and example – as to the precise meaning and emergence of unintended consequences. More specifically, in the context of this paper, unintended consequences are the results of how we manage regulatory induced business dilemmas – extreme but interdependent choices – from an industry, organizational and professional perspective.

Building a framework — regulators and financial managers
There is a regulatory process afoot – see Figure 1 on the regulatory cycle. It needs to be fully understood if regulators and financial managers are to work together in adding appropriate value to the soundness of the financial system and providing requisite business performance. The process demonstrates the roles and responsibilities of both regulators and financial managers and how such roles and responsibilities give rise to consequences, both intended and not. It indicates a logical flow of progression from regulatory initiatives, to intended and unintended consequences and then back again to regulatory measures.

---

1 Consequence: something that happens as a result of a particular action or set of conditions: something produced by a cause or necessarily following from a set of conditions, Merriam Webster Dictionary, www.merriam-webster.com
2 “The law of unintended consequences, often cited but rarely defined, is that actions of people—and especially of government—always have effects that are unanticipated or unintended. Economists and other social scientists have heeded its power for centuries; for just as long, politicians and popular opinion have largely ignored it. The concept of unintended consequences is one of the building blocks of economics. Adam Smith’s ‘invisible hand,’ the most famous metaphor in social science, is an example of a positive unintended consequence.... Most often, however, the law of unintended consequences illuminates the perverse unanticipated effects of legislation and regulation. The Concise Encyclopedia of Economics, www.econlib.org
initiatives; with the latter opening possibilities for additional regulatory scrutiny (or not) and new dimensions of regulatory control (or not). Let us look at the process more closely and the roles and responsibilities implied by it.

Specifically, for purpose of discussion, the process begins with the regulators. Regulators by and large approach the regulatory process from the perspective of solving a problem, or series of problems having one or more correct and independent solutions. For example, under Basel III the requisite amount and composition of Tier-1 capital is prescribed. Banks are given a number of acceptable ways they can comply with the regulatory capital requirement, e.g., common equity, retained earnings, qualified minority interests. In other words, from a problem-solving perspective regulators provide a single or series of independent answers as to how a bank can meet its Tier-1 capital requirements. Moreover, there is an intended consequence at work here; namely, to foster a financial system supported by a specified amount of regulatory capital with a higher quality risk absorption capacity. No doubt, regulations as problem-solving initiatives have an important role to play. They allow us to know what the rules are, putting us all on the same page of understanding and communication. But they are only one part of the consequences process or equation – the intended part.

In contrast to the above let us look at financial managers; and the second part of the regulatory process or consequences equation – the unintended part.

Total consequence (TC) = Intended consequences (IC) + Unintended consequences (UC)

No doubt, financial managers must solve problems; but, for purpose of this paper, the greatest challenges they face are the significant dilemmas that arise from prescribed regulatory actions.³ For example, what dilemmas – extreme but interdependent choices – do we face as financial managers when trying to comply with Basel III and its Tier 1 capital bolstering regulations? How do we manage those dilemmas knowing each will give rise to largely unintended consequences?

For this part of the regulatory process, there is no single or correct answer. We are faced with regulatory induced dilemmas. For example: do we sell-off what we believe to be non-core assets to raise capital and abandon our strategy of one-stop shopping? Do we rationalize/cull customer segments to optimize capital allocation and damage relationships that took years in the making? Do we begin to favor cost-cutting over growth initiatives, focusing more on capital efficiency and less on capital effectiveness? Do we begin to focus more on customer needs/product suitability and abandon our emphasis on product sales/increased market share?

Each of these questions expresses a regulatory-induced dilemma to be managed, but how each interdependent dilemma is managed is what actually gives rise to unintended consequences. Regulators have traditionally paid far less attention to the business dilemmas – extreme choices – that arise from complying (or attempting to comply) with their regulatory initiatives; and, similarly to the way these dilemmas – when managed – give rise to unintended and unforeseen business practices.

³ The inferences in this paper to problem-solving and the management of dilemmas has been drawn from, Johnson, B., 1998, “Polarity management: a summary introduction,” Polarity Management Associates.
Put differently, regulatory oversight and control has two goals. On the one hand, there is need “to significantly bolster the safety and soundness of the financial system,” and on the other hand, there is need to do so while “preserving as much as possible of the industry’s efficiency, innovativeness and competitiveness.” This is not a “this versus that” problem to be solved; rather it is a “this and that” dilemma to be managed. In essence, the regulatory process is a balancing act of sorts between intended and unintended consequences with regulators focused more on the former and financial managers the latter.

**Regulatory drivers**

Clearly, in the above discussion the regulatory driver of the observed consequences was one of a prescribed quantity and quality of regulatory capital. But there are other regulatory drivers that influence managerial choices and business outcomes, such as liquidity, stable-funding and financial leverage requirements.

Most often, such drivers have consequences in the same financial sector. This has just been illustrated by the Basel III banking example. However, we increasingly find that intended consequences in one financial sector often spill over as unintended consequences into another as financial managers make choices affecting the way such regulatory drivers affect operational efficiency, innovativeness or competitiveness.

By way of example, as regulators increasingly squeeze large banks to retreat from making what is perceived to be less liquid or more risky loan commitments, such intended consequences are manifesting themselves in new opportunities and challenges in the asset/alternative fund industry; or by way of extension between the banking sector and those financial organizations operating in the shadows of banking. Clearly, the interconnectivity of consequences across financial sectors is an unintended consequence of regulatory actions. Such consequences are not necessarily bad or good; however, one thing is for certain that regulators will be assessing those consequences for implications of further regulation.

Importantly, regulatory drivers come in other forms too. They are not simply found in financial ratios or prescribed rules of reporting and behavior. For example, one of the key regulatory drivers noted by alternative asset fund managers is the cost of regulatory compliance, ranking only second to demands for greater transparency. The increased cost of compliance in a performance-based industry can be a tremendous force in creating management dilemmas and unintended consequences. Such performance pressures are only intensified under a regulatory regime that comes with significant and rising compliance costs, not to mention the uncertainty of further regulation. Here too there are unintended consequences, such as the transforming impacts on business models and the furthering of implications for greater interconnectivity of the global financial system.

It is important to stress that unintended consequences are part and parcel of the process of shaping the financial world we live in. Financial managers must be interactive participants in that process. Certainly, regulators may largely choose not to focus on the business dilemmas we face as a result of having to comply with their initiatives. However, regulators do keep a keen eye open for unintended consequences and their implications for the furthering of their regulatory reach.

**Unintended consequences — the rubber hits the road**

Regulatory business strategy is not a matter of “this versus that” thinking; rather, it is representative of the growing tide of “this and that” business planning and competitive business differentiation, and the role it is playing in some of the most innovative and competitive organizations in the world. Put another way, there is an increased need by financial managers to look at regulatory induced dilemmas not as problems to be solved but instead as interdependent choices or dilemmas to be managed, and as will be seen, frequently extending their business choices and models to be more inclusive of inter-organizational relationships (including regulatory agencies). While there are many specific and unintended consequences of this observation, one of the single most important is that of increased interconnectivity of organizations and, in many cases, resulting systemic implications.

---


6 “There is significant competitive advantage for organizations that can both solve problems and manage polarities. The research is very clear on this.” Johnson, B., 1998, “Polarity management: a summary introduction,” Polarity Management Associates.
For purposes here, we will focus on those unintended consequences that can be observed in the business models and behaviors of the asset/alternative fund industry and their relationship to the broader banking and financial services industry. There are lessons to be learned here for all financial managers, including asset/alternative fund managers, bankers, investors and those employed directly by the corporate sector.

**Operational efficiency**

No doubt, an important regulatory driver of operational efficiency has been the increased costs of regulatory compliance. In an industry facing greater institutionalization and investor demands for increased performance, rising compliance costs are giving rise to managerial choices that are transforming business models and the landscape that financial organizations do business on. For example, in the asset/alternative fund industry both operating leverage and the institutionalization of the business have resulted in an emphasis on scale economies. This is true for the banking sector, in general, and for much of the asset/alternative fund industry.

More to the point, the costs of being regulatory compliant are driving the trend toward outsourcing as a way of facilitating even greater operational efficiency. Studies show that there is direct link between the rising costs of compliance and the unintended consequence of increased and significant outsourcing in the industry. Moreover, such outsourcing has its own unintended consequences arising from managerial decisions as to what activities are to be outsourced and what kept in-house, and how outsourced activities are going to be managed from a risk perspective? Vendor (or third-party risk) management has become a center-piece of managerial attention at many of the largest financial organizations in the world. This is no doubt fueled by (a) the need to allocate increasingly expensive regulatory capital to core – as opposed to non-core – business activities and (b) the negative impacts that rising regulatory costs have on business performance and rising investor demands for such.

This appears to be an area where the interests of the system may be at odds with the performance expectations and requirements of the industry. While some imply that it is only non-core activities that are being outsourced, leaving managers with greater time and resources for what they do the best. Such actions, nevertheless, only increase the interdependence of the financial industry and its systemic implications. Put another way, the very actions that industry observers seem to think are value-creating – in this case the focus on operational efficiency through the outsourcing of non-core business activities – are the ones that have made the financial system as a whole even more complex and interconnected.

The Federal Reserve Board of Governors and the U.K.’s Financial Services Authority have already weighed in on this shifting business model—no doubt reflecting on the past effects that the outsourcing and the interconnectivity of key business activities had on increasing the fragility of the global banking/financial system. The FSA has been particularly articulate about this trend and especially regarding the asset management industry. Notably: "...Our initial discussions and research have identified that the asset management industry outsources a growing number of activities... Our concern is that if an outsourcing provider were to face financial distress or severe operational disruption, UK asset managers would not be able to perform critical and important regulated activities, thereby causing detriment to customers." They go on to note: “Based on our findings so far we are not confident that across the industry, effective recovery and resolution plans are in place for the asset management sector as a whole... We recognize that there may be more robust contingency plans...we are not yet aware; but in all cases we expect firms to have devised adequate contingency plans which are viable, robust and realistic and set out a clearly defined exit strategy in the event of a termination of outsourced activity under any circumstances, including stressed market conditions.”

---

7 State Street Corporation, 2013, “The next alternative: thriving in a new fund environment”  
8 Ibid.  
9 “Around the globe and across our business lines, we engage about 12,000 vendors for a wide range of products and services – from technology hardware, to cash clearing, to facilities management. Working with third-party providers is often more efficient and cost-effective than performing certain activities in-house. But it also involves risks that must be thoughtfully managed.” Environmental, social and governance report, p. 25, Goldman Sachs & Co., 2011.  
10 State Street Corporation, 2013, “The next alternative: thriving in a new fund environment”  
13 Ibid.  
14 Ibid.
Such letters from the FSA and similarly from the Board of Governors are revealing in and of themselves. For example, regulators have traditionally always been more reactive in their problem-solving role, becoming the norm or expected behavior. But there appears to be a change in the offing; namely, in many cases regulators are getting ahead of their constituencies in the regulatory process, becoming more anticipatory or proactive in their communications. Such pro-activity can be positive; but only if it leads to greater interactivity between regulators and financial managers.

As just noted, regulatory “warnings” are on the rise—anticipating, if not foreshadowing, the unintended consequences of managerial choices yet to be made. With such warnings financial managers will need to act more quickly and be more convincing as to the non-perversion outcomes of managerial choices – some made and some yet to be made. More importantly, it is increasingly important to think of regulatory influence or leverage (the competency to amplify the impact of one’s efforts) more as a core competency of the firm than ever before. This is also an unintended consequence of the increased regulatory world we live in.

**Innovativeness**

Rising compliance costs, the need to make regulatory influence a core competency and demands for increased regulatory interactivity will no doubt continue to demonstrate unintended consequences in terms of operational efficiency. But such an observation carries over into the innovativeness of financial organizations as well. Increased innovativeness should be built on a foundation of transparency and trust. Specifically, studies point out that a commitment to being regulatory compliant, while raising operational costs, can pay off in the ability to attract new capital, and perhaps open the doors to new investors and fund-raising opportunities. Clearly, an unintended consequence of rising regulatory costs has been the increased credibility that such a displayed commitment has had on the investment community.17

Moreover, there is no doubt that increased regulatory oversight and control, especially of the banking sector, has given rise to the unintended consequence of increased innovativeness by the asset/alternative fund industry. Here, regulatory pressures on the banking sector are spilling over into a myriad of funding implications and opportunities for asset/alternative fund managers, and once again demonstrating the interconnectivity of financial businesses and markets. Additionally, we find the unintended consequence of transforming business models, the shifting channels of capital formation and the risk transference of financial assets from the banking sector to capital market investors. Finally, one finds the unintended consequence of increased reliance on business partnerships between financial firms operating in different financial sectors.

**Case one**

Sometimes, as already noted, the consequences of regulations in one financial sector create unintended consequences in another. Innovativeness, in the form of both recognizing and acting upon windows of opportunity, is frequently at the heart of this consequence transference process. For example, tougher capital and liquidity regulations are leading banks to curtail the tenure of their loan commitments. Forcing banks to become more liquid is clearly an intended consequence of Basel III regulatory actions, with such regulatory drivers as liquidity and stable funding ratios shortening loan maturities and reducing lending activities. However, as a result of these intended consequences, we see private equity and asset fund managers seeking to play a larger role in closing the resulting funding gaps – clearly, when looked at entrepreneurially, a window of opportunity. More structurally, such opportunities have unintended consequences on transforming business models and the process of risk transference.

For example, regarding the latter, asset fund managers have been busy establishing infrastructure debt funds.18 Some funds are being launched by the asset management arms of large

---

15 “Due to perceived regulatory failure of the banks during the financial crisis of 2007-2008, the UK government decided to restructure financial regulation and abolish the FSA. On 19 December 2012, the Financial Services Act 2012 received royal assent, abolishing the FSA with effect from 1 April 2013. Its responsibilities were then split between two new agencies (the Prudential Regulation Authority and the Financial Conduct Authority) and the Bank of England.” www.wikipedia.org

16 For a general discussion of such competencies see, Militello, F. C. and M. Schwalberg, 2002, Leverage competencies: what financial executives need to lead, Financial Times/Prentice Hall.


insurers,\textsuperscript{19} while others are being launched by large independent asset management firms. For example, one such infrastructure debt fund initiative was recently launched in late 2012.\textsuperscript{20} These funds invest primarily in investment grade project finance loans and bonds, including primary, refinancing and secondary deals. The funds are marketed to insurance companies and pension funds without the in-house capabilities to invest directly in project finance debt, thus taking advantage of a niche in the marketplace and differentiating themselves from funds owned by major insurance companies. As part of its business model, the fund manager will continue to work closely with its extensive banking relationships both in terms of benefiting from their expertise in structuring infrastructure projects and extending its deal sourcing network of opportunities, demonstrating once again the growing interdependence, or partnership, between financial organizations operating in different regulatory spaces.\textsuperscript{21} Clearly, regulatory induced innovativeness is an unintended consequence with implications again for increased interconnectivity and inter-organizational partnerships.

**Case two**

Regulatory drivers are impacting financial sectors in other ways; here, too, with an emphasis on innovativeness as a primary unintended consequence.

Specifically, recently the U.S. Department of the Treasury issued guidance\textsuperscript{22} to large banks to limit their lendings to businesses that are too highly leveraged. Unlike the previous case, here we find a situation where regulators are extending their reach into supposedly “unregulated” territory, namely private equity (PE) firms. Reportedly, a number of banks, falling under the guidelines, have already pulled out of private equity sponsored transactions partly as a result, with such decisions presumably demonstrating the intended consequences of reigning-in the activities of private equity organization.\textsuperscript{23}

If there should be any doubt, the achievement of such consequences has been boldly articulated by the Office of the Comptroller and Currency (OCC), specifically, its senior deputy controller who reportedly has been quoted as saying,\textsuperscript{24} “The impact on private equity, a significant driver of what we see as risky practices, is an intended consequence of our actions.” The senior official reportedly goes on to say,\textsuperscript{25} “As regulators, we certainly hope to change bad business practices and remove the extraordinary froth that is experienced at the peak of a credit cycle.”

In contemplating these quotes there are clearly some important implications. Perhaps, most notably is the observation that just like financial managers look at regulatory drivers (and the management dilemmas and unintended consequences resulting from such) we now have the regulators looking at their own set of regulatory drivers, with the latter focusing on which financial sectors will be subject (or not) to increased regulatory initiatives, each with their own intended consequences. There is also an element of extended regulatory reach here. Certainly, the impact of the guidelines in the banking sector is intended to have consequences on the private equity business, but the responses made by fund managers and private equity sponsors may have a far greater range of unintended implications on the private equity business model than is expected.

For example, middle market companies may find private equity sponsors less receptive. The cost of financing such deals will almost certainly rise and their valuations are almost certainly to fall. In many cases new sources of financing will need to be found; perhaps creating opportunities for non-bank lenders? Exit strategies, based on high leverage and special dividends, will have to be rethought and there will be a demonstrated transference of credit risk from the banking sector to the capital markets. In some cases this transference will be accomplished in the bond market, which is relatively more credit sensitive and expensive as compared to the bank markets.\textsuperscript{26}

Here we find “intended consequences squared” (IC\textsuperscript{2}) and yet another example of increased regulatory reach. Specifically, a


\textsuperscript{21} CFI.co., 2013, “BlackRock: bridging the gap—the rise of infra funds in privately financed infrastructure,” October 29.

\textsuperscript{22} The guidance was issued (Federal Register/Vol. 78, No. 56) on March 22, 2014—at a US interagency level including the Department of the Treasury, Federal Reserve System, Federal Deposit Insurance Corporation, The Office of the Comptroller of the Currency (OCC).


\textsuperscript{24} Ibid.


\textsuperscript{26} Ibid.
Financial perspective: the unintended consequences of regulatory oversight and control – lessons from the banking and the asset/alternative funds industries

situation where regulators are seeking intended consequences in not just one financial sector but simultaneously in two; and, leaving the unintended consequences of their regulatory actions to be dealt with by financial managers, both in the banking and alternative funds industries. Such a trend, if such be the case, would no doubt increase even further the regulatory burdens placed not only on asset/alternative fund managers but the broader financial services industry in general.

Competitiveness
This paper would be remiss if it did not include a commentary on the rise of strategy – perhaps one of the most debated and unintended consequence of regulatory oversight and control? With regulations on a continued projectile upward, it is not surprising to find that there is currently much debate about strategy in the financial services industry. Regulatory oversight and control impacts both problem-solving and the need to manage interdependent dilemmas or choices, and making choices is the essence of strategy. No doubt, there is much to consider and debate here.

Some argue that there is too much thinking today about strategy, at the expense of organizational innovation. On the other hand, some argue, including myself, that there has been too little strategy and organizational differentiation in the financial sector, with insufficient thinking about customer needs and product suitability. Such voices also hold that it was largely the lack of strategy that contributed to the global financial crisis. If true, then the re-emergence of strategy in the form of organizational differentiation is a mitigating force of systemic risk.

There is certainly much to debate here, but one thing is clear, namely that there most definitely would not be a debate about the existence or non-existence of strategy without regulatory oversight and control. This outcome, perhaps appearing at first rather ironic, is certainly another unintended consequence of a regulatory world.

Summary of key observations
While perhaps not originally intended, this paper has noted that regulators in their quest to mitigate systemic risks in one sector are frequently at the heart of their creation in many unintended financial sectors and practices. There is a paradox of regulation. On the one hand, in a regulatory world financial organizations seem to refocus their attention on strategic differentiation having a mitigating impact on risk concentration and its systemic implications. But, on the other hand, regulatory initiatives, as noted in this paper, have the impact of increasing the interconnectivity of financial organizations and industry sectors. Such regulatory induced interconnectivity only brings to the surface the potential for increased systemic risk and self-perpetuating regulatory reach.

For example this paper has shown:

- Today, the burdens of increased compliance are leading to an emphasis on operational efficiency, through outsourcing and other “partnership” arrangements, and, the emergence of a world of “this and that” thinking and behavior. Partnership initiatives, based on comparative core competencies, are on the rise. Such “partnerships” imply a transformation of organizational business models. These business models have the potential to focus organizational activities on core competencies; but, carry with them implications for increased interconnectivity and systemic risk. All of the above are unintended consequences of regulation.

- Today, financial organizations are recognizing the post-regulatory effects on their comparative strengths and weaknesses. This is true regarding comparative levels of expertise (such as in the structuring of transactions), comparative depths of customer relationships and business networks and comparative capabilities to fund various stages and tenors of projects and business transactions. These regulatory induced reflections certainly point to the ability of financial organizations to engage themselves in increased competitiveness and innovativeness. However, they also bring with them increased interconnectivity and systemic implications, demonstrating the power of regulatory induced unintended consequences.

