

The future of central bank data

Journal of Banking Regulation (2013) **14**, 185–194. doi:10.1057/jbr.2013.7; published online 3 July 2013; corrected online 14 August 2013

The article introduces this special issue of the *Journal of Banking Regulation* dedicated to data. Most of the articles published here are based on presentations delivered at a seminar organised by the Statistics and Regulatory Data Division through the Centre for Central Banking Studies (CCBS) at the Bank of England. The seminar included participants from many central banks and regulators across the world. The article and this special issue do not necessarily express the views of the Bank of England. Views expressed are exclusively those of the author.

INTRODUCTION

The word ‘clue’ has its origins in the word ‘clew’, a ball of yarn. The linguistic evolution occurred during the nineteenth century when Victorians redefined ‘clue’ to mean a tiny thread of fabric that could unravel the whole. This change in meaning can be read within context as itself a clue, as typifying a broader sharpening of the Victorian mind to detail. From Sherlock Holmes to the development of modern science, the most tedious minutiae started acquiring new and potentially fateful significance in this period.¹

However, while ‘clue’ still connotes mystery and elicits excitement, mentioning ‘data’ in polite conversation runs the risk of inviting boredom. This is so despite the fact that these two terms have nearly the same meaning; both clue and data refer to evidence on the basis of which decisions are made and problems are solved.

It would be a mistake to assume disinterest with data as a topic for discourse reflects merely

popular prejudices. On the contrary, it sums up sentiments expressed by even the noblest of minds. Consider Lord Cobbold, Governor of the Bank of England from 1949 to 1961. Reacting to the Radcliffe Committee report on banking (1960), which was light in terms of monetary policy recommendations but heavily advocated that the Bank collect more data and publish more statistics, Governor Cobbold purportedly retorted that, ‘A central bank is a bank, not a study group’.² Far from an outlier, Cobbold’s tepid reaction to the Radcliffe Commission’s recommendations encapsulates a patterned preference among central bankers historically for parsimonious data collections. From the Monetarist emphasis on measures of the money stock to the New Keynesian focus on output gaps, central bankers usually have taken a tailored approach to data collection, limited to measuring, monitoring and controlling a small set of aggregate policy variables.

However, this approach to data may be changing given that its misreporting has been central to the plot of a host of recent financial scandals from Enron to LIBOR. The increasing volume of regulation in response to these scandals has in turn generated the need for increasingly micro-level data to monitor compliance with them. This is evident in a number of recent initiatives. At an international level, for example, the Financial Stability Board (FSB) has proposed a common reporting template for globally systemically important banks (GSIBs) that will capture detailed data on their counterparty and instrument exposures. Meanwhile, across Europe, new reporting standards such as the common reporting (CoRep) templates

issued by the European Banking Authority (EBA) are changing the detail of the data that firms report. Suffice to note that the EU Solvency II directive mandates insurers report their assets to regulators on a security-by-security basis. At the same time, within the United Kingdom, the newly formed Financial Policy Committee has indicated it will seek to improve and expand data collections to better execute macro-prudential policy.³

Taken as a whole, the direction of regulatory travel towards more granular data may be viewed as a conjectural phenomenon, that is, as a reaction to the recent financial crisis. However, it may also register a fundamental structural shift, reflecting a more general embrace by society of so-called 'big data', that is, data sets that are large, frequently changing and multimedia.⁴ IBM, for instance, estimates that 90 per cent of the data in the world today has been created in the last 2 years.⁵ The precondition for this striking expansion in data volume is technological progress that has increased storage capacity and processing power while containing costs. However, technological progress has been spurred on by changing cultural habits and consumer demand. In the age of the Internet and reality TV, many individuals now live their lives virtually through a steady stream of texts, Tweets and Googles. These practices generate an enormous volume of electronic records that can be mined for new insights on individuals and society as a whole. The full implications of these developments with respect to individual privacy, public policy and commercial opportunity are just starting to be appreciated.⁶

CENTRE FOR CENTRAL BANKING STUDIES (CCBS) SEMINAR

It is against this backdrop that the Bank of England's Statistics and Regulatory Data Division (SRDD) convened a 2-day seminar through the CCBS entitled 'The Future of Regulatory Data and Analytics'. The conference was held 17 and 18 January 2013. Seminar

speakers from academia, government, for-profit and non-profit organisations addressed around 75 bank staff and approximately 25 delegates from foreign central banks. Delegates included central bankers from Austria, Azerbaijan, Chile, Croatia, the Czech Republic, Germany, Hungary, India, Israel, Italy, Nepal, the Netherlands, Romania, Russia, Saudi Arabia, Spain, South Korea, Sweden and the United States.

