The Data Standardization Challenge

Richard Berner
New York University

Kathryn Judge
Columbia University and ECGI

© Richard Berner and Kathryn Judge 2019. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
http://ssrn.com/abstract_id=3323719

www.ecgi.global/content/working-papers
Abstract

Data standardization offers significant benefits for industry and regulators alike, suggesting that it should be easy. In practice, however, the process has been difficult and slow moving. Moving from an abstract incentive-based analysis to one focused on institutional detail reveals myriad frictions favoring the status quo despite foregone gains. This paper explores the benefits of and challenges confronting standardization, why it should be a top regulatory priority, and how to overcome some of the obstacles to implementation. The paper also uses data standardization as a lens into the challenges that impede optimal financial regulation. Alongside capture and other common explanations for regulatory failures, this paper suggests that coordination problems, delayed benefits, and other banal, but perhaps no less intractable, challenges are often the real impediments to better financial regulation.

Keywords: data standards, financial regulation

JEL Classifications: G18, G28, K2, L51
The Data Standardization Challenge

Richard Berner and Kathryn Judge

Working Paper No. 598


January 27, 2019

Pleas do not quote or cite without author's permission.

To view other articles in the Columbia Law & Economics Working Paper Series, see:

http://web.law.columbia.edu/law-economic-studies/working-papers
The Data Standardization Challenge
Richard Berner\textsuperscript{a} & Kathryn Judge\textsuperscript{b}

Data standardization offers significant benefits for industry and regulators alike, suggesting that it should be easy. In practice, however, the process has been hard and slow moving. Moving from an abstract incentive-based analysis to one focused on institutional detail reveals myriad frictions favoring the status quo despite foregone gains. This chapter explores the benefits of and challenges confronting standardization, why it should be a top regulatory priority, and how to overcome some of the obstacles to implementation.

The paper further argues that data standardization provides a lens into the nature of the challenges that often impede optimal financial regulation. Alongside capture and other common explanations for regulatory failures, this essay suggests that coordination problems, delayed benefits, and other banal, but perhaps no less intractable, challenges are often the real impediments to better financial regulation.


\textsuperscript{a} New York University Leonard N. Stern School of Business, Center for Global Economy and Business; FinRegLab.
\textsuperscript{b} Columbia Law School; European Corporate Governance Institute.
Quality information is the lifeblood of strong, vibrant markets. Without it, investor confidence erodes. Liquidity dries up. Fair and efficient markets simply cease to exist. As the quantity of information increases exponentially through the Internet and other technologies, the quality of that information must be our signal priority.¹

SEC Chairman Arthur Levitt

Data standards are the rules by which data are described and recorded.²

United States Geological Survey

Introduction

Information is the lifeblood of finance. The role the financial system plays in moving capital from savers to borrowers and projects is inherently informational. Banks, investment banks, asset managers, insurance companies, and other financial firms add value in part by producing, verifying, and disseminating the information needed to determine who should get financing, on what terms and the risks involved. Financial markets, and the institutional structures like exchanges, play essential roles in aggregating and translating information into price signals. Institutional investors aggregate capital and further contribute to the information-generating role of the financial system. Many significant financial innovations, like securitization, exist not just to redistribute or manage risks, but to do so in a manner that alters the information sensitivity of the financial instruments available in the market to meet participants’ needs.³ Likewise, other financial-system functions, such as pooling and reallocating risks and providing liquidity depend on the sector’s information-related roles.⁴

The integral role of information in finance also helps to explain much financial regulation. Securities regulation, for example, reduces the costs investors incur gathering, analyzing and verifying information.⁵ Similarly, cornerstones of


² United States Geological Survey, Data Standards, online: USGS Data Management, <https://www2.usgs.gov/datamanagement/plan/datastandards.php>. That this definition comes from an entirely different discipline from finance illustrates the universality of the need for standardized data.


bank regulation, like capped deposit insurance coupled with supervision, can be framed as efforts to re-allocate monitoring and other information-production to the parties best suited to undertake it. Regulators also require information to fulfill many of their mandates, particularly with respect to promoting systemic stability and resilience.