---

27 Porter, M. E., 1996, “Strategy is the creation of a unique and valuable position, involving a different set of activities. If there were only one ideal position, there would be no need for strategy.” On Competition, HBS Press.


30 This has been observed by others as well. Tisa, S. K., 2014, “Regulators are encouraging shadow bankers,” American Banker, January 15.
Today, innovativeness can also be found in more contentious situations, namely where regulators purposely seek to achieve intended consequences in two or more financial sectors simultaneously. This might be indicative of a trend, whereby regulators begin to place themselves at the interconnectivity of organizations/sectors and seek to target what they consider to be “bad” business practices by financial organizations outside their immediate regulatory control. While possibly being effective, the unintended consequences here can be quite significant, unless offset by the innovativeness of financial organizations. No doubt, much of this interconnectivity of financial organizations and markets may exist with or without regulation. However, one thing is true for sure, the right balance between free enterprise and regulation is not a problem to be solved. There is not a single or series of independent correct answers for any of the observations immediately noted above or throughout this paper. Rather, experience, and hopefully the messages contained herein, teaches us that regulation and freedom are highly interdependent choices that we must actively manage, seeking out the best of each for the benefits of both business performance and the systemic safety of our global financial system. This is the challenge that lies ahead. It is formidable, but optimistically still manageable.
Risk management insights from Markowitz optimization for constructing portfolios with commodity futures

D. Sykes Wilford
The W. Frank Hipp Distinguished Professor of Business and Finance, The Citadel

Abstract
Based upon a Markowitz methodology, this paper considers use of commodities futures as assets to enhance portfolio diversification. A very simplistic methodological approach is chosen to allow various portfolio issues to be highlighted with data from the recent financial crisis. Naïve long-only diversification (with commodities) does not add a great deal of value relative to the event risk that can (and often does) occur. The opposite is concluded for the long/short approach. We find commodities can be key drivers in risk mitigation over time and in crisis conditions when utilized in long/short portfolios. Focus upon crisis risk exposes many of the fallacies inherent to simplistic portfolio creation and management.

1 Although this paper is focused on commodity risk allocations in particular, it is based on ideas resulting from several years of implementing and managing Markowitz-based portfolios. As such, it owes a great deal to others with whom the author has worked over the years. Many are coauthors of referenced papers. Assistance by Bluford H. Putnam and Samantha Azzarello must be acknowledged. The author thanks the CME Foundation for its support of the research and Richard Levich for his comments. All errors are those of the author.
1. Introduction: objectives of the study
Most money manager agree that commodities have been a very useful diversifier over the past few years, or do they? If they do, then when and how should commodities be used as diversifiers? The word diversification conjures up images of Nobel Prizes as well as the lexicon of “don’t put all your eggs in one basket”. Professor Harry Markowitz, the father of modern theories of diversification, is just one name associated with the accepted literature on portfolio diversification. Capital Asset Pricing Models have become de rigueur for all portfolio managers, classrooms and fund advisors. And with the crash of 2008 skeptics have come out blasting the concept of simple Markowitz approaches. Even when unfair, in their assessment of simplistic Markowitz diversification, such banter raises the issue of whether portfolio diversification in general is as important as we (including accepted regulatory guidelines) have been led to believe.

As DeMiguel et al. (2009) have pointed out, naïve models often outperform the most sophisticated forms of portfolio diversification in simulations; thus, when considering how to utilize commodities and commodity futures in a portfolio to achieve diversification one is faced with evidence that simple assumptions about its usefulness must be considered with care. In this research, we revert to the basic Markowitz model (believing that working with the original theory will give us more insights). The goal is to examine whether or not straightforward risk-return optimization that includes various commodity futures as assets will lead to more efficient – better or safer – portfolios (this is not to imply that the Markowitz approach is simplistic but rather that the forms that it takes in practice are often simplified versions of the theory). In doing so, we will focus on some key issues to better understand the pluses and minuses of optimized portfolios, which contain commodity futures, especially for periods of risk, such as the recent crisis when robust diversification was most needed.

Further, focus is on implicit assumptions in basic Markowitz optimization and the potential problems of breaking those inherent assumptions in liquidity-challenged asset classes, especially during a crisis. We examine how history can fool one into accepting diversification’s existence as safety and show that safety can evaporate during a crisis. Finally, the advantages and pitfalls of long-only versus long/short portfolios are considered when allocating to commodity futures.

In conclusion, we recommend that commodities be utilized to diversify risk but only in a long/short context. Naïve long-only diversification (with commodities) does not add a great deal of value relative to the event risk that can (and often does) occur. The opposite is concluded for the long/short approach. We find commodities as key drivers in risk mitigation over time and in crisis conditions when utilized in long/short portfolios.

2. Why commodities: the standard lesson
As is commonly heard on most “talking head” programs about investing, “adding commodity assets to portfolios help diversify risk because of their low correlations to other assets.” Examining this comments more carefully, Table 1 shows a typical correlation matrix, which includes oil, natural gas, gold, copper, corn and wheat, as well as the S&P 500, U.S. Treasuries (the 10 year) and cash, using commodity futures prices for nearby contracts from CME Group exchanges. A casual observation suggests that commodities are low correlation assets, both with each other in many circumstances and with the primary indices for equities and bonds. In Table 2, we break out the relationships with the S&P 500 and the U.S. 10-year Treasury correlations to the various commodities. The data are for the period from 2002 to 2012 (which purposely includes the recent crisis).

In a simplistic way, low correlations are a sufficient condition to suggest that diversification can be achieved by adding commodities to a portfolio. We must, however, actually utilize these data in portfolio optimizations to see if there are other issues that must be considered before we conclude whether what appears as obvious is correct or not.3

---


3 A very thorough discussion of commodity portfolio management is provided by Al Janabi (2012). Whereas Al Janabi is focused on the variance-covariance risk management methodologies utilized within commodity markets, we are ignoring these very important issues to focus on some other, more straightforward problems of simplistic optimization designed to create greater diversification of standard portfolios.
3. Optimization and risk issues
Initially, we create simple Markowitz (optimized) diversified portfolios. We use the term simple because we are focused on the basic points of utilization of commodities in a portfolio; the goal here is not to create a tradable portfolio. Doing so implies breaking some basic assumptions of Markowitz (1952) during the process. Assumptions in conflict with the basic Markowitz approach are highlighted. Sadly, many of the assumptions we are making in our modeling are not that different to those made in standard texts. The implications of the assumptions apropos this modeling and the potential implications for making such assumptions in actual portfolio optimization are discussed in Wilford (2012), Quintana et. al. (1998), and Norland and Wilford (2002). The three assumptions that we will examine in our optimizations, and which are dubious in a true Markowitz sense, are:

- In-sample optimization
- Long-only exposures (no shorting allowed)
- Utilization of one period's correlation matrix throughout

In-sample optimization should never be used for creating portfolios that are the basis of actual investment because they typically overstate expected return-risk ratios and are totally dependent on the past period being a good projection of future correlations and volatilities. Nevertheless, it is fairly standard for most financial advisors to do so. Long-only portfolios are the norm in the traditional asset management industry. Indeed, for most regulated entities long-only is all that is allowed (try to go short a normal security in a regulated Individual Retirement Account, for example). And, assuming a static correlation matrix is not all that unusual in practice. Dynamic correlation matrix estimation is fairly sophisticated and beyond the capability of

<table>
<thead>
<tr>
<th>Estimated correlation matrix</th>
<th>S&amp;P 500 Index</th>
<th>U.S. Treasury 10-year</th>
<th>Cash (U.S.$)</th>
<th>Oil</th>
<th>Natural gas</th>
<th>Gold</th>
<th>Copper</th>
<th>Corn</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>1</td>
<td>-0.6048</td>
<td>0</td>
<td>0.5253</td>
<td>0.0599</td>
<td>0.1093</td>
<td>0.5056</td>
<td>0.1871</td>
<td>0.1985</td>
</tr>
<tr>
<td>U.S. Treasury 10-year</td>
<td>-0.6048</td>
<td>1</td>
<td>0</td>
<td>-0.3689</td>
<td>-0.0572</td>
<td>0.0628</td>
<td>-0.3272</td>
<td>-0.1299</td>
<td>-0.148</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oil</td>
<td>0.5253</td>
<td>-0.3689</td>
<td>0</td>
<td>1</td>
<td>0.1135</td>
<td>1</td>
<td>0.1595</td>
<td>0.4204</td>
<td>0.2109</td>
</tr>
<tr>
<td>Natural gas</td>
<td>0.0599</td>
<td>-0.0572</td>
<td>0</td>
<td>0.1135</td>
<td>1</td>
<td>0.1595</td>
<td>0.4204</td>
<td>0.2109</td>
<td>0.241</td>
</tr>
<tr>
<td>Gold</td>
<td>0.1093</td>
<td>0.0628</td>
<td>0</td>
<td>0.3504</td>
<td>0.1595</td>
<td>1</td>
<td>0.4204</td>
<td>0.1664</td>
<td>0.1613</td>
</tr>
<tr>
<td>Copper</td>
<td>0.5056</td>
<td>-0.3272</td>
<td>0</td>
<td>0.5487</td>
<td>0.0462</td>
<td>0.4204</td>
<td>1</td>
<td>0.2109</td>
<td>0.241</td>
</tr>
<tr>
<td>Corn</td>
<td>0.1871</td>
<td>-0.1299</td>
<td>0</td>
<td>0.2605</td>
<td>0.1285</td>
<td>0.1664</td>
<td>0.2109</td>
<td>1</td>
<td>0.7332</td>
</tr>
<tr>
<td>Wheat</td>
<td>0.1985</td>
<td>-0.148</td>
<td>0</td>
<td>0.2571</td>
<td>0.1541</td>
<td>0.1613</td>
<td>0.241</td>
<td>0.7332</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Correlation matrix

<table>
<thead>
<tr>
<th>Estimated correlation matrix</th>
<th>S&amp;P 500 Index</th>
<th>U.S. Treasury 10-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>1</td>
<td>-0.6048</td>
</tr>
<tr>
<td>U.S. Treasury 10-year</td>
<td>-0.6048</td>
<td>1</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oil</td>
<td>0.5253</td>
<td>-0.3689</td>
</tr>
<tr>
<td>Natural gas</td>
<td>0.0599</td>
<td>-0.0572</td>
</tr>
<tr>
<td>Gold</td>
<td>0.1093</td>
<td>0.0628</td>
</tr>
<tr>
<td>Copper</td>
<td>0.5056</td>
<td>-0.3272</td>
</tr>
<tr>
<td>Corn</td>
<td>0.1871</td>
<td>-0.1299</td>
</tr>
<tr>
<td>Wheat</td>
<td>0.1985</td>
<td>-0.148</td>
</tr>
</tbody>
</table>

Table 2: Correlations to S&P 500 and U.S. Treasuries 10-year

4 Markowitz et al. (2000) discuss how simulations and iterations of models from data tend to overestimate the usefulness of results and Putnam (2000) defines this problem as "death by simulation." What is being done in this paper ignores these warnings initially only to later remind the reader of the dangers. In order to examine a simple approach to placing commodities into long-only optimizations, even long-short simulations, and to highlight how commodity risk may or may not (if modeled in a traditional framework) provide diversification those warnings outlined in Markowitz et al. (2000) will be ignored.
most fund advisors and many managers as well.\textsuperscript{5}

With these issues considered, three portfolios are constructed for consideration, a pre-crisis portfolio from 2005 through September 2008, a crisis portfolio from September 2008 to March 2009 (the idea being if you knew there was to be a crisis what would you have built), and a post-crisis portfolio from March 2009, which we will call the “new normal”.\textsuperscript{6} During this new normal period, we will use in-sample data to consider what the portfolio will look like under our general assumptions. We begin with the new normal, and allow for long/short portfolios and then utilize the same correlation matrix and volatility assumptions with pre-crisis and crisis estimations.

First, consider the post-crisis, new normal, dataset of expected returns and annualized standard deviations presented in Table 3. Next we utilize the post-crisis correlation matrix presented in Table 4.

Applying the same volatilities and correlations assumptions, but allowing for different return estimations via in-sample estimations of expected returns from all three periods, the portfolio optimizations are presented in Table 5.

\textsuperscript{5} For a good discussion of dynamic estimation of the variance-covariance matrix see Quintana and Putnam (1994).

\textsuperscript{6} In all fairness to the reader it should be understood that the author dislikes this term greatly. It implies that there is such a thing as a normal and that previous periods were old normal or not normal. It is a throwaway term that conveys little meaning, but since it is often used in portfolio discussions we will adopt it here.
sign of a portfolio being very dangerous. The key here is that the portfolio is designed knowing that the returns were extreme for all categories during the crisis, while the volatilities applied were those of the post-crisis period. That is, with perfect estimations of in-sample returns and the assumption of more normal (remember we are using the new normal) volatilities, the optimizer concludes that there is safety in these forecasts no matter the inconsistency in the forecast and the variance-covariance matrix (that in theory is estimated simultaneously). The optimization mechanism now “believes” there is much less risk than there actually is. Risk is greatly underestimated since the volatilities themselves are well below what actually occurred in the crisis (data from which the actual returns - estimated in theory - were drawn).

The third portfolio, pre-crisis returns with post-crisis volatilities, is more like what one would expect. Still the knowledge of the actual returns for the period ex-ante allows for very good returns relative to the risk that the portfolio takes. It is certainly less bold than the crisis portfolio that worked under the assumption (note using the wrong assumption) that there were good forecasts of returns (this was correct since we cheated to give it good forecasts by using in-sample data to create them), but also that there was little risk relative to these large negative return forecasts for the S&P, oil and gas and so forth. This clearly illustrates the problem of how underestimation of volatility can provide confidence to an optimizer (and thus a manager) that is unwarranted. It also demonstrates how easy it is to forget that optimizers, by their nature, are functioning based on a belief that the data entered is correct and consistent with the assumptions underlying the process. The optimizer “believes” it has the theoretically correct input information.7

Some observations are apropos at this point. If one has good insights and the assumptions of volatility are low enough, then good Sharpe ratios or information ratios can be produced, but they may be very misleading. Extreme positions for the crisis period highlight this issue. Finally, an often ignored point is highlighted, which we believe in the context of commodities allocations must always be considered. Knowledge about the consistency of volatilities is absolutely as critical as being a “good forecaster.” Otherwise the portfolios can have extreme positions, as was the case above where a short position in treasuries and S&P were disproportionally large with respect to the portfolio, thereby allowing it to gain so radically from a very long position in gold and other positions that were shorts. The volatility estimates were much too low. But there is also another lesson: volatilities do not drop in a crisis. They rise sharply, and this shift in the volatility regime must be considered if a crisis situation is a concern to the portfolio manager.

Table 6 displays the comparative volatilities for the periods. In all cases, crisis volatilities are higher, with the exception being

<table>
<thead>
<tr>
<th>Assets</th>
<th>Portfolio weights: post-crisis data</th>
<th>Portfolio weights: crisis returns and post volatilities</th>
<th>Portfolio weights: pre-crisis returns and post volatilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>50.01%</td>
<td>-52.81%</td>
<td>-13.33%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>118.11%</td>
<td>-128.51%</td>
<td>-8.83%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>-88.10%</td>
<td>275.85%</td>
<td>88.86%</td>
</tr>
<tr>
<td>Oil</td>
<td>-5.27%</td>
<td>-11.77%</td>
<td>5.46%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>-3.54%</td>
<td>-11.02%</td>
<td>5.54%</td>
</tr>
<tr>
<td>Gold</td>
<td>26.91%</td>
<td>58.41%</td>
<td>8.40%</td>
</tr>
<tr>
<td>Copper</td>
<td>-9.73%</td>
<td>-24.70%</td>
<td>14.07%</td>
</tr>
<tr>
<td>Corn</td>
<td>10.11%</td>
<td>-11.48%</td>
<td>1.70%</td>
</tr>
<tr>
<td>Wheat</td>
<td>1.48%</td>
<td>6.03%</td>
<td>-1.87%</td>
</tr>
<tr>
<td>Expected return of the portfolio</td>
<td>19.62%</td>
<td>54.60%</td>
<td>8.59%</td>
</tr>
<tr>
<td>Expected variance of the portfolio</td>
<td>0.0101</td>
<td>0.0218</td>
<td>0.0034</td>
</tr>
<tr>
<td>Expected standard deviation of the portfolio</td>
<td>10.03%</td>
<td>14.78%</td>
<td>5.86%</td>
</tr>
<tr>
<td>Expected excess return-risk ratio</td>
<td><strong>1.956</strong></td>
<td><strong>3.6947</strong></td>
<td><strong>1.4655</strong></td>
</tr>
<tr>
<td>Leverage (1 or less = no leverage)</td>
<td>1.88</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Sum of non-cash positions</td>
<td>188.10%</td>
<td>-175.00%</td>
<td>11.14%</td>
</tr>
<tr>
<td>Borrowings (if any)</td>
<td>88.10%</td>
<td>Placement of cash</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 5: Post-crisis variance-covariance matrix: optimizations

7 Again we refer the reader to Putnam (2000) to grasp the misunderstanding that can come from simulations, even those not so clearly theoretically compromised as the ones used in this paper.
natural gas. By comparison, the actual volatility for the S&P is three times higher, and for gold and copper it is twice as high. Treasury volatilities doubled and this was a safe haven. Safe havens will also see massive increases in volatility during a crisis. Understanding the interplay of volatility risk, quality of forecasted returns and portfolio optimization is critical when considering utilization of commodities for diversification.

Building portfolios with commodities that are used to diversify risk ex-ante must also consider the “which volatility” question that always arises during a crisis. The fact is, even in less extreme situations, volatilities and jumps in pricing in commodity markets are a concern, as pointed out by Al Janarbi (2012). Since major issues for using commodities arise in the volatility space, one must further examine resulting portfolios to better understand why “normal” may be misleading.

To do so, consider first what the portfolio for the crisis would have looked like if the correct volatilities were used to match up with these extremely good forecasts (in-sample forecasts by definition are nearly perfect). Following that approach, a portfolio that is typical of most wealth managers is constructed; it is forced to be long-only (we do not use constraints, but we do adjust expected excess return assumptions, thus the word forced). Third, we shock our portfolios to see how they would behave in a crisis, both the long/short and the long-only.

Table 7 presents the “crisis portfolio” with its own volatilities and expected returns. The information ratio falls to 1.31 from over 3, something much more realistic. The return expected from the portfolio falls to one fourth of what was expected under the “normal” volatility assumptions. The factual volatility adjustment had major consequences. It still suggests a very well-behaved portfolio, and this is certainly misleading because the optimizer knew the returns ahead of time.