Although the seminar had international participants, its impetus was domestic developments, namely the assumption by the Bank of supervisory and regulatory responsibilities for UK financial firms through its legal subsidiary the Prudential Regulatory Authority (PRA). This expansion in the Bank's responsibilities carries with it an expansion in the data for which SRDD has become responsible. Where in the past, SRDD collected data through the medium of approximately 40 statistical forms, the creation of the PRA means the division has inherited around 250 regularly reported and *ad-hoc* forms from the now defunct Financial Services Authority (FSA). Growth in the number of reporting forms poses potential operational challenges. In the past, one or two SRDD staff members would become specialists in a particular form such as the balance sheet return (Form BT) or analysis of deposits from UK residents (Form AD). However, the sheer number of forms now means that operational model may not be scalable. Furthermore, as in many cases the data collected by different legacy FSA forms overlap with each other and with existing Bank forms, current data collection methods may not be optimally efficient.

The immediate challenge for SRDD has therefore been to develop a new data strategy that eliminates redundancies while closing gaps, an approach that integrates micro-prudential data with macro-financial statistics. These are not only or primarily practical challenges. Rather, they touch upon a larger set of theoretical issues posed by the general transition to a 'big data' society. What are the benefits and who pays the costs for larger, faster and more accurate data? How is balance to be

struck between individual privacy and public transparency? And on what basis can heterogeneous data be harmonised? These are the timely questions that this special issue of the *Journal Banking of Regulation* try to answer.

OVERVIEW OF PAPERS

In the opening article ‘Leveraging Data for Financial Stability Monitoring’, Dessa Glasser, Deputy Director of the Data Center, Office of Financial Research (OFR) in the United States, observes that data are increasingly seen by financial firms as a strategic resource requiring its own specialist stewards. This is evidenced by the increasing employment of executive level Chief Data Officers and the establishment of dedicated data functions within financial firms. In many respects, such developments echo the rise of risk management as a separate specialisation during the 1980s and 1990s.

The increasing importance of data management in the private sector also parallels the heightened significance attributed to data by regulators post-financial crisis. As Glasser argues, the recent crisis revealed that the authorities did not have data sufficient for assessing rising risks in the financial system. According to Glasser, gaps in regulatory data collections arose for a number of reasons. In some cases, the data simply was not recorded by firms or was documented in formats that could not be easily conveyed to regulators. In other instances, even when the data existed, differences in how firms defined data made calculating accurate sectoral aggregates difficult, if not impossible. Glasser goes on to highlight how regulators are responding to these challenges through a variety of new initiatives, of which the creation of the OFR’s Data Center is a prime example.

The scene then shifts from the United States to the United Kingdom. As a key driver of changes in data management at the Bank of England is its expanded responsibilities for financial firm supervision as executed by the PRA, the next three papers in this collection

reflect on the distinguishing characteristics of the new regulatory regime and the role data might play in it. In her paper ‘Judgment-led regulation: What it is and how to make it work’, Andromachi Georgosouli argues that the increasing emphasis given to judgement suggests regulators now have greater scope for discretionary decision making. According to Georgosouli, this implies regulatory action will be now based not only on what the current data shows, but also by the past experiences and forward-looking imagination of micro-prudential supervisors.

