Article: On being the right size

The Journal of Financial Perspectives

EY Global Financial Services Institute

March 2014 | Volume 2 - Issue 1



Building a better working world

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, 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 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*.

Editorial

Editor Shahin Shojai EY LLP

Advisory Editors

Dai Bedford EY LLP Shaun Crawford EY LLP Carmine DiSibio EY LLP

Special Advisory Editors

David Cole Swiss Re Paul Feeney Old Mutual Wealth Paolo Fiorentino UniCredit

Editorial Board

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

Ratan Engineer EY LLP David Gittleson EY LLP Bill Schlich EY LLP

Mark Hartigan Zurich Robert W. Jenkins London Business School Thomas C. Wilson Allianz

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

Executive summary

On being the right size

by Andrew G Haldane, Executive Director, Financial Stability and Member of the Financial Policy Committee, Bank of England

Commentary: Is big dangerous?

by Philip Booth, Editorial and Program Director, Institute of Economic Affairs and Professor of Insurance and Risk Management, Cass Business School, City University This paper discusses the too-big-to-fail problem from the perspective of individual banks and the financial system as a whole. It starts by exploring the potential implications of recent financial deepening and concentration, which has generated escalating expectations of state support, thereby encouraging further expansion and concentration. The paper then explores three policy approaches to tackling the too-big-to-fail problem. The first is the imposition of systemic surcharges of additional capital, which have the effect of reducing expected system-wide-losses in systemically important banks, but not materially so at current levels of the surcharge. Second, new resolution regimes are being put in place to allow banks to fail safely – though the market still has doubts about their credibility for the biggest banks. Finally, structural reform of banks is taking place, through proposals by Volcker, Vickers and Liikanen. Despite this policy progress, expectations of state support remain high. This paper proposes potential additional reforms to tackle too-big-to-fail, such as placing limits on bank size and market share and increasing competition.

On being the right size

Andrew G Haldane

Executive Director, Financial Stability and Member of the Financial Policy Committee, Bank of England

Commentary: Is big dangerous?

Philip Booth

Editorial and Programme Director, Institute of Economic Affairs and Professor of Insurance and Risk Management, Cass Business School, City University

Abstract

This paper discusses the too-big-to-fail problem from the perspective of individual banks and the financial system as a whole. It starts by exploring the potential implications of recent financial deepening and concentration, which has generated escalating expectations of state support, thereby encouraging further expansion and concentration. The paper then explores three policy approaches to tackling the too-big-to-fail problem. The first is the imposition of systemic surcharges of additional capital, which have the effect of reducing expected system-wide losses in systemically important banks, but not materially so at current levels of the surcharge. Second, new resolution regimes are being put in place to allow banks to fail safely – though the market still has doubts about their credibility for the biggest banks. Finally, structural reform of banks is taking place, through proposals by Volcker, Vickers and Liikanen. Despite this policy progress, expectations of state support remain high. This paper proposes potential additional reforms to tackle too-big-to-fail, such as placing limits on bank size and market share and increasing competition. While existing initiatives are a step in the right direction, there may be some distance to travel before banking is the right size.

¹ This article is based on a speech delivered at Institute of Economic Affairs' 22nd Annual Series, The 2012 Beesley Lectures, at the Institute of Directors, Pall Mall, 25 October 2012. The views are not necessarily those of the Bank of England or the Financial Policy Committee. I would like to thank Andrew Gracie, Varun Paul, Kirsty Rodwell, Timothy Richards, Vicky Saporta, Rhiannon Sowerbutts and Belinda Tracey for their comments and contributions.

1. Introduction

In 1928, evolutionary biologist J. B. S. Haldane wrote an important article whose title I have borrowed, *On Being the Right Size*². The essential point was simple. The sheer size of an object, institution or animal determined their structure. In particular, as their size rose, their structure needed to strengthen more than proportionately if they were to remain robust and resilient. This principle is sometimes enshrined in the so-called "square-cubed" law. A proportional rise in an object's size causes its area to rise by the square, and its volume by the cube, of that rise. At one level, this is simple mathematical geometry. Yet in the real world, it carries fundamental implications for evolutionary structure.

Take the animal kingdom. The square-cubed law explains why a flea, even if it were the size of a man, would not be capable of jumping to the moon. It explains why a hippopotamus cannot turn somersaults. And it explains why King Kong and Godzilla were physiological impossibilities – the weight transfer associated with a single step would have shattered their thigh bones.

When the world's biggest banking beasts took a step too far in 2008, they too folded under their own weight. Their physiological structure proved inadequate to make them robust and resilient. That is the essence of the "too-big-to-fail" problem. In the language of Haldane, international policymakers have concluded that many of the world's largest banks are not the right size given their existing physiological make-up.

Over the past few years, initiatives to solve the too-big-to-fail problem have come thick and fast. At root, each has aimed to strengthen the structure of the world's biggest banks. That is the good news. Claims that they have solved the too-bigto-fail problem appear to me, however, premature, probably over-optimistic. Worse, they risk sending a false sense of crisis comfort. That is the bad news.

To see why such a cautious conclusion is warranted, we begin by tracking the structural evolution of the financial system over the past few decades. We then consider the three most prominent policy initiatives aimed at tackling too-big-to-fail – systemic



Figure 1: Bank assets/GDP in selected countries Source: Jorda et al. (2011)

surcharges, resolution regimes and structural reform. In the final section, we consider what supplementary policy options might be necessary to ensure banking is right-sized.

2. Evolution of the financial system

The past 50 years have seen seismic shifts in the structure, size and composition of the global financial system. These changes gave birth to the too-big-to-fail problem. Figure 1 plots the ratio of banking sector assets-to-GDP, and its cross-country dispersion, for a set of 14 advanced countries over the past 140 years.³

For the better part of a century, between 1870 and 1970, financial deepening in these countries followed a modestly upward trend. Over this period, the average bank assets-to-GDP ratio rose from 16% to over 70%, or less than six percentage points per decade. Since 1970, this trend has changed trajectory. The ratio of bank assets-to-GDP has more than doubled over the past 40 years, rising from around 70% to over 200%, or over 30 percentage points per decade. In other words, since 1970 financial deepening has occurred five times faster than in the preceding century. For some individual countries, the rise has been more dramatic still – in the U.K., the ratio has risen five-fold.