It is interesting to note, however, that even with such certain knowledge of the returns, the portfolio is able to only generate an information ratio of 1.31. This is highlighted because ratios that are higher, often reported to highlight the brilliance ex-post of a manager, will almost certainly have an imbedded bias that actually hides the real risk inherent in a strategy. For example, option writing strategies suggest low volatility in returns, that is, until the option is called. The same can be said for most credit-based hedge funds (that is until the debt defaults) whether Markowitz model based or not. Highlighting this point is the fall in the ratio from the first example, with underestimated volatility, by almost a factor of three. The options writing strategy

<table>
<thead>
<tr>
<th>Assets</th>
<th>Pre-crisis volatilities</th>
<th>Crisis volatilities</th>
<th>Post-crisis volatilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>16.13%</td>
<td>57.55%</td>
<td>18.88%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>6.25%</td>
<td>12.02%</td>
<td>6.85%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>0.09%</td>
<td>0.09%</td>
<td>0.01%</td>
</tr>
<tr>
<td>Oil</td>
<td>34.25%</td>
<td>95.92%</td>
<td>30.09%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>61.03%</td>
<td>54.74%</td>
<td>45.13%</td>
</tr>
<tr>
<td>Gold</td>
<td>16.99%</td>
<td>38.25%</td>
<td>18.21%</td>
</tr>
<tr>
<td>Copper</td>
<td>28.29%</td>
<td>69.56%</td>
<td>27.79%</td>
</tr>
<tr>
<td>Corn</td>
<td>25.64%</td>
<td>51.48%</td>
<td>33.03%</td>
</tr>
<tr>
<td>Wheat</td>
<td>28.99%</td>
<td>49.24%</td>
<td>37.97%</td>
</tr>
</tbody>
</table>

Table 6: Comparing volatilities

<table>
<thead>
<tr>
<th>Assets</th>
<th>Calculated weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>-7.71%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>-23.85%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>135.63%</td>
</tr>
<tr>
<td>Oil</td>
<td>-1.29%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>-8.48%</td>
</tr>
<tr>
<td>Gold</td>
<td>13.56%</td>
</tr>
<tr>
<td>Copper</td>
<td>-5.34%</td>
</tr>
<tr>
<td>Corn</td>
<td>-5.18%</td>
</tr>
<tr>
<td>Wheat</td>
<td>2.65%</td>
</tr>
<tr>
<td>Expected return of the portfolio</td>
<td>13.30%</td>
</tr>
<tr>
<td>Expected variance of the portfolio</td>
<td>0.0102</td>
</tr>
<tr>
<td>Expected standard deviation of the portfolio</td>
<td>10.11%</td>
</tr>
<tr>
<td>Expected excess return-risk ratio</td>
<td>1.3148</td>
</tr>
<tr>
<td>Leverage (1 or less = no leverage)</td>
<td>None</td>
</tr>
<tr>
<td>Sum of absolute values of exposures</td>
<td>203.69%</td>
</tr>
<tr>
<td>Sum of non-cash positions (shorts cancel longs)</td>
<td>-35.63%</td>
</tr>
<tr>
<td>Borrowings (if any) above 100% capital usage</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 7: Crisis portfolio with own volatilities and expectations
example reminds one that simple measures of overall portfolio volatility may be misleading, and thereby underestimates risk; for certain asset classes non-linear measures of risk may be more important.8

4. Long-only and realism
In reality, most portfolios are not long/short; 200% absolute exposures would frighten many investors (Table 7). Realism creeps into the analysis (and the usefulness) of commodities in a portfolio when it is constrained to be long-only. In this section, we “fool” the optimizer into creating long-only portfolios. Why fool the optimizer? Why not simply constrain the optimizations to be long-only, à la many advisors and managers? IRAs are constrained to be long-only to protect the consumer, thus should we not do the same when allowing commodities to be utilized as diversifiers? The simple answer is no. Constrained portfolios may be very misleading in their results.9 We will not list them here, but it is suffice to state that constrained optimization techniques hide the underlying problems in a straightjacket. Without examining the underlying reasons that an optimized portfolio does look reasonable to a seasoned portfolio management, one will certainly fail to appreciate the useful insights into the proper use of commodities in a portfolio.10

First for consistency, post-crisis volatilities and correlations will once again be applied in our estimations. The optimizer will be “fooled” by changes in the set of return expectations necessary to “force” a long-only portfolio. The purpose of this activity is to highlight the importance of mis-estimation of either the variance-covariance matrix (volatility as above and the correlation matrix as discussed) or the set of expected returns generated to create the portfolio.10

8 Having been involved in examination, management and analysis of many hedge fund strategies, the author is continually amazed by managers who purport to have information ratios of over 2.0, not recognizing the useless nature of the measure if applied to certain strategies.

9 Many authors have addressed this issue. We believe the general discussion on this topic in Wilford (2012) highlights most of the issues and provides a good set of references for the topic.

10 At this juncture, it is always good to remember that the optimizer assumes the inputs are theoretically consistent and correct. Garbage in will be maximized garbage out.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Long-only adjusted</th>
<th>Long-short</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>12.00%</td>
<td>12.97%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>1.00%</td>
<td>4.99%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Oil</td>
<td>26.12%</td>
<td>8.55%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>33.37%</td>
<td>-9.13%</td>
</tr>
<tr>
<td>Gold</td>
<td>17.36%</td>
<td>18.88%</td>
</tr>
<tr>
<td>Copper</td>
<td>30.86%</td>
<td>7.48%</td>
</tr>
<tr>
<td>Corn</td>
<td>12.21%</td>
<td>26.39%</td>
</tr>
<tr>
<td>Wheat</td>
<td>15.00%</td>
<td>23.51%</td>
</tr>
</tbody>
</table>

Table 8: Excess return comparisons – historical post-crisis versus adjusted

<table>
<thead>
<tr>
<th>Assets</th>
<th>Long-only allocations summary – modified expected returns</th>
<th>Long-short allocation summary – original expected returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>10.91%</td>
<td>50.01%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>49.47%</td>
<td>118.11%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>11.82%</td>
<td>-88.10%</td>
</tr>
<tr>
<td>Oil</td>
<td>3.78%</td>
<td>-5.27%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>5.66%</td>
<td>-3.54%</td>
</tr>
<tr>
<td>Gold</td>
<td>6.56%</td>
<td>26.91%</td>
</tr>
<tr>
<td>Copper</td>
<td>11.74%</td>
<td>-9.73%</td>
</tr>
<tr>
<td>Corn</td>
<td>0.10%</td>
<td>10.11%</td>
</tr>
<tr>
<td>Wheat</td>
<td>0.00%</td>
<td>1.48%</td>
</tr>
<tr>
<td>Expected return portfolio</td>
<td>9.45%</td>
<td>19.62%</td>
</tr>
<tr>
<td>Expected Variance Portfolio</td>
<td>0.0038</td>
<td>0.0101</td>
</tr>
<tr>
<td>Expected standard deviation portfolio</td>
<td>6.15%</td>
<td>10.03%</td>
</tr>
<tr>
<td>Expected excess return-risk ratio</td>
<td>1.5370</td>
<td>1.9560</td>
</tr>
<tr>
<td>Leverage (1 or less = no leverage)</td>
<td>None</td>
<td>1.88</td>
</tr>
<tr>
<td>Sum of non-cash positions</td>
<td>88.18%</td>
<td>188.10%</td>
</tr>
<tr>
<td>Borrowings (if any)</td>
<td>None</td>
<td>88.10%</td>
</tr>
</tbody>
</table>

Table 9: Long-only and long-short expectations modifications
The Journal of Financial Perspectives

Risk management insights from Markowitz optimization for constructing portfolios with commodity futures

Table 8 shows the adjustments to expectations that are made. The U.S. Treasury return expectations are lowered by 80%, while the return to oil is almost tripled, and natural gas is assumed to provide a 33% return, when we know it actually declined by 9 percent during the period. Other changes are also extreme. This exercise allows us to generate long-only allocations with a seemingly very good return risk ratio (information ratio) of 1.53%. And, the portfolio looks conservative; indeed it may be one that your wealth management advisor would like. There is allocation to commodities, but not too much. And with 50% treasuries, it is conservative. The standard deviation of the portfolio is low. Who would not like this portfolio as illustrated in the first column of Table 9?

By comparison, the long/short allocations appear “scary” even with the expected returns from in-sample period estimation. It does have a higher information ratio, unrealistic for certain, with what could be perceived as extreme long and short positions. This exercise is not intended to suggest one portfolio over the other. They both are flawed in their own ways. It is, however, designed to set up what I believe is the most important issue when allocating to commodities. When crises occur, prices gap, risk changes and correlations move to extremes and portfolios need to be able to withstand these types of shocks.

5. Portfolio shocks - robustness

It is critical to remember that one will always be proved wrong, ex-post, from what was expected based on any optimization process. Actual returns, volatilities and correlations are never what one expects. Only perfect foresight would imply this and if one had perfect foresight, why would there be a need for a portfolio? This was the insight of Markowitz (1952, 1959) many of us forget. As such, a critical element in portfolio construction and management is the stability of any portfolio under crisis conditions. The lack of prescience knowledge that Markowitz observed then leads to the logical next question. What happens when we are not just wrong about our expectations, but “really wrong”? How does a portfolio behave under severe stress? When using commodities, the answer to the stress question may be the most important in determining how to use commodity-linked assets in a portfolio, whether actual commodities or as in our case commodity futures.11

To examine the shock issue for our portfolios, two portfolios are first created:

1. A post-crisis long/short portfolio as per above
2. A long-only portfolio as per above

These are then subjected to the actual returns that occurred during the crisis period to see how they both behave. The results are reported in Table 10.

The results are startling in contrast. The long/short portfolio, which would appear fanciful at best and terrifying to most, only lost 2.5% in return. These results are not bad considering that the S&P lost some 50% from its peak to trough. Contrast a 2.5% loss to the long-only portfolio losing more than 12%. “Safe appearance” was not safe at all during the crisis (one might say it was a wolf in sheep’s clothing). Indeed, the portfolio that seemed more outlandish was the one that was safer, ex-post.

This does not tell the whole story, however. The long/short portfolio started with a much higher expected return; thus, one should scale the results based upon some comparative criteria.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Long-short optimization</th>
<th>Long-only optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Index</td>
<td>-2.87%</td>
<td>-4.82%</td>
</tr>
<tr>
<td>US Treasury 10-year</td>
<td>0.19%</td>
<td>1.59%</td>
</tr>
<tr>
<td>Cash (U.S.$)</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Oil</td>
<td>0.22%</td>
<td>-1.86%</td>
</tr>
<tr>
<td>Natural gas</td>
<td>-0.15%</td>
<td>-2.70%</td>
</tr>
<tr>
<td>Gold</td>
<td>0.88%</td>
<td>1.14%</td>
</tr>
<tr>
<td>Copper</td>
<td>0.35%</td>
<td>-5.69%</td>
</tr>
<tr>
<td>Corn</td>
<td>-0.90%</td>
<td>-0.04%</td>
</tr>
<tr>
<td>Wheat</td>
<td>-0.09%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Portfolio return</td>
<td>-2.36%</td>
<td>-12.36%</td>
</tr>
</tbody>
</table>

Table 10: Shocked portfolios: returns

Table 8

This literature is becoming more important following the 2008 banking crisis. An understanding of robustness, the misapplication of Value at Risk (VaR), why SIVs should not have had risk estimation with VaR, why simplistic Monte Carlo simulations can be misleading, as well as a whole range of issues are now being considered. For example, see Putnam and Wilford (1998), Norland and Wilford (2002), and Putnam et al. (2002) for some work that forewarned of these problems.

11 This literature is becoming more important following the 2008 banking crisis.
For example, suppose the portfolios were created seeking the same expected return (scaling is always essential in comparing portfolios ex-ante and should always be done ex-post), allowing the long-only portfolio to be scaled to seek that return. Doing so illustrates that the resulting (ex-post) risk of the long-only portfolio is actually greater than that of the long/short one and that the losses of return are some 11 times as large. The long-only portfolio does not have better risk characteristics and loses a huge amount of money in comparison to the long/short portfolio. Arguably, the robustness issue is extremely important in evaluating a portfolio, how the assets are to be utilized, and chosen for allocation.\(^{12}\)

Robustness must always be at the forefront of any allocation methodology. It is often ignored and portfolios are designed to “look the part”, as opposed to withstand a crises. When using commodities in a portfolio, this issue becomes even more important due to the volatile nature of the underlying assets’ price structure, liquidity difficulties with the asset classes, and price gaps that may occur during a crisis situation, either in general or within the asset class itself. These issues are more common to commodities than, say, to treasuries, the S&P, or FX if for no other reason than the depth of the markets for these other assets, as pointed out by Janabi (2012).

6. Summary points and conclusions
To determine the implications of putting commodities into simple portfolios, many issues need to be considered.

- How much risk can one live with, should a difficult, crisis environment occur?
- Commodity volatilities are not just large but can also change abruptly.
- Heteroscedasticity (i.e., the reality that there are volatility regime shifts) is a problem that cannot be assumed away.
- For those using an options theoretical approach to consider the risk in a portfolio this implies fairly large Vega risk must be measured, not something typically done by many money managers.
- Portfolios able to absorb large shocks – robust portfolios – may be very different from what common knowledge suggests is “safe-appearing.”
- Gapping of commodity prices can happen quickly; for example, in mid-April, 2013 gold prices gapped sharply in two trading days by some 15%, with little news, no crisis, but some information on the margin that suggested a major change in technical patterns would occur.
- In a crisis, all volatilities rise sharply; commodities more so, however.
- Liquidity becomes important in a shocked situation and can be critical even in situations that are not subject to dramatic economic shocks.
- Commodities provide diversification, but do not make the error of thinking they can simply fit into a typical long-only portfolio.
- The typical low correlation criteria employed by many managers for long-only portfolios may be very misleading.
- A crisis can undo some of the positive benefits of simplistic low correlation portfolios.
- Risk can be underestimated during periods of intense market uncertainty such as the crisis of 2008.
- Long/short portfolios can mitigate some of the problems of using commodities.
- Long/short optimized portfolios are theoretically more correct, have the characteristic to create the intended diversification, and when volatilities increase substantially for the individual asset classes this intended diversification can have a significant dampening effect on the overall portfolio volatility and risk.
- When using commodity risk in a portfolio based upon a Markowitz optimization seek ones that have very robust characteristics.

From this exercise, it can be deduced that putting commodities in portfolios does offer diversification, although not necessarily in a simplistic manner. Long/short portfolios that are created properly have a much greater chance of delivering the expected returns being sought by investors, than simple allocations that include a nod to commodities as asset classes. In the case of most foundations, pension funds or other “conservative” investors, it is recommended that commodities type long/short portfolios (funds or managers) be chosen when seeking diversification from commodities. Allocation to a commodity for low correlation reasons only can actually increase risk during a crisis situation; for these types of investment entities concern about large

\(^{12}\) Scaling could have been done by equalizing volatilities; that is, assume a targeted volatility and then measure the resulting gains or losses from each strategy. Again the ex-post losses of the long-only portfolio are increased to some 9 times that of the long/short (scary) portfolio.
Risk management insights from Markowitz optimization for constructing portfolios with commodity futures

tail events should be a point of focus. In a long/short portfolio approach potential stability in a crisis certainly offsets many of the negative characteristics that gapping and liquidity issues can create for a portfolio constructed with commodities, while leaving in place the fundamental positive nature of diversification through commodity investment.

References

Markowitz, H. M., 1959, Portfolio selection: efficient diversification of investments, New York: Wiley
Markowitz, H. M., 1987, Mean-variance analysis: portfolio choice and capital markets, Oxford: Blackwell
Roudier, F., 2007, “Portfolio optimization and genetic algorithms” at the Department of Management, Technology and Economics - DMTEC, Swiss Federal Institute of Technology (ETH) Zurich, May

Japanese patent index and stock performance

Takao Kobayashi
Professor of Finance, Aoyama Gakuin University

Yasuhiro Iwanaga
Aoyama Gakuin University & Sumitomo-Mitsui Trust Bank, Limited

Hideaki Kudoh
Nomura Asset Management Co., Ltd.

Abstract
In the global economy, high technology serves as a source of competitive advantage for Japanese companies. In Japan, there is a patent value indicator which is unique among other patent value indicators developed in the U.S. The uniqueness lies in its focus on measuring the exclusivity and technological competitiveness of each patent using data based on the number of actions taken by third parties against the patent. The construction of such a patent value metric became possible thanks to the Japanese Government's active disclosure of information on patent attacks. This paper is our first attempt to study the relationship between technological competence and firm performance using this technology indicator. In particular, we demonstrate how this technology indicator may be used to forecast company stock performance. We construct long/short strategies based on (1) the patent indicator, (2) R&D expenditure and (3) a combination of the two. The third strategy was the best performer. Combining R&D expenditure, which is readily available in financial statements, with the patent indicator enhanced portfolio expected return and reduced risk considerably. The best-performing strategy generated an annual mean return of 11.50%, standard deviation of 9.25%, and a Sharpe ratio of 1.23. The return is not attributable to the Fama-French three factors. Technology indicators should not work in some industry groups, hence, the result is even more striking given that the portfolio was constructed with a universe of stocks covering all industry groups except financials.
1. Introduction
As the industrial economy has been supplanted by the knowledge economy, intellectual property (IP) assets, such as patents, trademarks and copyrights, are gaining growing importance in corporate value creation. In the global economy, high technology serves as a source of competitiveness for Japanese companies. In particular, Japanese companies account for at least 20% of all international patent applications in the world and produce innovative products using patented technology as leverage. In the 2012–2013 edition of the Global Competitiveness Report of the World Economic Forum, Japan is ranked second (next to Switzerland) in the “innovation and sophistication” factor of global competitiveness.1

Our basic interest is to find out whether the accumulation of these IP assets is sufficiently incorporated in the stock market’s pricing of firms. If they are, information about companies’ R&D activities and/or the level of patent accumulation would not provide arbitrage opportunities. If, on the contrary, the information predicts future stock returns and serves as a useful signal for screening companies, the investment community needs an easily accessible source of information on the value of each firm’s IP assets. Not only does it satisfy the investors’ appetite for higher returns, but also promotes the efficiency of the stock market. The latter means that a firm with highly valuable IP assets would be able to raise equity capital at a lower cost. In Japan there is a patent value indicator, which is unique among other patent value indicators developed in the U.S. In this paper, our focus is mainly on this patent value index.

Thomas (2001) used a quantitative measure of patent assets based on the citation frequency. The index is available via an online service called TechLine and provided by Chi Research Inc. He showed that there is a strong relationship between this index and stock market valuation. He further showed that by fitting a model to this relationship, underpriced stocks relative to the model price tend to perform excellently and overpriced stock perform poorly in the subsequent period. Cardoza et al. (2008) used another measure called PatentRatings, which is provided by Ocean Tomo. This index includes a number of qualitative measures for a patent, which they call the patent’s “economic value.” They constructed a portfolio of 300 stocks selected based on this patent index and used it as a benchmark performance index of the technology sector of the U.S. economy. They showed that this portfolio outperforms the S&P500 Index.

R&D expenditure carries information on R&D intensity. This information is readily available in financial statements. Consequently, researches on the relationship between intangible IP assets and stock market performance started with a focus on R&D expenditure. The empirical results in this area are mixed. Titman et al. (2004) found that intensiveness in physical investment has a negative relationship with stock returns. The literature calls it the “investment paradox.” These authors measure investment intensity by physical investment relative to the amount of tangible assets. Li and Liu (2010) showed that the relationship between R&D intensity and stock return is similar. Namely, when R&D intensity is measured by R&D investment-to-intangible asset ratio, R&D intensity has a negative relationship with stock returns. In contrast, Chan et al. (2001) showed that R&D-intensive firms earn higher returns. Their measure of R&D intensity used market capitalization as the scale. Lev and Sougiannis (1996, 1999) and Chambers et al. (2002) obtained similar results. In addition, Jung (2005) reported that R&D-intensive companies generate exceptionally high positive returns in Japan.

We will examine the relationship between the value of patents (YK) and R&D expenditure to future stock performance. In so doing, our objective is to distinguish between R&D efficiency (or, R&D productivity) and R&D intensity, with the conjecture that they carry disparate sets of information about IP assets. We measure “R&D efficiency” by the ratio of YK value to R&D expenditure. We measure “R&D intensity” by the ratio of R&D expenditure to market capitalization. We will show that both measures have independent predictability for future stock returns. We further show that combining these two measures significantly enhances investment performance.

The rest of the paper is organized as follows: section 2 illustrates how the YK value is constructed. We also report some descriptive statistics of the YK value. In section 3, we investigate the relationship between IP signals and stock returns in a portfolio context. We construct portfolios of stocks using a different set of IP signals and compare the performance of portfolios. In 3.1, we describe the data sources. In 3.2, we use the YK value scaled

1 The other factors are “basic requirements” and “efficiency enhancers.”
by market capitalization as the screening variable. In 3.3, we use YK value scaled by R&D expenditure (i.e., R&D efficiency), and independently, R&D expenditure scaled by market capitalization (i.e., R&D intensity). In 3.4, we combine YK value and R&D expenditure to make two-dimensional sorts of stocks. In 3.5, we compare results of our portfolio strategies, and in 3.6, we report the industry groups in which our strategy works well and the industry groups in which it works poorly. Section 4 summarizes the paper.

2. The YK value
In recent years, there have been many attempts to measure the value of patents. Chi Research Inc.'s TechLine [see Thomas (2001)] and Ocean Tomo's Patent Ratings [see Cardoza et al. (2008)] represent two of them.

TechLine's measure is based on the following patent indicators:

1. Number of patents held by a company
2. Growth in the number of patents held by a company
3. Current impact index: the number of citations a company's patents receive within other parties' patents issued in the most recent year
4. Scientific linkage: the number of references to scientific papers that a company's patents makes
5. Technology cycle time: the median age of patents that a company lists on the front page of the company's patent report

These are all quantitative measures that count the number of patents, number of citations, the age (obsolescence) of patents and so on. In addition to these indicators, Ocean Tomo’s Patent Ratings includes a qualitative measure, which they call the IPO Score. Patent owners must pay periodic fees to maintain their patents in force. By analyzing patent maintenance data, Ocean Tomo constructs a score that predicts the probability that a patent will be maintained or abandoned. They regard this score as a measure of the economic value of each patent.