Like other authors in this volume, Georgosouli observes that data do not speak for themselves. In other words, data becomes meaningful only within context and the same data is open to different interpretations depending upon our theoretical priors. For instance, whether an increase in the growth rate of the money stock is a consequence of greater demand for assets and consumables or will cause such demand is not strictly given by the facts. In the field of regulatory reporting, the issue of multiple interpretations extends in both directions. It is not only that regulators may have differing interpretations of the same data. It is also the case that reporting firms can have different interpretations of reporting requirements. Consider, for example, the requirement to report the monetary value of real estate used as collateral to secure loans.⁷ In the absence of more precise guidance, some firms might report the gross current market price of the real estate, whereas others may report historic appraisals netting out more senior debt obligations.

The potential for differences in the interpretation of data and data requirements underscores the importance for regulators and reporters to develop consistent definitions of data, an issue that other contributors to this volume discuss in-depth. For Georgosouli, however, this inscribes a tension at the core of the new regulatory regime. At the same time supervisors are being encouraged to exercise judgment and work in non-standard ways, there are simultaneous calls to standardise data.

While on the surface these two projects may appear to pull in opposite directions, they could be viewed as complementary. In order to see how, a critical distinction can be drawn between metrics and data *sensu stricto*. Currently, what firms often report to regulators are metrics such as their regulatory tier 1 capital or their liquidity coverage ratio. Firms construct these metrics by aggregating their numerous financial exposures. However, it is at the exposure level that we have data in the pure sense of empirical facts expressed via conceptual categories intrinsic to any financial position. Thus, while the specific mix of assets that compose the liquidity coverage ratio metric might change over time, the categories defining an asset as such are more standard. These include a counterparty name; a nominal amount for the asset denominated in a particular currency; the legal system through which an asset is constituted; and the time horizon over which it matures. This suggests an affinity between efforts to promote data standardisation and data collection at a highly granular level. Such standardised, granular data might in turn enable forward-looking regulatory judgment. Specifically, regulators would not be analytically constrained by a limited set of reported metrics but could instead combine and reconfigure the underlying granular data to continually construct new measures to assess the state of the financial system. Such a hypothetical regulatory regime could be data-intensive but light in terms of *ex-ante* prescriptive rules. In such a scenario, regulators might not specify in advance how much capital a firm needs to hold against losses, but instead arrive at such judgments over the cycle on the basis of data they received.

Of course not everyone is convinced that bank capital regulation or discretionary supervision can prevent financial crises.⁸ As Rosa Lastra notes in her paper 'Defining forward looking, judgment based supervision', the current challenge of fleshing out the meaning of judgment-based supervision echoes a longer debate over the balance to be struck in central bank operations between rules and discretion.

In benign economic conditions, central banks have tended to prefer operating according to rules. For most of the twentieth century, this principally involved pegging central bank liabilities (notes) to a growth rule, either nominally fixed or floating by reference to an external standard (gold or another currency). However, crises tend to provoke the rewriting of rules and greater discretion in the form of lender-of-last resort facilities. This includes greater discretion regarding which institutions fall within the regulatory reporting perimeter and are thus eligible to receive such support.

Joanna Gray and Peter Metzing consider this topic in their article 'Defining and delivering judgment-based supervision: The interface with the legal system'. According to Gray and Metzing, the post-crisis approach to supervision might include gathering data from third parties such as city lawyers who engineer financial transactions. There are clear challenges to collecting data at the detailed level of legal documentation. Most significantly, there is the issue about whether and to what extent such data are protected by conventions of lawyer-client confidentiality. There are practical challenges as well. These include how to transform millions of pages of unstructured legal agreement text into machine-readable data that can be easily queried by regulators.⁹

The reward for solving this practical challenge could be great. Currently many financial analyses focus attention on the static numbers recorded in profit and loss, balance sheet, and cash flow statements. Like the economic impact of monetary policy, the reporting of these numbers is subject to long and variable lags. Moreover, these discrete values are better described by a probability distribution because they are dynamic.¹⁰ The dynamic nature of this data is not just the result of market fluctuations. On the contrary, it is an intrinsic feature of the legal structure of the exposures themselves. Thus the economic value of an exposure can change because of covenants defining default, termination, and margin call events. As Gray and Metzing conclude, 'the granularity of the

legal issues involved and hence the challenge of collecting component data to such product codes cannot be overstated'.¹¹