Designing and implementing appropriate data standards is an important mechanism through which the government can fulfill these dual roles of enhancing the efficiency of private activity and obtaining the high-quality information necessary to better serve the public at large. Standardization is essential to compare and aggregate data and can enable firms and regulators to produce more accurate and timely information about a host of issues at lower expense. Firms benefit by better understanding their clients’ needs, and in pricing and managing risk. Regulators benefit by being better able to identify trends across firms and markets, and potentially being better able to assess the local and systemic risks associated with those changes. Standardization also facilitates sharing data so that firms and regulators can use and understand the same data within and across jurisdictions, reducing redundancy and misunderstanding. At its best, data standardization can serve as a crucial ingredient in building and sustaining the trust, accountability, adaptability and efficiency that are essential finance.

Well-designed and broadly used data standards also have the potential to promote systemic resilience. The financial crisis of 2007-2009, commonly known as the Great Financial Crisis, vividly illustrates the potential for how information gaps that grow in the absence of quality data contribute to systemic fragility. When subprime mortgage-backed securities proved to be more risky than their ratings suggested, the extent of the market dysfunction that followed far exceeded that which could be readily explained by the potential actual losses on those instruments. The averted failure of Bear Stearns and actual failure of Lehman Brothers entailed similar dynamics, as knowable unknowns (that is, information gaps) exacerbated fragility and impeded the ability of government actors to intervene in a timely and proportionate way. Had firms and regulators been able to use standardized, reliable information to assess more accurately the origins and impact of these developments, the overall course of the crisis might have been far less severe.

Hence, as explained by the Financial Stability Board and the International Monetary Fund, “the recent crisis has reaffirmed an old lesson—good data and good analysis are the lifeblood of effective surveillance and policy responses at both the national and international levels.” Data standardization alone obviously cannot produce a stable financial system, but in improving data quality, it represents a critical tool to build and sustain a resilient system, that is, one that can take a few

---

knocks and still get back on its feet without significant spillover on the real economy.

This chapter’s first contribution is to show why data standardization deserves a higher place on the post-crisis regulatory priority list. Although there has been some progress towards the development of and mandates to employ data standards in the decade since the GFC, improvements in data standardization have been slow, hard won, and, in many areas, elusive. The examination here reveals numerous, often unexciting, but consequential, impediments to data standardization. Collective action problems, positive externalities, and governance and incentives challenges are among the many factors impeding progress.

It might seem like an odd time to be making data standardization such a priority. With the massive technological changes of the last few years, some might believe that these issues will solve themselves. We think otherwise. A core assumption underlying our analysis is that it is not enough to increase the scope and quantity of information; if anything, today we are overwhelmed by the quantity. This makes efforts to address the qualitative characteristics of the input data, whether it is sufficiently reliable, adequately defined, and well understood, all the more important. These aims are at the core of data standardization. Improving data standards and promoting more widespread standardization has the potential to unleash real gains for public and private stakeholders and society more generally. By illuminating those gains and myriad impediments, this chapter helps explain the mixed progress over the past decade and how to do better in the decades ahead.

The chapter’s second contribution is to use data standardization as a lens for understanding the challenges that impede effective financial regulation more generally. Data standardization has the potential to benefit banks, other financial firms, and the public generally. Why then haven’t we seen more progress? The answer suggested here is that mundane details can matter as much as abstract incentives. The benefits of standardization accrue slowly, are spread widely and are only partially captured by financial firms, while the costs are borne upfront and mainly by those firms.

Also significant, and telling, is that data standardization doesn’t fit neatly into either a “tough on finance” or a deregulatory agenda. Oftentimes, financial regulation is depicted as a battleground. In this frame, the public is assumed to benefit from more regulation and is situated on one side, with banks and other financial firms, who are assumed to oppose regulation, on the other. Close discussions between regulators and financial firms, of the kind that are critical to producing high-quality standards, are seen as troubling evidence of regulatory capture. At the same time, because of coordination and other challenges, the efficacy of data standards often depends on broadly applicable government mandates. As a result, standardization is not deregulatory, as that term is commonly understood.

---

7 See Part III infra.
The good news is that standardization is now underway in a number of domains. Standards that emerged in response to weaknesses revealed by the GFC are slowly becoming more widely required, and there is new leadership in efforts to develop more thoughtful data standards in other areas. Given the clear value of thoughtful standardization and the many impediments slowing adoption, examining data standardization illuminates systemic challenges that impede both the efficiency of financial firms and the quality of financial regulation.

I. Background: What we learned from the crisis

This part explores what the GFC revealed about the importance of data standardization and why, post-crisis developments accentuate the public and private value of well-designed data standards.