² Haldane (1928). Haldane is, regrettably, no relative.

³ The data is from Jorda et al. (2011).

In cross-country studies, financial deepening of this type has generally been found to have a positive effect on medium-term growth [Beck and Levine (2004)]. Taken literally, this would suggest that the rise in banking scale over recent decades has provided a significant tailwind to medium-term growth in advanced countries. So it seemed in the pre-crisis period.

But that conventional wisdom has recently been called into question. IMF research has suggested that financial deepening can indeed be growth-positive – but within limits. There is a threshold at which private credit-to-GDP may begin to have a negative impact on GDP growth [Arcand et al. (2012)]. That threshold is found to lie at a private credit-to-GDP ratio of around 80%-100%. This finding is consistent with earlier cross-country evidence suggesting that, at credit-to-GDP ratios above unity, output volatility tends to increase [Easterly et al. (2000)]⁴. This threshold lies significantly below current levels of financial depth in most advanced economies. In other words, taken at face value this evidence suggests that, at its current scale, banking could be acting as a headwind to medium-term growth.

Accompanying this dramatic rise in banking scale has been an equally dramatic rise in banking concentration. Figure 2 plots the evolution of the asset share of the three largest banks in the U.S., U.K., Germany and Switzerland over recent decades. The U.S. has undergone the most dramatic upward shift, with the share of the top three banks rising from around 10% to 40% between 1990 and 2007. For other countries, there is a less dramatic rise in concentration, but from a much higher starting point, with the top three banks accounting for between two thirds and threequarters of assets in the U.K., Switzerland and Germany.

This dramatic rise in banking scale and concentration has been driven by financial liberalization. But it has also spawned an acute problem for society due to escalating expectations of state support for the banking system. These expectations generate lower funding costs, in particular for the largest banks, which in turn encourages further expansion and concentration, worsening the too-big-to-fail dilemma. There is a self perpetuating "doom loop" [Alessandri and Haldane (2010)].



Figure 2: Banking concentration in selected countries

Source: Capie and Webber (1985), FDIC, Bundesbank, Swiss National Bank, Bank of England calculations

The size of the resulting "implicit subsidy" from the state to the big banks has been the subject of recent study – and no little controversy [Noss and Sowerbutts (2012)]. There are a number of possible methods for estimating the subsidy. Perhaps the simplest is found by comparing the "standalone" and "support" ratings assigned to debt issued by large banks. The difference between these ratings gives an estimate, used by the market when pricing debt, of the probability of state support. Figure 3 plots the difference between these ratings for the 29 institutions deemed by the Financial Stability Board (FSB) last year to be the world's most systemically-important. In the pre-crisis period, this difference averaged 1.3 notches. This sounds material, if not breathtaking. It suggests that, while non-trivial, too-big-to-fail may not have been a first-order driver of the rising scale and concentration in banking.

Yet even small notches of support can translate into big implicit subsidies if balance sheets are large. So it is for the world's largest banks. Over the period 2002 to 2007, the implied annual subsidy to the world's largest banks averaged U.S.\$70b per year using a ratings-based measure (Figure 4). That is roughly 50% of the average post-tax profits of these banks over the period.

As the crisis struck, this implicit promise became explicit. Financial support was extended to the banking system in the

⁴ This has strikingly close parallels with recent econometric evidence suggesting that growth is harmed at debt-to-GDP ratios above unity, for both the public sector [Reinhart and Rogoff (2010)] and private sectors [Cechetti et al. (2011)].



Figure 3: Ratings uplift for systemic institutions Source: Moody's, Bank of England calculations

form of capital injections, guarantees and liquidity insurance. On some estimates, this support rose to three-quarters of annual GDP in some countries. In response to these interventions, there has predictably been a further ratcheting-up in ratings-implied degrees of state support to banks.

By 2009, the ratings difference had more than doubled to above three notches, with the implied monetary subsidy over U.S.\$700b per year. That was well in excess of average annual pre-crisis profits of these firms. Even if an over-estimate, the scale of the implied subsidy signaled something dramatic was at play. Too-bigto-fail had become hard-wired into the structure and pricing of the financial system.

3. Systemic surcharges

One way of interpreting these implicit subsidies is as the market's best guess of how much a policymaker would be willing to pay each year to avoid the failure of the world's biggest banks. They proxy the expected social costs of big bank failure. In the jargon, they capture a systemic externality.

This notion of a systemic externality has underpinned recent academic and policy efforts to solve the too-big-to-fail problem. Brunnermeier et al. (2009) use this framework to motivate levying a tax – a "Pigouvian tax" – on institutions posing systemic



Figure 4: Implicit subsidy for systemic institutions Source: S&P Capital IQ, SNL Financial, Bank of America, Merrill Lynch⁵, Moody's, Bank of England calculations

risk externalities. This tax would be set at levels which offset the effects of the bank's actions on wider society. A number of academics have since proposed measures along broadly Pigouvian lines [Archarya et al (2010)].

Rather remarkably, policy reforms in practice have followed closely in the spirit of these proposals. In 2010, the FSB announced its intention to introduce a "systemic surcharge" of additional capital on the world's largest banks. In July 2011, the Basel Committee published a methodology for measuring global systemic importance based on indicators of bank size, connectivity and complexity, with additional capital of up to 2.5% depending on the scores for these indicators. This methodology was fully finalized in 2013.