If the obstacles to implementing such improvements seem daunting, the financial sector can take comfort from the historical evolution of data management in other industries. This is the hopeful message conveyed by Alistair Milne in his paper 'The Rise and Success of the Barcode: Some lessons for Financial Services'. Although the barcode is now a taken-for-granted technology, Milne reminds us it only became widely adopted in the 1980s. It has generated significant benefits. By providing data that populates near real-time trade repositories, the barcode enables firms to flexibly manage their supply chain in order to more quickly respond to changes in market demand. Moreover, the barcode not only allows for the more efficient management of existing supply chains, but has also helped forge new ones. As commodities can now circulate simply through scanning, the native language of market participants poses less of an obstacle to international trade.¹² If we believe Milne, universal adoption of legal entity identifier (LEIs) and product identifier (PIs) codes might have similarly beneficial effects. For example, one of the key costs incurred as part of mergers and acquisitions in the financial industry is the integration of different information technology (IT) systems. If the IT architecture and data models of firms were more similar, this could reduce transaction costs in the market for corporate control.

Indeed Milne's paper highlights the need for regulators to help market participants find the profit opportunity in better data management, if initiatives such as the LEI are to be successful over the long run. Ironically, although economic theory often explains financial firms' *raison d'être* as their information advantage, a growing body of empirical literature suggests the contrary, that the sector as a whole lags behind other industries when it comes to efficiently exploiting data. This is so because the pre-crisis revenue model of many

financial firms emphasised originating new business, through active liabilities management and sophisticated sales practices, with less attention given to improving existing business processes that did not yield immediate return on investment.¹³ This has meant financial firms have tended to under-invest in data management. Exacerbated by frequent mergers and acquisitions in the industry, the net effect is that data in many firms are stored in disconnected silos.¹⁴

The opportunity costs of this patchwork approach to data may be large. For example, a recent study undertaken by Clearstream Banking with Accenture found that financial firms do not have a holistic view of their unencumbered collateral. This is because such data are warehoused by different operational divisions in IT systems that do not interface with each other. The report concluded that around 10–15 per cent of collateral available for discount is left idle, translating into an annual cost of more than £4 billion to the industry.¹⁵

Reducing opportunity costs from such data fragmentation requires that financial firms develop a single view of their financial exposures. However, according to the Bank for International Settlements, many firms lack the ability to quickly and accurately aggregate their risk exposures across their various legal subsidiaries.¹⁶ In theory, if such a consolidated view existed and was shared with regulators, this could provide the basis for more meaningful dialogue. A major obstacle here is the development of enterprise and industry wide definitions of data. This is because the financial sector often uses the same word to mean different things. Thus 'shares' on the stock market mean publicly traded securities that confer on their owners rights to a variable dividend, while 'shares' in a credit union context signify non-tradable claims for a nominally fixed amount, equivalent to a bank deposit. Likewise, the financial sector often uses different words to mean the same thing. Thus, what are called gilts in the United Kingdom are called government

bonds elsewhere. In order for regulators to make meaningful peer comparisons and compile accurate aggregates, firms would ideally share a common language and report their data at the same frequency.

Mike Bennett is uniquely positioned to discuss the challenges of constructing a common financial language and elaborates on them in his paper titled 'The Financial Industry Business Ontology: Best Practice for Big Data'. Bennett is the Head of Semantics and Standards at the Enterprise Data Management (EDM) Council. In partnership with financial institutions, EDM has been developing the Financial Industry Business Ontology (FIBO), a data dictionary that defines terms and relationships between the range of instruments and business entities found in financial services.

There are two features of FIBO that make it potentially attractive as a data standard. First, unlike LEIs or PIs that are regulatory artifices grafted on to business realities, FIBO is based on natural language terms already used in finance. The familiarity of FIBO terms makes its acceptance more likely. Second, FIBO is a taxonomy in which terms are related to each other. Thus, a mortgage is defined as a type of debt instrument, which is defined as a kind of contract. These relational descriptions bring logical order to an otherwise bewildering array of bespoke financial products.

