

Chapter 2

Monitoring Financial Stability in a Complex World

Mark D. Flood, Allan I. Mendelowitz and William Nichols

Abstract We offer a *tour d’horizon* of the data-management issues facing macroprudential supervisors. Traditional financial oversight has been very firm-centric, with strong respect for the boundaries of the firm. Even in this firm-oriented context, financial information has been expanding much faster than traditional technologies can track. As we broaden to a macroprudential perspective, the problem becomes both quantitatively and qualitatively different. Supervisors should prepare for new ways of thinking and larger volumes of data.

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Office of Financial Research or the U.S. Treasury.

M. D. Flood (✉) · W. Nichols

U.S. Department of the Treasury, Office of Financial Research, Cambridge, MA, USA
e-mail: Mark.Flood@treasury.gov

W. Nichols

e-mail: William.Nichols@treasury.gov

A. I. Mendelowitz

Committee to Establish the National Institute of Finance, Washington, DC, USA
e-mail: amendelowitz@yahoo.com

2.1 Introduction

This paper outlines a network approach to monitoring threats to financial stability and some of the strategic data-management challenges that will confront regulators and market participants as they implement the Dodd-Frank Act (DFA).¹ Because of the need to monitor the large and growing data volumes from disparate sources across the financial system, a financial stability supervisor will require specialized techniques for risk measurement and data capture, and expansive capacity for risk analysis.

We identify three strategic forces affecting data management for financial supervisors. First, financial market data volumes are growing exponentially. One should thus expect traditional data-management technologies to fail, and they have. In particular, back offices of trading firms have not kept up with their own front office processes (in terms of volume and complexity of data created), nor with evolving practices in other industries to manage growing data volumes and changes in source types and persistence mechanisms. Second, systemic monitoring requires a new focus on the relationships among firms and markets across the financial system. The most important of these are the contractual relationships created by financial transactions. To assess threats to financial stability one must quantify the bilateral and multilateral relationships—and the chains and networks of relationships—between entities (including investors and issuers). Third, the possibility for diverse contract types to create very similar economic exposures, and the large volume of data needed to monitor the entire system, require a supervisor to build cognitive capacity. All these are especially important in a macroprudential context, where the data may help inform regulatory decisions that affect the whole system. However, our goal in this paper is simply to call attention to the scope of the macroprudential supervisor's data-management issues—issues that are too often ignored—for a nonfinancial audience.

2.2 Legacy Financial Supervision

Before turning to the problems of complexity and supervision at the systemic level, we first consider the issues for data management at a microprudential scale.

¹ The DFA is officially the Dodd-Frank Wall Street Reform and Consumer Protection Act; see U.S. Congress (2010). The OFR provisions of the DFA were based on an earlier bill introduced by Sen. Jack Reed; see U.S. Senate (2010). Among many other things, the DFA created the Financial Stability Oversight Council (FSOC) and Office of Financial Research (OFR) to monitor threats to financial stability in the U.S. The Federal Reserve Board established a new Office of Financial Stability Policy and Research. The Federal Deposit Insurance Corporation (FDIC) established a new Office of Complex Financial Institutions. Similar significant initiatives exist at other central banks, regulatory agencies, and multilateral institutions worldwide.

2.2.1 Firm-Level Supervision and Disintermediation

Traditional accounting still is the data-management framework most widely used for monitoring risks in financial institutions, especially for regulatory purposes. In general, a firm's risk exposures enter through its assets and liabilities, which appear on the balance sheet; a straightforward, well-defined reporting format that has been refined over centuries. There are off-balance-sheet exceptions, of course, such as exposures through unconsolidated subsidiaries or assets under management, but these only reinforce the issues we describe here. Financial standards are quite explicit and intentional in their focus on the reporting entity and the "boundaries of the firm" (see, for example, FASB 2008), as this is the managerial locus of decision-making and control, and comprises the scope of legal obligation. This intense distinction between intraorganization activities versus interorganization transactions has a long history, covered most famously in Coase's (1937) essay on transaction costs, and surveyed more recently by Holmström and Roberts (1998). The boundaries of the firm are also important for identifying and policing anticompetitive behaviour across entities (see, e.g. Gaspar et al. 2006).

Firm-level accounting measures are central to prudential supervision. Banks, broker-dealers, investment companies, and other market participants are all supervised at the level of the corporate entity. Capital requirements apply at the corporate level for both individual entities as well as their parent holding companies. The observation frequency for generally accepted accounting principles (GAAP) reporting is typically quarterly or annual: the state of a nonfinancial firm typically changes only gradually over time as sales are completed and expenses incurred; thus, a quarterly reporting cycle is usually adequate. Indeed, for firms practicing just-in-time (JIT) inventory or manufacturing, quarterly filing schedules are trivial to meet. Most financial regulatory reporting has adopted this same frequency.² In contrast, large financial firms with significant trading operations are able to modify their valuation and risk profiles much more rapidly, and large banks are therefore generally subject to continuous onsite supervision.

The firm-centric conception of risk inherited from accounting also appears in many of the modelling abstractions that are commonplace to applied risk management; value at risk (VaR), economic value of equity (EVE), risk-weighted assets (RWA), and other familiar metrics are good examples. As we argue below, there are important risk-management and data-management implications of an exclusive focus on firm-level exposures. An individual financial firm exists in a volatile marketplace with only limited visibility into the views, preferences, and constraints that guide the behaviour of its counterparties. From the perspective of a manager or regulator it can often be a plausible and convenient simplification to regard the firm as an "island," treating the market values of assets and liabilities

² Workhorse regulatory data collections in the U.S. include the SEC's 10-K (annual) and 10-Q (quarterly) reports, bank Call Reports (quarterly), and Thrift Financial Reports (quarterly). While it is difficult to generalize, reporting abroad tends to be less frequent than in the U.S.

myopically as the outputs of some random process. Assuming some stability in the data-generating process, one can estimate the parameters of the price distribution, calculate confidence intervals, and then use them to set position limits and allocate capital. This is the essence of the VaR methodology. This procedure is well-suited to programmatic risk management within the firm as well as standardized capital regulation from outside, partly because it links firm-level managerial objectives to microprudential regulatory objectives.³ Unfortunately, the firm-as-island conceptualization ignores important system-level phenomena. For example, the financial crisis of 2007–2009 amply demonstrates the importance of systemic effects. Danielsson and Shin (2003) highlight the fallacy of composition inherent in a strictly microprudential supervisory regime by arguing that the whole is not the sum of the parts: firm-level risk management and supervision alone are inadequate.⁴ An important practical manifestation is the so-called “volatility paradox,” (see, e.g., Brunnermeier et al. 2011) whereby an episode of low volatility such as the 2003–2007 period combines low firm-level risk measures (e.g., VaR) with growing systemic risk as aggregate imbalances are accumulated.

Accounting data have other limitations when used as a source of risk information. Traditionally, valuations were recorded at historical cost, with the advantage—one extremely useful for contract enforcement—of being unambiguous. However, historical cost is a backward-looking measure, and therefore a very poor choice for risk accounting. Relatively recent changes to “fair value” standards are more forward-looking, but ironically make GAAP financial statements more difficult to interpret.⁵ There are intricate rules for distinguishing “held to maturity” versus “trading” or “available for sale” securities. The former are recorded at amortized cost, while the latter are typically recorded at fair value. This has the potential to mislead by confounding two measurement frameworks in a single balance sheet. Because up-to-date market prices are not always available, determination of fair value introduces significant discretion and ambiguity into the measured valuations. For example, Benston (2006) recounts the use and abuse of discretionary “Level 3” mark-to-model fair valuations by Enron, and their contributions to its demise.

For a static regulatory regime to provide consistent system-level supervision over time implicitly requires some stability in the institutional structure of the regulated sector. However, as emphasized by Williamson (1991), the locus of

³ Blundell-Wignall and Atkinson (2010) describe many of the issues that arise in this approach. Adrian and Shin (2010) discuss ways in which myopic VaR implementations can exacerbate boom-bust leverage cycles in financial markets.

⁴ The literature on systemic risk measurement is large and growing. Bisias et al. (2011), IMF (2009), and ECB (2010b) provide overviews.

⁵ The final rule on fair value measurement was adopted by the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) in 2011; see IASB (2011). This harmonizes the FASB and IASB approaches, and replaces earlier, very similar guidance under FASB Topic 820, formerly the Statement of Financial Accounting Standards (SFAS) 157.

activity within organizations compared to that between organizations has evolved endogenously over time in response to transaction costs. Kaufman and Mote (1994) or Boyd and Gertler (1994) updated by Feldman and Lueck (2007) document that financial activity has systematically disintermediated away from traditional banking institutions over many decades. The fact that many of the new markets are relatively lightly regulated compared to banks has been a significant enticement to disintermediate. While securities regulators have typically focused on disclosure and transparency over the more intrusive regulation faced by banks, many firms in the so-called “shadow” banking system, such as hedge funds, now operate with little scrutiny at all.⁶ At the same time, the shadow banking system maintains close ties and interactions with traditional intermediaries, so their activities cannot be isolated.

2.2.2 Financial Innovation and the Complexity of Data Management

The trend toward disintermediation has also been facilitated by the opportunities created by financial innovation: especially noteworthy are the enormous growth in derivatives markets since the late 1970s, the expansion of trading systems and securitization markets since the late 1980s, and advances in the modelling and management of portfolio credit risk since the late 1990s.⁷ Innovating firms typically view new contract types favourably. Because they face limited competition, innovative contracts earn larger economic rents for the seller, typically in the form of higher spreads. Some securities conceal embedded short puts or other contingent losses in order to entice investors to overpay. For example, Ingersoll et al. (2007) document the ways in which contingent exposures can be used to manipulate standard investment performance metrics published to investors. In banking firms, a similar problem emerges in the “opaqueness” of assets. For example, Flannery et al. (2004, 2010) show that, prior to the crisis, there was little evidence from equity trading characteristics that investors in large banks were significantly deprived of information about the credit quality of bank assets. During the crisis, however, many of these same large institutions were cut off from funding altogether due to uncertainties about their solvency, suggesting that they held large contingent exposures that came into the money in the context of the crisis.

⁶ Pozsar et al. describe the shadow banking sector in greater detail. Lo (2011b) asserts that outsiders know almost nothing about the nature and magnitude of the risk exposures of the hedge fund industry, and are forced to resort instead to plausible “conjectures.” Based on data from Institutional Investor, he emphasizes that the size of the now defunct LTCM is an order of magnitude smaller (in 1998 dollars) than a number of current hedge funds.

⁷ The general literature on technological innovation is largely beyond our scope. See Antonelli (2009) for a general overview. See Tufano (2003) for an economist’s overview of the literature on financial innovation.