In November 2011, the FSB endorsed this methodology and announced 29 global systemically important banks (G-SIBs). It has updated the list of G-SIBs on an annual basis since then. The

⁵ The BofA Merrill Lynch indices are provided for informational purposes only and may not be used without the written permission of BofA Merrill Lynch. The BofA indices are not intended to be used for investment or other commercial purposes. BofA Merrill Lynch is licensing the BofA Merrill Lynch indices "as is," makes no warranties regarding same, does not guarantee the suitability, quality, accuracy, timeliness and/or completeness of the BofA Merrill Lynch indices or any other data included in, related to, or derived therefrom, assumes no liability in connection with their use, and does not sponsor, endorse, or recommend (BOE), or any of its products or services.



Figure 5: Expected system-wide loss Source: Bank of England calculations

surcharges will be phased in from 2016. Legislation is already in place in the E.U. to implement the systemic surcharge.

These proposals are clearly a practical step in the right direction. By boosting levels of capital in the system, the probability of big bank failure will be reduced. You would have got good odds back in 2007 that something as seemingly elliptical as a Pigouvian tax on systemic risk would have found its way onto the regulatory statute books. Now we have it. "We are all Pigouvians now," even if most of us cannot spell it.

The practical question is how far systemic surcharges take us in tackling the systemic risk externality. In other words, at current levels, by how much have systemic surcharges reduced expected losses for the financial system? That is an empirical question. To assess it, consider the impact of capital surcharges on the expected losses facing the 29 institutions identified by the FSB.

Using each of these banks' balance sheets, we generate a measure of default probability using the Merton (1974) contingent claims model. We assume that the base level of equity for these banks is 7% (the new Basel III minimum) and that, in the event of default, they suffer losses on their assets of around 30%⁶. To keep things simple, banks' assets are assumed to be (log-)normally distributed, in line with Merton (1974), and default occurs only when a bank's capital is fully exhausted. Assume initially that default risks across banks are independent. This is a highly conservative assumption, as in practice bank default probabilities are highly correlated in times of stress. Indeed, bank correlation coefficients in crisis often head toward one. For that reason, this thought-experiment provides a lower bound on expected losses across the financial system.

Figure 5 shows expected losses across the 29 banks at different levels of the systemic surcharge. In the absence of any surcharge, expected losses across the system are just less than U.S.\$200b per year. Were every large bank instead to face the maximum capital surcharge of 2.5%, then expected system-wide losses would fall by around 60% relative to their base level. And to remove 90% of the systemic externality – expected losses of around U.S.\$5b – a surcharge of over 7% would be needed.

A more plausible experiment would be to assume a non-zero correlation among bank defaults. For example, the failure of a large bank, which caused it to fire-sale assets could impose externalities on other large banks holding these same assets [Wagner (2009)]. To place an upper bound on expected losses in the face of these fire-sale externalities, assume instead a correlation coefficient of one. Figure 5 illustrates the impact on expected system-wide losses of a high degree of default correlation among the big banks. The expected system-wide loss increases to around U.S.\$750 billion per year – similar in size to the implicit subsidy at its peak. A 2.5% surcharge now only reduces expected system-wide losses to around \$350 billion per year. To lower expected system-wide losses to be around U.S.\$5 billion per year would require a surcharge of around 15% – six times its current upper limit.

If anything, these thought-experiments probably produce conservative estimates of system-wide losses and the necessary systemic surcharge. For example, bank asset returns are in practice much fatter-tailed than the log-normal distribution. And in practice, banks are likely to default well before their capital is fully exhausted. Relaxing either assumption would push up

⁶ James (1991). We also assume bank asset volatility of 4% per year, in line with previous empirical studies.

estimates of expected losses and the surcharge necessary to curtail these losses [Schanz et al. (2012)]. Nonetheless, if expected system-wide losses are a reasonable proxy for the system-wide externalities large banks pose, this analysis delivers a rather gloomy prognosis. At current levels of the surcharge, a large chunk of the systemic externality would remain untouched. If too-big-to-fail is the problem, then systemic surcharges seem to offer only a partial solution.

4. Resolution regimes

Capital surcharges lower systemic externalities by lowering default probabilities for the world's largest banks. An alternative way of lowering those externalities would be to reduce the collateral damage associated with their failure. This has been a key motivation for a second strand of the reform debate – the design of effective resolution regimes.

Few examples better illustrate the costs of getting this wrong than the spiraling queues outside branches of Northern Rock in September 2007. Despite being a medium-sized retail bank, Northern Rock's failure caused systemic disruption and put taxpayers' money at risk. In response, the U.K. put in place in 2009 a special resolution regime for banks, providing the Bank of England with tools for winding down a failed bank.

As the crisis illustrated, financial failure which causes systemic disruption is not confined to banks. And to avoid taxpayer bailout, losses may need to be imposed on a wide class of bank creditors, including holders of debt as well as equity – for example, by "bail-in." Over recent years, these resolution lessons have been enshrined in banking legislation. For example, in the U.S. Title II of the Dodd-Frank Act was passed in July 2010. It creates a new regime for the liquidation of financial companies, banks and non-banks, which pose a systemic financial stability risk. It enables losses to be imposed on creditors in resolution, while also prohibiting state bailouts.

Internationally, in November 2011, the G20 endorsed the FSB's Key Attributes of Effective Resolution Regimes, developed by an international working group chaired by my colleague Paul Tucker. Efforts are now underway to align national resolution regimes with these principles. As part of that, in Europe, a draft Directive on bank recovery and resolution was published in June 2012.

These initiatives are an important practical step in the right direction, lowering the societal costs of bank failure. As with systemic surcharges, it is striking how much progress has been made in so short a space of time on so complex an issue. The practical question is how far this takes us toward removing the too-big-to-fail externality.

During a bank resolution, one way of ensuring continuity of banking services is by transferring assets and/or liabilities of a failing firm to a third party. But the only entity with sufficient financial and managerial resource to absorb a large asset or liability portfolio, without suffering chronic indigestion, is another big bank. So it was during the crisis – for example, Bear Stearns was swallowed by JP Morgan Chase and Merrill Lynch by Bank of America.