It may be significant that FIBO is pitched at the level of instruments and entities. This again suggests an affinity between data standardisation and data collection at a granular level. This affinity exists because, in spite of the apparent complexity and diversity of financial products, they share core characteristics. Here, a distinction can be borrowed from cultural anthropology between the *emic* and *etic* – in this case, between the myriad of terms used by financial market participants to describe their activities and a description of these same activities in an analytically more abstract form. As the American economist John Kenneth Galbraith once observed, although 'the world of finance hails the invention of the wheel over and over

again ... all financial innovation involves, in one form or another, the creation of debt secured in greater or lesser adequacy by real assets'.¹⁷

Consider, for instance, the parallels between retail pawnbrokers and security lenders in wholesale markets, the roughly basic identity between financial transactions predominantly engaged in by lower income households on the high street, and financial transactions between firms in the realm of haute finance. In pawnbroker shops, one party (the retail client) pledges their assets (often though not exclusively jewellery) to the counterparty (the pawnbroker) who temporarily acquires possession of the asset in return for cash. On the basis of the precedent established in case law (*Donald v. Suckling*, 1866), the pawnbroker might then legally re-pledge their clients' property in order to secure funds themselves.¹⁸ Likewise, in security-lending agreements, one party (often pension funds) pledges their assets (say stocks) to another party (think banks) that temporarily acquires possession of the asset in exchange for cash. Like pawnbrokers, the bank can then use the asset for its own transactions, often to settle trades in which the bank may have sold the stocks before they possessed them.¹⁹

Therefore, although the monetary value of pawnbroking and security lending transactions differ, as do their meaning for the people involved, there are structural similarities between them. Both types of transactions involve counterparties, currencies, re-hypothecation rights and collateral possessing characteristics discounted by some haircut. This data could be captured in an integrated database built around these key categories. The benefit of such an approach is that it could flexibly adapt to expansions or contractions of the regulatory perimeter. For example, although the regulation of doorstep lenders does not currently fall within remit of the UK central bank, such outfits could be easily brought within scope if an integrated data structure existed, as their business is not fundamentally different from other, already regulated lenders such as banks.

This would save the central bank the costs of developing custom-made forms and warehousing solutions simply for doorstep lenders.

An integrated, granular approach to data is discussed in-depth by Erich Hille in his article 'Recent Developments in Restructuring the Austrian Bank Reporting System'. Hille is at the forefront of developing such an approach as part of his work in the statistics division of the Oesterreichische Nationalbank (OeNB), Austria's central bank. Although Hille's article focuses on developments in Austria, the issues he raises with respect to regulatory reporting have wider resonance. According to Hille, the key reason for redundancies and gaps in regulatory data sets is that collections have traditionally been conducted by subject matter specialist teams working in isolation of each other. This fragmented approach to data means central banks have not often benefitted from economies of scale and scope; different divisions and individuals have had varying access to different data sets. The alternative is for central banks to develop a central data management function to rationalise various data requirements from across the enterprise. Here, a distinction can be drawn from debates on central bank independence during the 1990s between goal and operational independence. In this case, which data to collect (goals) should be disentangled from the issue of how to best collect it (operations). While other areas of the central bank might set data requirements, the actual methods employed to meet those requirements might be best determined by a specialist data function. As Jean Claude-Trichet, former President of the European Central Bank (ECB) has stated it is now 'considered best practice for a central bank to integrate the overlapping parts of the various statistics it produces and to concentrate the production of all of its statistics within a single statistical organizational unit'.²⁰

In Austria, the OeNB is implementing such an approach in collaboration with their reporting institutions. The centre-piece of the Austrian approach is what Hille calls a 'basic cube'. Basic cube is essentially a relational

database. Like Excel, the cube has rows and columns. Rows list exposures on a contract-by-contract basis, while columns contain attributes of those contracts such as reference rate, maturity date and valuation procedures. Although the broad structure of the basic cube has been agreed, the exact form of regulator-reporter collaboration is still being worked through. At its most ambitious, the OeNB and the industry will jointly own a legal subsidiary that will operate a common data platform. At the very least, the industry will adopt a common financial language. In either case, reporters will fully disclose their financial data to the regulator.