This volume commemorates (if such a word can be used for such an event) the 2007-2009 financial crisis, so that is our starting point. The decade that has passed since the crisis reached its nadir in the wake of the failure of Lehman Brothers in 2008. Most agree that adverse information dynamics played a central role contributing to the scope of the market dysfunction and the adverse spillover effects on the real economy that made the crisis so tragic. We have each addressed the important role of information gaps in contributing to the crisis elsewhere. Instead of reiterating that work here, we will illustrate how information gaps contribute to fragility with a focus on the failure of Lehman Brothers, and on how the government’s information production and dissemination helped restore stability through the market’s positive response to the first round of stress tests.

With the benefit of hindsight, it is clear that the actual losses incurred as a direct result of Lehman Brothers’ failure were not sufficient to compromise directly the vitality of any major financial institution or key market structure. Nonetheless, Lehman’s failure triggered widespread market dysfunction. One reason was the inability of market participants to quickly assess their total exposures to Lehman, including its numerous subsidiaries and affiliates, or to know the exposures of their other counterparties. Financial regulators were also in the dark. A market reaction that is outsized relative to the content of new information is the very definition of systemic fragility. Had standardized, reliable information been available to assess the origins and impact of these developments, the overall course of the crisis might have been far less severe.

---


The massive runs on money market mutual funds (MMMFs) that followed offer examples of information-gap spillovers. One fund, the Primary Reserve Fund, which held $785 million in commercial paper issued by Lehman Brothers, was forced to “break the buck” – to redeem shares below the $1.00 net asset value typically assigned to MMMF shares. At the lowest, shares in the fund were redeemed at 97 cents, and most holders eventually received 99 cents on the dollar. No other MMMF ended up breaking the buck. Not one.

Nonetheless, these events triggered a massive ($400 billion, or 20%, in a few days) run on MMMFs holding privately issued debt. MMMF investors had only limited information about the actual assets and their creditworthiness backing their claims, so it was easier, cheaper, and safer to run than to learn more about the actual risks of sticking around. As Ricardo Caballero and Alp Simsek demonstrate formally, most financial firms may have been able to assess their rough exposure to Lehman Brothers, but they could not readily ascertain the exposure and soundness of other financial firms. So they withdrew from activities that might expose them to credit risk, which impaired liquidity and market functioning. The resulting significant ripple effects on short-term money markets were halted only through government intervention.

The positive impact of the first-round of stress tests illustrates the flipside -- the important role of more information in facilitating market functioning. By early 2009, the government had engaged in massive efforts to stabilize the financial system. Among them: implementing a range of nonbank liquidity facilities and guarantees of deposits and new senior debt, providing massive support to AIG, converting key nonbanks to bank holding companies, creating the Troubled Asset Relief Program, offering capital to many banks through a voluntary Capital Purchase Program, and announcing that stress tests would be used to assess banking system capital needs. These steps had brought about significant improvements in market functioning, but normalcy remained out of reach.

Then, on May 7, 2009, the Federal Reserve and other bank regulators announced the results of the Supervisory Capital Assessment Program (the stress tests). The tests provided market participants much needed information that they could verify and trust. The Fed deliberately disclosed the results of the SCAP. In providing an unprecedented level of detail regarding the methodology and inputs used in reaching those results, the Fed turned on its head the assumption that bank supervision should always be secret. Both gambles paid off. As then-Fed Chair Bernanke later observed: “The SCAP stands out ... as one of the critical turning points in the financial crisis. It provided anxious investors with something they craved: credible information about prospective losses at banks.”

---

11 Board of Governors of the Federal Reserve System, Lessons from the Failure of Lehman Brothers (2010), online:
These contrasting episodes reveal that lack of information contributed to the fragility that became manifest during the crisis, and the provision of information played an important role in helping to restore stability. Although there is a gap between data and information, high-quality data standards are one of the most effective ways of narrowing that gap, and enabling timely information production.

The remainder of this essay will explore why data standardization might seem to be easy, and why it nonetheless has proven difficult and slow.

II. Why it should be easy

Obstacles to data standardization would be understandable if the only gain was enhanced systemic stability at some point in the future. This part shows why well-designed and comprehensively implemented data standardization also confers benefits outside of crisis periods. In so doing, it shows why data standardization seems like it should be easier to implement than post-crisis reforms that impose only costs outside of crisis periods or that benefit the public at the expense of regulated institutions.