While innovation is widely encouraged on trading desks, the ramifications for data management within the innovating firm are typically neglected. Anecdotally, sell-side firms are reported to systematically underinvest in the back-office infrastructure needed to support their front-office innovations properly. Gottfredson and Aspinall (2005) demonstrate that this pathology is not limited to the financial sector. Incentive schemes, such as the alternative investment management industry's standard "2-and-20" rule, that reward increases in gross revenues incentivize innovations that boost measured performance. Gottfredson and Aspinall (2005) argue that it is commonplace for firms of all types to fail to account for the managerial and operational complexity implied by their product innovations, resulting in a phenomenon of excessive complexity and over-innovation. In short, the introduction of new products necessitates the costly development of specialized data-management infrastructure to track transactions. Failing that, the burden of data integrity falls upon back-office personnel, resulting in an inevitable incidence of operational errors. Since the operational sluggishness engendered by an innovation tends to also affect existing product lines as well, many of the costs are inframarginal. Most of these complexity costs also fall on the back office. Notably, Gottfredson and Aspinall (2005) propose the count of a firm's distinct SKUs (stock-keeping units) as a basic operational complexity metric. Unlike most manufacturing and retail sectors, there is as of yet no comprehensive, shared SKU system—i.e., a globally standard set of instrument type identifiers—in finance.

Securitization innovations have helped to supplant traditional portfolio lending with an originate-to-distribute business model, fundamentally altering the lender's production of data and information in subtle and not-so-subtle ways. So-called "soft" information about creditworthiness, derived in part from a loan officer's subjective experience of working directly with the borrower, is discarded when the loan is sold into a securitization. Instead, all information on loans intended for securitization is reduced to a set of "hard" information defined by the inputs for an automated underwriting calculator.⁸ A series of recent papers explores how lending, including the types of credits underwritten, differs systematically between small and large banks as a result of distillation of the underwriting information to a set of strictly hard criteria by larger institutions.⁹ At one extreme, mortgage securitization underwriters submitted predefined pool characteristics to mortgage bundling operations which then accumulated newly originated loans to fulfill the

⁸ The terms hard and soft are close to their usage in the knowledge management (KM) literature, and related to the notions of tacit and explicit knowledge (see Hildreth and Kimble 2002). Very roughly, knowledge is soft (hard) if it cannot (can) be fully and accurately written down. Examples of soft knowledge are technical skills or the capacity for rapid cognition; examples of hard knowledge include loan amounts or monthly income. Knowledge is tacit (explicit) if it is not (is) written down. In practice, knowledge typically exists in a continuum or mixture of hard and soft, tacit and explicit. See Choo (2006) for a survey of the field of KM in organizations.

⁹ See, for example, Berger et al. (2005), Agarwal and Hauswald (2010), Liberti and Mian (2012), and Petersen and Rajan (2002). In the mortgage industry, the two most common automated underwriting systems are Freddie Mac's Loan Prospector and Fannie Mae's Desktop Underwriter.

specifications. In such cases, information about loans other than the predefined attributes was not collected. In short, information loss along the supply chain is a function of provenance and lineage practices; the fact that there are no requirements to guard against information loss means that no firm willingly incurs the costs to maintain this information. Indeed, under pressure to fill the pools, tolerance for occasional missing values expanded naturally via “ignorance creep” to create a large specialized market for low-doc and no-doc loan pools.¹⁰ Loan officer incentives matter in this context: Dell’Ariccia et al. (2008), for example, show that mortgage lending standards declined more in regions where loans were most likely to be sold into securitization.

Paradoxically, the rise of the originate-to-distribute model has increased the overall volume of data in the financial system by fracturing the traditional origination process into a sequence of highly specialized transactions. Many parties along the securitization pipeline have an active interest in the performance of each loan. As financial products such as mortgages have been systematically securitized—with those securitizations then structured and repackaged—loan details that once might have been recorded only by the borrower and the originating lender (holding the whole loan in portfolio), are shared with the borrower, originating bank (for recourse commitments), loan servicer, securitization trust, securitization bondholders, and the buyers and sellers of credit protection derivatives. On the one hand, assuming that individual participants have enough information in one place to support reconciliation, the data reconciliation process necessitated by interfirm contracting should improve data quality by focusing more sets of eyes on the same set of facts. On the other, having so many consumers of the information multiplies the validation burden more than proportionally.¹¹

2.2.3 Scalability of Data Management

The steady expansion of new market segments has moved significant portions of the financial system to the fringes of regular supervision. We lack accurate, up-to-date

¹⁰ For example, Cordell et al. (2011, p. 25, emphasis in the original), citing IOSCO (2008, p. 2), note that, “clearly data quality was a problem, fueled as it was by declining underwriting standards. One very valid point on the data is that the quality of the data being provided deteriorated significantly in the buildup to the crisis because of declining underwriting standards, by the IOSCO’s reckoning, ‘beginning in late 2004 and extending into early 2007.’”

¹¹ See Flood (2009) on the costs of data reconciliation across multiple schemas and systems. The discussion here implicitly assumes that the information collected at origination comprises a stable and well-defined set of attributes. Because relational databases are costly to change in production systems, back-office practices typically require static data models describing instruments that traded repeatedly. Front-office dealers, on the other hand, frequently prefer customized deal structures built from one-off or semi-standardized components with idiosyncratic data representations. This can overwhelm back-office data modelling.

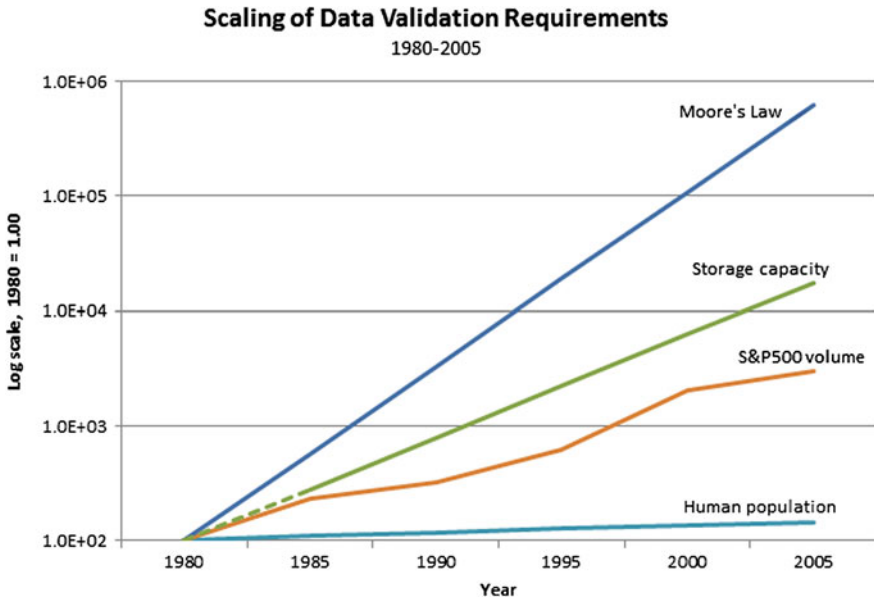


Fig. 2.1 Differential exponential growth rates for data validation requirements. Moore's Law is estimated as a linear regression of transistor densities (in logarithms) against year of introduction over the 1971–2011 period; data were downloaded from Wikipedia (2011). Storage capacity is based on the average annual growth estimate (23% per year) of Hilbert and López (2011a) for the 1986–2007 period, extrapolated back to cover the 1980–1985 interval. S&P500 trading volume was downloaded from Yahoo Finance (2011). Human population is based on total midyear world population, from the U.S. Census Bureau (2011)

estimates of the total size of many over-the-counter markets.¹² As a result, we know surprisingly little about the simple scale of certain market segments, which limits our understanding of the overall data-management problem in the financial system. Because of technological innovation, the problem is growing in size. Individual innovations tend to be disruptive events, but they also accumulate over longer periods into a smoother high-level growth trajectory. A look at the basic orders of magnitude is helpful (see Fig. 2.1). Similar to Moore's law for transistor densities, data volumes (proxied by aggregate digital storage space or Internet throughput) have grown globally at an exponential rate.¹³ Hilbert and López (2011a, b), for

¹² Pozsar et al. (2010) provide estimates of the size of the shadow banking markets. On the other hand, other segments, such as the hedge fund industry are much murkier; see Lo (2011b).

¹³ The numbers provided here are intended to be suggestive of the situation in financial markets, rather than conclusive. The growth in processing power represented by Moore's Law is particularly relevant as a benchmark for the growth in storage requirements in finance, since advances in processor power help enable the development of new market segments. Valuation of structured securitizations, for example, makes frequent use of CPU-intensive Monte Carlo analyses; see, for example, Berthold et al. (2011). Similarly, while high-frequency trading is

example, estimate 1986–2007 average annual growth rates for storage capacity (23 %/year). Koh and Magee (2006) first estimated the long-range exponential growth rates in data storage and transport-layer bandwidth, with Nagy et al. (2011) later reprising the study, and arguing that growth had in fact been super-exponential. Financial activity is data- and information-intensive, and exemplifies this type of rapid growth. Data validity is critical for financial activity; this is not true of much of the generic traffic on the Internet. At 25 frames per second, a downloaded video (even heavily compressed) incorporates a great deal of signal redundancy. A few corrupted bits or even the loss of an entire frame would seldom be noticed, let alone provoke misinterpretation. Signal redundancy is much less common in a financial context, in part because contractual ambiguity is potentially very costly. As a result, corrupting a few digits in a transaction confirmation or payment instruction could easily be cause for significant concern. Flipping a single bit might mean the difference between paying and receiving millions of dollars: Nick Leeson’s billion-dollar rogue trading loss in the Barings Bank scandal began with an innocent clerical error of this sort (see Bookstaber 2007, pp. 38–39).

Figure 2.1 suggests the nature of the problem. Starting with double-entry book-keeping, participants deploy a range of technologies and processes to scrub and validate their data. Traditionally, these techniques have relied heavily on human diligence and attentiveness. Even in processes involving a large degree of automation, a “human in the loop” will typically be a binding constraint, so we use global population as a rough proxy measure of aggregate capacity for processes that depend significantly on manual input. Population has more than doubled in the last half-century, while stock market trading volume has increased almost a thousand fold.¹⁴

(Footnote 13 continued)

typically latency-dependent, it nonetheless benefits from high-performance processing power; see, for example, Intel (2010).

¹⁴ For at least two reasons, the S&P 500 trading volume depicted here represents a lower bound on the growth in data generated by the financial system. First, it does not encompass the vast increase in derivative markets that has occurred since 1980. Comprehensive data on outstanding balances (not trading volumes) for OTC derivatives are available only since 1998; see BIS (2010). These have shown a roughly order-of-magnitude increase over the past decade, with approximately \$600 trillion notional outstanding in June 2010 (ca. \$25 trillion in market value), dominated by interest-rate swaps. The growth in trading is also reflected in and compounded by the growing “financialization” of the economy: the share of GDP represented by the U.S. financial sector (including insurance) has tripled since World War II, and nearly doubled since 1980 (see Philippon 2008, Fig. 1, p. 36). Second, each transaction generates a number of internal and external versions of the trade information for financial reporting, regulatory compliance, risk management, etc. These ancillary data sets should all be kept consistent, but the number of reconciliations required does not typically scale linearly with the number of positions or transactions (see Flood 2009). Note that time scales in financial markets have also been shrinking, evidenced by the growth of algorithmic trading; see Castura et al. (2010) or Hendershott et al. (2011). Because more must happen faster, the consequences of process failure are correspondingly larger.