An important question is what can regulators do with such detailed data? Even if the authorities received position-by-position data on a real-time basis, regulatory interventions might be already 'too late' as the exposure would have been originated. In this respect, better data is not a substitute for better incentives and institutional structures in the financial sector. Furthermore, there is longstanding skepticism in some quarters about the good that can possibly come from central authorities collecting data at all. Such skepticism at root draws its sustenance from more general critiques of government regulation, particularly Friedrich Hayek's argument that central authorities are incapable of developing a holistic view of the economy and distort market processes when they try to do so.²¹ But as even Hayek himself conceded, 'As soon as it is realized that, owing to the existence of banks the equilibrating forces of the economic system cannot bring about the automatic adjustment of all its branches to the actual situation, which is described by classical economic theory, it is justifiable even from a liberal point of view that the banks should be subjected to degrees of publicity as to the state of their business which is not necessary in the case of other undertakings'.²²

Sheri Markose extends Hayek's insights in her paper 'Systemic Risk Analytics: A Data Driven Multi-Agent Financial Network (MAFN) Approach'. According to Markose, financial

crises are created by banks and other credit issuing institutions in the course of their normal operation. Although it is profitable for individual institutions to create interest-bearing credit, they collectively have a tendency to create debt in excess of central bank money in which such obligations are ultimately settled. In other words, these institutions generate systemic risk: negative externalities potentially paid for by the non-bank public.

As a remedy, Markose proposes levying a tax on financial institutions. The tax would be used to prefund a pot of money for resolving financial firms when they fail. According to Markose, such a tax should be progressive, proportional to the risk an individual firm poses to the stability of the system as a whole. This requires detailed data on the bilateral liabilities of each firm to assess their overall size and interconnectedness to other firms to determine precisely from where systemic risk is likely to arise and spread. In particular, Markose argues that the systemic riskiness of a firm can be determined by their 'eigen-vector centrality', a measurement whose meaning is made clear in the course of her article.

According to Markose, such data-intensive mapping of financial firms' liabilities is required because of the apparent inadequacies of conventional key risk indicators as early warning signals of the recent crisis. Consider Northern Rock. At the start of 2007, the same year as the run on the Rock, the firm had the highest rated stock of all European banks and had low spreads on its bonds. The firm still reported mortgage arrears at or below half the British bank industry average and recorded similarly low repossession rates. These impressive results were the product of Northern Rock's generally solid underwriting standards. The company had limited exposure to loans in what are conventionally viewed as high risk sectors, such as business lending and credit cards. In terms of residential lending, the company diversified nationally and had a limited number of buy-to-let and second charge mortgages on its books.²³

Yet the firm still failed. Indeed, even after it was nationalised, Northern Rock's credit rating

never dipped below A-²⁴. Perhaps then, the lesson to be learned from Northern Rock is that regulating firms according to a static set of risk metrics is insufficient because our interpretation of these metrics changes over time, to echo Georgosouli. Suffice to recall that Northern Rock and Halifax Bank of Scotland (HBOS) were celebrated by equity analysts in the lead up to the crisis for having the lowest cost-to-income ratios of all British banks. What was then interpreted as an indicator of efficiency appears now in retrospect as a sign of underinvestment in risk management.²⁵

Furthermore, the potential problem with limiting regulatory data collections only to the data needed to construct currently en vogue regulatory metrics is that the utility of these metrics diminishes with time. This is not only because analytical fashions exogenously change. Rather, it is an endogenous effect: firms may modify their behavior in response to the fact that they know they are being monitored by regulators via specific metrics.²⁶ As the British anthropologist Marilyn Strathern has noted, 'when a measure becomes a target, it ceases to be a good measure'.²⁷ Thus, for example, supervising the banking system simply on the basis of a leverage ratio and collecting data only fit for that purpose could encourage banks to increase their risk per unit of assets, thus reducing the usefulness of the leverage ratio over time as a measure of systemic risk.²⁸