This makes for a rather uncomfortable evolutionary trajectory, with rising levels of banking concentration and ever-larger toobig-to-fail banks. Levels of banking concentration have risen in many countries since 2007, precisely because of such shot-gun marriages by over-sized partners. In other words, resolving big banks may have helped yesterday's too-big-to-fail problem, but at the expense of worsening tomorrow's.

One way of avoiding this problem is to re-capitalize a bank by bailing-in its creditors, rather than transferring its assets. But resolution rules of this type are not problem-free either. Like all policy rules, they face what economists call a time-consistency problem. Whether a rule is followed in practice depends on the balance of costs and benefits at the time crisis strikes, not at the time the rule is written. That is why policy might in practice lack consistency over time – hence time-inconsistency.

Consider that trade-off when a big, complex bank hits the rocks. On the one side is a simple, but certain, option – state bail-out. On the other is a complex, and less certain, option – resolution. Policymakers face a trade-off between placing losses on a narrow set of taxpayers today (bail-in) or spreading that risk across a wider set of taxpayers today and in the future (bailout). If governments are risk-averse and wish to smooth the pain across taxpayers and across time, then bailout may look attractive on the day.

Financial history certainly suggests so. The history of big bank failure is a history of the state blinking before private creditors



Figure 6: Bank ratings uplift in the U.S. Source: Moody's, Bank of England calculations

[Haldane (2011)]. Recent crisis experience has written another chapter in this history. Next time may be different. For example, the public backlash against future bailouts could reinforce government's resolve to impose losses on creditors. And recent U.S. legislation in principle locks the taxpayer cashbox and throws away the key. Looking forward, the issue is whether this ex-ante rule is ex post credible.

As Figure 6 illustrates, implied levels of support for the U.S.'s biggest banks are much higher than before the crisis. More telling still, the passage of Dodd-Frank appears to have had little impact on levels of implied state support. It is early days for this new resolution regime and credibility may take time to be earned. Nonetheless, at present, the market believes the time-consistency problem for big banks is as acute as ever.

Even if it might appear the expedient option on the day of crisis, it is questionable whether bailout is the optimal response over the medium term. Figure 7 looks at the response of bank and sovereign CDS spreads around the time of bank bailouts in a selection of crisis countries. While bailouts lowered bank CDS spreads, as might be expected, bailout came at the expense of a rise in sovereign CDS spreads.

It is not difficult to see why. The financial crisis has caused huge damage to the balance sheets of governments in advanced



Figure 7: Sovereign and bank CDS during bailout announcements⁷ Source: Markit Group Ltd, BoE calculations

economies. For the G20 countries, the IMF forecast that the debtto-GDP ratio will rise by almost 40 percentage points between 2007 and 2016, to almost 120%. At these levels, public sector debt may be a significant drag on medium-term growth [Rogoff and Reinhart (2010)].

For economies with large banking systems and without a credible resolution regime, this leaves policymakers caught between a rock and hard place. When the call comes to ride to the banking rescue, government may be unable to afford not to. But nor, at least over the medium term, can they afford to. This is just the dilemma facing advanced countries today.

Disclaimer of liability notice. This may contain information obtained from third parties. including ratings from credit ratings agencies such as Standard & Poor's. Reproduction and distribution of third party content in any form is prohibited except with the prior written permission of the related third party. Third-party content providers do not guarantee the accuracy, completeness, timeliness or availability of any information, including ratings, and are not responsible for any errors or omissions (negligent or otherwise), regardless of the cause, or for the results obtained from the use of such content. Third party content providers give no express or implied warranties, including, but not limited to, any warranties of merchantability or fitness for a particular purpose or use. Third party content providers shall not be liable for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including lost income or profits and opportunity costs or losses caused by negligence) in connection with any use of their content, including ratings. Credit ratings are statements of opinions and are not statements of fact or recommendations to purchase. hold or sell securities. They do not address the suitability of securities or the suitability of securities for investment purposes, and should not be relied on as investment advice.

5. Structural reform

One way of lessening that dilemma, and at the same making resolution and bail-in more credible, is to act on the scale and structure of banking directly. Perhaps not surprisingly, recent regulatory reforms have sought to do just that. They have taken seriously the maxim that, if a bank is too big to fail, then it is too big. The result has been detailed proposals for structural reform in a number of advanced countries.

In the U.S., the "Volcker rule" has been introduced. This prohibits U.S. operating banks from undertaking proprietary trading and restricts private equity activity. The rule was tabled in October 2011 and became law in July 2012. Banks have two years to comply. In the U.K., the proposals of the "Vickers Commission" include placing a ring-fence around retail banking activities, supported by higher levels of capital. The final version of these proposals was tabled in September 2011, and legislation to enact them is planned by 2015. Banks will have until 2019 to comply. Most recently in Europe, the "Liikanen plan" was announced in October 2012. It proposes that the investment banking activities of universal banks be placed in a separate entity from the remainder of the banking group. There are at present no plans to legislate these proposals.

As with the other reform strands, it is remarkable how quickly radical structural reform proposals are finding their way onto the statute book. And, although different in detail, these proposals share a common motivation: separation of certain investment and commercial banking activities. In theory, such a separation delivers financial stability benefits of two distinct types [Boot and Ratnovski (2012)].

First, separation reduces the risk of cross-contamination. Riskier investment banking activities, when they go wrong, can pollute and dilute the financial resources of the retail bank. This potentially inflicts losses (or fear of losses) on depositors. It may also constrain banks' ability to make loans to the real economy when it might most need them. This is a crisis-time benefit of separation.

Second, separation can secure an improved pre-crisis allocation of financial resources from a societal perspective. High private return investment banking activities may crowd-out the human and financial resources devoted to high social return commercial banking activities. Investment banking activities might also piggy-back on the cheaper cost of deposit funding. In effect, universal banking allows privately optimal, but socially sub-optimal, cross-subsidization. This is a normal-times benefit of separation.