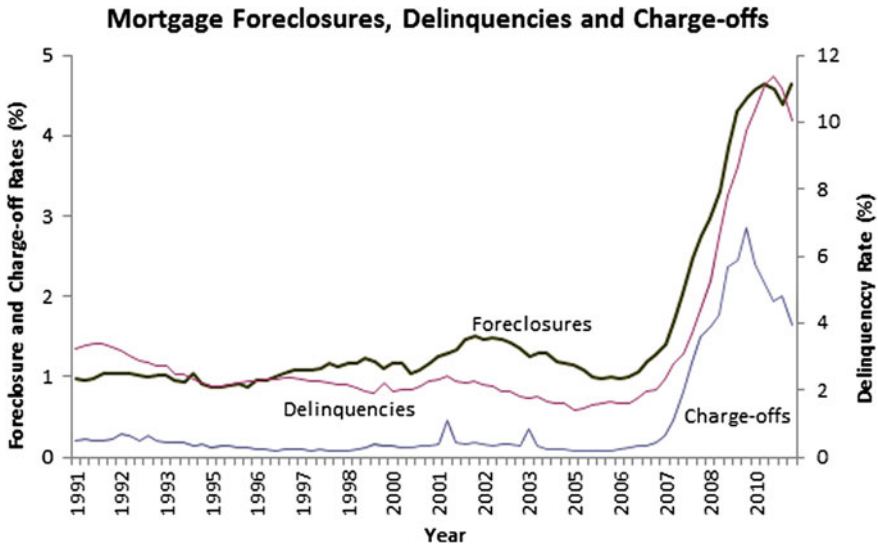


Fig. 2.2 Overwhelming the foreclosure-processing infrastructure. Delinquencies represent balances on delinquent loans for single-family residential mortgages held by all U.S. commercial banks as a percent of all such loans; data are taken from Federal Reserve Bank of St. Louis (2011a), series DRSFRMACBS. Charge-offs represent balances on charged-off loans for single family residential mortgages held by all U.S. commercial banks, as a percent of all such loans; data are taken from Federal Reserve Bank of St. Louis (2011b), series CORSFRMACBS. Foreclosures represent total foreclosed loans as a percent of all mortgage loans, and are taken from Bloomberg, series DLQTFOR Index (PX_LAST)

In turn, the trend in trading volume is broadly consistent with Hilbert and López's (2011b) estimates of the growth in aggregate storage volumes.¹⁵

These measures are also consistent with recent evidence that the trade settlement process is increasingly staggering under the activity load. Exception management accounts for a large fraction of the total cost of trade processing. For example, Bradley et al. (2011) (Fig. 2.2) note that overall settlement fails have generally been increasing since at least 1996. The failure rate series is volatile, with occasional severe spikes. Trimboth (2008) finds that, prior to the financial crisis, settlement failures in U.S. bond markets rose over the last decade, with the trend interrupted by regulatory and market actions. In some cases, back-office behaviour has been chastened by losses and invigorated regulation in the wake of the 2007–2009 crisis, rendering pre-crisis evidence obsolete or suspect. For

¹⁵ Extrapolating from their 23 % approximate annual growth rate over the 1986–2007 period—and assuming it applies at least equally to the financial services sector—we see that data storage requirements are on the order of 10,000 times greater in 2005 compared to 1980. For comparison, they estimate annual growth rate for worldwide computing capacity at 58 %/year, and telecommunications traffic at 28 %/year. At the same time, advances in processing power are also creating engineering challenges as applications impose heavier demands on legacy database technologies; see, for example, Stonebraker et al. (2007) and Pandis et al. (2010).

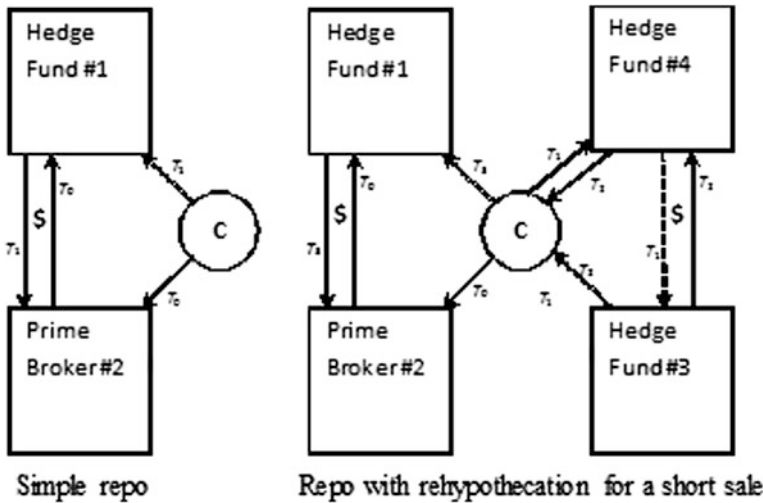


Fig. 2.3 The impact of rehypothecation on interconnectedness

example, Bradley et al. (2011) (Fig. 2.3) show that settlement fails in the market for U.S. Treasuries dropped sharply after imposition of a penalty by the Federal Reserve for fails in this market in May 2009. However, the same chart indicates that unpenalized fails in the mortgage-backed securities (MBS) market have continued to grow steadily over the same time period.¹⁶ To be effective, regulation must be applied and enforced; it does not occur automatically.

The practical implications of pushing too much data through a capacity-constrained process can be disastrous. For example, mortgage foreclosure rates have skyrocketed since the collapse of the market in 2007–2009. Figure 2.2 shows foreclosure, delinquency, and charge-off rates for residential mortgage loans jumping abruptly above historical precedent starting in 2007. While the delinquency rate roughly quintupled during this episode, the charge-off rate at the peak was roughly 20 times higher than its 1990s average. Foreclosure of mortgage loans has historically been handled on a case-by-case basis, with much manual processing.¹⁷ A natural consequence of an unanticipated increase in the foreclosure

¹⁶ TMPG (2011) offers main mechanisms for settlement fails: miscommunication, operational problems (e.g., the September 11, 2001 disruption), “daisy chain” fails in which failure to receive collateral on one deal leads to failure to deliver on another (this is an example of “tight coupling” as described by Perrow 1999 and Bookstaber 2007), and “strategic” fails in which the “short” counterparty intentionally reneges, typically because the cost of borrowing securities to fulfill his commitment approaches or exceeds the time-value opportunity cost of postponing delivery. Strategic fails are thus exacerbated by episodes of low interest rates.

¹⁷ Note that this is the delinquency rate for mortgages overall, including both prime and subprime loans. The delinquency rate for subprime loans in isolation was much worse, peaking at over 15.5 % in the final quarter of 2009. Prime mortgage borrowers are easy for mortgage servicers to handle: until the crisis, defaults and foreclosures were rare, and loans typically had

processing throughput was an acceleration of the legacy (largely manual) processes to accommodate the new volume. One of the practical manifestations of this has been “robo-signing” of foreclosure documents.¹⁸ As Kaufman et al. (2010), Holland (2011), Wallace (2011) and Hunt et al. (2011) all make clear, this is not an isolated problem, but emblematic of a range of institutionalized and partially manual processes throughout the mortgage industry.¹⁹ As other parts of the securitization plumbing have increased their throughput, the narrower pipes are often overwhelmed, provoking process failures. The de facto inability to perform proper diligence at this scale results in a dilemma between costly type I (excessive foreclosure) and type II (excessive forbearance) errors. In principle, the information for accurate decisions is available, but the processing power is not.

Poor incentives and externalities also plague data management. The cost of remediating backlogs and building new processes are borne directly by the firms involved, while many of the risks involved are by nature systemic, and therefore not internalized. This creates a natural role for a supervisory intervention. For example, in 2005 the largest New York dealers in the market for credit default swaps (CDSs) were admonished by regulators for their enormous paperwork backlog, and then agreed to clean it up (see Senior Managements 2005). As with the above mentioned settlement failures in the Treasuries market, this was less a question of inadequate technology, and more a question of industry leadership and new regulatory incentives.²⁰ More famously, increases in stock market trading

(Footnote 17 continued)

very standard structures. As a result, the mortgage servicing business became concentrated in a handful of specialized banks that invested in the relevant information technology infrastructure. In contrast, subprime mortgages employed a variety of innovative terms ostensibly intended to constrain the monthly mortgage payment to a level expected to be sustainable for the borrower. In addition to a more complex servicing process, subprime loans exhibit very different default rates. In hindsight, it is apparent that both underwriting standards and credit pricing were too lax for an enormous number of subprime mortgages, especially those originated after 2005. Dungey (2007a) provides a good introduction to the mechanics of the mortgage servicing process. Dungey (2007b) is a similar overview of the foreclosure process on the eve of the crisis.

¹⁸ Robo-signing is the practice of attaching signatures to affidavits and other foreclosure documents so quickly that it is inconceivable that a reasonable review occurred. This is a data-validation issue on two levels: first, the signature is an attestation, based on (supposedly) diligent human review, of the validity of the information in the affidavit. Second, because it seems in many cases that the task was delegated to unauthorized and unqualified shills as an expedient, the signatures themselves become data requiring subsequent validation.

¹⁹ Noll (2009) chronicles a much earlier instance of robo-signing, perhaps the first in history. As the U.S. Treasury systematically shrank the denominations of the new greenback currency it was issuing to finance the Union efforts in the Civil War, the number of signatures required quickly outstripped the capacity of Treasury staff, even after signature authority was broadly delegated. By 1863, the Treasury building was home to industrial printing operations, including engraved signatures.

²⁰ Fleming and Garbade (2005) provide a contemporary analysis of settlement fails in the Treasuries market. The Counterparty Risk Management Policy Group Report (CRMPG 2005), a statement by participants of industry best practices, was a catalyst for change at the time. When the operational costs are small and/or not internalized, unilateral remediation is difficult to justify.

volume overwhelmed the industry's back-office capacity during the Paperwork Crisis of 1968–1970. Settlement fails and “DK” (don’t know) transactions proliferated. The New York Stock Exchange (NYSE) began closing early to help clear the backlog, and the markets closed entirely on Wednesdays for a period of several weeks in July–August 1968. The episode culminated in the Securities Investor Protection Act of 1970, which created the Securities Investor Protection Corporation (SIPC).²¹

Because manual processes are the most vulnerable to input surges, they are gradually being replaced by straight-through-processing (STP) architectures.²² Figure 2.1 depicts steady exponential growth in data throughput volumes. Few processes—even automated ones—scale well when they are pushed orders of magnitude beyond their designed capacity. In this context, the transition to STP simply moves the automation boundary. That is, after the shift to automated application of data-validation rules (i.e., the shift to STP) has extracted its efficiency gains, additional efficiency will again be demanded. This requirement will perhaps be satisfied by techniques for the automated generation of the data-validation rules themselves, or thereafter by the automated generation of domain-specific languages for specifying data-validation rules.²³ However, because risk is a central concern of the supervisory process, seemingly straightforward outlier-detection rules that are useful for low-intelligence bulk validation in other domains are likely inappropriate in this context; see, for example, Ghoting et al. (2008) or Han et al. (2011). For risk applications, the most interesting facts are very often ones that appear as “outliers” in the data.