The YK value of the individual patent n at time t is defined by

$$YK_n(t) = \sum_i P_i N_i \gamma_i$$

(formula 1), where i denotes an individual action, P_i is the score given to action i, N_i is the number of parties who took action i, and \(\gamma_i\) is a decay factor reflecting the speed with which its patents are expected to lose value or be replaced by superseding technologies. About 80% of patent abolishment is due to unpaid patent maintenance fees and the rest is due to the legal termination of the patent right (20 years). Very few patents are abandoned during the first 4 years, and on average they are abandoned 13 to 19 years after the date they were granted. The decay factor is computed by fitting a probabilistic model of technology obsolescence to the data on patent abandonment and termination.

The YK value of a company at time t is the sum of YK_n(t):

$$YK(t) = \sum_n YK_n(t)$$

(formula 2). The data has been created by Kudo & Associates on a monthly basis since January 1988, and it covers all listed firms in Japan. Information about third-party
actions is available from a publication by the National Center for Industrial Property Information and Training. The public release of detailed information about third-party actions against patents is peculiar to Japan. Hence, the availability of the YK value adds uniqueness to the current research.

3. Returns to IP-driven strategies

If all the information about companies’ IP competitiveness is fully reflected in stock prices, YK value would not provide any profitable opportunities for equity investors. The purpose of our analysis is to examine whether this is true. We demonstrate that the YK value provides quite a useful signaling mechanism for screening stocks when constructing highly performing equity portfolios.

3.1 Data sources

We include in our study all firms listed in the first section of the Tokyo Stock Exchange (TSE) except banks, securities companies, insurance companies, and other financials. Stock prices and returns are from QUICK-Astra and accounting data are from the Nikkei Economic Electronic Databank System (NEEDS). The YK value was provided by Kudo & Associates. We computed the Fama–French three factors for all the stocks in our study using the method proposed by Kubota and Takehara (2007). The sample period is between September 2002 and December 2012.

3.2 The default portfolios selected by YK/ME

Figure 1 shows the relationship between the YK value and market capitalization (later denoted by ME, shorthand for market equity) for firms in the electric appliances industry. The x and y axes are both in logarithmic scale. The correlation coefficient is 0.72.

Table 1 shows the correlation coefficient between these two variables for industries within the 33 TSE industry classifications where the correlation coefficient is larger than 0.6. In Table 1, we also report the mean, the maximum value, and the cross-sectional standard deviation of the YK value in these industries. As evident, the correlation is very high: larger firms tend to have larger YK values. This suggests that YK values should not be used as a stand-alone measure for selecting stocks. We chose to scale

<table>
<thead>
<tr>
<th>Industry</th>
<th>Maximum</th>
<th>Mean</th>
<th>St.dev</th>
<th>Correlation with market cap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision instruments</td>
<td>4,614</td>
<td>536</td>
<td>991</td>
<td>0.85</td>
</tr>
<tr>
<td>Land transportation</td>
<td>436</td>
<td>174</td>
<td>200</td>
<td>0.83</td>
</tr>
<tr>
<td>Rubber products</td>
<td>3,602</td>
<td>901</td>
<td>1,131</td>
<td>0.80</td>
</tr>
<tr>
<td>Textiles and apparels</td>
<td>7,853</td>
<td>943</td>
<td>2,006</td>
<td>0.75</td>
</tr>
<tr>
<td>Chemicals</td>
<td>16,407</td>
<td>1,231</td>
<td>2,447</td>
<td>0.73</td>
</tr>
<tr>
<td>Electric appliances</td>
<td>30,356</td>
<td>1,308</td>
<td>3,610</td>
<td>0.72</td>
</tr>
<tr>
<td>Pharmaceutical</td>
<td>2,181</td>
<td>314</td>
<td>459</td>
<td>0.70</td>
</tr>
<tr>
<td>Glass and ceramics products</td>
<td>3,663</td>
<td>594</td>
<td>868</td>
<td>0.69</td>
</tr>
<tr>
<td>Machinery</td>
<td>5,099</td>
<td>401</td>
<td>796</td>
<td>0.69</td>
</tr>
<tr>
<td>Transportation equipment</td>
<td>7,160</td>
<td>659</td>
<td>1,306</td>
<td>0.68</td>
</tr>
<tr>
<td>Metal products</td>
<td>2,708</td>
<td>310</td>
<td>631</td>
<td>0.67</td>
</tr>
<tr>
<td>Nonferrous metals</td>
<td>4,703</td>
<td>971</td>
<td>1,412</td>
<td>0.67</td>
</tr>
</tbody>
</table>

We thank Quick Corp. for providing data on monthly stock returns for this research.


3 We thank Quick Corp. for providing data on monthly stock returns for this research.

4 We thank Mr. Ichiro Kudo of Kudo & Associates for enabling us to access the YK value database for this research.

5 The minimum value is zero for most industries.
the YK value by market capitalization (YK/ME). Since it makes little sense to compare the patent competitiveness of a chemical company with that of an auto manufacturer, stocks’ YK/ME comparisons were undertaken within each industry.

At the end of each month, stocks are separated into quintiles based on YK/ME and equal-weight portfolios are formed in each quintile. We select stocks from each of the 33 TSE industry classifications. Portfolio 1 is the portfolio of stocks from the highest YK/ME quintile of each industry and portfolio 5 is the portfolio of stocks from the lowest YK/ME quintile of each industry.

Panel A of Table 2 presents the properties of the quintile portfolios. BE/ME is book-to-market (the ratio of book value of equity to market value of equity), E/P is the earnings yield (the ratio of earnings per share to stock price), ROE is return on equity (the ratio of net income to book value of equity), and log(ME) is the log of market value of equity measuring the size of the firm, with ME denominated in yen.

One can see some “value-tilt” for higher YK/ME portfolios. Book-to-market is higher and ROE is lower for higher YK/ME portfolios, and these relationships are monotone. This should be a little surprising as growth firms, rather than value firms, tend to spend more on R&D activities. We will scrutinize this observation later by segregating the effects of YK/ME into R&D efficiency and R&D intensity.

Panel B reports the portfolio performance. Mean return is the time-series average of the annualized portfolio returns in percent. St.dev is the annualized standard deviation of the returns. Sharpe ratio is the mean excess return (over the risk-free rate) per unit risk (measured by the standard deviation), which is a standard risk-adjusted measure of performance. We use the one-month T-bill rate for the risk-free rate. The column “H-L” refers to the return of a hedge portfolio, which is long in portfolio 1 and short in portfolio 5.

The highest YK/ME portfolio generates a mean return of 9.34% annually, while the lowest YK/ME portfolio generates 1.52%. Consequently, the mean return on the H-L portfolio is 7.81% annually. It has a t-statistic of 3.82, which shows that the mean return of the H-L portfolio is significantly positive. The standard

---

6 We tested other measures of size, such as book value of assets, book value of equity, and sales as the scaling variable. Market capitalization was best among these variables.
deviation of return is higher for higher YK/ME quintiles. The monotonicity of the Sharpe ratio indicates that higher YK/ME quintile portfolios yield higher risk-adjusted returns.

For each quintile portfolio and the H-L portfolio, we run a time-series regression of monthly excess return to the Fama–French three factors:

\[ R_{i,t} = \alpha_i + \beta_{i,\text{MKT}} R_{\text{MKT},t} + \beta_{i,\text{SMB}} R_{\text{SMB},t} + \beta_{i,\text{HML}} R_{\text{HML},t} + \epsilon_{i,t} \]  

(formula 3), where, for each month \( t \), \( R_{i,t} \) is the return of portfolio \( i \), \( R_{\text{MKT},t} \) is the risk-free rate, \( R_{\text{SMB},t} \) is the Fama–French size factor (the return to being long in small ME stocks and short in big ME stocks), and \( R_{\text{HML},t} \) is the Fama–French value factor (the return to being long in high BE/ME stocks and short in low BE/ME stocks). We created the Fama–French three factors for the Japanese stock market following the procedure proposed by Kubota and Takehara (2007).7

The row labeled “FF3-alpha” reports the estimate of the intercept term \( \alpha_i \) and its t-statistics (in parentheses). If any portfolio has a significantly positive \( \alpha_i \), it means that the portfolio has additional risk-adjusted performance relative to its risk exposure to the Fama–French three factors.

The FF3-alphas of the highest YK/ME portfolio and the H-L portfolio have t-statistics of 2.86 and 3.32, respectively. Hence, we can conclude that these two portfolios generate additional risk-adjusted returns. This result confirms that YK/ME serves as a useful signal for equity investment.

Table 3: Portfolios selected using R&D efficiency (YK/R&D)

<table>
<thead>
<tr>
<th>Panel A: Properties of portfolios selected using YK/R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (High)</td>
</tr>
<tr>
<td>YK/R&amp;D</td>
</tr>
<tr>
<td>BE/ME</td>
</tr>
<tr>
<td>E/P</td>
</tr>
<tr>
<td>ROE</td>
</tr>
<tr>
<td>Log(ME)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Return of portfolios selected using YK/R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (High)</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>St.dev</td>
</tr>
<tr>
<td>Sharpe ratio</td>
</tr>
<tr>
<td>FF3-alpha</td>
</tr>
</tbody>
</table>

The MKT factor was computed using (1) the cum-dividend monthly return series of TOPIX for the market return and (2) the one-month T-bill rate for the risk-free rate. The risk-free rate is 0.2 percent during our period of study, reflecting the fact that most of the years were under the “zero interest rate policy regime” by Bank of Japan. TOPIX is the market index for the first section of the TSE. Accordingly, we computed the SMB and HML factors on the same universe of stocks. The selection of stocks for each portfolio was done annually, at the end of August. Small, big, high and low portfolios were value-weighted portfolios in which weighting was based on the number of “free-float” shares (shares available for trading).

The MKT factor was computed using (1) the cum-dividend monthly return series of TOPIX for the market return and (2) the one-month T-bill rate for the risk-free rate. The risk-free rate is 0.2 percent during our period of study, reflecting the fact that most of the years were under the “zero interest rate policy regime” by Bank of Japan. TOPIX is the market index for the first section of the TSE. Accordingly, we computed the SMB and HML factors on the same universe of stocks. The selection of stocks for each portfolio was done annually, at the end of August. Small, big, high and low portfolios were value-weighted portfolios in which weighting was based on the number of “free-float” shares (shares available for trading).

A variable that complements R&D intensity is the productivity of R&D activities, which we call R&D efficiency. Noting that YK/ME can be decomposed to become: YK/ME = (YK/R&D) × (R&D/ME)
(formula 4), we will regard the first term, $Y_K/R&D$, as a measure of R&D efficiency.

Using these variables, we now investigate how the R&D efficiency measure, $Y_K/R&D$, and the R&D intensity measure, $R&D/ME$, work independently in selecting stocks for portfolios. As in the previous case, selection is undertaken for each industry.

We start with $Y_K/R&D$. We repeat the same process as in Table 2 except that the R&D efficiency, $Y_K/R&D$, is used for the screening signal. The results are reported in Table 3. As can be seen from Panel A, there is no discernable difference across the $Y_K/R&D$-selected quintile portfolios in terms of book-to-market, earnings yield, ROE, and size. Panel B reports the portfolio performance. The highest $Y_K/R&D$ portfolio generates a mean return of 7.02% annually, while the lowest generates 3.79%. Consequently, the mean return on the H-L portfolio is 3.23% annually. It has a t-statistic of 2.79, which shows that the mean return is significantly positive. The standard deviation of returns is almost identical across the quintile portfolios. The Sharpe ratio shows that higher $Y_K/ME$ quintile portfolios yield higher risk-adjusted returns, but the relationship is not monotone.

The FF3-alphas of the highest $Y_K/R&D$ portfolio and the H-L portfolio have t-statistics of 3.31 and 2.70, respectively. Hence, we can conclude that these two portfolios generate additional risk-adjusted returns. This results confirm that R&D efficiency provides a useful signal for equity investment.

We repeat the same process except that the R&D intensity, $R&D/ME$, is used for the screening signal. The results are reported in Table 4.

Panel A shows a certain degree of “value-tilt” for higher $R&D/ME$ portfolios. Book-to-market is higher and ROE is lower for higher $R&D/ME$ portfolios; and these relationships are mostly monotone. We observed this value-tilt for portfolios selected by $Y_K/ME$ in Table 2. The bottom row shows that log($ME$) is smallest for the highest $R&D/ME$ portfolio. In Japan, most of the value firms are smaller in size. Since we measure R&D intensity in terms of R&D expenditure per market capitalization, the highest $R&D/ME$ portfolio may have a majority of value firms.

Panel B reports the portfolio performance. The highest $R&D/ME$ portfolio generates a mean return of 10.13% annually, while the lowest generates 1.92%. Consequently, the mean return on the

<table>
<thead>
<tr>
<th>Panel A: Properties of portfolios selected using R&amp;D/ME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>R&amp;D/ME</td>
</tr>
<tr>
<td>BE/ME</td>
</tr>
<tr>
<td>E/P</td>
</tr>
<tr>
<td>ROE</td>
</tr>
<tr>
<td>Log(ME)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Return of portfolios selected using R&amp;D/ME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>10.13</td>
</tr>
</tbody>
</table>

| Sharpe ratio | 0.43 | 0.25 | 0.24 | 0.12 | 0.10 |
| FF3-alpha    | 0.46 (2.25) | 0.17 (1.41) | 0.18 (1.71) | -0.01 (-0.11) | 0.00 (0.03) | 0.45 (2.12) |

Table 4: Portfolios selected using R&D intensity ($R&D/ME$)
H-L portfolio is 8.21% annually. It has a t-statistic of 2.73, which shows that the mean return of the H-L portfolio is significantly positive. Unlike Panel B of Table 3, the mean return increases monotonically for higher R&D/ME quintiles. On the other hand, the standard deviation of returns is also higher for higher R&D/ME quintiles. The Sharpe ratio shows that higher R&D/ME quintile portfolios yield higher risk-adjusted return. FF3-alphas of the highest R&D/ME portfolio and the H-L portfolio have t-statistics of 2.25 and 2.12, respectively. We can, therefore, conclude that these two portfolios generate returns in excess of what can be expected by their exposures to the three risk factors of Fama and French. This result confirms that R&D intensity also provides a useful signal for equity investment.

### 3.4 Two-dimensional selection

We now scrutinize the result of Table 3 that the mean returns and alphas are higher for stocks with high R&D efficiency. We first sort stocks into quintiles based on their R&D intensity. We then sort stocks into quintiles based on their R&D efficiency and average across the R&D-intensity quintiles. This way we make sure that in each of the five YK/R&D-selected portfolios stocks are spread evenly over R&D intensity. We call these five portfolios “YK/R&D-selected portfolios controlling for R&D intensity.”

At the end of each month, stocks are allocated into quintiles based on R&D/ME. Then, within each R&D/ME quintile, we separate stocks into five subgroups based on YK/R&D. This way we create 25 groups of stocks each month and in each group we form an equal-weight portfolio. As before, we do the grouping of stocks within each of the 33 TSE industry classifications. The results are reported in Table 5. Column “H-L” refers to the return of a hedge portfolio that is long portfolio 1 and short portfolio 5 for each R&D/ME quintile.

For each of the 30 monthly time series of portfolio excess returns (for the 5×5 portfolios and the 5 H-L portfolios), we run a regression to the three Fama-French factors. The table reports the estimate of the intercept term $\alpha$, and its t-statistics (in parentheses).

In each R&D/ME quartile, the alpha is higher for higher YK/R&D portfolios, although there are some exceptions. As we conjectured, among firms with similar R&D intensity, firms with higher R&D efficiency generate higher alphas. The rightmost column shows that the FF3-alpha of all H-L portfolios are positive, although some of the t-statistics are not large enough to endorse statistical significance. The FF3-alpha is highest with 0.51 (t-statistic of 2.20) for the highest YK/R&D portfolio belonging to the highest R&D/ME quintile (i.e., cell (1,1)). At the other end of the FF3-alpha spectrum is the lowest YK/R&D portfolio belonging to the lowest R&D/ME quintile (i.e., cell (5,5)). Its FF3-alpha is -0.20 and its t-statistic is -1.24.

In the bottom row are the FF3-alphas and the t-statistics for the YK/R&D-selected portfolios controlling for R&D intensity.

Again, the results indicate a monotone relationship between R&D efficiency and risk-adjusted portfolio performance. In the H-L column, we see that after controlling for R&D intensity, the FF3-alpha of the H-L portfolio remains significant with 0.25% per month with t-statistic of 2.43. This shows that the long/short portfolio produces additional risk-adjusted performance relative to its risk exposure to the Fama-French three factors.
In the same spirit we reexamine the results of Table 4, in which we saw that portfolios with higher R&D intensity generate higher mean returns and alphas. At the end of each month, stocks are separated into quintiles based on YK/R&D. Then, within each YK/R&D quintile, we separate stocks into five subgroups based on R&D/ME. This way we create 25 groups of stocks each month and in each group we form an equal-weight portfolio. As before, we do the grouping of stocks within each of the 33 TSE industry classifications. This way we make sure that in each of the five R&D/ME-selected portfolios stocks are spread evenly over R&D efficiency. We call these five portfolios “R&D/ME-selected portfolios controlling for R&D efficiency.” The results are reported in Table 6.

In each YK/R&D quintile, the alpha is higher for higher R&D/ME portfolios. This time the monotonicity is violated only by one cell (cell (3,4)). Among firms with similar R&D efficiency, more R&D intensive firms generate higher alphas. The rightmost column shows that the FF3-alpha of all H-L portfolios are positive. The FF3-alpha is highest with 0.62 (t-statistic of 2.70) for the portfolio of most R&D-intensive firms in the highest YK/R&D quintile (cell (1,1)). At the other end of the FF3-alpha spectrum is the portfolio of the least R&D-intensive firms in the lowest YK/R&D quintile (cell (5,5)). Its FF3-alpha is -0.12 and its t-statistic is -0.82.

The bottom row reports the FF3-alphas and their t-statistics for the R&D/ME-selected portfolios controlling for R&D efficiency. The results again indicate a monotone relationship between R&D intensity and risk-adjusted portfolio performance. In the bottom H-L column we see that after controlling for R&D efficiency, the FF3-alpha of the H-L portfolio remains significant with 0.51% per month with t-statistic of 2.62. This shows that this long/short portfolio produces additional risk-adjusted performance relative to its risk exposure to the Fama-French three factors.

In conclusion, the results of Tables 5 and 6 suggest that the effects of R&D efficiency and R&D intensity are mutually independent. In addition, we can obtain returns more efficiently by two-dimensional selections based on these two aspects of R&D activity.

### 3.5 Comparison of long/short strategies

We proposed that the best use of the YK value as a signal for stock selection might be in combination with the data on firms’ R&D expenditure, which is readily available in financial statements. It enables us to measure firms’ IP competitiveness in two dimensions: the R&D efficiency (YK/R&D) and the R&D intensity (R&D/ME). We found in the control experiments of Tables 5 and 6 that these two factors carry independent information.

Thus, our best candidate for a portfolio strategy is to take a long position in the cell (1,1) portfolio and a short position in the cell (5,5) portfolio in either of the two-dimensional selections of Tables 5 or 6.

Table 7 reports the mean return and standard deviation of these long/short strategies in the first and the second column. Strategy 1 is the long/short strategy based on the two-dimensional selection of Table 5, and strategy 2 is the long/short strategy based on the two-dimensional selection of Table 6. The t-statistics are shown in parentheses for testing whether the mean return is significantly positive. We also report the Sharpe ratio.
Strategy 1 generates a mean return of 11.85% annually with a standard deviation of 10.23%. Strategy 2 generates a mean return of 11.50% annually with a standard deviation of 9.25%. Thus, strategy 1 has slightly higher mean return and higher risk than strategy 2. The Sharpe ratios are 1.14 and 1.23, respectively.

To compare the performance of these strategies with our “default strategy” of selecting stocks based on YK/ME, we separated stocks into 25 groups based on YK/ME and constructed an equal-weight portfolio in each group. The procedure is identical to the one used for Table 2 except that we now sort stocks into 25 groups rather than 5. We then take a long position in the highest YK/ME portfolio and a short position in the lowest YK/ME portfolio. This long/short portfolio has a mean return of 9.22% and standard deviation of 10.83%, yielding a Sharpe ratio of 0.84. The mean return is lower than strategy 1 and strategy 2. The standard deviation is higher than strategy 1 and strategy 2. Thus, we confirm a considerable performance improvement (increasing mean return and reducing risk) by decomposing IP competitiveness into two dimensions.

We also report the results for long/short portfolios based solely on R&D efficiency (YK/R&D) or R&D intensity (R&D/ME). The procedure is again identical to the one we used for Tables 3 and 4, respectively, except that we separated stocks into 25 groups.