A. Firms benefit

Firms benefit from the quality data standards provide by better understanding their clients’ needs, and in pricing and managing risk. Effective risk management requires coordination. The capacity of senior management to oversee and coordinate actions among various divisions within their organizations depends on standardized information.

While key players are accountable for important functions, everyone working in a financial institution should be attuned to business risks. Such a structure requires that all members of the team communicate effectively. To do so requires that they have appropriate access to the same high-quality, accurate facts as the basis for risk detection, understanding, monitoring, remediation and recovery, and that the firm assures policies for standardization that are followed and that create accountability for data quality throughout the firm. If a firm has an operational incident that impairs, corrupts or loses data, it will be easier to recover and recreate client and transaction data if they are standardized. Standardized, high-quality data are also critical for effective use of third-party vendors for risk management or compliance functions.

Standards can also help reduce regulatory compliance costs. As standards become more widely understood and used, regulatory demands for duplicate information in slightly different forms should decline. Standardization could make it easier for firms to communicate with regulators, ensuring that they precisely define what it is they are measuring and why.

B. Supervisors and other regulators benefit

Regulators benefit from quality by being better able to spot risks in firms, markets and across the system, and to design effective, efficient and appropriately tailored remedies to mitigate those risks. The interests of prudential regulators and regulated firms should be well aligned: Each wants firms to be profitable, well managed and well capitalized, taking risk knowingly and managed appropriately.

But a core challenge for bank and other supervisors is that a regulated firm inevitably knows far more about its operations and risk exposures than its supervisors. Data standardization will not eliminate but it can help to mitigate this asymmetry. Particularly when bank supervisors regularly rotate among firms, as they now do, standardized data can enable regulators to compare risk management practices across firms and quickly assess a new firm’s practices. Standardized data could, in time, help regulators to develop or improve the robustness of their techniques to assess firms’ risk and risk-management capabilities.

The Fed’s stress tests are an example of a domain where greater standardization enhances the accuracy of supervisory efforts. Clearly, both regulators and firms have made progress in this area since the GFC. The challenge is to do that more effectively than in current practice. Were the Fed to embrace data standards in stress testing and other supervisory activities, they and the firms they supervise would benefit.

Using standards facilitates sharing data appropriately and securely so that firms and regulators can use and understand the same data within and across jurisdictions, and so that duplication of data reporting can be reduced if not eliminated. Both benefit by building and sustaining the trust, accountability, adaptability and efficiency in finance essential to its functioning. These efficiency gains could help boost economic productivity by allocating funds to their highest uses and could also prove crucial in the ever-more-competitive and ever-more-global financial markets.

C. Society benefits

The microprudential benefits of standardization are significant. However, the most important public value served by greater standardization is the potential for enhanced system-wide resilience. Financial crises are informational events. Many of the classic forms of market dysfunction, from runs by short-term creditors like depositors, to the refusal of liquidity providers to continuing playing that role, are, at least in part, the byproduct of incomplete or asymmetric information about events or changes in circumstances.

As a consequence, when it comes to thinking about the system as a whole, the alignment mentioned above between the interests of prudential regulators and regulated firms is asymmetric. System-wide, financial stability or macroprudential policies are needed because neither firm risk management nor microprudential policies are sufficient to deal with the system-wide, externalities and market failures that can arise from asymmetric information and mispriced guarantees. And, as a practical matter, market participants and regulators are constantly operating with only a subset of the information that could be pertinent to the decisions they are
making. Given the costs and other frictions inhibiting data compilation, distribution and analysis, this is inevitable. The critical point is that these information gaps are an increasingly important source of systemic fragility.

The flip side of the challenge posed by information gaps is the opportunities that they create. When market participants are reluctant to transact because of a lack of information that recent events have made pertinent, filling in those gaps can be an important mechanism through which regulators help restore functionality. High-quality information about the location of capital deficiencies, interconnections among firms, and other vulnerabilities can also help regulators identify and address fragilities in a timely fashion. Data standardization is one component of a broader information infrastructure that can and must be built in “peacetime” if we expect regulators to have any chance of producing timely and relevant information once panic sets in.

Of course, the benefits of risk management and financial stability activities are often not obscured during periods of calm. More generally, systemic resilience is a good that benefits far more than the firms engaged in standardization, creating the typical challenges that arise with positive externalities.