Both of these costs were evident ahead of, and during, the recent crisis. Ahead of crisis, resources gravitated to the investment banking side of the fence. Between 2000 and 2007, U.K. banks' trading books rose six times as fast as their banking books. Human capital made the same journey, helped by investment banking salaries rising four times as fast as commercial bank salaries since 1980.

In the teeth of crisis, risk cross-contamination became a potent factor. Basic banking services in universal banks were often subject to severe disruption from trading book losses, which exceeded by many multiples the capital allocated to them. That is why national deposit insurance schemes were extended and, in some cases, became temporarily unlimited. It is also why repeated attempts have had to be made to resuscitate weak credit growth over the past few years.

So, how far will existing structural proposals take us in harnessing these benefits? Volcker, Vickers and Liikanen seek legal, financial and operational separation of activities. So, in principle, each ought to prevent cross-contamination at crisis time. Whether they do so in practice depends on loopholes in, or omissions from, the ring-fence. And each of the existing proposals has open questions on this front.

For example, the Volcker rule separates only a fairly limited range of potentially risky investment bank activities, in the form of proprietary trading. The Vickers proposals mandate only a limited range of basic banking activities to lie within the ringfence, namely deposit-taking and overdrafts. And the Liikanen plans allow a wide range of derivative activity to lie outside of the investment banking ring-fence.

It could be argued that these loopholes are modest. But as the history of the Glass-Steagall Act demonstrates, today's loophole can become tomorrow's bolthole, today's ring-fence tomorrow's string vest. At a minimum, this suggests the need for full and faithful implementation of the spirit as well as the letter of the Volcker, Vickers and Liikanen plans, if risk cross-contamination is to be avoided.

A larger question-mark still hangs over whether these proposals will lead to a sea change in the allocation of resource to retail and investment banking. The cultures of investment and retail banking are quite distinct. Retail banking relies on forming long-term relationships, while investment banking is inherently shorter-term and transactional. Housing these sub-cultures under one roof makes achieving the necessary separation of cultures and capital a significant operational headache.

At a minimum, such a separation of culture and capital is likely to require entirely separate governance, risk and balance sheet management on either side of the ring fence. Without that, human and financial resource allocation to either side of the ringfence will become blurred. For example, without separate debt issuance for retail and investment banking, the cost of debt for a big bank will be a blended mix. The implicit subsidy in funding costs would then remain, and with it one of the main distortions associated with too-big-to-fail.

Only time will tell whether cultural separation can be achieved under the existing structural reform proposals. In the go-go years, will these reforms be sufficient to prevent the grass always appearing greener on the riskier side of the (ring-)fence? This is the acid test of the structural reform agenda.

6. Where next?

Progress has been made over the past few years toward eliminating too-big-to-fail, with further progress on implementation planned. But today's task is even more daunting than before the crisis. The big banks are even bigger. The system itself is more concentrated. And despite reform efforts, the market's best guess today about tomorrow's implicit subsidy is far larger than before the crisis, at over U.S.\$300b per year (Figure 3). The market believes that illicit state promise is even more likely to be kept.

The wrong conclusion to draw would be that existing reforms have failed or are unnecessary – quite the contrary. Rather it is that these reform initiatives, while necessary, may be insufficient to eliminate the too-big-to-fail externality. If so, what are the alternatives? Several have been mooted. **Resizing the capital surcharge is one possibility.** This would further reduce default probabilities among the biggest banks, thereby lowering the expected system-wide losses associated with big bank failure. Taking the earlier illustrative example, to reduce materially expected system-wide losses for the world's largest banks would require a capital surcharge several times larger than its current upper limit. Interestingly, this would take bank capital ratios to levels not dissimilar to recent quantitative estimates of their optimal value [Miles et al. (2011), Hellwig et al. (2011)].

Placing limits on bank size is a second option. By reducing balance sheet exposures, this measure would reduce directly system-wide losses in the event of big bank failure. The Dodd-Frank Act includes an explicit limit on the maximum deposit market share of U.S. banks, capping it at 10%. But this does not prevent banks rising to a scale, relative to GDP, at which they could imperil state solvency. For that reason, limiting bank size relative to GDP has recently been proposed by a number of commentators and policymakers [Fisher (2011), Hoenig (2012), Johnson and Kwak (2010), Tarullo (2012)].

Full structural separation of investment and commercial banking. A modern-day Glass-Steagall Act, has continued to attract support. The main benefit this would bring, relative to structural ring-fencing, is that it would eliminate loopholes from the ring-fence and better ensure that the distinct cultures of retail and investment banking were not cross-contaminated. That would lessen the risk of basic banking activities being starved of human or financial capital, both ahead of and during crisis. Full separation may also be operationally simpler to implement than the existing structural proposals.

Finally, **enhanced banking competition** would potentially help to reduce some of the problems of too-big-to-fail by reducing the degree of banking concentration. Greater exit from banking through enhanced resolution regimes can help. But, a bigger problem still is bank entry: the U.K. went 100 years without a new retail bank being set up. One potential barrier to banking entry is the difficulty of switching deposit accounts and loan contracts. A shared banking platform, containing customer account details, would dramatically reduce the frictions in search and switch for deposit and loan products for customers [Leadsom (2012)]. It could lower materially barriers for new banking entrants. A powerful counter-argument to all of these more radical proposals is that they could erode the economies of scale and scope associated with large banks. These economies clearly do exist in banking, as they do in other industries. For example, fixed costs are large in finance and spreading them widely ought to deliver productivity improvements.

The interesting question is at what point these economies of scale are exhausted. Indeed, informational and managerial diseconomies of scale are likely at some scale, whatever the business line. In his classic theorem, Ronald Coase tells us that firms will seek a privately optimal size which balances the benefits of economies of scale against these diseconomies [Coase 1937)].