The column YK/R&D is for stocks selected based on YK/R&D. The mean return is 4.46% annually and standard deviation is 6.70%.
The column R&D/ME is for stocks selected based on R&D/ME. The mean return is 11.38% annually and standard deviation is 15.32%. The Sharpe ratios are 0.65 and 0.73, respectively. It is surprising that merely using R&D expenditure to construct a long/short portfolio yields an annual mean return of 11.38%, which is comparable to our strategies 1 and 2. On the other hand, its risk is considerably higher so that the Sharpe ratio is much lower than the two outperforming strategies.

3.6 The industry factor
The importance of R&D activities varies across industries. There should be industries in which our investment strategy based on IP signals work very well and industries in which our strategy is not effective.

To answer this question, we examine how selections based on YK/ME are useful for each of the 33 TSE industry classifications. At the end of each month stocks are separated into quintiles based on YK/ME and equal-weight portfolios are formed in each quintile. Portfolio construction is done for each industry. The results are reported in Table 8. Portfolio 1 is the portfolio of stocks in the highest YK/ME quintile and portfolio 5 is in the lowest YK/ME quintile. Mean return is the time-series average of the annualized portfolio returns in percent. The rightmost column “H-L” refers to the return of a long/short portfolio which is long in portfolio 1 and shorts portfolio 5. The t-statistics for the mean return of the H-L portfolio is shown in parentheses.

Panel A lists industry groups with t-statistics greater than 2.0. chemicals, glass and ceramics products, machinery, electric appliances, transportation equipment, and land transportation are the industry groups in which significant mean return can be generated by the long/short strategy based on YK/ME. Panel B lists industries with negative mean return on the H-L portfolio. fishery, agriculture and forestry, pulp and paper, and oil and coal products are the industries in which the YK/ME signal does not seem to work, which is not surprising.

In the foregoing analysis, the universe of stocks covered all industries except financials. One can easily imagine our strategies working much more strikingly if one is allowed to be selective in industries.

4. Conclusion
The analysis in this paper suggests that the accumulation of the IP assets of Japanese firms is not sufficiently incorporated in the stock market’s pricing of firms. The YK value provides useful signals to predict future stock returns. We constructed long/short strategies based on (1) the YK value, (2) R&D expenditure, and (3) a combination of the two. The third strategy enhanced portfolio expected return and reduces risk considerably. The best-performing strategy generated an annual mean return of 11.50%, standard deviation of 9.25%, and a Sharpe ratio of 1.23.

Thus, making the YK value available to the investment community will satisfy the investors’ aspirations for higher returns. But more importantly, it will promote the efficiency of the Japanese stock market. The latter means that a firm with highly valuable IP assets will become able to raise equity capital at a lower cost. Constructing a technology-driven index such as Ocean Tomo 300 for use as a benchmark index will serve the investment community and simultaneously contribute to enhancing the IP competitiveness of the economy.

---

8. For pharmaceuticals, the mean return of H-L portfolio was 2.37% with a t-statistic of 0.49.
References
An intraday event study methodology for determining loss causation

Alex Rinaudo
Senior Vice President, Compass Lexecon

Atanu Saha
Executive Vice President, Compass Lexecon

Abstract
We set out an intraday event study methodology relying on minute-by-minute data and formulate an analytical framework to determine the window of time, i.e., the event window, over which stock prices fully reflect relevant new information. While the traditional daily price data-based event study approach assumes the event window to be a full day, this paper’s methodology allows the data to determine the length of the event window. This is particularly relevant because many prior studies have shown that stock prices reflect new information within a matter of minutes. Our event study model not only provides an accurate measurement of the stock price impact of a relevant event, but also determines whether the impact is statistically significant. Our study has obvious implications for institutional investors and asset managers engaged in securities litigation where loss causation, materiality and quantification of damages are at issue.

1 We wish to thank Senior Analysts, Nicolas Shea and George Wang, for their invaluable research and programming assistance throughout this research project and Senior Research Assistant Mary Shen for her research and editorial assistance.
An intraday event study methodology for determining loss causation

1. Introduction
In Rule 10b-5 causes of action, a central issue is establishing that a material misrepresentation or omission and the subsequent corrective disclosure was causally responsible for a drop in the stock price, resulting in economic losses to the shareholders. Typically, for 10b-5 damages, shareholder losses are measured through share price drop resulting from corrective disclosures. The price drop serves as a proxy for economic losses suffered when purchasing shares at inflated prices due to the alleged misrepresentation or omission.

This requirement of establishing “loss causation” has long been part of the common law. It has become increasingly emphasized by circuit courts, especially in light of the Supreme Court’s landmark decision in Dura Pharmaceuticals vs. Broudo. Prior to Dura, it was not uncommon for claimants in Rule 10b-5 matters to seek damages based on the entire stock price drop in the time period between the beginning of the alleged misrepresentation and the corrective disclosure. Since Dura, however, damage claims have typically relied on the stock price drop attributable to the corrective disclosure, net of market and industry effects. Dura underscored the importance of correctly identifying the effect of the relevant news (e.g., corrective disclosure) as distinct from other factors contributing to the share price drop: “that lower price may reflect, not the earlier misrepresentation, but changed economic circumstances, changed investor expectations, new industry-specific or firm-specific facts, conditions, or other events, which taken separately or together account for some or all of that lower price. Other things being equal, the longer the time between purchase and sale, the more likely that this is so, i.e., the more likely that other factors caused the loss.”

Event study analysis is a widely used method to disentangle these “other factors”, including market and industry effects, from the effect of the corrective disclosure in explaining a share price drop; it was and remains an integral tool for establishing loss causation and materiality in Rule 10b-5 matters. However, while the theory of loss causation has evolved based on the specificity provided by Dura, the event study methodology typically used by experts has not. Over the last three decades, event studies by and large have continued to use daily stock price changes (that is, the differences in daily closing prices). By relying on daily stock price changes, existing event studies implicitly assume that the impact of an event is not reflected in stock prices for many hours, and in some cases, as long as 24 hours. The latter would be true if an event occurred right after a trading day’s close at 4 pm; in that case, a daily closing price-based event study would assume that the event’s impact was not fully reflected in the relevant stock price until almost 24 hours later, in the next day’s closing price.

The problem is no less severe for events that occur during regular trading hours. Consider the following hypothetical: a firm announces at 2:00 PM that it will have to restate its financials for the preceding four quarters, and this news is characterized as a corrective disclosure by the claimants. Also assume on the same day, before the market opens, an analyst issues a downgrade of the firm’s stock; this downgrade precedes the restatement announcement and it is unrelated to it. In this hypothetical, the usage of close-to-close stock prices in an event study would not only include the price movement after the restatement announcement at 2:00 PM, but critically, would also include that day’s price movements before 2:00 PM, which contains the effect of the analyst’s downgrade, an event unrelated to the corrective disclosure. In other words, the closing price at 4:00 PM on that day would reflect the impact of both: the event at issue, i.e., the restatement announcement, and the confounding event of an analyst’s downgrade. The usage of close-to-close prices precludes disentanglement of the effect of the corrective disclosure from the effect of other confounding events that occur on the same day.

More generally, close-to-close price-based event study analyses potentially introduce the effects of extraneous events and, as a result, can provide an inaccurate measure of the true impact of the event at issue. This methodology, while widely accepted and used, is contrary to the spirit and the letter of Dura, which

---

2 The first circuit opinion to mention the requirement of “loss causation” in a Rule 10b-5 action was at the very start of the development of Rule 10b-5 jurisprudence in the Second Circuit’s opinion in Schlick v. Penn-Dixie Cement Corp., 507 F.2d 374 (1974). See Pasley v. Freeman, 3 T.R. 5:1, 100 Eng. Rep. 450, 457 (1789) (if “no injury is occasioned by the lie, it is not actionable . . .”); see also Dura Pharmaceuticals v. Broudo, 544 U.S. 336, 344 (2005) (collecting cites to common law requirement of loss causation). The requirement of loss causation for Rule 10b-5 causes of action was codified in the Private Securities Litigation Reform Act of 1995 which requires plaintiffs to “prove[] that the act or omission of the defendant alleged to violate [Section 10b] caused the loss for which the plaintiff seeks to recover damages.”


cautions us against conflating the impact of a corrective disclosure with “other factors [that] caused the loss” in share prices.

In the past, data limitations left event study practitioners with little choice but to use close-to-close daily stock price data. However, with the widespread availability of intraday stock price data, it is rather surprising that event study methodology has not evolved to rely on minute-by-minute data to accurately measure the impact of events. Importantly, as we discuss in more detail below, there is voluminous evidence and a large body of literature suggesting that for liquid stocks, new information is fully reflected in prices in a matter of minutes.

The key contribution of our paper is to set out an analytically consistent intraday event study methodology relying on minute-by-minute stock price data. We formulate an analytical framework to statistically determine the window of time over which stock prices fully reflect the relevant new information (i.e., the “event window”). While the traditional close-to-close data-based event study approach naturally assumes the event window to be a full day, this paper’s proposed methodology allows the data to determine the length of the event window. Our event study model not only provides an accurate measurement of the stock price impact of the relevant event, but also determines whether the impact is statistically significant. Thus, our intraday event study framework has obvious implications for institutional investors and asset managers engaged in securities litigation where loss causation, materiality and quantification of damages are at issue.

The remainder of the paper is organized as follows. In section 2, we provide additional motivation for the proposed intraday approach and discuss the relevant literature. Section 3 contains a discussion of the analytical framework. Finally, section 4 contains a few examples, which use actual intraday data to illustrate our approach and its application.

2. Motivation and literature review

2.1 A recent event

Before reviewing the relevant literature, we discuss a recent market event that received widespread publicity; this event provides compelling evidence on the remarkable speed with which news is assimilated in stock prices, and underscores the need to formulate an event-study framework consistent with these market dynamics. The Wall Street Journal provided a succinct summary of the event that occurred on 11 June 2013: “The Dow Jones Industrial Average plunged within minutes early this afternoon, after somebody hacked an AP Twitter account and posted a bogus tweet saying the White House had been attacked. The Dow, which had been up about 130 points, fell into the red within two minutes, and then bounced back just as quickly as it became obvious that the “news” was false, and a prank.”5 (emphasis added).

Figure 1 contains the minute-by-minute data on the movement of the Dow Jones index on that day. The figure shows that the Dow fell by 143 points within two minutes of the false AP Tweet; equally remarkably, the index fully recovered from the loss also within a few minutes after it became clear to the market that the news was false. The index moved a total of 280 points in less than five minutes. This rapid move elicited interesting press commentary regarding the role of various market participants. For example, on that day The Guardian observed: “Others attacked the allegedly pernicious influence of high-frequency trading algorithms that comb news and execute trades in nanoseconds … [A] well-known critic of HFT [high-frequency trading]...”

An intraday event study methodology for determining loss causation

Trading] … told The Guardian: “There’s a substantial business by high-frequency trading hedge funds reading machine-readable news sold to them for big bucks by brand-name news organizations....” (emphasis added).

The incident on June 11, 2013 illustrates two important aspects of the market dynamics: the effects of HFT based on computer algorithms and the increasingly dominant role of social media in disseminating information in the marketplace.

2.2 Changing market dynamics
While both phenomena – HFT and social media – have become increasingly important in recent years, the shrinkage of reaction time to news cannot solely be attributed to these recent developments. 24/7 news channels, including financial news channels, have been in existence for decades. The widespread usage of the Internet, particularly since the mid-1990s, has also led to broader and more rapid dissemination of news. With the availability of cheaper and faster computing power, algorithmic trading has been playing an increasing role for years. The advent of electronic exchanges, which also dates back a few decades, marked a sea change in the speed at which orders are placed and trades executed. Execution times in matter of seconds or less have been the norm for large institutional traders for decades.

All of these changes have led to increasingly rapid price response to new information in the equity markets.

Studies dating back to the 1980s have observed that stock prices incorporate new information within 15 minutes or less. For example, Patell and Wolfson (1984) is one of the first studies to examine intraday stock returns. They found that stocks react to new earnings or dividend information within 5 to 15 minutes after a news release. Many other prior studies also show that stock prices respond to news events in less than 15 minutes and in some cases in as short a time as a few minutes. While these studies have examined intraday returns, the methods they utilize are not well suited to quantify the impact of a particular event

---

on a particular company's stock price, which is the focus of an event study in the context of securities litigation. Prior studies noted above, utilizing intraday price data, have typically examined different types of events across many firms and determined an average time window over which stock prices react to a particular type of event.

2.3 Event study framework
The event study framework commonly used by experts in security litigation matters in determining loss causation and damages in the context of a single firm is well-established in the academic literature. Kothari and Warner (2007) review five leading journals from 1974 to 2000 and find more than 500 published articles containing event studies. The event study methodology, widely used today, was first introduced by Fama et al. (1969). Primarily due to the difficulty of collecting data, this paper along with many early event studies considered a minimum event window of one month. Since then, the most common event window has contracted from one month to a day as data has become more readily available and as computing power in processing of large datasets has become ubiquitous; however, the methodology and the interpretation of event study results have remained relatively unchanged.

While the event study method has remained virtually unchanged over the past 40 years, equity market dynamics, as noted earlier, have changed significantly, especially over the last 15 years. Despite the structural changes in the equity markets, and notwithstanding numerous studies documenting that stock prices react to events in a matter of minutes, the vast majority of event studies proffered in securities litigation matters use close-to-close stock prices. We believe one possible reason for this is the absence of a well-established intraday event study methodology in the academic literature. Our paper seeks to fill this void by proposing a method to determine the duration and magnitude of the effect of an event or events on a single firm's stock price, without a priori assuming the event window to be a full day. In doing so, we extend the framework of traditional event study analysis in another key area. In a typical event study framework, it is extremely common to assume that the daily changes in the stock price (i.e., stock's daily return) are normally distributed. This distributional assumption plays an important role in the determination of whether an event's stock price reaction is statistically significant, and thus, is critical in ascertaining materiality in the context of securities litigation.

The assumption that a stock’s daily returns are normally distributed implies that large stock price movements – more than twice the average daily moves in either direction – are unlikely; specifically, there is less than a 5% chance of observing moves of that magnitude. Consequently, if a sufficiently large market-adjusted stock price response to an event is observed, relying on the normality assumption, the typical event study practitioner concludes that the price change cannot be explained by random chance and is attributable to the event. The fundamental problem with this widely used approach in determining statistical significance of events is that numerous studies in the academic literature have established that stock price movements do not follow a normal distribution. Large moves in stock prices are far more frequent than those predicted by the normal distribution. Consequently, imposing the assumption of normality could lead one to erroneously conclude that an event at issue is statistically significant (i.e., material), when, in fact, data suggests that

---


14 The graphical representation of the distribution of data points of a normally distributed random variable is the familiar bell-shaped curve. The area below the curve up to any value along the horizontal axis (starting from the left end) represents the probability of observing a data point of that or lower value. The curve has a single peak at the average value of the random variable and the width of the “bell” depends on the volatility of the random variable, which is commonly measured by the standard deviation or “sigma” of the variable. The bell-shaped curve is symmetric around it peak, and drops off at both ends, implying that there is a lower probability of observing data points considerably higher or lower in value than the average. For example, that there is only a 5% chance of observing a data point that is approximately two standard deviations (i.e., a 2-sigma observation) away from the average value. Under normality, the chance of observing a 6-sigma event is less than one in a billion. Thus, if the stock market’s daily returns followed a normal distribution, then one would expect to observe a 6-sigma single-day market move (in either direction) only once in a billion trading days, which is more than 3.95 million years. Yet, in the last 110 years alone, the Dow Jones Index has experienced moves higher than six-sigmas in more than 55 days. (See, Sahal, A., B. G. Malkiel, and A. Greco, 2009, “The clustering of extreme movements: stock prices and the weather,” Journal of Investment Management 7(1), 5-14; Mandelbrot, B., and R. L. Hudson, 2006, The misbehavior of markets: a fractal view of financial turbulence, Basic Books; Karoglou, M., 2010, “Breaking down the non-normality of stock returns,” European Journal of Finance 16(1), 79-95. 15 See, for example, Fama, E. F., 1976, Foundations of finance, Basic Books; Brooks, C., 2002, Introductory econometrics for finance, Cambridge University Press. 16 This means that the true distribution of the stock price changes is “fat-tailed” (Karoglou (2010)). That is, there is a much higher probability of observing large price changes than predicted by the normal distribution, as noted in preceding footnote.
large price moves can happen by chance alone and should not necessarily be attributed to the event.

The deviation from normality is particularly pronounced for intraday stock returns. In this paper, we have tested for and rejected the assumption of normality in each of the intraday returns datasets used in the illustrative applications of our event study approach. Furthermore, studies of intraday volatility of stock prices have shown that volatility often is highest at the open of the trading day, declines through the early afternoon and then rises again near the end of the day.\(^\text{17}\)

In this paper, we propose an alternative measure of the determination of statistical significance without relying on the assumption of normality of returns. Our approach of determining statistical significance relies on the empirical distribution of intraday returns. This approach also allows the volatility of the intraday stock returns to vary by the time of the day. While our proposed method has not been used in traditional event study models, it is grounded, however, in the scientific econometrics literature. In particular, the usage of empirical distribution for determination of statistical significance is a well-accepted scientific method in academic studies.\(^\text{18}\)

3. Intraday event study methodology
In this section, we set out the analytical framework for the proposed intraday event study approach. Specifically, we first describe the method to determine the length of the event window; this determination, in turn, depends on ascertaining statistical significance of the intraday returns, which is discussed next. Finally, we discuss the computation of the stock price response in dollars (from the estimated return response) to the event at issue.

3.1 Estimation of cumulative abnormal returns
A principal objective of an event study model is to determine the market adjusted impact of an event on a firm's stock prices. In a traditional event study framework, this is accomplished through the use of a regression analysis in which the daily returns of the security at issue is regressed against a market and/or industry index, thereby allowing one to ascertain the residual return (attributable to company-specific news or events) on the event day. The typical event study model is as follows:

\[
\frac{\ln(p_t)}{p_t} = \beta_0 + \beta_1 M_t + \beta_2 I_t + \sum_{i=1}^{k} \alpha_i D_i + \epsilon_t
\]  

where \(p_t\) is the security closing price; \(M_t\) is the market index return; \(I_t\) is the industry index return; \(D_i\) is the ith dummy variable that takes the value of one on the ith event day, zero otherwise; \(\epsilon_t\) is the error term; and the subscript t denotes the tth day [Ferrell and Saha (2007)].

In an intraday event study framework, the first few steps are broadly similar to those in the traditional model set out above. The first is the choice of an appropriate control for market and industry effects. Typically this is accomplished through the use of an industry index.\(^\text{19}\) However, intraday (minute-by-minute) data are not readily available for many industry indices. We, therefore, propose the usage of industry Exchange Traded Funds (ETFs), which are traded throughout the day, and for which intraday data are readily available.\(^\text{20}\) Second, one must choose a control period

---


\(^{19}\) For simplicity, we use just one control variable, an industry index. In our experience, an additional, separate market index is generally insignificant after the addition of an appropriate industry index. However, each situation is unique and an individual determination regarding the appropriate control variables should be made in each situation. Our method is easily generalized to more than one control variable.

\(^{20}\) In some cases, Industry ETFs are too broad for the stock at issue in which case it may be appropriate to use a control constructed from a small set of comparable companies or even a single major competitor.
over which to estimate the relationship between the returns of the security at issue and returns of the industry. We examined control periods of varying lengths and found minimal difference in our intraday results for the examples discussed in the next section. We have explored control periods as short as 30 days to as long as 60 days. After determining an appropriate industry control and period, the next step is to run the following regression model using minute-by-minute returns data over the chosen control period:

\[
r_t = \alpha_0 + \alpha_1 \cdot I_t + \epsilon_t
\]  

(2)

where \( I_t \) is the industry index (ETF) return; \( \epsilon_t \) is the error term, and the subscript \( t \) denotes the \( t \)th minute. After \( \alpha_0 \) and \( \alpha_1 \) in equation (2) have been estimated using the control period's intraday data, we utilize the following formula to compute minute-by-minute abnormal return (i.e., \( AR_t \)):

\[
AR_t \equiv \bar{\epsilon}_t = r_t - \hat{\alpha}_0 - \hat{\alpha}_1 \cdot I_t
\]  

(3)

Using the per-minute abnormal returns given by (3), the cumulative abnormal return (CAR) is computed next. Specifically, the CAR from time 0 to \( T \) is defined as:

\[
CAR_T = \sum_{t=0}^{T} AR_t
\]  

(4)

where \( T \) denotes a period of time in minutes. This CAR is computed both for each period from time 0 to \( T \) of the control period (e.g., 60 days prior to the event day) and for each period from time 0 to \( T \) of the period immediately following the news at issue on the event day.22

For an event or news announcement to be material, it first has to have at least a period of statistically significant CAR following the event. To determine the "event window" (i.e., the period of time within which the impact of the event at issue is fully reflected in stock prices), we propose the following three steps: (a) compute the CAR for the first few minutes (e.g., 5 minutes) after the event at issue using the per-minute abnormal returns on the event day; (b) assuming this initial CAR is found to be statistically significant, continue computing the CAR incorporating each subsequent minute (i.e., \( CAR_{t+1}, CAR_{t+2}, \ldots \)) on the event day and test the statistical significance of each of these CARs; (c) let \( T^* \) denote the time interval in minutes following the event at which the CAR (denoted by \( CAR_{T^*} \)) ceases to be statistically significant.