²¹ See Markham (2002, pp. 362–367) and SEC (1972, pp. 3–6).

²² On the need for straight-through processing, see CPSS-IOSCO (2001) and CPSS (2008). For an example of an implementation perspective, see Ciulla et al. (2010). CPSS (2011) identifies five main categories of financial market infrastructure, each of which encompasses a multitude of processes, and each of which might benefit from STP:

- payments systems
- securities and other settlement systems (SSSs)
- central securities depositories (CSDs)
- central counterparties (CCPs)
- trade data repositories (TRs)

Ironically, technological advances may also encourage novel practices—such as transacting via text messages from wireless devices—that place further demands on data management and validation.

²³ Madnick and Zhu (2006) offer some concrete examples of the role of semantic context in defining the quality of a data set, as well as suggestions for effective management of that semantic context to improve data quality. Fuerber et al. (2011) indicate a similar path forward, defining a data quality constraints language targeted at the Semantic Web.

2.3 Systemic Supervision and the Network of Counterparty Claims

Information management is more challenging still at a systemic level. In addition to a proliferation of institutions and institution types, there is a separate set of interfirm considerations that only apply to macroprudential supervision.

2.3.1 *Networks and Information*

While the implementation of the DFA is fundamentally redefining the supervisory process, some implications for systemic risk monitoring are already becoming clear. First, because it is systemic, data-validation challenges are likely to be severe. The broad scope—all financial sectors and institutions—implies very large data volumes. Systemic supervision also implies more kinds of data (e.g., accounting, macroeconomic, contractual terms and conditions, etc.) from multiple markets sectors (e.g., equities, over-the-counter (OTC) derivatives, commercial loans, etc.). Tools and techniques for reducing the data-reporting burden and streamlining the data-validation process will be especially welcome.

Moreover, the notion of risk changes as the context broadens to the systemic level. As described above, it is commonplace for both firm-level risk managers and microprudential supervisors to regard the firm (expediently) as an island buffeted by unpredictable random shocks. Individual firms typically try to evaluate their immediate counterparties, but cannot peer more deeply into the network beyond that. Portfolio positions and risk exposures are closely held business secrets. Andrew Lo highlights this problem of myopia in his own attempts to understand the behaviour of hedge funds²⁴:

... you know for a fact that there are people out there that know what actually happened, but they're not talking. So, in fact, this entire paper could be science fiction or it could be dead on, we have no idea. To this day, we don't know, because nobody's talking. They're not allowed to talk, because that would disadvantage their shareholders.

In contrast, a view of the full financial network provides additional conditioning information relative to what is available to the individual firm. Price events that appear to the myopic participant to be deep in the tail of the unconditional distribution—so-called “black swans”—might be much less surprising with knowledge of the connections and imbalances observable in the web of counterparty claims. Macroprudential supervision could well focus on the network of contractual relationships. This is conditioning information with high marginal value.

As noted above, disintermediation has been one important influence in the overall growing significance of macroprudential factors—especially the network

²⁴ See Lo (2011a, at 13: 18). The study he refers to is Khandani and Lo (2011).

of claims—with important implications for information and data management. Securitization in particular moves lending away from the originating lender, a relationship which traditionally maintained extensive hard and soft information about the borrower, distributing responsibility for it across a variety of specialized agents, including the loan servicers, bond investors, CDS protection sellers, etc. To support this web of interests and relationships, data about the loans is compartmentalized and replicated across a range of participants. The issues are particularly acute for tranching or structured products, such as collateralized debt obligations (CDOs). Judge (2011) refers to this process as fragmentation, and coins the term “fragmentation node” to describe a counterparty where cash flows and associated data are parcelled into subsets and repackaged for sharing with others in the network. As discussed above, such data fragmentation is a “lossy” conversion, in the sense that most of the soft information from the origination process is lost as a loan enters a securitization. In other words, the pre-fragmentation information set is typically greater than the sum of the post-fragmentation parts. Securitization distills it all down to a narrow subset of hard information, with the responsibility for collecting and maintaining the information distributed across a range of participants.

Moreover, fragmentation per se is an obstacle to the comparison and aggregation of information. Cordell et al. (2011), for example, compare subprime MBSs to “structured finance asset-backed securities collateralized debt obligations (SF ABS CDOs)” based on the same subprime MBSs. Ordinary MBSs have a relatively simple senior/subordinated structure, while CDOs—because they typically combine multiple tranches from each of many MBSs—have a much more intricate subordination scheme mapping the individual loans through the MBSs to the particular tranches of the higher-level CDO structure. After examining write-downs on the universe of publicly traded ABS/MBS securities and SF ABS CDOs issued between 1999 and 2007, Cordell et al. (2011, p. 24) highlight an extraordinary difference between subprime MBS and the more structured ABS CDOs: “only 4 % of the AAA-rated subprime MBS securities issued from 2005 to 2007 were impaired or are expected to become impaired. By our calculations, 98 % of the AAA-rated SF ABS CDOs issued between 2005 and 2007 will suffer write-downs.” In order to accept the AAA rating on these SF ABS CDOs, investors and rating agencies required either highly implausible assumptions for loss experience, or, more likely, failed to perform the analysis at all.

In some cases, contractual complexity can render diligence impossible. It is easier to create certain pricing problems—for example, constructing an intricately structured derivative security—than to solve those problems. A recent paper by Arora et al. (2011) illustrates these difficulties.²⁵ The standard argument, presented by DeMarzo (2005), is that issuers can reliably signal the quality of newly issued security by taking a first-loss position (junior tranche). In contrast, Arora et al.

²⁵ Flood et al. (2010) also discuss some of the implications of financial complexity for information management.

(2011) show how a CDO issuer can “boobytrap” a subset of its CDOs by strategically hiding the worst-performing assets in them. This creates a natural information asymmetry, in which the creator of the contract inherently knows more about its value than prospective buyers. In extreme cases, it is literally impossible, due to computational bounds, for the seller to prove that the offering price is reasonable, and likewise impossible for the buyer to verify the seller’s claims. Because information asymmetries in financial markets are typically profit opportunities, complex securities tend to arise endogenously; they are not accidents of nature. While the boobytrap example demonstrates the impossibility of full diligence in certain cases, it also suggests that issuers strategically deceive investors, implying recurring investor na and begging the question of why, in a repeated game, deceptive issuers are not ultimately driven from the market. Even if the impossibility of diligence defeats the usefulness of signalling via a first-loss position, reputation should discourage manipulative behaviour (see, for example, Hartman-Glaser 2011). However, the dynamics of selection do not require intentional deception in order for the market to prefer complex securities: any product for which diligence and reasoning are imperfect, and for which errors in analysis tend to favour the issuer—the so-called “winner’s curse” (see Thaler 1988)—will have an “evolutionary advantage.”

At the network level, the web of claims helps to obfuscate because important system-level patterns are not visible to individual, myopic participants. Indeed, this is an important justification for government supervision of the system. Moreover, shocks can propagate in surprising ways. For example, Bookstaber (2007) offers the example of the LTCM failure in 1998, in which Russia’s sovereign bond default ricocheted through the network of claims and hit the market for Danish mortgage bonds. The latter had no immediate connection to Russian debt, but simply happened to be among the more liquid assets in large portfolios that were otherwise exposed to Russia. Although this connection is surprising—certainly it was for LTCM—in principle, such indirect linkages may be foreseeable, since portfolio holdings are a matter of fact, while the behaviour of portfolio managers in a panic is likely more tightly constrained and predictable than otherwise.

2.3.2 An Example: Rehypothecation of Repo Collateral

We offer the example of rehypothecation of repo collateral to illustrate the importance of monitoring the threat of investor myopia to financial stability amid the network of contractual relationships. Rapid deleveraging in the repo markets was an important crisis propagation channel in the wake of the Lehman Brothers failure in the fall of 2008. As discussed below, feedback and contagion among leveraged institutions can produce nonlinear responses to exogenous shocks at the system level.

A “repo” is a sale of securities (i.e., collateral) combined with a simultaneous commitment to repurchase them at a later date, usually in the near term.²⁶ A relatively simple example is a hedge fund that wants the risk and return profile of a particular security (e.g., corporate bonds) for its portfolio, but wants to boost returns by leveraging its capital. In this example, the hedge fund buys the bonds on the open market and immediately sells them into a repo transaction with its prime broker.²⁷ The hedge fund gets the desired bonds for its portfolio, but is effectively using borrowed money to pay for them. Of course, the hedge fund does not receive the full value of the bonds in the front leg of the repo; a haircut is assessed to protect the prime broker against fluctuations in the value of the collateral. The net effect is one of leveraging, as the hedge fund can use the cash proceeds from the repo sale to purchase additional bonds. It is common for the prime broker in a repo transaction to take absolute title to the collateral. This facilitates the sale of collateral by the prime broker in the event the collateral pledger fails to repurchase it as promised at the maturity of the repo.

However, depending on the jurisdiction and the details of the prime brokerage agreement, the collateral pledgee will have a “right to use” the collateral.²⁸ Among other things, a prime broker with a right to use may rehypothecate (re-lend) the pledger’s collateral to third parties for other purposes. For example, another hedge fund might pay to borrow the collateral to use in a short sale transaction. Gorton and Metrick (2009, p. 8) note that collateral is a scarce resource in securitization markets, so that there are strong incentives to leverage it through rehypothecation. Deryugina (2009, p. 257) observes that both the pledger and pledgee can benefit from the additional revenues generated by this reuse.

These relationships are depicted in Fig. 2.3, which shows both a simple repo transaction on the left and a repo involving rehypothecated collateral on the right. Note that rehypothecation has the effect of connecting two subgraphs, which significantly complicates the topology in the counterparty network graph.²⁹ We emphasize that rehypothecation occurs invisibly to the original pledger of collateral (“Hedge Fund #1” in the figure); although pledgers are aware that rehypothecation goes on, they do not in general observe when their own collateral is rehypothecated or to whom. This lack of transparency about the network of

²⁶ Taub (2008), IMF (2001), and Copeland et al. (2010) describe the mechanics of the repo markets in greater detail. The repo markets are very large, and there are naturally numerous variations.

²⁷ A prime broker is a specialized firm that provides a range of related services to hedge funds and other investment managers. Typical services include custody, securities settlement, tax accounting, and account-level reporting. Lehman Brothers acted as prime broker for a number of large hedge funds at the time of its demise. In the example here, the hedge fund is the “collateral pledger” and the prime broker is the “collateral pledgee.”

²⁸ Deryugina (2009) describes the structure of rehypothecation transactions and related legal considerations in detail. She emphasizes the importance of the relatively lenient U.K. rules on rehypothecation in attracting prime brokerage business to London.