How does all of this apply in banking? The empirical evidence on economies of scale and scope in banking is surprisingly patchy. Early studies, using data from the 1980s, failed to find scale economies much beyond bank asset sizes of around U.S.\$100m [Pulley and Braunstein (1992)]. Empirical studies in the 1990s nudged up the optimal bank scale to around U.S.\$10 billion [Amel et al. (2004), Mester (2005)].

Most recently, a small number of studies using data from the 2000s have pointed to scale-economies at much higher asset thresholds. For example, Wheelock and Wilson (2012) find scale economies for banks with assets up to U.S.\$1 trillion and Feng and Serilitis (2009) for bank with assets up to around U.S.\$1.5 trillion. Using data on banks with assets in excess of U.S.\$100b, Mester and Hughes (2011) not only find scale economies, but argue that these may increase with bank size.

At face value, these findings pose a real challenge to policy options which re-size banks. Or do they? Bank of England research has re-looked at the evidence on economies at different banking scales [Davies and Tracey (2013)]. They show that in standard models, there is evidence of scale economies for banks with assets above U.S.\$100b. Indeed, these economies tend to rise with banking scale.

But this finding is based on estimates of banks' funding costs which take no account of the implicit subsidy associated with too-big-to-fail. Removing this subsidy raises banks' funding costs, lowers estimates of bank value-added, and thereby reduces measured economies of scale. Davies and Tracey (2013) show that once an allowance is made for the implicit subsidy, the picture changes dramatically. There is no longer evidence of economies of scale at bank sizes above U.S.\$100b.

This evidence reconciles Coase's Theorem with the too-big-to-fail phenomenon. In line with Coase, banks have chosen the size that maximizes their private value. But implicit subsidies may have artificially boosted the privately optimal bank size. Subtracting this subsidy, removing the state crutch, would suggest a dramatically lower socially optimal banking scale. Like King Kong and Godzilla, these giants would arguably then be physiological impossibilities.

What about economies of scope? A recent study by Boyd and Heitz (2011) conducts a simple, but compelling, thoughtexperiment. They compare the lowest-available estimate of the social cost of the crisis with the highest-available estimate of the private benefit of scale and scope economies in banking. The social costs of too-big-to-fail exceed the private benefits of scale economies by an order of magnitude.

7. Conclusion

In his 1928 article, J B S Haldane observed that when "you drop a mouse down a thousand-yard mine shaft it walks away, a rat is killed, a man is broken, a horse splashes." When big banks disappeared down the mineshaft in 2008, their splashes generated a tsunami. To prevent that, their physiology needs to change. Existing change initiatives are right in direction, but may be insufficient in degree. There may be a distance to travel before banking is the right size.

Commentary by Philip Booth: Is big dangerous?

Andrew Haldane begins his paper with a lesson from biology. I think that there are other lessons we can learn from evolutionary biology as well as the one cited by Andrew. The family of the economist F. A. Hayek, whom I shall quote later, were also biologists and Hayek was very familiar with the subject himself. Hayek drew analogies between the market process and evolution. A market process is a process of evolution; of trial and error; of the copying of good ideas; and of the failure of bad ideas. This requires that banks fail. Clearly to ensure that we have an orderly market, they must fail in an orderly fashion. But, failure is essential if competition is to be functional. The second lesson that Hayek took from the biological sciences is the way in which markets were extended networks of individuals and institutions in which entrepreneurs would respond to the information that was relevant to them. That information cannot be centralized. As such, the detailed regulation of market and the replication of market processes by government agencies, is not just difficult or undesirable, but impossible.

I will come back to these issues in a moment, but first I want to focus on the specific issues brought up in Andrew's paper. The paper focuses, to a large degree, on the problem of too-big-to-fail banks. It is argued that these have a large implicit subsidy from the taxpayer – and I entirely agree with that. It was mentioned that certain reforms have been put in place and possible further reforms are then suggested by Andrew.

However, it is worth noting that it is not entirely clear that a concentrated banking system – at least before the regulatory reforms that we have seen in Britain – was more dangerous than a less concentrated one. If you compare the U.K. and the U.S. banking systems with each other over the last century or so, the U.K.'s has consistently been extremely concentrated and that in the U.S. has not. Arguably, the U.K.'s has been more stable – certainly that was the case during the Great Depression. Also, the failed savings and loans sector in the U.S. was not a concentrated sector either, but that demanded a huge bailout. I think there are three nuances that we have to add to Andrew's argument:

- In the traditional analogy between the payments' system and the plumbing system, the system breaks at its weakest link – not necessarily its biggest link.
- 2. If similar mistakes are made throughout the system by market practitioners for whatever reason then the impact (as in the savings and loans crisis) could be just as great as if one big bank failed in a concentrated system.
- 3. In a concentrated banking system it can be easier to coordinate a response to the failure of one link in the system. Indeed, the remaining banks may have a strong incentive to coordinate a "lifeboat" without state help because the externalities of failure are more easily contained within a small group of institutions that can coordinate effectively.

In addition to this, larger banks might be more diversified, and therefore less prone to failure. We have, it seems, implicitly accepted that "bigness" and/or concentration go together with danger in the banking system because of the events of 2008, but this has not always been the experience.

Andrew's point about concentration and size, though, is more pertinent under the new regulatory regime. Under that regime, a resolution process to ensure orderly failure is the key. Our legal system for dealing with a failing bank in 2008 was not fit for purpose. Therefore, the failure of any bank could cause chaos (even the failure of a small bank). If that legal system becomes fit for purpose, it will become much easier to resolve smaller and simpler financial institutions and big, while complex institutions could still cause problems.

Is structural change and capital regulation the answer?