The "event window" for a statistically significant event, therefore, is the time interval of \( T^* \) minutes after the event. Thus, we do not impose a priori impose a predetermined event window (e.g., a full day) but allow the returns data to reveal the length of time during which the relevant new information is fully reflected in the stock price. Needless to say, if the initial CAR (denoted by \( CAR_0 \)) is found to be statistically insignificant, then the information contained in the event at issue would not be deemed material.

As is clear from the foregoing, in our proposed approach, the determination of statistically significant CARs is critical for determining whether or not a relevant event has a material impact on the firm's stock price and, assuming it has, for the determining the length of the event window. We now turn to the method of determining statistical significance of CARs.

3.2 Testing for statistical significance of CARs

In traditional event study analysis, the most commonly used test for determining significance of daily stock price returns is the \( t \)-test; if the \( t \)-statistic is greater than 1.96 in absolute value then the event day's abnormal price move is considered to be statistically significant. However, this test statistic relies on the assumption that stock price returns are normally distributed. As discussed earlier, this assumption is contradicted by voluminous evidence in numerous prior studies. In this paper, we propose the testing of statistical significance of CARs following the event at issue using the empirical distribution of the time-matched CARs from the control period. Specifically, we compare event day's CAR for a specific time window after the event at issue with the threshold level of CAR, which is derived from the values of all CARs over the same time window within the control days. If value of the CAR of interest (following the event) is less than the 5th percentile of these time-matched control period's CARs then the event's CAR is considered to be statistically significant at the 95% confidence level.23

---

21 The usage of CAR for analysis of stock price response is well established in the econometric literature; see, for example, Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, The econometrics of financial markets, Princeton University Press.

22 The control period CARs are used to compute statistical significance as described in more detail below.

23
Observe, in proposing this method, we have not relied on the assumption of normality regarding the distribution of the stock price returns; instead, we have relied on the empirically observed distribution of the time-matched CARs during the control period. This approach allows the threshold level for statistical significance to change over the course of a day. This is particularly important because, as noted earlier, the volatility of intraday returns is more pronounced in the first and last thirty minutes of the trading day. As a result, large stock price moves are more likely to occur by random chance within these two periods than other periods within the day. Therefore, the threshold level to determine statistical significance of price move is likely to be different for these two half-hour periods than during the rest of the trading day.

Our proposed method is best clarified using an example. Suppose we are interested in determining the duration and magnitude of a firm's stock price reaction to a news event that occurs at 10:00 AM of a particular day. Also, assume, the practitioner has chosen the prior 60 days as the control period. In this case, the practitioner would first compute the CARs incorporating each additional minute after 10 AM using all 60 days in the control period. The empirical distribution of the control period's 60-days' CARs for each time interval after 10:00 AM provides the thresholds for determining statistical significance of the price response over the corresponding time intervals after 10:00 AM on the event day. For example, the CAR on the event day between 10:00 AM and 10:05 AM will be compared to the 5th percentile of the observed CARs for the same 5-minute time window (i.e., between 10:00 AM and 10:05 AM) for all 60 control days; next, the event day's CAR between 10:00 AM and 10:06 AM will be compared to the 5th percentile of the same 6-minute CARs for all the control days. This process continues until the event day's CAR is no longer statistically significant, that is, no longer less than the thresholds, which are the time-matched 5th percentiles of the control period's CARs. The moment in time when this occurs defines the event window, that is, the time period in minutes when the event is fully reflected in stock prices.

A graphical example of the time-varying empirical thresholds (based on actual intraday data discussed in more detail in the next section) is shown as the solid line in Figure 2. As is evident from this figure, the threshold level zigzags during the course of the trading day; we, therefore, propose a further refinement of smoothing these time-varying threshold levels. The dotted line in Figure 2 shows the smoothed time-varying threshold levels. A simple smoothing approach is to fit a polynomial function of the interval of time in minutes.

3.3 Dollar impact of event
Once the event window has been determined, the observed event day CAR during this time interval (denoted by CAR\text{\textsubscript{T*}} for the period between t=0 and t= T*, where t=0 is the time of the event at issue) can be used to compute the market-adjusted dollar impact of the event. The abnormal dollar impact (\Delta) is then given

\[ \Delta = D_{t} \times \text{market-adjusted price change} \]

23 We consider the specific case where the researcher wishes to test whether the drop in share prices is statistically significant. Our approach can be easily modified if the relevant question is whether the rise in share prices is statistically significant. In that case, one would examine whether the CAR of interest (following the event) is greater than 95% of the time-matched control period's CARs.

24 We propose to increase the order of the polynomial until the fit (as captured by Adjusted R\textsuperscript{2}) is no longer meaningfully increasing (e.g., in the second place of decimal). For example, the equation for the 3rd order polynomial would be: \[ TH = \gamma_{1} + \gamma_{2} x + \gamma_{3} x^{2} + \gamma_{4} x^{3} \] where \( TH \) denotes the time-varying per minute threshold level, and \( x \) denotes the elapsed time in minutes over which the threshold level is being computed. In general we found that no more than a fifth order polynomial was necessary, and often a third order polynomial showed sufficient explanatory power.
by the following formula:\textsuperscript{25}
\[ \Delta = P_{t^*} \cdot (1 - \exp(-\text{CAR}_{t^*})) \]  
where $P_{t^*}$ denotes the stock price at time $T^*$ minutes after the event occurs.

4. Illustrative applications of the intraday event study methodology

In this section, we illustrate our proposed intraday event study framework through its application using actual data on three firm’s intraday stock prices. In the first example, a single unexpected event occurred during regular trading hours. In the second example, we demonstrate the method’s ability to disentangle the effect of two events occurring on the same day. Our third example shows the method is capable of distinguishing between the effects of market conjectures and a company press release. All minute-by-minute data used in these examples are publicly available and have been obtained from TickData.com. News reports related to the events discussed were obtained from Bloomberg L.P.

It is important to emphasize that, to our knowledge, none of these firms whose names have not been divulged, face any 10b-5 claims or securities litigation matters based on these events on the chosen three days. These events have been chosen purely to illustrate our approach.

4.1 Example 1 - a single event

Our first example considers a single unexpected event that occurred during the trading day. Given that this is the only event occurring during this day, we would expect the intraday method to find an event impact estimate similar to one found using a daily price change based event study. The event at issue is the announcement of a lawsuit against a major publicly quoted firm, company A. The lawsuit was filed at 1:11 PM. We find news articles commenting on the filing as early as 1:19 PM. As can be seen in Figure 3a, which depicts the intraday prices of company A, this news had an almost immediate negative impact.

We begin our analysis by selecting an appropriate industry control. In this case, company A is a diversified company with revenue sources spanning many industries. Thus, for simplicity, we utilize the S&P 500 ETF (the ticker symbol is SPY) to control for market and industry movement. An event study utilizing close-to-close prices shows that the price response on this day is statistically significant, with a market-adjusted impact of

negative U.S.$7.55. To evaluate the impact of this event in an intraday framework, we start by estimating the CARs starting at 1:11 PM. We also determine the time-varying thresholds for significance based upon the empirical distribution of CARs during the control period of 60 trading days prior to the day at issue. Both the event day CAR after 1:11 PM (the solid line) and the smoothed time-varying thresholds for significance (dotted line) are shown Figure 3b.

The event day’s CARs after the event falls below the threshold of significance almost immediately and remains there for the rest of the trading day. Thus, the event window is 1:11 PM to 4:00 PM. Using equation (5), we estimate the abnormal dollar impact to be negative U.S.$7.36, which is slightly lower than but consistent with the close-to-close based estimate of U.S.$7.55. Our intraday estimate is lower because it does not contain the price decline that occurred prior to the news announcement. Note that the event day CAR is below the threshold level at the end of the day, which suggests that the event impact might continue into the next trading day.

Our second example involves the share price reaction of company B to two separate events. The company released its earnings at 8:07 AM and held an investor conference call at 10:00 AM. The intraday stock prices of company B are shown in Figure 4a. A key advantage of the intraday event study approach, as noted earlier, is the ability to disentangle the separate impact of events occurring on the same day. In the above example, let us hypothetically assume that there is an allegation that company B had falsified its earnings for the previous year, and it disclosed the restatement as well as the corrected lower earnings at 8:07 AM on that day. Using close-to-close prices, one would attribute the entire day’s price drop to the earnings restatement. It is clear from Figure 4a that, while there is some downward movement after the open of trading, there is a pronounced price drop during the earnings call, which started at 10:00 AM. Therefore, one would need to determine what information was released at this point during the earnings call and whether or not it was related to the last year’s earnings. For illustrative purposes, hypothetically assume that the review of the earnings transcript reveals that the earnings call was not related to last year’s restated earnings but was focused on future earnings. Also assume that the management provided lower guidance for future quarters during the earnings call. In this scenario, the 8:07 AM restatement
announcement is the corrective disclosure and the event at issue; but the earnings call is not. Close-to-close price based event study would consider the full day as having one event. The intraday method, by contrast, has the potential to disentangle these two events and determine their impact and significance individually.

Before any assessment of abnormal price impact can be made, a suitable industry control must be chosen. We considered several leisure industry indices, but found that all were too broad to serve as a suitable industry control. We then decided to use the closest publicly traded competitor as the industry control. After controlling for industry effects, a close-to-close price based event study finds the price change of company B to be statistically significant, with industry-adjusted dollar impact of -U.S.$0.53. As seen in the Figure 4b, intraday event study analysis reveals that, after the open of trading, company B’s stock price response to the first event (i.e., restatement) is statistically significant and negative only for the time window between 9:31 AM and 9:46 AM.

To examine the impact of the second event (i.e., earnings call), we create CARs and thresholds for significance starting at 10:00 AM, when the earnings call began. This analysis shows the stock price reaction to the earnings call is clearly statistically significant, and it continues to be so for the rest of the trading day. Thus, our intraday event study methodology is able to disentangle the effects of two separate events: (a) the one at issue, the hypothetical restatement announcement at 8:07 AM; its impact

### Table 1: Close-to-close versus intraday event study estimates of impact

<table>
<thead>
<tr>
<th></th>
<th>Close-to-close method</th>
<th></th>
<th>Intraday method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Market-adjusted return</td>
<td>Market-adjusted dollar impact</td>
<td>Market-adjusted return</td>
</tr>
<tr>
<td>Company A</td>
<td>-13.99%</td>
<td>-U.S.$7.55</td>
<td>-13.66%a</td>
</tr>
<tr>
<td>Company B</td>
<td>-6.53%</td>
<td>-U.S.$0.53</td>
<td>-3.08%b</td>
</tr>
<tr>
<td>Company C</td>
<td>-4.15%</td>
<td>-U.S.$0.51</td>
<td>0.00%c</td>
</tr>
</tbody>
</table>

Note: An Intraday event CAR is from 13:11–16:00. b Intraday event CAR is from 9:31–9:46. c Intraday event period is after 10:14 AM

Source: Bloomberg, L.P; Tick Data
An intraday event study methodology for determining loss causation

is found to be statistically significant between 9:36 AM and 9:46 AM; and (b) the confounding event of the earnings call which dealt with future earnings and it had a statistically significant impact due to lowered guidance.

4.3 Example 3 - market speculation versus company announcement

Our final example shows the effectiveness of our intraday event study method to separate the effects of rumors and market speculation from the impact of a firm's announcement. This example relates to company C's stock price movements on a select day. On that day, the market opened with speculation swirling about the company's exposure to European Sovereign debt. At 10:14 AM, the company released a statement providing clarity regarding its exposure to Europe. Figure 5a contains its stock prices over the course of the day.

The event day's CAR and the threshold of significance for company C, both before and after 10:14 AM, are shown in Figure 5b. This figure shows that the rumors and market speculation regarding the company's exposure to European Sovereign debt clearly had a significant impact soon after market opened, and the CAR fell below the significance threshold. By 10:00 AM, however, the company's share prices began to recover as is evident from the rising CAR. After the release of statement at 10:14 AM, the price reaction is positive and the CAR stays above the threshold of significance for the rest of the day.

4.4 A summary of our findings

In Table 1, we present a summary of our findings from the illustrative examples. We compare the results from the traditional close-to-close event study analysis with those from the intraday approach. As noted earlier, for company A, since there is a single event during the trading day, the two approaches yield fairly similar results. By contrast, the event impact estimates are markedly different for the remaining two examples. For both of these, the close-to-close approach fails to disentangle the impact of the event at issue from confounding news or information: in the case of company B, the confounding news is the earnings guidance for future quarters, which occurs after the event at issue (the hypothetical restatement announcement); and for company C, the confounding news is the market's conjectures about the firm's exposure to European debt before the event at issue (the company's press release). Because of its inability to disentangle confounding effects, the close-to-close approach for the company B example yields an impact estimate that is 96% higher (U.S.$0.53 versus U.S.$0.27 on a per share basis) and for company C it finds an impact of U.S.$0.51 per share, when, in fact, the impact is zero.

These results illustrate that the close-to-close event study approach may, in many cases, particularly in the presence of confounding events, yield erroneous estimates of event impact, and lead to incorrect inferences regarding loss causation and materiality in securities litigation matters. While these differences measured on the per share basis may seem small, they have a magnified effect when it comes to settlement or damages. For example, company B has approximately 200 million shares outstanding. Assuming they are all eligible for a damages recovery, the daily event study would estimate damages at U.S.$106m ($0.53 x 200 million shares) while the intraday event study method would estimate damages at U.S.$54m ($0.27 x 200 million shares).

We view the implications of the intraday method as relevant to two separate groups: investors bringing lawsuits, and corporations defending lawsuits. Investors (including pension funds, asset managers and financial institutions) who serve as plaintiffs in securities litigation cases can use this method to (1) more accurately estimate the impact associated with an event of interest leading to better cost-benefit analyses; and (2) separate the impact of different events within the same 24 hour period, which the daily event study method does not allow and would otherwise render the analysis of the event “confounded.” Meanwhile, corporations (including many financial institutions) are often the subject of securities litigation. This method would allow them to narrow the time period of inquiry and potentially reduce damages as a result.

5. Concluding comments

In this paper, we have set out an intraday event study methodology relying on minute-by-minute data and have formulated an analytical framework to determine the window of time, i.e., the event window, over which stock prices fully reflect the relevant new information. Our study extends the rich body of literature on event studies based on closing prices of stocks. The principal motivation for our paper comes from many academic studies' observation that stock prices assimilate new information...
Are negative P/E ratio firms different from positive P/E firms? 
The case of interlisted vs. non-interlisted firms in Canada

George Athanassakos
Ben Graham Chair in Value Investing, Ivey Business School, Western University, London, Ontario, Canada

Abstract
Using separately interlisted and non-interlisted Canadian stock market data for the period 1985-2010, the main purpose of this paper is to examine whether negative P/E stocks are really different from positive P/E firms, and whether the former outperform, on average, the latter. The paper also purports to examine (a) whether interlisted and non-interlisted firms behave similarly or there are distinct differences between them and (b) whether there are differences in relation to this paper's key questions only in one group of stocks or whether differences are equally driven by both. We find that firms with negative multiples are indeed different from firms with positive multiples in that (a) a relatively small number of firms with negative multiples experience high forward stock returns, even though in the majority of cases there isn't a large difference between mean and median returns and (b) the value, size, liquidity and business risk premiums behave differently for negative P/E as opposed to positive P/E firms. This indicates that prior academic research was right in excluding negative multiple firms from their analyses. Moreover, the paper also shows that there are key differences between interlisted and non-interlisted firms both in the positive and negative P/E space. As a result, not only must negative P/E firms be segregated from positive multiple firms, but also interlisted firms ought to be segregated from non-interlisted firms in related research as aggregation would undermine the clarity and generality of findings, affect the homogeneity of the sample, and dilute findings and tests of significance.
1. Introduction

Academic papers examining the performance of value versus growth firms (i.e., low versus high P/E ratio firms) and the so-called value premium exclude from their analyses negative P/E firms as they are believed to behave differently from the positive multiple firms. To keep their sample homogenous, these studies segregated the negative multiple firms from the positive ones [see, for example, Basu (1977), Chan et al. (1991), Fama and French (1992, 1993, 1995) and Lakonishok et al. (1994)]. But is this the right thing to do? Moreover, empirical research in Finance has documented that the lowest P/E stocks tend to outperform [see, for example, Chan and Lakonishok (2004) and Athanassakos (2009, 2011, 2013)]. As negative P/E firms can be considered to belong to the lowest P/E group of stocks, it begs the question whether they will also have superior performance.

Additionally, research in finance has aggregated together Canadian interlisted and non-interlisted stocks. But is this right? First, interlisted stocks tend to be larger, more glamorous and well known than non-interlisted stocks and are followed by more analysts [Athanassakos et al. (2010)]. If we wish to keep our sample as homogeneous as possible, then interlisted must be segregated from non-interlisted stocks. Second, there is evidence that the marginal trader for Canadian interlisted stocks is an American investor [see Booth and Johnston (1984)]. If stocks that are most likely to be traded by Canadians behave differently than those that are not, then, again for homogeneity sake, interlisted must once more be segregated from non-interlisted stocks.

Separating interlisted from non-interlisted Canadian stock market data for the period 1985–2010, we will address two key questions in this paper: first, are negative P/E stocks really different from positive P/E firms, and second, as one can view the negative P/E stocks as part of the lowest P/E stocks, do they outperform, on average, the universe of positive P/E stocks? Answering these questions will be of importance to both institutional and individual investors, particularly those who follow a value approach, as well as academics who need to have clear evidence justifying the exclusion of negative P/E stocks from studies of the value premium. We would additionally like to examine whether there are distinct differences in negative P/E stocks depending on whether they are interlisted or not and how negative P/E firms differ from those that have a positive P/E, be they interlisted or not.

We will ask the following questions: do interlisted and non-interlisted firms behave similarly or are there distinct differences between them that grouping them would bias results and produce unfounded generalizations and conclusions? And, are there differences in relation to this paper’s key questions only in one group of stocks or are differences equally driven by both? Findings will have implications not only for the Canadian, but also for the U.S. markets.

While previous studies [see, for example, Fama and French (1992, 1993, 1995), Lakonishok et al. (1994), and Chan and Lakonishok (2004)] derive trailing price-to-earnings (P/E) ratios using price as at the end of June of year (t) and earnings per share as of fiscal end of year (t–1), this study will derive trailing ratios where price is as at the end of April of year (t), given that our sample only includes firms that already have reported financials (for fiscal year (t–1)) by the end of April of year (t). We see no reason to wait until June given that a stock selection strategy can be implemented at an earlier time.

This paper shows that firms with negative multiples are indeed different from firms with positive multiples in that (a) a relatively small number of firms with negative multiples experience high forward stock returns even though the majority of them does not, resulting in a large difference between mean and median returns, and (b) the value, size, liquidity and business risk premiums behave differently for negative, as opposed to positive, P/E firms. This indicates that prior academic research was correct in excluding negative multiple firms from their analyses. Moreover, the paper also shows that there are key differences between interlisted and non-interlisted firms both in the positive and negative P/E space. As a result, not only must negative P/E firms be segregated from positive multiple firms, but also interlisted firms ought to be segregated from non-interlisted firms in related research as aggregation would undermine the clarity and generality of findings, affect the homogeneity of the sample and dilute findings and tests of significance.

The rest of the paper is structured as follows: section 2 discusses the data sources, sample selection and methodology, section 3...
reports the empirical results and compares the performance of positive and negative multiples stocks and that of interlisted and non-interlisted stocks, and section 4 concludes the paper and discusses the implications of findings.

2. Data and methodology
Our sample includes all interlisted and non-interlisted companies that traded on Canadian Stock Exchanges for the period 1 May 1985–30 April 2010, as well as their financials for the period 1984-2008.