²⁹ Pozsar and Singh (2011) further explore the complexities introduced by rehypothecation of collateral.

relationships played an important role in the recent crisis. Deryugina (2009, pp. 274–275) notes that when Lehman Brothers International Europe (LBIE) failed in London in September 2008, it had rehypothecated or commingled over \$20 billion worth of client collateral, much of which LBIE could not identify immediately. Most of those pledgers became general creditors in the subsequent bankruptcy; Deryugina (2009, pp. 274–75, note 111) quotes from the court's response to pledgers' petition for information about the whereabouts of their collateral:

[I]t would be necessary to investigate particular records held by LBIE and to obtain data and records from relevant third party custodians, depositaries and other parties.... [T]he difficulties that this process faces, not least the refusal of a number of custodians and others to comply with demands for information and that, in the meantime, the administrators are only able to call upon limited LBIE resources.

The flip side of the financial boom sustained by increasing leverage of collateral is the self-reinforcing deleveraging cycle that ensues when the value of the collateral is called into question. In such a cycle, redemption of collateral at fire-sale prices depresses the value of remaining collateral, forcing additional margin calls and subsequent redemptions. Gorton and Metrick (2009) and Singh and Aitken (2010) describe this process in detail in the context of the Lehman failure; if it were simply a question of investor disclosure, an obvious fix would be to impose tighter restrictions on pledgees' ability to reuse collateral without explicit permission of the pledgers.

However, this would not remove the incentives for pledgees to reuse scarce collateral. Their gains from leveraging collateral are internalized, but the risks of a contagious deleveraging are externalized, suggesting a possible role for prudential supervision. Because of the intrinsic myopia of individual participants, supervisory visibility into the full network of relationships is especially valuable. Kashyap et al. (2011) survey the economic literature on fire-sale contagion during the crisis, and argue that the fire-sale problem fits naturally into the broader framework of macroprudential policy.³⁰ They draw a straightforward but powerful conclusion from a sketch of a simple three-sector (households, financial institutions, and a central bank) model of the economy. Just as an airplane pilot has three sets of control surfaces to manage roll, pitch, and yaw, a regulator charged with managing defaults, credit crunches, and fire-sale contagion in financial markets requires three policy tools to do the job effectively. Capital requirements and liquidity requirements are two such instruments (supplemented with backstop capital and liquidity facilities during the emergency phase of the crisis). Evidence is strong that fire-sale contagion is a third significant threat, and minimum collateral margin (or "haircut") requirements are a plausible tool to address it. From a data-management perspective, tools such as regulatory haircut requirements demand that policy-makers be able to observe and measure emerging patterns amid the contractual network. From an accounting perspective, this will mean tracking financial

³⁰ Shleifer and Vishny (2011) survey the issues surrounding fire sales and contagion.

relationships as objects each with its own explicit identity in the system, rather than simply as attributes that describe the legal entities. In other words, a graph consists of both nodes and edges, and both are important.³¹

2.3.3 *Implications for Supervisory Implementation*

The above paints a daunting picture of the data requirements facing macroprudential supervisors; there are at least three major technical challenges. First, there is the exponential growth in data volumes. Second, there is the need to monitor financial relationships, especially contractual relationships and ownership hierarchies. Collecting contractual terms and conditions is a prerequisite for forward-looking cash flow and risk analysis; terms and conditions are not systematically collected by supervisors today. Contracts are also a key ingredient in mapping the network of contractual relationships for systemic modelling. Measuring the edges—i.e., financial contracts—in the counterparty network graph will require the capture of much more detail about those contracts than is possible with traditional firm-centric accounting systems. Supervisors need to know who is connected to whom. As a first step, this requires a reliable system of legal entity identifiers (LEIs) to unambiguously identify the parties to any contract; see OFR (2011). Third, there is the issue of complexity, which can occur both at the level of the individual contract as well as in the network of contractual relationships. We propose that intelligently collecting contract-level terms and conditions can balance these challenges.

Financial contracts have several characteristics that make them desirable digital records of the financial system. First, by definition, contracts connect the individual entities in the system, creating something beyond a simple aggregation of its constituent parts. In particular, the potential for feedback effects and spillovers explains the inadequacy of strictly microprudential (i.e., firm-centric) supervision.³² Second, there are strong incentives to make the contracts valid, complete, and unambiguous statements of the promises and commitments being made. Parties to the contract benefit directly from this transparency, while basic legal principles like the parole evidence rule and contractual “integration” clauses encourage clarity to be built into the contract from the start, since it cannot be added after the fact.³³ This helps in defining foundational truths that support

³¹ A “graph” is an abstract mathematical formalism of a set of elements, called “nodes” (or vertices or points), and a corresponding set of “edges” (or lines) that connect the nodes. Graph theory has developed a large body of proved propositions describing the nature of graphs. See, for example, Diestel (2006) for further details.

³² The literature on network models of systemic risk is large and growing. For recent overviews, see Haldane (2009), ECB (2010a), or Moussa (2011). Engle and Weidman (2010) specifically consider the technical capabilities needed for supervising systemic financial risk.

³³ See, for example, Gooch and Klein (1997), especially pp. 63–64.

datum-level validation, as well as the internal consistency needed for contract-wide data integrity rules. Third, many, but not all, financial contracts already exist in well-understood digital representations; in these cases the data representation problem is largely solved. To facilitate large-scale processes for trade confirmation, settlement, corporate actions, etc., firms encode most contracts in highly structured and well-documented public messaging schemas, such as ISO20022 (2011), FIX (2011) or FpML (2011). Lastly, and most importantly, contracts define the contingent cash flows that constitute the financial essence of the relationship. The details of who pays whom, how much, when, and under what circumstances are the key to calculating valuations and understanding risk exposures. A fundamental capability is to capture and understand each contract's cash flow commitments—often contingent on other factors—between the counterparties. Understanding the cash flows is crucial because it is possible for two contracts or portfolios to generate substantially identical cash flow patterns, even when their legal or accounting representations differ widely. Much of financial engineering is devoted to repackaging a fixed set of cash flow commitments into a different contractual configuration, perhaps to manage or lay off risk, avoid taxable events, reduce the market impact of a trade, or simply to obfuscate the activity.

Monitoring risks from across the financial spectrum implies comparing and aggregating seemingly disparate exposures, such as a structured mortgage-backed security and a subordinated corporate debenture. Doing it in individual portfolios is one thing, however, to do it at the scale and scope of the full financial system would require additional automation and analytics, even if the monitoring frequency is not continuous. The upshot is a need for robust instrument type identification, including standardized, structured, machine-readable representations of financial contracts, and data integration technologies that build on top of them. These technologies should include the capability to project any contract into the financial space of state-contingent cash flows, abstracting from other details that do not affect the contractual cash flows. Brammertz et al. (2009) suggest a solution along these lines that collapses the seemingly disparate universe of financial contracts into a manageable number of cash flow patterns. Hence, two contracts with the same state-contingent cash flows appear as identical contracts for the purposes of this approach, irrespective of whether they are called loans, bonds, or derivatives, etc. A limited number of cash flow patterns can be used as building blocks to assemble more complicated patterns, so that the state-contingent cash flow obligations from the vast majority of financial contracts can be handled in a standardized and manageable way. Projections of this sort would create a set of equivalence classes that implicitly define instrument types based on financial considerations (i.e., cash flows) rather than legal, accounting, or regulatory distinctions.

While collecting contract-level details for the full financial system is a powerful supervisory approach, it is a major challenge that will take a long-term sustained effort to execute. It will also take careful design and structuring to avoid overwhelming the macroprudential supervisor with data storage, security, and validation burdens. Other industries have been innovative in this area where finance has not: for example, retail merchandising has deployed “eventually correct”

architectures with distributed processing.³⁴ Techniques for resolution reduction are another obvious response which should also support systemic risk monitoring in the nearer term. While resolution reduction originated in the visualization community as a set of techniques to compress images while still retaining important patterns and features, it has broader applicability to other domains where data compression is useful. For example, in defining the “optimal granularity” of supervisory reporting for counterparty credit risk on OTC derivatives, Mutnikas and Zerbs (2011) propose that supervisors collect contingent exposures only from the 50 largest firms, for five to 10 future value dates, and under chosen set (ca. 200) contingent scenarios. Moreover, this reporting would collect aggregated gross and net bilateral exposures. Duffie (2011) suggests a similar subset-and-aggregate approach to resolution reduction. However, surveillance requirements depend intensely on the state of the world. During a crisis, or in the aftermath of a firm’s failure, the supervisor’s need for information will be much more extensive and urgent than ordinarily. For example, state-contingent data collection is a central motivation for the “living-will” requirements of the DFA. FDIC (2011) describes the role of its new Office of Complex Financial Institutions (OCFI) thus:

A critical component of successfully addressing a distressed SIFI [systemically important financial institution] is having sufficient information and clear strategic options at the time of failure to enable decision makers to reasonably foresee the outcomes of alternative scenarios. One of the FDIC’s biggest challenges during the fall of 2008 was not having the information necessary to make informed decisions. Robust pre-planning—which entails understanding how and where these enterprises operate, as well as the structure of their business lines, counterparties, business risks, their role in the financial system, and their place in financial intermediation—is essential in giving regulators viable resolution options other than a bailout in the midst of a crisis. OCFI’s monitoring activity of these systemic enterprises will be the principal mechanism for validating the entities’ resolution plans and informing the FDIC on the development of Title II resolution plans.

“Robust pre-planning” should include the technical ability to ingest fully granular terms and conditions on financial contracts held by the relevant firms. The capacity for state-contingent resolution enhancement should be available for the supervision of the counterparty network as well, with the important extension that the network graph also has a role in the early warning toolkit. IMF (2009), for example, highlights an empirical model of the financial network with some ability to foreshadow systemic events. Alternatively, supervisors might simulate shocks to the network to learn how different topologies propagate disruptions.

Even a very low-resolution instance of the network graph could prove to be a powerful supervisory tool. Consider a graph that identifies all of the contracts in the system (or some subsystem of interest), but with only a very minimal set of attributes for each contractual edge in the network—for example, counterparties,

³⁴ The experience of the MERS system (see Hunt et al. 2011) is emblematic of the difficulties and unintended consequences endemic to the automation of long-standing processes. See Gilbert and Lynch (2002) on eventually consistent architectures and the so-called “CAP theorem.” See Srivastava (2006) on other recent advances in data architectures.

notional amount, and some instrument type classification. Such a “thin graph” would reveal the contractual network topology, exposing accumulating imbalances and suggesting crisis propagation channels. By presenting limited information at the contract level, it would avoid the issues of aggregation (loss of information, programming effort/bugs, reconciliation, etc.) while nonetheless limiting the burdens of data validation, security, and confidentiality. At the same time, the thin graph would provide the basic scaffolding to support resolution enhancement in a crisis, by attaching a fuller set of terms and conditions as attributes of the edges in the network. As noted above, a basic requirement for building such a graph is consistent and reliable counterparty identification.³⁵ Large complex financial institutions may comprise hundreds or thousands of distinct legal entities. Because of this, building a network graph to monitor threats to financial stability will require data on such corporate ownership families. While not the primary focus of such an effort, an additional benefit of systematic issuance of counterparty identifiers is that it should yield significant operational cost savings for financial firms by materially reducing the number of failed trades caused by the inconsistent designation of counterparties. Finally, the thin graph would provide a baseline scoping of the size and coordinate dimensions of the financial system: how many firms and instruments exist, and what types. Such a perspective is crucial for prioritizing the various options for research and supervision. To avoid looking “only under the streetlights” requires new sources of insight and light for the broader market.