Such considerations do not necessarily imply that there are no stable predictors of financial crises. As Claudio Borio argues in text published from his keynote address, 'The Great Financial Crisis: setting priorities for new statistics', a strong case can be made that a joint deviation of the ratio of credit-to-GDP and asset prices from historical trends are leading indicators of financial crises that are robust across history.²⁹ Nor do such considerations imply that regulators should not set priorities when deciding which data to analyse. According to Borio, regulators could greatly benefit simply from having internationally comparable statistical series on property prices, and comprehensive balance sheet and profit-loss information for firms on a globally consolidated basis.

CONCLUSION

In 2007, the amount of digital data created in the world exceeded the creation of new data storage capacity for the first time.³⁰ As this introduction has argued, this development should not be understood in narrowly technological terms. Rather, it bears witness to broader economic and cultural transformations that some commentators have dubbed ‘datafication’: an exponential growth in the recording and subsequent analysis of previously undocumented aspects of everyday life, everything from friendships via social networking platforms like Facebook, to our dining habits codified in the payment data generated from credit card transactions. As Victor Mayer-Schönberger and Kenneth Cukier have argued, these developments may be as epoch-defining as other inflection points in the history of recording data, from the invention of writing to the Gutenberg Revolution.³¹ At a minimum, the collection and analysis of ‘big data’ raises a number of thorny issues about the potential tradeoffs between personal privacy and public transparency; efficiency rewards from data sharing against risks posed to its security; and questions about who should have access to data and for what purposes. In financial services, these issues are acute because what is at stake is the visibility of large amounts of money.

The papers in this volume do not offer easy resolution of these issues. However, they do offer considered reflections on the challenges. If one theme stands out from the papers, it is their implication that financial data should be understood as people in digital disguise.³² Viewed through this lens, financial institutions do not own data. Rather, they are its custodians and should be judged according to how well they execute this trust.

ACKNOWLEDGEMENTS

I am grateful to Mark Robson and Dan Opreescu for supporting the idea for the CCBS seminar; Sarah Pegg, Gill Hammond, Céline Gondat-Larralde, Grant Alexander and Yee Li for help

in its planning; Andy Haldane, Andrew Bailey, Lyndon Nelson, Perttu Korhonen, Ben Dubow, Robleh Ali, Damien Charles and Liz Dixon Smith for their participation; and Dalvinder Singh, Liz Holwell, Amy Shackleton, Maria Baiye and Marjan Al-Dejaily for helping publish this special issue.

REFERENCES AND NOTES

- 1 Comaroff, J. (2011) *Divine Detection: Crime and the Metaphysics of Disorder*. Lecture at the University of Chicago. Chicago, IL, 17 May.
- 2 Quoted in Goodhart, C. (2010) *The Changing Role of Central Banks*. Basel, Switzerland: Bank for International Settlements. BIS Working Paper no. 326.
- 3 Bank of England (2013) *The Financial Policy Committee's Powers to Supplement Capital Requirements: A Draft Policy Statement*. London: Bank of England.
- 4 I thank Keith Saxton for this definition.
- 5 IBM (2012) What is big data? <http://www-01.ibm.com/software/data/bigdata/>, accessed 21 February 2013.
- 6 *BBC News* (2012) EU proposes ‘right to be forgotten’ by internet firms. 23 January, <http://www.bbc.co.uk/news/technology-16677370>, accessed 21 February 2013.
- 7 Hille, Recent Developments in Restructuring the Austrian Bank Reporting System, this volume.
- 8 Capie, F. and Wood, G. (2013) Do we need regulation of bank capital? Some evidence from the UK. IEA Current Controversies no. 40, London, UK.
- 9 I thank Michael Will for bringing this to my attention.
- 10 Harris, I., Mainelli, M. and Onstwedder, J. (2012) *Confidence Accounting: A Proposal*. London, UK: Association of Chartered Accountants, Chartered Institute for Securities & Investments, and Long Finance.
- 11 Gray-Metzing, Defining and delivering judgement-based supervision: The interface with the legal system, this volume.
- 12 Ali, R., Haldane, A. and Nahai-Williamson, P. (2012) Towards a common financial language. Speech to the Securities Industry and Financial Markets Association ‘Building a Global Legal Entity Identifier Framework’ Symposium. New York, 14 May.
- 13 Oliver Wyman (2013) *A Money and Information Business: The State of the Financial Services Industry 2013*. I thank Liz Dixon Smith for this reference.
- 14 Intellect UK, *Biting the Bullet — Why Now is the Time to Rebuild the Foundations of the Financial System*. London, UK: Intellect UK.
- 15 Robert-Nicoud, M. (2012) Securities markets: A European central securities depository perspective. In: B. Summers (ed.) *Payment Systems: Design, Governance and Oversight*. London: Incisive Media, pp. 43–55.
- 16 Basel Committee on Banking Supervision (2012) *Principles for Effective Risk Data Aggregation and Risk Reporting*. Basel, Switzerland: Bank for International Settlements.
- 17 Quoted in Engelen, E., Erturk, L.A., Froud, J. and Williams, K. (2010) Reconceptualizing financial innovation: frame, conjuncture and bricolage. *Economy and Society* . 39(1): 33–63.