The paper uses data from COMPUSTAT, from which trailing price-to-earnings (P/E), total returns, stock liquidity, market cap and firm fundamentals are derived. For the trailing P/E ratios, the price (P) is as of the end of April of year (t) and E is the fully diluted annual earnings per share for companies with fiscal year end in year (t–1), as reported in COMPUSTAT.

Firm fundamentals, derived from company financials, are defined as follows: CASH is cash plus marketable securities over assets. EBIT MARGIN is EBIT over revenues (i.e., operating margin). TURNOVER is revenues over assets. CURRENT RATIO is the ratio of cash plus short-term investments, inventories and accounts receivable to current liabilities. DEBT is short- and long-term debt to equity. EPS GROWTH, EBIT GROWTH and REV GROWTH are the annual growth rates of EPS, EBIT and revenues, respectively for fiscal year (t–1). Market metrics are defined as follows: MARKET CAP is derived by multiplying price per share by shares outstanding at the end of April of year (t). LIQUIDITY is the annual stock trading volume of the year prior to May of year (t) over shares outstanding as at April of year (t). BUSINESS RISK is the industry code that captures an industry’s business risk (1=low business risk, 2=medium business risk and 3=high business risk). BUSINESS RISK assigns industry groups from COMPUSTAT to business risk categories which are based on results reported by Athanassakos (1998).

The industries belonging to each of the three business risk categories referred to above are from Athanassakos (2013).

To eliminate likely data errors (see, for example, Griffin and Lemmon (2002) and Cohen, Polk and Vuolteenaho (2003)), we have excluded firms with P/E values over |500|. Finally, to be included in our sample, a stock had to have a price of at least $1 and to have reported financials in COMPUSTAT.

After all aforementioned screenings, for the positive P/E stocks, we end up with 1,043 unique companies (6,479 firm-year observations) belonging to the non-interlisted sample and 219 unique companies (1,502 firm-year observations) belonging to the interlisted sample.

For the negative P/E stocks, the non-interlisted sample includes 8,059 firm-year observations for 1,322 unique firms, while the interlisted sample includes 864 firm-year observations for 124 unique firms.

Annual total stock returns are calculated as the price change plus the dividend from 1 May of year (t) to 30 April of year (t+1) over the price in 1 May of year (t). In other words, total stock returns are calculated for the year following the formation of the P/E ratios (i.e., May year (t) to April year (t+1)). Equally weighted mean (and median) returns for each group of stocks are then derived (see, for example, Fama and French (1992), Lakonishok et al. (1994)).

Non-overlapping forward annual stock returns, which are adjusted for stock splits and stock dividends, are thus obtained over the period 1 May 1985–30 April 2010. Trailing company fundamentals, as defined earlier, are for the period 1984 to 2008.

3. Empirical results
3.1 Summary statistics
Tables 1 and 2 report, respectively, the summary statistics for key variables of interlisted and non-interlisted firms with positive and negative P/E ratios. The tables include the mean, median and first and third quartile statistics for each variable.

In these tables, we observe that there is a significant discrepancy between mean and median values. This is particularly true for the negative P/E sample. As a result, we will focus our discussion on the median values.

In Table 1, Panel A, we see that, for interlisted firms, the median
Are negative P/E ratio firms different from positive P/E firms? 
The case of interlisted vs. non-interlisted firms in Canada

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
<th>Panel C: Median tests (p-values)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>25%</td>
</tr>
<tr>
<td>EBIT margin</td>
<td>0.17</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>Current ratio</td>
<td>2.65</td>
<td>1.66</td>
<td>1.09</td>
</tr>
<tr>
<td>Cash</td>
<td>0.12</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Debt</td>
<td>0.79</td>
<td>0.43</td>
<td>0.14</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.66</td>
<td>0.5</td>
<td>0.29</td>
</tr>
<tr>
<td>EPS growth</td>
<td>0.43</td>
<td>-0.02</td>
<td>-0.56</td>
</tr>
<tr>
<td>EBIT growth</td>
<td>0.6</td>
<td>0.11</td>
<td>-0.21</td>
</tr>
<tr>
<td>Rev growth</td>
<td>0.43</td>
<td>0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>Liquidity</td>
<td>0.77</td>
<td>0.27</td>
<td>0.06</td>
</tr>
<tr>
<td>Market cap</td>
<td>6389.9</td>
<td>1895.3</td>
<td>376.7</td>
</tr>
<tr>
<td>Return</td>
<td>0.15</td>
<td>0.1</td>
<td>-0.15</td>
</tr>
<tr>
<td>P/E</td>
<td>35.5</td>
<td>19.7</td>
<td>13.1</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics of interlisted firms for 1 May 1985–30 April 2010

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
<th>Panel C: Median tests (p-values)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>25%</td>
</tr>
<tr>
<td>EBIT margin</td>
<td>0.18</td>
<td>0.12</td>
<td>0.06</td>
</tr>
<tr>
<td>Current ratio</td>
<td>2.77</td>
<td>1.59</td>
<td>1.06</td>
</tr>
<tr>
<td>Cash</td>
<td>0.09</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Debt</td>
<td>0.67</td>
<td>0.38</td>
<td>0.09</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.97</td>
<td>0.75</td>
<td>0.30</td>
</tr>
<tr>
<td>EPS growth</td>
<td>0.13</td>
<td>0.01</td>
<td>-0.47</td>
</tr>
<tr>
<td>EBIT growth</td>
<td>0.23</td>
<td>0.10</td>
<td>-0.16</td>
</tr>
<tr>
<td>Rev growth</td>
<td>0.47</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>Liquidity</td>
<td>0.34</td>
<td>0.19</td>
<td>0.06</td>
</tr>
<tr>
<td>Market cap</td>
<td>760.0</td>
<td>154.7</td>
<td>54.1</td>
</tr>
<tr>
<td>Return</td>
<td>0.12</td>
<td>0.06</td>
<td>-0.16</td>
</tr>
<tr>
<td>P/E</td>
<td>27.9</td>
<td>15.7</td>
<td>10.3</td>
</tr>
</tbody>
</table>

Table 2: Summary statistics of non-interlisted firms for 1 May 1985–30 April 2010
EBIT margin and turnover for the positive P/E firms are 15% and 0.50, respectively. The median annual growth rates of revenues and EBIT have all been positive over the sample period, but not that of EPS. The median firm is not overleveraged as indicated by the debt to equity ratio of 0.43 and has a market cap is CAD$1895.3m. Median values for cash (and marketable securities) to assets and current ratio are 6% and 1.66, respectively. Moreover, the median firm trades about 27% of the shares outstanding over the previous year. Finally, the median stock return of firms with positive P/Es is 10%.

Comparing Panels A and B of Table 1, we see that firms with negative P/Es have negative median EBIT margin and EBIT growth rate as opposed to positive ones for firms with positive P/Es. Negative P/E firms also have lower market cap, debt to equity and turnover, but higher liquidity than firms with positive multiples. Median tests, reported in Table 1 (Panel C) and based on CHI-SQUARE tests for testing the null hypothesis that median values for the variables of Panels A and B (i.e., of positive vs. negative multiple firms) are equal, show that median values of Panel A variables are statistically different from the median values of the same variables in Table 2, Panel B, at conventional levels of significance.

In Table 2, Panel A, we see that, for non-interlisted stocks, the median EBIT margin and turnover for the positive P/E firm are 12% and 0.75, respectively. The median annual growth rates of revenues, EPS and EBIT have all been positive over the sample period. The median firm is not overleveraged as indicated by the debt to equity ratio of 0.38 and has a market cap is CAD$154.7m. Median values for cash (and marketable securities) to assets and current ratio are 3% and 1.59, respectively. Moreover, the median firm trades about 19% of the shares outstanding over the previous year. Finally, the median stock return of firms with positive P/Es is 6%.

Comparing Panels A and B of Table 2, we see that firms with negative P/Es have negative median EBIT margin and EPS and EBIT growth rate as opposed to positive ones for firms with positive P/Es. Negative P/E firms also have lower market cap, debt to equity and turnover, but higher liquidity than firms with positive multiples. Median tests, reported in Panel C, show that median values of Table 2, Panel A variables are statistically different from the median values of the same variables in Table 2, Panel B, at conventional levels of significance.

In Panel B, we also see that there is a larger difference between mean and median returns when firms have negative P/E (6% and -2%, respectively) vis-à-vis corresponding numbers when firms have positive multiples (12% versus 6%, respectively (see Panel A)). Similar to the evidence from the interlisted stocks, this may indicate that there are relatively more high positive return outliers for the negative P/E sample than there are in the positive P/E sample and that while the majority of negative P/E firms have negative returns, many high positive return stocks can still be found within the negative P/E stocks. Non-interlisted firms with negative multiples are thus also different than non-interlisted firms with positive multiples.

### 3.1.2 Univariate analysis

**A. Low versus high P/E ratios: is there a value premium in negative P/E stocks?**

Tables 3 and 4 show, respectively, that negative P/E interlisted and non-interlisted firms are different than positive P/E firms in another dimension.

In Table 3, Panel A, we see that interlisted firms with below median positive P/E values outperform those with above median P/E values (median return of 13% vs. 5%), i.e., there is a value premium consistent with prior research (see, for example, Athanassakos (2009, 2011)). However, in Panel B, we see that the below median negative P/E firms have a median return of -3% vs. -2% for the above median P/E firms. Unlike positive P/E firms there is no value premium in negative P/E interlisted firms.
Are negative P/E ratio firms different from positive P/E firms?  
The case of interlisted vs. non-interlisted firms in Canada

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Liquidity</td>
<td>1.04</td>
<td>0.43</td>
</tr>
<tr>
<td>Market cap</td>
<td>5060.3</td>
<td>1680.4</td>
</tr>
<tr>
<td>P/E</td>
<td>58.3</td>
<td>36.1</td>
</tr>
</tbody>
</table>

Table 3: Annual returns of low P/E versus high P/E firms for 1 May 1985–30 April 2010: interlisted firms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Liquidity</td>
<td>0.37</td>
<td>0.22</td>
</tr>
<tr>
<td>Market cap</td>
<td>834.3</td>
<td>202.1</td>
</tr>
<tr>
<td>P/E</td>
<td>46.0</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Table 4: Annual returns of low P/E versus high P/E firms for 1 May 1985–30 April 2010: non-interlisted firms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>0.144</td>
<td>0.13</td>
</tr>
<tr>
<td>Liquidity</td>
<td>0.98</td>
<td>0.32</td>
</tr>
<tr>
<td>Market cap</td>
<td>12218.9</td>
<td>6807.3</td>
</tr>
<tr>
<td>P/E</td>
<td>30.5</td>
<td>18.5</td>
</tr>
</tbody>
</table>

Table 5: Annual returns of small versus large cap firms for 1 May 1985–30 April 2010: interlisted firms

In Table 4, Panel A, we see that non-interlisted firms with below median positive P/E values outperform those with above median P/E values (median return of 10% versus 2%). In Table 4, Panel B, however, we see that while there is a value premium when comparing median returns (below median negative P/E firms have a median return of -1% versus -6% for the above median P/E firms), there is no value premium in mean returns – both groups of stocks have a mean return of 6%. This is unlike the case of positive P/E non-interlisted stocks where there is a value premium no matter how one measures it.

B. Small cap versus large cap firms: is there a size premium for negative P/E stocks?  
This section shows that there is another difference between
positive and negative P/E stocks in our sample of interlisted and non-intelisted companies that relates to the size premium.

In Table 5, Panel A, we see that, when comparing below and above median size firms, there is no size premium in the positive P/E interlisted firms (below median market cap firms have a return of 9% versus a return of 13% for the above median market cap firms). But there is one for the negative P/E stocks, as shown in Panel B. Above median market cap firms with negative P/Es have a median return of -4% while the below median firms with negative P/Es have a median return of -1%. Although the difference is not statistically significant, it is in the right direction, as opposed to the case of positive P/E interlisted firms, where the direction is opposite (and statistically significant) from the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Liquidity</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Market cap</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>P/E</td>
<td>Mean</td>
<td>Median</td>
</tr>
</tbody>
</table>

Table 6: Annual returns of small versus large cap firms for 1 May 1985–30 April 2010: non-interlisted firms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Liquidity</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Market cap</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>P/E</td>
<td>Mean</td>
<td>Median</td>
</tr>
</tbody>
</table>

Table 7: Annual returns of low liquidity versus high liquidity firms for 1 May 1985–30 April 2010: interlisted firms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel A: Positive P/E firms</th>
<th>Panel B: Negative P/E firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above median</td>
<td>Below median</td>
</tr>
<tr>
<td>Return</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Liquidity</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Market cap</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>P/E</td>
<td>Mean</td>
<td>Median</td>
</tr>
</tbody>
</table>

Table 8: Annual returns of low liquidity versus high liquidity firms for 1 May 1985–30 April 2010: non-interlisted firms
Are negative P/E ratio firms different from positive P/E firms?  
The case of interlisted vs. non-interlisted firms in Canada

one documented in the literature [see, for example, Kothari et al. (1995)].

In Table 6, Panels A and B, we see that there are size premiums in both the positive and negative P/E non-interlisted firms with a stronger size premium for the latter firms. Above median market cap firms have a median return of 4% while the below median firms have a median return of 7% for the positive P/E firms, with the corresponding numbers being –7% and 0.0% for the negative P/E firms. We see that segregating interlisted from non-interlisted stocks brings to life important differences in the behavior of the size premium between positive and negative P/E firms.3

C. Low liquidity versus high liquidity firms: is there a liquidity premium for negative P/E stocks?  
This section highlights yet another difference between positive and negative P/E stocks in our sample of interlisted and non-interlisted companies that relates to the liquidity premium. In Table 7, Panel A, we see that, when comparing below and above median liquidity firms, there is no liquidity premium in the positive P/E interlisted firms. But there is one in for the negative P/E stocks as shown in Panel B. Above median liquidity firms with negative P/Es have a median return of –5% while the below median firms with negative P/Es have a median return of –1%.

In Table 8, Panels A and B, we see that there is a liquidity premium in both positive and negative P/E non-interlisted firms with a stronger liquidity premium for the latter non-intelisted P/E firms. Above median liquidity firms have a median return of 4% while the below median firms have a median return of 8% for the positive P/E firms, while the corresponding numbers are –10%

3 Kothari et al. (1995) find that smaller firms tend to perform unequivocally, economically and statistically better than larger firms. But they did not segregate the positive from the negative P/E firms and the interlisted from the non-interlisted firms. Segregation allows us to zero in and find more finite differences between the various groups of stocks and avoid generalizations that may be unwarranted or not applicable in all cases.
and 0.0% for the negative P/E ratio firms. This is another proof of the need to segregate interlisted from non-interlisted stocks as important differences in the behavior of the liquidity premium between positive and negative P/E firms are discovered.4

D. Low business risk versus high business risk firms: is there a risk premium for negative P/E stocks?
In Tables 9 and 10, we observe more differences between positive and negative P/E stocks in our sample of interlisted and non-interlisted companies that relate to the business risk premium. In Table 9, Panel A, we see that the high business risk stocks have the same median return as the low business risk group for the positive P/E interlisted firms. For the negative P/E stocks, as shown in Panel B, the high business risk group has clearly lower median return than the low business risk group. For example, the median return for the high business risk negative P/E group is –5%, while that for the low business risk negative P/E group is 8%.

---

4 Baker and Stein (2004) and Fang, et al. (2008), among others, find that less liquid firms tend to perform better than more liquid firms. But again, they did not segregate the positive from the negative P/E firms and the interlisted from the non-interlisted firms. This further exemplifies the importance of segregating negative from positive multiple firms and interlisted from non-interlisted firms in related research as inclusion would undermine the clarity and generality of findings and dilute the significance of empirical evidence.
In Table 10, Panels A and B, we see that the high business risk non-interlisted stocks have lower median return than the low business risk group for both the positive and negative P/E non-interlisted firms. For the negative P/E stocks, as shown in Panel B, the median return for the high business risk group is -6%, while that for the low business risk group is 4%. For the positive P/E firms shown in Panel A, the corresponding numbers are 4% and 9%. Here too, we see that there are differences in the behavior of positive and negative P/E firms when it comes to the business risk premium.

### 3.1.3 Regression analysis

In this section, we further examine differences between positive and negative P/E firms in our sample of interlisted and non-interlisted stocks looking, more formally, at the relationship between forward stock returns and historical firm fundamentals and market metrics by estimating the following regressions:

**Interlisted**

\[
R_{it} = a_0 + a_1 \cdot \text{CASH}_{it-1} + a_2 \cdot \text{EBITMARGIN}_{it-1} + a_3 \cdot \text{MARKETCAP}_{it-1} + a_4 \cdot \text{LIQUIDITY}_{it-1} + e_{it},
\]

**Non-interlisted**

\[
R_{it} = a_0 + a_1 \cdot \text{PE}_{it-1} + a_2 \cdot \text{EBITMARGIN}_{it-1} + a_3 \cdot \text{MARKETCAP}_{it-1} + a_4 \cdot \text{LIQUIDITY}_{it-1} + a_5 \cdot \text{TURNOVER}_{it-1} + e_{it},
\]

where the dependent variable, \(R_{it}\), is the annual return for firm \(i\) at time \(t\). The independent variables include the stock liquidity (LIQUIDITY), firm size (MARKET CAP), EBIT/Sales (EBIT MARGIN), cash plus marketable securities to assets (CASH), P/E ratio (PE) and revenues/assets (TURNOVER). All independent variables are as at time \(t-1\).

Diagnostic tests showed no evidence of multicollinearity in regressions (1) and (2).

Tables 11 and 12 report, respectively, estimates of the coefficients of models (1) and (2) using two estimation techniques and annual data from 1 May 1985 through 30 April 2010. The first column reports estimates of the model using pooled ordinary least squares (OLS). The second column reports estimates of a “fixed effects” model.6

### A. Explaining returns

The coefficient estimates of model (1) for positive and negative P/E interlisted firms are reported respectively in Table 11, Panels A and B. Panels A and B show that while most of the variables are significant for the positive P/E firms, none of the fundamental variables are significant for the negative P/E firms, irrespective of the estimation procedure. This seems to indicate that the negative P/E firms move to their own tune and behave in a way different than that of the positive P/E firms.

The coefficient estimates of model (2) for positive and negative P/E non-interlisted firms are reported respectively in Table 12, Panels A and B. Panels A and B show that while liquidity is significant for the positive and negative P/E firms, the P/E ratio is significant only in the positive P/E firms. Typically, the P/E ratio is negatively related to forward returns, but here this is only the case for positive P/E firms. Moreover, while the P/E ratio is significant for the positive P/E non-interlisted firms, it is not significant at all for the positive P/E interlisted firms. Finally, EBIT margin and turnover have diagrammatically opposite signs for positive and negative P/E firms. The above findings seem to provide further evidence that negative P/E firms behave differently than positive P/E firms and that interlisted stocks behave differently than non-interlisted stocks.

### 4. Conclusions

Using separately interlisted and non-interlisted Canadian stock market data for the period 1985–2010, the main purpose of this paper was to examine whether negative P/E stocks were really different than positive P/E firms, and whether negative P/E stocks outperformed, on average, the universe of positive P/E stocks. The paper additionally purported to examine (a) whether interlisted and non-interlisted firms behaved similarly or there were distinct differences between them, and (b) whether there were differences in relation to this paper’s key questions only in one group of stocks or differences were equally driven by both.

---

5 We employed the VIF option in the REG SAS procedures to carry out diagnostic tests.

6 To estimate a fixed effects model, year and industry dummies were added to regressions (1) and (2).
We found that firms with negative multiples are indeed different from firms with positive multiples in that (a) a relatively small number of firms with negative multiples experience high forward stock returns even though the majority of them does not result in a large difference between mean and median returns, and (b) the value, size, liquidity and business risk premiums behave differently for negative and positive P/E firms. This indicates that prior academic research was right in excluding negative multiple firms from their analysis. Moreover, the paper also shows that there are key differences between interlisted and non-interlisted firms both in the positive and negative P/E space. As a result, not only must negative P/E firms be segregated from positive multiple firms, but also interlisted firms ought to be segregated from non-interlisted firms in related research as aggregation would undermine the clarity and generality of findings, affect the homogeneity of the sample and dilute findings and tests of significance.