Nonetheless, one lesson of the financial crisis is that vigilance – and the investments needed to achieve it – in peacetime is essential. Because the next crisis is unlikely to replicate those of the past, we do not know in advance what information we will need, expanding the infrastructure that must be built in place in advance. Data standardization is critical to that undertaking. This is one of the core lessons of 2008 that has only been partially realized thus far.

III. Why it is often quite hard

Having argued that market participants and regulators both stand to gain from improved data standardization, the natural question to ask is why we haven’t seen more of it? That’s even more perplexing given that industry has long recognized and supported the need for standardization. To quote from the seminal “Linchpin” paper on entity identification: “The financial services industry has been exploring the issue of unique entity identification for decades. More recently, several [industry] efforts have been made to advance the idea of a standard [Legal Entity Identifier, or] LEI, but competing priorities, funding issues, and an evident lack of industry focus have kept such a standard, and the benefits it could have yielded, from being implemented.”¹²

To be sure, in the decade since the crisis, there has been meaningful progress towards the development of and mandates to employ data standards. In derivatives markets, for example, regulators globally have started to require using the Legal

Entity Identifier (LEI) to identify parties to financial transactions (“who is who”) and the relationships among those parties (“who owns who”). LEIs have the potential to make precise entity and counterparty identification frictionless, thus improving the quality of the information that firms use to manage risk and report to supervisors. There has also been some progress made in efforts to standardize transactions and products (“who owns what”), which is equally important in interconnected, global markets. At the same time, improvements in data standardization remain sluggish and are a long way from addressing even those problems revealed to be pressing in the GFC. LEI adoption remains incomplete and progress on other fronts has been even slower. The use of LEIs in derivatives and, increasingly, in other markets is a good example of a post-crisis reform that has been slow but meaningful. Work is also progressing to create coherent reference

The LEI is a 20-digit, alpha-numeric code that clearly and uniquely identifies parties to financial transactions; specifically, the legal entities within companies participating in global financial markets. Thanks to public and private collaboration, use of the LEI system is now widespread—to date, more than 1.3 million LEIs are in use—but it is far from ubiquitous. Global Legal Entity Identifier Foundation LEI Statistics (2018), online: https://www.gleif.org/en/lei-data/global-lei-index/lei-statistics.

These are established with so-called Level 2 LEI data; see https://www.gleif.org/en/lei-data/access-and-use- lei-data/level-2-data-who-owns-whom#


See infra Part II.

See, e.g., Committee on Payments and Market Infrastructures & Board of the International Organization of Securities Commission, Harmonisation of Critical OTC Derivatives Data Elements (other than UTI and UPI) (2018), online: <https://www.bis.org/cpmi/publ/d175.pdf>.


13 The LEI is a 20-digit, alpha-numeric code that clearly and uniquely identifies parties to financial transactions; specifically, the legal entities within companies participating in global financial markets. Thanks to public and private collaboration, use of the LEI system is now widespread—to date, more than 1.3 million LEIs are in use—but it is far from ubiquitous. Global Legal Entity Identifier Foundation LEI Statistics (2018), online: https://www.gleif.org/en/lei-data/global- lei-index/lei-statistics.

14 These are established with so-called Level 2 LEI data; see https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-2-data-who-owns-whom#


17 See infra Part II.

18 See, e.g., Committee on Payments and Market Infrastructures & Board of the International Organization of Securities Commission, Harmonisation of Critical OTC Derivatives Data Elements (other than UTI and UPI) (2018), online: <https://www.bis.org/cpmi/publ/d175.pdf>.

data – the data that help define financial instruments and the taxonomy or ontology to establish how instruments relate to one another.²⁰ Work is likewise progressing to move from a sometimes duplicative documents-plus-data financial reporting regime to data-centric reporting.²¹

That slow progress is not from lack of industry support. Since the “Linchpin” paper was published in 2011, industry has strongly supported standards initiatives and provided leadership in this area. The Data Coalition, a non-profit supported by industry, has supported the need for data standards, both in financial services and in the Federal government.²² Market participants on the Financial Research Advisory Committee of the Office of Financial Research have urged broader adoption.²³

Moreover, industry and regulators have worked together on these initiatives. The industry supported the creation of the Global Legal Entity Identification Foundation (GLEIF).²⁴ Robin Doyle recently noted on behalf of JP Morgan that recent efforts at data standardization have entailed “a unique collaboration among industry, regulators, and other standard-setting bodies to develop data and reporting standards and to create global systems to manage and maintain the standards.”²⁵

Despite this meaningful progress and support, however, the improvements have far lagged the need. It has been more than a decade since the GFC hit its nadir and not one of the initiatives just described has been fully implemented; some have been poorly implemented, requiring rework; and far more work remains to be done in other domains. Part of what makes data standardization a useful lens is that progress is heterogeneous. It’s not that it’s not happening; but it is happening inconsistently across substantive domains and across jurisdictions. Understanding where it is proceeding more quickly and the challenges that have impeded effective standardization shows that significant improvements are possible, but they are not assured.