- 18 I thank Michael Bridge of the London School of Economics for clarifying the law surrounding re-pawning.
- 19 Dive, M., Hodge, R. and Jones, C. (2011) Developments in the global securities lending market. *Bank of England Quarterly Bulletin* 51(3): 224–233.
- 20 Trichet, J. (2008) A Strategic Vision on Euro Area Statistics: The ECB's view. Speech to 4th ECB Conference on Statistics. Frankfurt, Germany, 24 April.
- 21 Hayek, F. (1991) *The Fatal Conceit: The Errors of Socialism*. Chicago, IL: University of Chicago Press.
- 22 Hayek, F. (2012) Monetary theory and the trade cycle. In: H. Klausinger (ed.) *The Collected Works of F.A. Hayek*. Vol. 7: *Business Cycles Part I*. Chicago, IL: University of Chicago Press, p. 165.
- 23 Bholat, D. and Gray, J. (2012) Organisational Form as a Source of Systemic Risk. Economics Discussion Papers, advance online publication, 24 July, <http://www.economics-journal.org/economics/discussionpapers/2012-35>, accessed 21 February 2013.
- 24 This is because credit ratings reflect credit risk and sovereign guarantees; so a nationalised firm such as Northern Rock necessarily scores high on the latter, bringing up its overall rating. I thank Helen Clark of Fitch Training for this observation.
- 25 I thank Murray Stephens for this observation.
- 26 This is the so-called Hawthorne effect in sociology, also termed 'Goodhart's law' in economics.
- 27 Wikipedia (2013) Goodhart's law. http://en.wikipedia.org/wiki/Goodhart's_law#cite_note-5, accessed 5 March 2013.
- 28 Haldane, A. and Madouros, V. (2012) The dog and the frisbee. Paper presented at The Federal Reserve Bank of Kansas City's 36th Economic Policy Symposium; 31 August, Jackson Hole, USA, London: Bank of England.
- 29 Reinhart, C. and Rogoff, K. (2009) *This Time is Different: Eight Centuries of Financial Folly*. Princeton, NJ: Princeton University Press.
- 30 Manyika, J. et al (2011) *Big Data: The Next Frontier for Innovation, Competition, and Productivity* London: McKinsey Global Institute.
- 31 Mayer-Schönberger, V. and Cukier, K. (2013) *Big Data*. London: John Murray.
- 32 Neather, A. (2013) We must make hard choices in Google's brave new world. *London Evening Standard*, <http://www.standard.co.uk/comment/comment/andrew-neather-we-must-make-hard-choices-in-googles-brave-new-world-8509870.html>.

CORRECTION

The guest editor's corrections to spelling and grammar have been made to the original Advanced Online Publication version of this paper.

David M. Bholat
Statistics and Regulatory Data Division,
Bank of England,
London EC2R 8AH, UK