References


Trust in banks after the financial crisis

Santiago Carbo-Valverde
Professor of Economics and Finance, Bangor Business School

Abstract
This article analyzes recent evidence and data suggesting that trust in banks has declined significantly after the financial crisis in several countries. Even if rebuilding trust is not an explicit public policy objective, it is worthwhile noting that trust represents a key ingredient of financial stability which, in turn, is an essential element of social and economic stability and development. The evidence shown in this article suggests that changes in the levels of trust are mostly in the hands of the banks themselves. In this sense, relying upon recent empirical evidence for Spain, we show that by changing specific attributes of the services that they provide, financial intermediaries can improve trust from their customers and even offset the negative effects of the financial crisis.
1. Introduction

Financial crises are particularly hazardous times for the reputation of banks. Lack of credit, compensation schemes and the distribution of risky financial assets among uninformed customers are among the main determinants of the lack of trust in banks during a financial crisis. In this context, even if social responsibility has become a major issue in corporate management, banks have been among those enterprises that have found it tough to improve their reputation. The financial crisis has obviously not helped the situation. As has been the case with previous financial crises and economic recessions, a number of regulation initiatives have been launched with the aim of making the banking sector safer, thereby protecting consumers and enhancing trust. However, while regulations can set a variety of protection and prevention tools, financial institutions will need to undergo significant transformations in order to enhance the reputation of the banking sector. In other words, regaining customers' trust is in their own hands. As reported by Financial Times in February 2013, just after being appointed Governor of the Bank of England Mark Carney said, "A lack of trust in banks is holding back the global economic recovery (...). Integrity cannot be legislated, and it certainly cannot be bought (...). It must come from within."\(^1\)

From a theoretical perspective, the level of trust in banks may be considered, at least to some extent, to be a function of the role of financial institutions within the so-called financial safety net. A nation's financial safety net can be seen as a government-owned enterprise that monitors and contains systemic risk. Systemic risk combines two kinds of risk-taking: calculated risk-taking by protected institutions (banks) and countervailing risk-management programs operated by safety net managers (regulators). During financial crises, safety net managers need to frequently apply exceptional measures, some of which are oriented toward rescuing banks. Such bailouts have also become a source of distrust in financial intermediaries, as bailouts frequently involve taxpayers' money. During the crisis, authorities exercised a loss-shifting “taxpayer put” that converted most of the losses incurred by insolvent banks into government debt; hence, using taxpayer money to bail out banks is, to some extent, a put option that banks can also benefit from.

Given these constraints, improving trust in financial institutions involves a transformation of the social perspectives of banks, which is particularly difficult to achieve in a post-crisis environment. It is also important to bear in mind that financial stability is a major national concern, and hence requires improving financial education.

Despite these external forces, it is the banks themselves that need to institute the changes needed to regain customers’ trust, as they are ultimately the ones providing the services and customers' trust is essential for their survival and growth. This is the focus of this article. We examine the status of trust in financial institutions during and after the financial crisis and analyze the possibilities that financial institutions have to rebuild and improve trust from their customers. This paper explores the different dimensions that define trust in banks and the ways that banks have been performing on this front. We also survey some recent empirical evidence, which suggests that by improving different aspects of their service, banks can potentially offset the negative effects of the financial crisis on their reputation.

The structure of the article is as follows. In section 2, we look at the evolution of trust in banks from an international perspective and discuss the main initiatives that are being instituted to reform the financial industry, which should, in turn, have an (hopefully positive) impact on trust. Section 3 looks at the empirical literature on how trust is built between the banks and their customers and the implications of these relationships on economic development. Some recent empirical results on the effects of the crisis on trust in financial institutions – taking Spain as a reference – are discussed in the section 4. The article ends with a summary of conclusions.

2. A public policy problem?

Trust is based on the customers' beliefs that financial institutions take actions and decisions on their behalf, with the expectation that these do not violate the “vote of confidence given” by the customers [Gill et al. (2006)]. If we follow this simple but illustrative definition, there are three main pillars that may determine the levels of trust in banks. The first and most important is the services that banks provide and the way they are provided. The second is the perception of the customer about the relationship it holds with a financial institution and on the banking sector in general. The third pillar is the role that authorities

---

1 http://www.ft.com/cms/s/0/a444e114-7f7b-11e2-8d96-00144feabdc0.html#axzz2wt4OhJt2
in a given jurisdiction may have in trying to ensure that these relationships are built in a sound manner and that financial stability is guaranteed.

The financial crisis that started in 2007 was the last episode when substantial changes in the trust in financial institutions were observed. In this section, we will refer to the second and third pillars of trust-enhancing mechanisms (customers’ perceptions and the role of financial authorities) and in the following sections we will focus on the role played by financial institutions, which is the main objective of this article.

As for customers’ perceptions, some surveys have been recently conducted upon which some trust indices have been built. Some of them compare the financial industry with other industries while a number of surveys concentrate specifically on the banking sector.

The surveys conducted annually by the public relations company, Edelman, is very illustrative on the effects of the crisis on trust in banks. The survey is conducted by research firm StrategyOne on behalf of Edelman, and involves 20-minute online interviews with 30,000 people in 25 countries. Edelman describes participants in the survey as members of the “informed public” — college graduates whose household income is in the top quartile for their age, and who follow newsworthy issues several times a week. An interesting output from the survey is the comparison of the level of trust in financial services in general — and in banks, in particular — with other industries. The latest results correspond to 2013 and reveal that financial services and banks are the least trusted sectors (Figure 1). Taking an index from 0% to 100%, trust in financial services was at 46% in 2013 and the figure for banks was 49%. The best considered sector in the study was Technology, at 70%. Other industries that were perceived more positively included media (50%), chemicals (51%), brewing and spirits (54%), consumer health companies (55%), pharmaceuticals (57%) and energy (57%).

The Edelman study also provides some information on the geographic dispersion of trust in banks. In particular, the survey asked the following question: “how much do you trust banks to do what is right?” The responses were scaled on a range 0%-100%. The results (Figure 2) show that in China (80%), Singapore (76%) and India (71%) customers have the highest levels of trust in financial institutions. Countries where financial institutions were viewed least positively were Russia (36%), Japan (41%), Spain (42%), Argentina (45%) and Ireland (46%). Other countries such as the U.K. (53%), France (54%), Germany (55%) and the U.S. (59%) fall within what Edelman describes as “neutral” levels of trust.
Additionally, a number of studies have been conducted in the U.K. and the U.S. on the topic of trust in banks. In the U.S., an interesting study was conducted by Gallup [Wood and Berg (2011)], which provides a long-term historical perspective on the evolution of trust in banks. The study allows us to identify the changes that can occur in trust in banks before, during and after a financial crisis. The study covers the period 1979–2010 and shows that by October 2010 Americans’ trust in banks fell to an all-time low of 18%. Gallup considers this score “to be a continuation of a free-fall that began in 2006.” The main question in the Gallup survey was “Please tell me how much confidence you, yourself, have in banks - a great deal, quite a lot, some or very little?” Figure 3 shows the percentage of respondent that responded by saying “a great deal” or “quite a lot” out of the total. In 1970, this ratio was 60% in the U.S. and it fell to 30% in 1991, after the crisis of the late 1980s. From there, it experienced a sort of recovery to reach 49% in 2005, until it fell to 22% with the crisis, and reached an all-time low of 18% in 2010.

Another recent and systematic survey conducted in the U.S. with the specific purpose of tracking the evolution of trust in banks is the “Trust in the Financial System Indicator” of the Chicago Booth/Kellogg School of Management. The school commissions Social Science Research Solutions to conduct a telephonic survey on a representative sample of 1,034 American households. One adult respondent in each household is randomly contacted and asked whether they were in charge of household finances, either alone or together with the spouse. Only individuals who claimed such responsibility are included in the survey. The first set of questions asked how much people trust certain types of people or institutions, on a scale ranging from 1 to 5, where 1 means “I do not trust at all” and 5 means “I trust completely.”

Figure 4 reports the results for this survey in 2013. The trust variables are rescaled to a range of 0%-100%. The results compare trust in banks with trust in the stock market, mutual funds, and large corporations. Banks appear to have gained some of their lost confidence in the U.S. as their level of trust in 2013 was at 35%, even if this is still a percentage that reveals a relatively large level of distrust. The index for banks is at least larger than for mutual funds (31%) and large corporations (17%) with the stock market registering the lowest mark (15%).

In the U.K., an interesting initiative is the “YouGov Cambridge Public Opinion Survey on Public Trust in Banking” conducted by the University of Cambridge and The Guardian in April 2013.
The research was undertaken in two stages. The first survey included “a number of broad questions about the performance of the banks.” After giving people information and inviting them to think more deeply about banking issues, they repeated some of the broad questions in a second survey, as a robustness check. Overall, they find some differences “but not enough to convert a majority of earthlings from critics to fans.” In the first survey, 57% of the U.K. respondents said the overall reputation of the banking sector was bad. In the second survey, the figure was 46%. The proportion saying “favorable” was up three points, from 12% in the first survey to 15% in the second survey. Figure 5 shows some of the results of this two-stage survey on professional and honesty attitudes, contribution to solve the crisis and compensation schemes in the banking sector, which are among the least valued characteristics of banks in the country.

At first glance, the surveys illustrate a deterioration in the reputation of banks during the financial crisis. This would require a response from what we have defined as the third pillar of the trust in banks, namely the role of public authorities. In fact, literally all financial crises in recent years have been followed by a wave of regulatory initiatives aimed at correcting, and preventing similar episodes. Consumer protection issues are among those that are frequently cited and they result in limitations, or restrictions, on a number of activities carried out by financial intermediaries. Between 2008 and 2010, there were numerous international meetings, in particular among the G-20 countries, where the “reform of the global financial system architecture” was a critical feature. The result has been a wave of regulatory reforms, the three most prominent of which are the Volcker’s rule for U.S. banks, the Vickers report for U.K. banks, and the Liikanen-E.U. bank report. Even though there are substantial differences between these new regulations, their main thrust has been to eliminate what are perceived to be the major deficiencies of the financial systems and mitigate systemic risk, which, in turn, are essential requirements to foster trust. Having said that, the legal considerations are beyond the scope of this paper.

Although these initiatives have been mainly oriented toward limiting or preventing bank activities, an important public policy objective is to rebuild trust in the banking system. However, regulations can only have an impact on trust indirectly, which means that financial institutions are solely responsible for regaining their own customers’ trust. In the next two sections, we will discuss what banks can do to rebuild their customers’ trust after the financial crisis.

3. Some theory and evidence from previous studies on trust

Most of the literature on this subject has focused either on the determinants of general trust [Alesina and La Ferrara (2002)] or trust in political institutions [Mishler and Rose (2001)]. The number of studies dealing with trust in financial institutions is much limited. As shown by Powell (1990) “trust should be considered as “a remarkably efficient lubricant to economics exchange [that] reduces complex realities far more quickly and economically than prediction, authority or bargaining”. In the case of the financial system, from a general perspective, Guiso et al. (2004, 2008) describe the (statistically significant and negative) long-run consequences of a fall in trust in financial markets and financial intermediation on economic output. They also show that, in the short term, another threat is the possibility of financial panics and bank runs. Along similar lines, Sapienza and Zingales (2012) connect the relevance of

2 A good summary of these regulatory initiatives can be found in an IMF report by ViVils et al. (2013).
trust to the developments surrounding the financial crisis and argue that “something important was destroyed in the last few months. It is an asset crucial to production, even if it is not made of bricks and mortar. While this asset does not enter standard national account statistics or standard economic models, it is so crucial to development that its absence - according to Nobel laureate Kenneth Arrow - is the cause of much of the economic backwardness in the world. This asset is trust (...). While trust is fundamental to all trade and investment, it is particularly important in financial markets, where people depart with their money in exchange for promises. Promises that aren’t worth the paper they’re written on if there is no trust”. Indeed, some studies have demonstrated that trust is positively correlated with access to credit, the use of other financial services and on investment decisions by households [Porta et al. (1997), Calderon et al. (2002), Guiso et al. (2004, 2008)].

An important caveat when studying trust in financial institutions is that, at least to some extent, trust in banks is dependent on the general levels of trust of the customers as citizens. Political and institutional trust is, therefore, one of the main conditions that may affect specific trust in financial institutions. In this sense, there are a number of studies that investigate the interaction of general and institutional trust with specific issues related to trust, such as regulatory actions [Carlin et al. (2009), Aghion et al. (2009)]. It is important to note that there are at least two alternative, but not necessarily mutually exclusive, explanations for the determination of institutional trust. On the one hand, the so-called cultural theories argue that trust is exogenous [Inglehart (1997)] and based on dyadic trust. As such, it is frequently viewed as being learned early in life. Institutional theories, on the other hand, argue that it is endogenous [Hetherington (1998), North (1990)] and influenced by institutional performance. We will base our study on the latter assumption when discussing what banks can do to modify customer trust in them.

Following the institutional approach to trust, and therefore, assuming a certain degree of endogeneity, there are at least three recent studies that specifically analyze the case of banks. Knell and Stix (2009) study whether trust in Austrian banks declined during the global financial crisis and assess what factors determine the level of trust in banks. They provide evidence that the degree of individual information does not affect trust and that trust in banking is a trend and is contagious. They also find that the extension of deposit insurance coverage in October 2008 had a positive effect on trust. They also show that trust in banks is mainly affected by “subjective” variables like the individuals’ assessment of the current economic and financial situation and their future outlook. A second related study, by Mosch and Prast (2008), provide evidence on trust in banks in the Dutch financial sector. Based on surveys over the period 2003-06, they find a statistically significant and positive relationship between confidence in the economy and trust in banks, in line with the results provided by Knell and Stix (2009). A third study, by Gill et al. (2006), analyze the impact of some personal level features related to the provision of banking services such as empathy, kindness and similarity or affinity on trust in banks. Their results suggest that these attributes of bank services generally favor confidence and trust to a significant extent.

The results by Gill et al. (2006) can also be related to another strand of the banking literature, which does not deal specifically with trust in banks but provides a number of relevant microeconomic insights on how trust relationships are built. The so-called “relationship banking” literature defines a banking relationship as the association between a bank and its customers that goes beyond the execution of simple and anonymous financial transactions [Ongena and Smith (1998)]. The ability of banks to maintain such relationship depends on a number of variables, such as the quality and price of the services provided by the bank and the competitive environment in which the bank operates. Over the course of a long-term relationship with their customers, banks accumulate significant and valuable information about their needs, allowing them to tailor their services to suit customers’ financial requirements, and as a consequence enhance trust. Following this principle, several studies have over the last 20 years or so explored a variety of dimensions of these relationships. In a seminal study, Petersen and Rajan (1994) elaborated on the informational properties of these bank-customer relationships and suggest that they can reduce the costs of borrowing. Similarly, Berger and Udell (1995) show that lenders offer firms with longer relationships lower loan rates and are less likely to require collateral for such loans.

A number of other studies have since been published that looked specifically at other dimensions of bank relationships. One of these dimensions is the “length of relationship,” which assumes that the longer the relationship between a borrower and a lender
the greater the information flow between the two parties and the more important the relationship becomes (Boot (2000), Degryse and van Cayseele (2000), Cole et al. (2004)). Another issue is whether the relationships of the customer with their banks are “single or multiple,” suggesting that the existence of multiple relationships reduces the value of information (Thakor (1996), Detragiache et al. (2000), Berger and Udell (1998)). Finally, an additional dimension that defines the nature of relationship banking is “geographic distance,” assuming that the collection of information and building a relationship with a customer is easier, the closer the distance between the bank and that customer.

On the whole, relationship banking has provided evidence on the significant effects that information production by banks and the repeated contact with customers may have on economic activity. However, as we have shown, studies that deal specifically with the more general issue of the level of trust in banks are much rarer, mainly due to the limited availability of data.

4. Recent empirical results: the case of Spain
An exception to the paucity of empirical studies on how financial institutions can act to modify the level of trust their customers have in them is the one recently conducted for Spain by Carbo-Valverde et al. (2013). In this section, we will discuss some of the findings of the study that we consider to be particularly relevant for our purposes. We find this study to be quite useful for three reasons. Firstly, it specifically analyzes the effects of the financial crisis on trust in banks. Secondly, it measures the extent to which banks can potentially offset the effects of the crisis on the trust levels by changing certain attributes of their services. Finally, it refers to Spain, a country that experienced a severe banking crisis recently and where banks have been shown to be experiencing very low levels of trust among their customers (Figure 2).

The dataset in Carbo-Valverde et al. (2013) corresponds to the survey conducted by the Instituto de Encuestas y Opinión Pública (IMOP) on trust in banks in Spain in 2009. The survey is based on a broad sample size of 1,601 bank customers covering the entire Spanish territory. The survey was carried out by computer-aided telephonic interviews using a computer-assisted telephonic survey technique. The field work was carried out in January 2009 by a team of IMOP telephone field network interviewers. The survey sample is nationally representative, and it contains rich information about bank customers’ perceptions of several bank attributes, which is not usually available in other data sources.

The specific level of trust in banks is measured by the question: “I trust the solvency of commercial banks/savings banks in general and of my commercial bank/savings bank in particular.” The respondents could respond by choosing one out of six levels of specific trust in banks: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree and do not know. As in other surveys, this method allows us to scale the main dependent variable. In the estimations presented in Carbo-Valverde et al. (2013), trust in banks is measured conditional on different sociodemographic characteristics of the customers as well as on the general levels of trust/distrust of the respondents not necessarily associated with financial services. Importantly, the study adds a number of attributes of bank services in the right-hand side of the equation that permits the authors to undertake simulations on how banks can act to change trust in the eyes of their customers. These attributes are banks’ sensitivity to customers’ problems, effectiveness in the provision of the services, quality of the services provided, involvement in social activities, comfort in the banking relationship at the bank branch and commitment to the needs of the customer.

In the baseline tests, the coefficients of bank customer sociodemographic characteristics are in line with previous studies. In particular, customer income appears to be the most important determinant among socioeconomic characteristics, while other characteristics, such as marital status or age, are only of limited statistical significance in a reduced number of cases. As for the attributes of bank services, most of them have a positive and significant effect on trust in banks. Sensitivity to customers’ problems seems to be the variable with a larger economic impact. In particular, for one unit change in this variable the trust variable is found to increase by 0.81. Similarly, for one unit change in the variable effectiveness in the provision of the services, the trust variable would increase by 0.45. Similar results are obtained for other bank attribute variables that determine customer’ perceptions, such as the social involvement of the banks (with an estimated marginal effect on trust of 0.62) and comfort at the bank branch (with a marginal impact on trust of 0.75).

Along with the baseline results, an interesting experiment is undertaken on the subsample of survey respondents that
considered that the behavior of Spanish financial institutions were changing for the worse. We calculate the marginal effects of each attribute of bank service on the level of trust in banks for these customers. Focusing on the most relevant results (Figure 6), for a unit change in the variable measuring bank sensitivity to customers’ problems, trust in banks would increase by about 34.7% (percentage change in the trust index). Similarly, for a unit increase in the variable measuring commitment with the needs of the customer, the variable measuring trust in banks would increase by 34%. The general observation in the different tests conducted by Carbo-Valverde et al. (2013) is that by modifying these attributes banks could offset the negative effect of the crisis on the levels of trust by their customers.

The evidence shown in this paper suggests that even if the three pillars are essential to improve trust in banks, changes in the levels of trust are mostly in the hands of the banks themselves. In this sense, relying upon recent empirical result for Spain, where the banking crisis has been among the most severe and long-lasting and where trust in financial institutions is at the rock bottom, we show that by changing specific attributes of the services that they provide, financial intermediaries can improve the trust that their customers have in them, and even offset the negative effects of the financial crisis on trust in banks.

**Conclusions**

This article analyzes recent evidence and data suggesting that trust in banks have declined significantly after the financial crisis in several countries. Although rebuilding trust is not an explicit public policy objective, it is worthwhile noting that trust represents a key ingredient of financial stability which, in turn, is an essential element of social and economic development. There are three pillars involved in the evolution of trust which interact with each other: changes in customer’s perception about financial services provided by banks, regulatory actions against the financial services sector and the transformations in the way banks provide their services to and establish relationships with their customers.
References
YouGov Cambridge public opinion survey on public trust in banking, 2013, Public trust in banking, April symposium, YouGov-POLIS programme for public opinion research, University of Cambridge and The Guardian
About EY

EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization, and may refer to one or more, of the 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 ey.com.

© 2014 EYGM Limited.
All Rights Reserved.

EYG No. CQ0146

In line with EY’s commitment to minimize its impact on the environment, this document has been printed on paper with a high recycled content.

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, or other professional advice. Please refer to your advisors for specific advice.

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 EY do not necessarily represent the views or opinions of EYGM Limited nor any other member of the global EY organization. The articles produced by EY 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 EY organization accepts responsibility for loss arising from any action taken or not taken by those receiving The Journal.

The views of third parties set out in this publication are not necessarily the views of the global EY organization or its member firms. Moreover, they should be seen in the context of the time they were made.

Accredited by the American Economic Association

ISSN 2049-8640