But, what else should we do if we accept the argument that there may be a residual problem of too-big-to-fail? I struggle with the idea of imposing structural change through legislation. It is a very blunt tool. I think that The Vickers' approach is better than the E.U. proposals and Dodd Frank, but I still have some concerns about Vickers' proposals and, indeed, the government has relaxed the rules proposed by Vickers in specific cases.

Further, capital regulation is another possibility and we will, of course, have even more capital regulation under the new regime. However, I reject that approach entirely for all the reasons that Andrew laid out in his "Dog and frisbee" speech at Jackson Hole [Haldane and Madouros (2012)]. We must get away from discretionary regulation designed to ensure that each bank has exactly a 0.5% probability of failure – or whatever the probability is. Simple capital rules get gamed and lead to complexity. Complex rules get gamed using complex methods and lead us to where we are now. Last year, there were 14,200 new banking regulations worldwide; before 2008, there were probably 1,000,000 paragraphs of FSA banking regulation; and the Dodd Frank bill will probably be 30,000 pages long with the associated regulations. This is completely crazy and also leads to the system being captured by a combination of the industry and a bureaucratic clique. Of course, any system of capital requirements designed to prevent failure also prevents competition from operating properly. Competition needs failure.

International capital regulation is even worse as it institutionalizes systemic risk: if the regulations incentivize particular types of behavior, the whole system is liable to go wrong at the same time. And this leads to an important question: if we are convinced that the resolution regimes will work, then why do we need a capital regime to stop failure happening? If we are not convinced that our resolution regimes will work, we should think again.

At the same time, for all the reasons we understand from public choice economics, and for the reasons laid out by Hayek, regulatory discretion – which is an alternative to detailed rules – is very dangerous. As Hayek (1989) said: "The curious task of economics is to demonstrate to men how little they really know about what they imagine they can design." Even the most wellmeaning and intelligent regulators should not be trusted with large amounts of regulatory discretion.

Stable legal frameworks: the E.U. elephant in the room

It is because Andrew understands this that his paper has discussed other possible solutions to the problem of too-big-to-fail. They may have some merit; but I worry that they would not last more than a generation as developments came along that transcended their relevance. So, I would like to add one more possibility to the mix. Indeed, I would like to make a proposal that could make all other banking regulation – except disclosure requirements – irrelevant.

I have come to the view from comparing systems of financial regulation in different sectors at different times in history that there is only one stable, just and effective solution to these problems. Everything must depend on the ability to ensure that a financial institution fails safely; and regulatory discretion must be extremely limited. Andrew partly accepts this starting point. However, in trying to achieve this, I think the focus on the separation of investment and trading and retail banking has been misplaced. There are too many grey areas, and an inherently incorrect assumption has been made that retail banking is somehow not risky. Separation of retail and trading functions may sometimes be a good idea, but it will not be necessary in some cases and will certainly not be a panacea.

In looking at the retail/invest/trading divide we have missed another aspect of bank structure that makes resolution much more difficult: international branching, especially when legal frameworks, deposit insurance schemes and the legal personality of subsidiaries do not match up. Perhaps nobody dares to speak out about this because of issues regarding E.U. regulation and its jurisdiction over U.K. law that would be raised. But, many financial firms do not use branching for their international operations and work through subsidiaries. There is an active international insurance market operating through subsidiaries too – even in the E.U. where branching is encouraged. You can have free trade without banning – as the E.U. does – requirements to have subsidiaries in different jurisdictions. It is often the international branching of banks that makes resolution so difficult – the Icelandic case being one important example. I would not go as far as requiring international subsidiarization of banks, but it ought to be an option that is available.

Specifically, I think that the most stable approach would be as follows. The legal duty of the Bank of England in the field of banking regulation should be to ensure that any bank can be wound up safely without it disrupting the rest of the banking system, and without undue calls on the deposit insurance system. The Bank of England should not enforce capital requirements to prevent failure, but should have the duty to ensure that failure is orderly if it happens.

This, of course, requires power. If it believes that a bank could not be wound up in an orderly fashion, the Bank of England should have the power to enforce structural change to ensure that it can be wound up so that the corporate structure is simplified, subsidiaries are set up and/or living wills are drawn up. This could involve separation of international business into subsidiaries, separation of trading and retail activities and so on, but it may not necessarily do so. I would not regard this as regulation as such. The Bank of England would simply be placing requirements on those banks to which it provided lender of last resort support that were deemed necessary in order for banks to have the principles of their special bankruptcy law applied to them. This would be complemented by depositor preference and risk-based deposit insurance levies, if deposit insurance remains.

Obtaining a banking license, therefore, would no longer depend on whether a bank was safe but on whether it could fail safely. This certainly would facilitate new entry as new entrants would hardly need to be regulated at all. There would, of course, need to be a right of appeal to an independent technical panel and then to a court, in case the Bank of England was excessive in exercising these legal powers, but the judge would decide on the basis of a simple principle – what structure is necessary to ensure that a bank can be wound up so that as many as possible of its contractual obligations can be met in an orderly way.

Those systems of financial regulation that have succeeded in the past have relied on good legal principles and the incentives that the market itself provides. In my view, there should be less emphasis on regulatory control, preventing failure and on structural reform. I have considered lots of alternatives and I cannot think of another enduring and stable way of resolving the problems that Andrew has talked about in this paper as well as in his dog and the frisbee speech. We must retreat from detailed regulation of all behaviors to simple legal frameworks. I should add that this leaves unconsidered whether regulation is necessary for consumer protection purposes. In my view, it is not, but that is a separate argument.

In conclusion, if we really do have a credible failure regime then we do not have to worry about whether the socially optimal size of banks is different from the privately optimal size or whether socially optimal capital levels are different from privately optimal levels. If we get the legal framework right, a lot of the problems with which we struggle are no longer relevant. I hope that Andrew might add this proposal to his list of possible ways of dealing with the too-big-to-fail problem.