²⁰ One example is the Financial Industry Business Ontology (FIBO). EDM Council, FIBO Primer (2018), online: <https://cdn.ymaws.com/edmcouncil.org/resource/collection/16D6DC67-430E-4F75-9E07-08B1EC228091/FIBO_Primer_v0.2.pdf>
²² Data Coalition https://www.datacoalition.org/issues/policy-agenda/
This Part attempts to answer the question of why data standardization has been so slow and difficult given its recognized importance and the apparent alignment of public and private interests.

A. Short-term costs, long-term gains

Standardizing data is not a free good. The costs of developing standards, testing them, retooling firm and regulatory systems to use them, and working out the kinks in implementation are considerable, and the costs of these investments are incurred early on. The benefits follow with a lag, and the benefits aren’t restricted to those who incur the bulk of the costs.

In addition, the connection between bad data and financial instability is far harder to understand than the moral hazard and unfairness of bailouts, and far less tangible than corrupt bankers who can be blamed when something goes wrong. So while some regulators are engaging in a time-consuming process of helping to design and implement well-crafted data standardization requirements, universal regulatory support for data standards, at least in the United States, is lacking. Few voters will demand that their elected representatives legislate data standardization (although a bipartisan trio of visionary Members did exactly that\textsuperscript{26}).

Collective action problems, positive externalities, and other classic but still meaningful challenges associated with governance and incentives compound the challenge.\textsuperscript{27} This externality creates a free rider issue -- gains aren’t equally shared so there’s a first mover disadvantage. As in other public finance projects which entail upfront costs and an uncertain stream of long-term benefits to all, mandates and incentives are required to promote action.

B. Coordination challenges

Another core challenge is that data standardization requires coordination from diverse parties, not only government and market actors, but various groups within those broad categories. Thus the problem is not only a coordination problem of the kind recognized by Mancur Olson decades ago, but also a cultural one.\textsuperscript{28} Effectively implementing data standardization requires groups who think about issues in slightly different ways--business leaders, risk managers, legal officers, data scientists, and financial economists--to understand how others see the world. They must find vocabulary and figure out collectively how to devise standards that can serve related but not identical aims.

An additional challenge: Private and public actors will only fully benefit from switching their current proprietary systems to new models that use standardized


\textsuperscript{27} Office of Financial Research, \textit{Collective Action: Toward Solving a Vexing Problem to Build a Global Infrastructure for Financial Information} (2017), online: <https://www.financialresearch.gov/briefs/files/OFRbr_2017_01_LEI.pdf>; see also infra Part II.

\textsuperscript{28} Mancur Olson Jr., \textit{The Logic of Collective Action} (Harvard University Press, 1971).
inputs when there is widespread adoption. So maintaining legacy systems while putting new ones in place adds to the costs. These network externalities underscore the need for government coordination and mandates.

C. Inadequate IT and data infrastructure at firms

In theory, private firms stand to enjoy an array of gains from well-designed data standards. To fully enjoy these gains, however, firms need the technology and enterprise-wide data management and governance practices that will allow them to use the data to better identify and manage their risks.

The evidence available suggests that many firms are not yet in a position to harness these gains. According to a 2013 Report by the Basel Committee on Banking Supervision:

One of the most significant lessons learned from the global financial crisis that began in 2007 was that banks’ information technology (IT) and data architectures were inadequate to support the broad management of financial risks. Many banks lacked the ability to aggregate risk exposures and identify concentrations quickly and accurately at the bank group level, across business lines and between legal entities.

Although there has been significant progress, these IT-related challenges remain. A 2018 Progress Report brings this point home. According to the Basel Committee, “banks have increasingly recognised the value of implementing the Principles and have stepped up their efforts to do so,” progress towards those goals has been insufficient (BCBS, 2018). Only three G-SIBS are fully compliant with the principles, and most have not taken a sufficiently comprehensive approach to assuring the principles’ goals. Further progress will be essential to produce comprehensive, high quality, interoperable data.