Addressing these challenges will depend on the overall cognitive capacity of the organization, which includes:

- situational awareness of the financial system;
- decision support for policymakers; and
- crisis response capability

In addition, there must be a research function to augment and refine each of the foregoing, as well as publication channels to inform external stakeholders.³⁶

A core task for situational awareness is data collection and ingestion. Data collections will typically revolve around regularly repeated inflows of structured,

³⁵ The DFA, at §154(b) (2) (A) (i), also requires the OFR to build a “financial company reference data base.” This will not be trivial because many individual obligors exist in parent-subsidiary hierarchies with de facto cross-guaranties. In some cases, these are de jure cross-guaranties: the DFA (at §616) reiterates and extends the “source of strength” doctrine that requires bank and thrift holding companies provide financial support to their subsidiary depository institutions.

³⁶ Situational awareness is a concept that originated in a military context to describe the outcome of a tactical process of perception, comprehension, and projection onto a near-term decision space; see, for example, Leedom (2001). The issues of organizational capacity for systemic surveillance are better developed and understood in certain other disciplines. See, for example, Wagner et al. (2006).

machine-readable numeric series, such as market prices or transaction reports.³⁷ Data ingestion is an important step in the process, since this is where a number of important structuring activities occur, including data validation, basic classification, application of retention and filtering rules, cataloguing, initial versioning, and establishment of provenance.³⁸ The resulting metadata will provide the core informational support for subsequent “functional accessibility” to the data; the ability to navigate, query, link, and define the data. For example, the machine representations of the contracts might be mapped to a semantic context (e.g., a semantics repository) to provide additional interpretive specificity; in this case, both the contract schemas and associated semantics should be explicitly versioned over time.³⁹ Metadata also matters for data dissemination: financial exchanges, regulators, and other participants share a wide range of information, including both raw data inputs and calculated outputs, with each other and with third parties. Standardization of term definitions, classification schemes, and methods to evolve them across the regulatory and industry communities will be critical; without them, the ability to sensibly aggregate information will be compromised. Because of the large volumes of data involved, it will likely not be possible to achieve perfection in data validation at a fully granular level. Resource constraints will imply a trade-off between quantity and accuracy.⁴⁰ This trade-off should be managed to avoid mistakes and to prioritize access to the most important data. For example, incoming data might be staged in its raw, pre-ingested state until demanded by some downstream process, effectively creating a just-in-time

³⁷ There are important exceptions, of course. Unstructured data, for example, articles from newspapers and the trade press or interviews with regulators or industry participants, will be an important source of information. The information on settlement fails—which by definition do not result in contracts—presented by Bradley et al. (2011) might provide the basis for a systemic key risk indicator. Bisias et al. (2011) identify a class of early warning models that are based solely on macroeconomic aggregates.

³⁸ Provenance is a technical term for the metadata to support repeatable collection or derivation of the data. In many cases where issues regarding chain of custody or data lineage apply, establishing accurate data provenance can be crucial. Data source tagging—i.e., citation of the source—is a basic technique. There are standard markup languages, such as the Data Documentation Initiative (see DDI 2009) for capturing provenance metadata in a structured format.

³⁹ Similarly, efforts to build a “semantic repository” for finance—a comprehensive set of standard, structured, and interrelated definitions to augment the data model and help specify the attributes of contractual relationships; for example, see Enterprise Data Management Council (EDMC) (2011) or Madnick and Zhu (2006)—are extremely useful, but not sufficient. A semantics repository is also only one input into the process of understanding, and not a full solution or a methodology. Other important techniques include object definition, unique entity symbology, information standardization, and business process flow; these are beyond the scope of the present paper.

⁴⁰ For example, Vogels (2009), in a discussion of the “eventual consistency” model of distributed and replicated data, cites Brewer’s (2000) “CAP (consistency, availability, partition-tolerance)” proposition that, “of three properties of shared-data systems—data consistency, system availability, and tolerance to network partition—only two can be achieved at any given time.” A formal proof is given by Gilbert and Lynch (2002).

inventory system. The prioritization analysis might itself be assisted by techniques for automated discovery, inference, and pattern recognition. Based on accumulated experience, and perhaps supported by machine learning, newly arriving data might contain easily detected features or anomalies of special interest.

Because statistical analysis and data visualization are powerful tools for data aggregation, pattern extraction, and dimensionality reduction, both should play an important role in decision support in this data-rich environment.⁴¹ Decision support is one of the most important applications for the assembled information resources. Given the vast amounts of data involved and the complexity of relationships, there must be techniques for systematizing, streamlining, and rationalizing the raw data into presentations tailored to the needs of policymakers and other stakeholders. Regarding statistical analysis, Bisias et al. (2011) survey a diverse range of economic models of threats to financial stability, which they classify into five broad categories based on the modelling techniques employed and financial phenomena considered: macroeconomic measures, illiquidity measures, probability distribution measures, network analysis measures, and contingent-claims/default measures. In addition, they organize the paper around an alternative breakdown into broad categories based on data inputs and outputs and analytical methods applied: macroeconomic measures, granular foundations and financial networks, forward-looking risk assessment, stress tests, cross-sectional measures, and liquidity/insolvency and crisis behaviour. Finally, they identify the particular data inputs required by the individual models examined (see Bisias et al. 2011, Table 1). Beyond traditional econometrics, well-designed dashboard graphics and animations can condense important information for rapid assimilation for decision support. Data exploration is another area where visualization tools can make a major contribution. Certain basic rules for data classification, analysis, and triage can be automated, but many others will require a human analyst. For example, analysis of anomalous market activity is an example of something that may be difficult to train a machine to do well. Graphics are a useful technique for aggregating data for broader publication, as important decisions are taken not only by regulators and policymakers, but also by investors and other market participants.

Finally, rapid response is a required capacity for what are perhaps the most interesting facts of all, namely news of major unanticipated events. The costs of poor decisions and deferred decisions can be large, with the benefits of good decision support correspondingly large. By nature, the data delivery mechanism in such cases is unpredictable: news of a large price jump could arrive through a regular data-ingestion process; alternatively, news or direct experience of a terrorist attack might arrive outside of normal channels. The ability to react appropriately will depend on having in place the right skills, computational capacity, and functional access to information when the news arrives. For example, the

⁴¹ Regarding data visualization, see Hansen et al. (2009), Johnson et al. (2007) and Lemieux et al. (2012).

CFTC's ability to provide timely and effective support for the preliminary report on the "flash crash" (see CFTC-SEC 2010a, b) was significantly enhanced by the data-ingestion infrastructure that was already in place when the event occurred. Rapid response capability implies a need for a very broad range of specialized expertise, some of which might be outsourced through networks of on-call analysts and researchers outside the agency. Like a triage in an emergency room, the first task will be to assess the nature of the event so it can be handed off to the proper expert or team for classification (diagnosis) and finally response (treatment). An example of a possible response is a "flash report," defined as a decision memo offering preliminary findings and policy options within 24 h of a significant market event. In a rapid-response context, even short-horizon early warning indicators from a risk dashboard can serve a useful function by escalating situational awareness and alerting the on-call network.

2.4 Summary

The preceding sections highlight important forces that shape the landscape for monitoring threats to financial stability. First, data volumes are growing at an exponential rate far exceeding the growth rate in human population. While this is a general phenomenon, it also appears to apply with even greater force to financial data flows. Traditional data-management processes are unsustainable in this environment. Second, monitoring the financial system will require much greater attention to the edges in the network—financial contracts—than is available with traditional accounting or supervisory techniques. Individual participants in the system will always have limited visibility beyond their own immediate relationships. This creates a natural role for a macroprudential supervisor to monitor the evolution of the counterparty network as a whole. Third, the complexity of the problem domain, combined with the volume of data involved and the pace of decisions and activity will create a very challenging information environment for a financial stability monitor. Significant attention and resources should be devoted to building cognitive capacity in the organization. The authors are grateful to Dick Berner, John Bottega, Con Crowley, Mark Flannery, Joe Langsam, Adam LaVier, Irina Leonova, Jim Rhyne, Jonathan Sokobin, Vicki Lemieux, and participants in the 2011 Records and Information Management for Financial Analysis and Risk Management Workshop at the University of British Columbia, participants at the 2012 Workshop on Privacy and Confidentiality Issues for Financial Data at Penn State University, and discussions of the Open Financial Data Group for numerous helpful comments. Any remaining errors pertain to the authors alone.

References

- Adrian T, Shin HS (2010) Liquidity and leverage. *J Financ Intermed* 19:418–437
- Agarwal S, Hauswald R (2010) Distance and private information in lending. *Rev Financ Stud* 23:2757–2788
- Antonelli C (2009) The economics of innovation: from the classical legacies to the economics of complexity. *Econ Innov New Technol* 18:611–646. doi:[10.1080/10438590802564543](https://doi.org/10.1080/10438590802564543)
- Arora S, Barak B, Brunnermeier M, Ge R (2011) Computational complexity and information asymmetry in financial products. *Commun ACM* 54:101–107
- Bank for International Settlements (BIS) (2010) Triennial and semiannual surveys: positions in global over-the-counter (OTC) derivatives markets at end—June 2010. Press release. http://www.bis.org/publ/otc_hy1011.htm
- Benston GJ (2006) Fair-value accounting: a cautionary tale from Enron. *J Account Public Pol* 25:465–484
- Berger AN, Miller NH, Petersen MA, Rajan RG, Stein JC (2005) Does function follow organizational form? Evidence from the lending practices of large and small banks. *J Financ Econ* 76:237–269. doi:[10.1016/j.jfineco.2004.06.003](https://doi.org/10.1016/j.jfineco.2004.06.003)
- Berthold J, Filinski A, Henglein F, Larsen KF, Steffensen M, Vinter B (2011) Functional high performance financial IT: the Hiperfit Research Center in Copenhagen—project paper. In: Peña-Mari R, Page R (eds) *Trends in functional programming (TFP'11)—12th international symposium, Madrid, Spain. Revised selected papers*. Springer, Berlin 16–18 May 2011
- Bisias D, Flood M, Lo A, Valavanis S (2011) A survey of systemic risk analytics. Working paper, number 0001, Office of Financial Research
- Blundell-Wignall A, Atkinson P (2010) Thinking beyond basel III: necessary solutions for capital and liquidity. *OECD J Financ Mark Trends* 1:1–23
- Bookstaber R (2007) *A demon of our own design: markets, hedge funds, and the perils of financial innovation*. Wiley, New York
- Boyd J, Gertler M (1994) Are banks dead? Or are the reports greatly exaggerated? Fed Reserve Bank Minneapolis Q Rev 18:2–23. <http://www.minneapolisfed.org/research/QR/QR1831.pdf>
- Bradley H, Fawls R, Litan RE, Sommers F (2011) Canaries in the coal mine: how the rise in settlement ‘fails’ creates systemic risk for financial firms and investors. Technical report, Ewing Marion Kauffman Foundation. <http://www.kauffman.org/uploadedfiles/canaries-in-the-coal-mine-report.pdf>
- Brammertz W, Akkizidis I, Breymann W, Entin R, Rustmann M (2009) *Unified financial analysis: the missing links of finance*. Wiley, New York
- Brewer E (2000) Towards robust distributed systems. In: *Proceedings of the 19th annual ACM symposium on principles of distributed computing (PODC'00)*
- Brunnermeier MK, Eisenbach TM, Sannikov Y (2011) *Macroeconomics with financial frictions: a survey*. Working paper, Princeton University
- Castura J, Litzenberger R, Gorelick R, Dwivedi Y (2010) Market Efficiency and microstructure evolution in U.S. equity markets: a high-frequency perspective. Working Paper, RGM Advisors, LLC. http://fnce.wharton.upenn.edu/news/Litzenberger_transient_vol5_2010.pdf
- U.S. Census Bureau (2011) Total midyear population for the world: 1950–2050. International Data Base (IDB). <http://www.census.gov/population/international/data/idb/worldpoptotal.php>. Accessed 26 July 2011
- Choo CW (2006) *The knowing organization: how organizations use information to construct meaning, create knowledge, and make decisions*. Oxford University Press, Oxford
- Ciulla T, Bloom D, Ages J (2010) Automating the OTC derivative collateral management function. *J Secur Oper Custody* 3:103–110
- Coase RH (1937) The nature of the firm. *Econ New Ser* 4:386–405
- Committee on Payment and Settlement Systems (CPSS) (2008) The interdependencies of payment and settlement systems. Technical report, Bank for International Settlements. <http://www.bis.org/publ/cpss84.pdf>