References

Acharya, V., L. Pedersen, T. Philippon, and Richardson, 2010, "A tax on systemic risk," NYU Stern Working Paper

Admati, A., P. DeMarzo, M. Hellwig, and P. Pfleiderer, 2011, "Fallacies irrelevant facts and myths in the discussion of capital regulation: why bank equity is not expensive," Rock Center for Corporate Governance at Stanford University Working Paper No. 86 Alessandri, P., and A. Haldane, 2010, "Banking on the state," Bank of England Amel, D., C. Barnes, F. Panetta, and C. Salleo, 2004, "Consolidation and efficiency in the financial sector: a review of the international evidence," Journal of Banking and Finance, 28(10), 2493-2519

Boot, A., and L. Ratnovski, 2012, "Banking and trading," IMF Working Paper WP/12/238

Boyd, J., and A. Heitz, 2011, "The social costs and benefits of too-big-to-fail banks: a bounding exercise," University of Minnesota working paper

Brunnermeier, M., A. Crockett, C. Goodhart, A. Persaud, and H. Shin, 2009, "The fundamental principles of financial regulation," London, Centre for Economic Policy Research

Cecchetti, S., M. Mohanty, and F. Zampolli, 2011, "The real effects of debt," BIS Working Papers 352, Bank for International Settlements

Coase, R. H., 1937, "The nature of the firm," Economica, New Series, 4(16), 386-405

Davies, R., and B. Tracey, 2013, "Too big to be efficient? The impact of too-big-tofail factors on scale economies for banks," Mimeo **Easterly, W., R. Islam, and J. Stiglitz, 2000**, "Shaken and stirred, explaining growth volatility," Annual Bank Conference on Development Economics. World Bank, Washington D.C.

Feng, G., and A. Serilitis, 2009, "Efficiency, technical change, and returns to scale in large US banks: panel data evidence from an output distance function satisfying theoretical regularity," Journal of Banking and Finance, 34(1), 127-138 Fisher, R., 2011, "Taming the too-big-to-fails: will Dodd-Frank be the ticket or is lap-band surgery required?" Remarks before Columbia University's Politics and Business Club Haldane, A., and V. Madouros, 2012, "The dog and the Frisbee," Speech delieverd at the Federal Reserve Bank of Kansas City's 36th economic policy symposium, "The Changing Policy Landscape", Jackson Hole, Wyoming, 31 August

Hayek, F. A., 1989, The fatal conceit: the errors of socialism, University of Chicago Press

Hoenig, T., 2012, "Back to basics: a better alternative to Basel Capital Rules," Speech to The American Banker Regulatory Symposium, 14 September. http://www.fdic.gov/news/news/speeches/chairman/spsep1412 2.html

Hughes, J., and L. Mester, 2011, "Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function," Working Papers 11-27, Federal Reserve Bank of Philadelphia

James, C., 1991, "The losses realized in bank failures," Journal of Finance, 46(4), 1223-1242

Johnson, S., and J. Kwak, 2010, "13 bankers: The Wall Street takeover and the next financial meltdown," Pantheon

Jorda, O., M. Schularick, and A. Taylor, 2011, "Financial crises, credit booms, and external imbalances: 140 years of lessons," IMF Economic Review, 59(2), 340-378 Leadsom, A., 2012, "How an old hand would change Barclays," Financial Times Opinion Editorial, July 4. http://www.ft.com/cms/s/0/fb094718-c5f7-11e1-b57e-00144feabdc0.html#axzz2A6YBHLDV

Levine, R., 2004, "Finance and growth: theory and evidence," NBER Working Papers 10766, National Bureau of Economic Research, Inc.

Merton, R. C., 1974, "On the pricing of corporate debt: the risk structure of interest rates," Journal of Finance, 29(2), 449-470

Mester, L. J., 2008, "Optimal industrial structure in banking," in Boot, A., and A. Thakor (eds.), Handbook of financial intermediation, Amsterdam: North-Holland Miles, D., J. Yang, and G. Marcheggiano, 2012, "Optimal bank capital," Economic Journal, 123, 1-37

Noss, J., and R. Sowerbutts, 2012, "The implicit subsidy of banks," Bank of England FS Papers Series, FS Paper No 15

Panizza, U., J.-L. Arcand, and E. Berkes, 2012, "Too much finance?," IMF Working Papers 12/161, International Monetary Fund

Philippon, T., and A. Reshef, 2009, "Wages and human capital in the U.S. financial industry: 1909-2006," NBER Working Papers 14644, National Bureau of Economic Research, Inc.

Pulley, L. B., and Y. M. Braunstein, 1992, "A composite cost function for

multiproduct firms with an application to economies of scope in banking," Review of Economics and Statistics, 74(2), 221-30

Reinhart, C., and K. Rogoff, 2010, "Growth in a time of debt," American Economic Review, 100(2), 573-578

Segoviano, M., and C. Goodhart, 2009, "Banking stability measures," IMF Working Paper 09/04, Washington: International Monetary Fund

Schanz, J., D. Aikman, P. Collazos, M. Farag, D. Gregory, and S. Kapadia, 2011, "The long-term economic impact of higher capital levels," Bank of England, Mimeo Tarullo, D., 2012, "Financial stability regulation," speech at the Distinguished Jurist Lecture, University of Pennsylvania Law School, Philadelphia, Pennsylvania Wagner, W, 2009, "In the quest of systemic externalities: a review of the

literature," mimeo Wheelock, D. C., and P. W. Wilson, 2012, "Do large banks have lower costs? New estimates of returns to scale for U.S. banks," Journal of Money, Credit, and Banking,

44(1), 171-199

EY | Assurance | Tax | Transactions | Advisory

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 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 Global Limited, each of which is a separate legal limited by guarantee, does not provide services to clients. For more information about our organization, please visit ey.com.

All Rights Reserved.



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

contained within *The Journal* are generic and represent the views and opinions of their authors. 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 or not taken by those receiving The Journal.

Accredited by the American Economic Association

ISSN 2049-8640