- Committee on Payment and Settlement Systems (CPSS) (2011) Principles for financial market infrastructures. Consultative report, Bank for International Settlements. <http://www.bis.org/publ/cpss94.pdf>
- Committee on Payment and Settlement Systems (CPSS), Technical Committee of the International Organization of Securities Commissions (IOSCO) (2001) Recommendations for securities settlement systems. Consultative Report, CPSS-IOSCO Joint Task Force on Securities Settlement Systems. <http://www.bis.org/publ/cpss42.pdf>
- Commodity Futures Trading Commission and Securities and Exchange Commission (CFTC-SEC) (2010a) Preliminary findings regarding the market events of May 6, 2010. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. <http://www.sec.gov/sec-cftc-prelimreport.pdf>
- Commodity Futures Trading Commission and Securities and Exchange Commission (CFTC-SEC) (2010b) Findings regarding the market events of May 6, 2010. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues, Sept 2010. www.sec.gov/news/studies/2010/marketevents-report.pdf
- U.S. Congress (2010) Dodd-Frank wall street reform and consumer protection act, bill summary and status 111th congress (2009–2010), H.R.4173. Enrolled Bill. <http://thomas.loc.gov/cgi-bin/query/z?c111:h4173>
- Copeland A, Martin A, Walker M (2010) The tri-party repo market before the 2010 reforms. Federal Reserve Bank of New York Staff Reports, number 477. http://www.copeland.marginalq.com/res_doc/sr477.pdf
- Cordell L, Huang Y, Williams M (2011) Collateral damage: sizing and assessing the subprime CDO crisis. Working Paper, number 11-30, Federal Reserve Bank of Philadelphia, Philadelphia, Pennsylvania. <http://www.philadelphiafed.org/research-and-data/publications/working-papers/2011/wp11-30.pdf>
- Counterparty Risk Management Policy Group (CRMPG) (2005) Toward greater financial stability: a private sector perspective. Report of the Counterparty Risk Management Policy Group II. <http://www.crmgroup.org/crmgroup2/docs/CRMPG-II.pdf>
- Danielsson J, Shin HS (2003) Endogenous risk. In: Field P (ed) Modern risk management: a history. Risk Books, London
- Data Documentation Initiative (DDI) (2009) Data Documentation Initiative (DDI) technical specification part I: overview, version 3.1. Technical report. <http://www.ddialliance.org/Specification/DDI-Lifecycle/3.1/>
- Dell’Ariccia G, Igan D, Laeven L (2008) Credit booms and lending standards: evidence from the subprime mortgage market. IMF Working Paper, WP/08/106. <http://www.imf.org/external/pubs/ft/wp/2008/wp08106.pdf>
- DeMarzo PM (2005) The pooling and tranching of securities: a model of informed intermediation. *Rev Financ Stud* 18:1–35
- Deryugina M (2009) Standardization of securities regulation: rehypothecation and securities commingling in the United States and the United Kingdom. *Rev Bank Financ Law* 29:253–288
- Diestel R (2006) Graph theory, 3rd edn. Springer, Berlin
- Duffie D (2011) Systemic risk exposures: a 10-by-10-by-10 approach. Working paper, Stanford University. <http://www.darrellduffie.com/uploads/policy/Duffie10By10By10July2011.pdf>
- Dungey D (2007a) Tanta: mortgage servicing for UberNerds. Calculated risk. <http://www.calculatedriskblog.com/2007/02/tanta-mortgage-servicing-for-ubernerds.html>. Accessed 27 July 2011
- Dungey D (2007b) Foreclosure sales and REO for UberNerds. Calculated risk. <http://www.calculatedriskblog.com/2007/04/foreclosure-sales-and-reo-for-ubernerds.html>. Accessed 27 July 2011
- Engle RF, Weidman S (2010) Technical capabilities necessary for regulation of systemic financial risk: summary of a workshop. National Research Council of the National Academies, National Academies Press, Washington, DC

- Enterprise Data Management Council (EDMC) (2011) Enterprise Data Management Council semantics repository. <http://www.hypercube.co.uk/edmcouncil/>. Accessed 29 July 2011
- European Central Bank (ECB) (2010a) Financial networks and financial stability. *Financ Stab Rev* 17:155–160. <http://www.ecb.int/pub/pdf/other/financialstabilityreview201006en.pdf>
- European Central Bank (ECB) (2010b) New quantitative measures of systemic risk. *Financ Stab Rev* 17:147–153. <http://www.ecb.int/pub/fsr/shared/pdf/ivefinancialstabilityreview201012en.pdf?abea0fb869314b1aca4c72f2d826b897>
- Federal Deposit Insurance Corporation (FDIC) (2011) Statement of Martin J. Gruenberg, Acting Chairman, Federal Deposit Insurance Corporation on enhanced oversight after the financial crisis: wall street reform at one year before the committee on banking, Housing and Urban Affairs, United States Senate; 538 Dirksen Senate Office Building, 21 July 2011. <http://www.fdic.gov/news/news/speeches/chairman/spjul2111.html>
- Federal Reserve Bank of St. Louis (2011a) Delinquency rate on single-family residential mortgages, booked in domestic offices, all commercial banks (DRSFRMACBS). FRED economic data. <http://research.stlouisfed.org/fred2/series/DRSFRMACBS/downloaddata?cid=32440>. Accessed 28 July 2011
- Federal Reserve Bank of St. Louis (2011b) Charge-off rate on single family residential mortgages, booked in domestic offices, all commercial banks (CORSFRMACBS). FRED economic data. <http://research.stlouisfed.org/fred2/series/CORSFRMACBS/downloaddata?cid=32439>. Accessed 28 July 2011
- Feldman R, Lueck M (2007) Are banks really dying this time? An update of Boyd and Gertler. *The Region—The Federal Reserve Bank of Minneapolis* 6-51
- Financial Accounting Standards Board (FASB) (2008) Preliminary views: conceptual framework for financial reporting: the reporting entity. Financial Accounting Series No. 1580-100. Financial Accounting Standards Board, Norwalk, Connecticut. http://www.fasb.org/pv_conceptual_framework_reporting_entity.pdf
- Financial Products Markup Language (FpML) (2011) What is FpML®? <http://www.fpml.org/about/factsheet.html>. Accessed 29 July 2011
- FIX Protocol (FIX) (2011) What is FIX? <http://www.fixprotocol.org/what-is-fix.shtml>. Accessed 29 July 2011
- Flannery MJ, Kwan SH, Nimalendran M (2004) Market evidence on the opaqueness of banking firms' assets. *J Financ Econ* 71:419–460. doi:10.1016/S0304-405X(03)00185-5
- Flannery MJ, Kwan SH, Nimalendran M (2010) The 2007–09 financial crisis and bank opaqueness. Working paper 2010-27, Federal Reserve Bank of San Francisco. <http://www.frbsf.org/publications/economics/papers/2010/wp10-27bk.pdf>
- Fleming M, Garbade K (2005) Explaining settlement fails. *Curr Issues Econ Financ Fed Reserve Bank N Y* 11:1–7
- Flood M (2009) Embracing change: financial informatics and risk analytics. *Quant Financ* 9:243–256
- Flood M, Kyle ASP, Raschid L (2010) Knowledge Representation And Information Management For Financial Risk Management. Report of a workshop, Arlington, VA, National Science Foundation and Pew Charitable Trusts. <http://irix.umiaccs.umd.edu/docs/FIWreport-FINAL.pdf>
- Fuerber C, Hepp M, Wischniewski M (2011) The data quality constraints library: Primer, Draft 2010-11-24. <http://semwebquality.org/documentation/primer/20101124/index.html>. Accessed 26 Novemb 2011
- Gaspar JM, Massa M, Matos P (2006) Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. *J Financ* 61:73–104
- Ghoting A, Parthasarathy S, Otey ME (2008) Fast mining of distance-based outliers in high-dimensional datasets. *Data Min Knowl Disc* 16:349–364. doi:10.1007/s10618-008-0093-2. <http://www.springerlink.com/content/x58067325kv41n26/>
- Gilbert S, Lynch N (2002) Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *ACM SIGACT News* 33:51–59. <http://dl.acm.org/citation.cfm?doid=564585.564601>

- Gooch A, Klein K (1997) A review of case law affecting swaps and derivative instruments: United States case law. In: Schwartz R, Smith C (eds) *Derivatives handbook: risk management and control*. Wiley, New York
- Gorton GB, Metrick A (2009) Haircuts. Working paper 15273, National Bureau of Economic Research, Cambridge, Massachusetts. <http://nber.org/papers/w15273.pdf>
- Gottfredson M, Aspinall K (2005) Innovation versus complexity: what is too much of a good thing? *Harvard Bus Rev* 83:62–70
- Haldane AG (2009) Rethinking the financial network. Speech delivered at the financial student association, Amsterdam. Bank of England, London. <http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf>
- Han J, Kamber M, Pei J (2011) *Data mining: concepts and techniques*, 3rd edn. Morgan Kaufmann, Waltham
- Hansen C, Johnson C, Pascucci V, Silva C (2009) Visualization for data-intensive science. In: Hey T, Tansley S, Tolle K (eds) *The fourth paradigm: data-intensive scientific discovery*. Microsoft Research, Redmond
- Hartman-Glaser B (2011) Reputation vs. signaling in a security issuance game. Working paper, University of California at Berkeley. <http://energy.uh.edu/departments/finance/documents/Barney%20main.pdf>
- Hendershott T, Jones C, Menkveld A (2011) Does algorithmic trading improve liquidity? *J Financ* 66:1–33
- Hilbert M, López P (2011a) The world's technological capacity to store, communicate, and compute information. *Int J Commun* (forthcoming). <http://www.martinhilbert.net/HilbertLopez%20InfoCapacityMethod.pdf>
- Hilbert M, López P (2011b) How to measure the world's technological capacity to communicate, store and compute information? *Science* 332:60–65
- Hildreth PM, Kimble C (2002) The duality of knowledge. *Inf Res* 8:8-1
- Holland P (2011) The one hundred billion dollar problem in small claims court: robo-signing and lack of proof in debt buyer cases. *J Bus Tech Law* 6:101–129
- Holmström B, Roberts J (1998) The boundaries of the firm revisited. *J Econ Perspect* 12:73–94
- Hunt J, Stanton R, Wallace N (2011) The end of mortgage securitization? Electronic registration as a threat to bankruptcy remoteness. Working paper, University of California at Berkeley. <http://faculty.haas.berkeley.edu/stanton/papers/pdf/mers.pdf>
- Ingersoll J, Spiegel M, Goetzmann W, Welch I (2007) Portfolio performance manipulation and manipulation-proof performance measures. *Rev Financ Stud* 20:1503–1546
- Intel (2010) Extreme speed, extreme throughput. Marketing brochure. http://www.thesystech.com/docs/3853-8_IN_ThesysTechCaseStudy-hi.pdf
- International Accounting Standards Board (IASB) (2011) IFRS 13 fair value measurement. Guidance document. <http://www.ifrs.org/Current+Projects/IASB+Projects/Fair+Value+Measurement/IFRS+13+Fair+Value+Measurement/IFRS+13+Fair+Value+Measurement.htm>
- International Monetary Fund (IMF) (2001) The macroeconomic statistical treatment of reverse transactions. In: Fourteenth meeting of the IMF committee on balance of payments statistics Tokyo, Japan, BOPCOM-01-16. International Monetary Fund Statistics Department, Washington, DC. <http://www.imf.org/external/pubs/ft/bop/2001/01-16.pdf>. 24–26 Oct 2001
- International Monetary Fund (IMF) (2009) Assessing the systemic implications of financial linkages. In: *Global financial stability report: responding to the financial crisis and measuring systemic risks*. International Monetary Fund, Washington, DC
- International Organization for Standardization (ISO20022) (2011) ISO 20022 universal financial industry message scheme. <http://www.iso20022.org/>. Accessed 29 July 2011
- International Organization of Securities Commissions (IOSCO) (2008) Report on the subprime crisis: final report. Technical report. http://www.iosco.org/library/pubdocs/pdf/IOSCO_PD273.pdf
- Johnson C, Ross R, Ahern S, Ahrens J, Bethel W, Ma KL, Papka M, Rosendale JV, Shen HW, Thomas J (2007) *Visualization and knowledge discovery: report from the DOE/ASCR*

- workshop on visual analysis and data exploration at extreme scale. Technical report. <http://www.sci.utah.edu/vaw2007/DOE-Visualization-Report-2007.pdf>
- Judge K (2011) Fragmentation notes: A study in financial innovation, complexity and systemic risk. Working paper 406 Columbia Law School, Center for Law & Economic Studies <http://ssrn.com/abstract=1894105>
- Kashyap A, Berner R, Goodhart C (2011) The macroprudential toolkit. *IMF Econ Rev* 59:145–161
- Kaufman GG, Mote LR (1994) Is banking a declining industry: a historical perspective. *Econ Persp Fed Reserve Bank Chicago* 18:2–21
- Kaufman T, Neiman R, Silvers D, McWatters JM, Troske K (2010) Examining the consequences of mortgage irregularities for financial stability and foreclosure mitigation. Congressional oversight panel November oversight report. U.S. Government Printing Office, Washington, DC
- Khandani A, Lo AW (2011) What happened to the quants in August 2007? Evidence from factors and transactions data. *J Financ Mark* 14:1–46
- Koh H, Magee CL (2006) A functional approach for studying technological progress: application to information technology? *Technol Forecast Soc Chang* 73:1061–1083
- Leedom DK (2001) Sensemaking symposium final report to the command and control research program, Office of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence, U.S. Department of Defense. http://www.au.af.mil/au/awc/awcgate/ccrp/sensemaking_final_report.pdf
- Lemieux V, Fisher B, Dang T (2012) The visual analysis of financial data. In: Brose M, Flood M, Krishna D, Nichols W (eds) *Handbook of financial data and risk information*, vol 2: implementation, forthcoming. Cambridge University Press, Cambridge
- Liberti J, Mian A (2012) Estimating the effect of hierarchies on information use. *Rev Financ Stud* 22:4057–4090
- Lo AK (2011a) Andrew Lo discusses systemic risk. Presentation video, UMD/CMU/NYU/UCB conference on systemic risk and data issues, Washington, DC. <http://www.youtube.com/watch?v=nuDloBeNwD0>. 5 Oct 2011
- Lo AK (2011b) In the valley of the shadow of hedge funds. Presentation slides, UMD/CMU/NYU/UCB conference on systemic risk and data issues, Washington, DC. <http://www.rhsmith.umd.edu/cfp/events/2011/confSystemicRisk1011/>. 5 Oct 2011
- Madnick S, Zhu H (2006) Improving data quality through effective use of data semantics. *Data Knowl Eng* 59:460–475
- Markham J (2002) A financial history of the United States, vol 2: from J.P. Morgan to the institutional investor (1900–1970). M.E. Sharpe, Armonk
- Moussa A (2011) Contagion and systemic risk in financial networks. PhD thesis, Columbia University. <http://academiccommons.columbia.edu/catalog/ac:131474>
- Mutnikas Y, Zerbs M (2011) Optimal granularity: a supervisory deep dive on counterparty credit risk. *Think November*: 32–37
- Nagy B, Doyne Farmer J, Trancik JE, Gonzales JP (2011) Superexponential long-term trends in information technology. *Technol Forecast Soc* 73:1061–1083
- Noll F (2009) Lincoln's greenback mill: civil war financing and the start of the bureau of engraving and printing, 1861–1863. *Fed Hist J* 1:25–31
- Office of Financial Research (OFR) (2011) Office of Financial Research issues statement on progress to date and next steps forward in the global initiative to establish a legal entity identifier (LEI). Press release. <http://www.treasury.gov/press-center/press-releases/Pages/tg1275.aspx>
- Pandis I, Johnson R, Hardavellas N, Ailamaki A (2010) Data-oriented transaction execution. *Proc VLDB Endow* 3:928–939
- Perrow C (1999) *Normal accidents: living with high-risk technologies*. Princeton University Press, Princeton
- Petersen MA, Rajan RG (2002) Does distance still matter? The information revolution in small business lending. *J Financ* 57:2533–2570

- Philippon T (2008) Why has the financial sector grown so much? The role of corporate finance. NBER Working paper, number 13405, National Bureau of Economic Research. <http://w4.stern.nyu.edu/emplib/philippon.pdf>
- Pozsar Z, Singh M (2011) The non-bank-bank nexus and the shadow banking system. IMF working paper, number WP/11/289. <http://www.imf.org/external/pubs/ft/wp/2011/wp11289.pdf>
- Pozsar Z, Adrian T, Ashcraft A, Boesky H (2010) Shadow banking. Staff report number 458, Federal Reserve Bank of New York, New York. http://www.newyorkfed.org/research/staff_reports/sr458.pdf
- Securities and Exchange Commission (SEC) (1972) Thirty-eighth annual report of the securities and exchange commission for the fiscal year ended June 30th. U.S Government Printing Office, Washington, DC. http://www.sec.gov/about/annual_report/1972.pdf
- U.S. Senate (2010) The national institute of finance act of 2010, bill summary and status 111th congress (2009–2010), S.3005. Introduced 4 Feb 2010. <http://thomas.loc.gov/cgi-bin/bdquery/z?d111:s3005>
- Senior Managements of: Bank of America NA, Barclays Capital, Bear, Stearns & Co, Citigroup, Credit Suisse First Boston, Deutsche Bank AG, Goldman, Sachs & Co, HSBC Group, JP Morgan Chase, Lehman Brothers, Merrill Lynch & Co, Morgan Stanley, UBS AG, Wachovia Bank NA (Senior Managements) (2005) Letter to Mr. Timothy Geithner, President, Federal Reserve Bank of New York. Federal Reserve Bank of New York, New York
- Shleifer A, Vishny R (2011) Fire sales in finance and macroeconomics. *J Econ Perspect* 25:29–48
- Singh M, Aitken J (2010) The (sizable) role of rehypothecation in the shadow banking system. IMF working paper, number WP/10/172. <http://www.imf.org/external/pubs/ft/wp/2010/wp10172.pdf>
- Srivastava UH (2006) Efficient query processing for modern data management. Doctoral thesis, Stanford University. <http://ilpubs.stanford.edu:8090/786/1/2006-20.pdf>
- Stonebraker M, Madden S, Abadi DJ, Harizopoulos S, Hachem N, Helland P (2007) The end of an architectural era (it's time for a complete rewrite). In: VLDB '07: proceedings of the 33rd international conference on very large data bases. VLDB Endowment, Vienna
- Taub L (2008) Borrowed securities: implications for measuring cross-border portfolio investment. In: IFC bulletin number 28, IFC's contribution to the 56th ISI session, Lisbon, 2007. Bank for International Settlements, Basel
- Thaler RH (1988) Anomalies: the winner's curse. *J Econ Perspect* 2:191–202
- Treasury Market Practices Group (TMPG) (2011) Understanding settlement fails in agency mortgage-backed securities. Technical report, Federal Reserve Bank of New York. http://www.newyorkfed.org/tmpg/tmpg_04292011.pdf
- Trimbath S (2008) Trade settlement failures in U.S. bond markets. Working paper, number STP2007_01 (v.2), STP Advisory Services. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1016873
- Tufano P (2003) Financial innovation. In: Constantinides G, Harris M, Stulz R (eds) *Handbook of the economics of finance*, vol 1a: corporate finance. Elsevier, Amsterdam
- Vogels W (2009) Eventually consistent. *Commun ACM* 52:40–44
- Wagner M, Moore A, Aryel R (2006) *Handbook of biosurveillance*. Academic Press, Waltham
- Wallace N (2011) Private-label residential mortgage securitization: recording innovations and bankruptcy remoteness. Presentation slides, University of California at Berkeley. http://www.law.berkeley.edu/files/bclbe/Wallace_MERS.pdf
- Wikipedia (2011) Transistor count. http://en.wikipedia.org/wiki/Transistor_count. Accessed 22 July 2011
- Williamson OE (1991) Comparative economic organization: the analysis of discrete structural alternatives. *Admin Sci Quart* 36:269–296
- Yahoo Finance (2011) S&P 500 INDEX, RTH (^GSPC). <http://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices>. Accessed 22 July 2011



<http://www.springer.com/978-3-642-32231-0>

Financial Analysis and Risk Management
Data Governance, Analytics and Life Cycle Management
Lemieux, V. (Ed.)
2013, XVI, 208 p., Hardcover
ISBN: 978-3-642-32231-0