D. Fragmented regulatory structure and path dependence

In the United States, the balkanized regulatory structure is an obstacle to adoption of data standards. Regulators are free to use whatever mode of collection they choose. The CFTC and the SEC did choose to require the LEI for reporting swap and securities-based swap data.29 But in most cases the agencies request, rather require its use.30 This may reflect a lack of regulatory understanding regarding the benefits or a failure to adequately value gains that accrue outside a particular regulator’s mandate.


Adding to the challenge, no higher authority can compel agencies to use a particular standard. The Financial Stability Oversight Council, which was created to solve collective action problems among regulators, can make recommendations but it has no authority to require member agencies to adopt data standards. The other post-crisis change, the newly formed Office of Financial Research, does have authority to require standards in data it collects or that member agencies collect on behalf of the Council, but it has no authority to require their use for other data those agencies collect.31

In Europe, the situation is the opposite: Despite the fragmentation resulting from 28 sovereigns in the EU, ESMA and other pan-European regulators have required the use of the LEI in all but a handful of reporting requirements.32 And even post-Brexit, it seems likely that the UK will retain those requirements. That Europe has moved more quickly than the United States to impose a broadly applicable requirement that parties use the LEI highlights the challenge of entrenched interests and the way newcomers can leapfrog over incumbents. More specifically, Europe is in the position of trying to create a Capital Market Union. In contrast to the United States where capital markets have long provided roughly two-thirds of all intermediation, Europe remains more reliant on banks and capital flows across countries thus remain lower than would seem optimal. This creates the motivation needed to adopt new rules, at least some of which—like the broad LEI requirement—are closer to current state of the art than the United States.

IV. A Look Ahead

We believe that data standards have benefits that greatly exceed the costs, but that, as a classic public good, government must solve the collective action problem by mandating their use. That said, data standards work well only when well designed, and that also requires meaningful participation from industry and public-private cooperation and partnerships to align actions and interests.

We close with two examples that illustrate both the potential of such participation and the need for further progress.

a. Derivatives.

New, post-crisis rules sometimes require and otherwise strongly incentivize central clearing of derivatives.33 These changes have brought a host of new

31 *Dodd-Frank Wall Street Reform and Consumer Protection Act*, 12 U.S.C. (2018), §5343(c)(2). Helpfully, the OFR has helpfully proposed a rule for collecting bilateral repo data that requires the use of standards.


33 See generally *Dodd-Frank Wall Street Reform and Consumer Protection Act* title VII, July 21, 2010
challenges and opportunities, only some of which are informational. One of those aims was to promote post-trade transparency in derivatives for both market participants and regulators through the establishment of swap data repositories (SDRs, TRs in Europe) and trade reporting requirements. As mentioned earlier, in the United States, the CFTC required the LEI in swap reporting, which improved the quality and the value of those data. The industry has also supported the use of standards. But implementation fell short. Initial efforts to collect swap data left it up to the CCPs or their SDRs to decide what data elements should be collected, and how they should be reported. So each SDR reported data in its own, idiosyncratic way, with different data elements in different formats. The result, in the words of a senior official, was a “meteor storm.”

There has been meaningful progress. According to a recent CFTC Report, “in 2014, roughly half of all reports for the highly standardized credit default swaps (CDS) lacked complete price information, and approximately 15% of all CDS trades lacked a legal entity identifier, making it difficult to identify the counterparty. By early 2018, roughly 95% of all CDS trades had complete counterparty and price information.”

Nonetheless significant challenges remain. Swap data repository requirements are not yet harmonized between the SEC and CFTC; nor are they harmonized between the United States and EU—another area in which public-private cooperation could help solve this coordination problem. The sheer volume of granular data and their lack of comparability has limited the ability to manage and process the data collected and to extract useful information. As a result, regulators cannot fully compare the data currently being produced with other datasets to effectively oversee or stress test CCPs.

In response, the CFTC announced in July 2017 that it was launching a new review of the swap data reporting regulations, and it subsequently announced a wholesale review of swaps regulations and reporting. Leadership is key, and the CFTC is now trying to provide that.


Id.; United States Commodity Futures Trading Commission, Roadmap to Achieve High Quality Swaps Data (2017), online: <https://www.cftc.gov/sites/default/files/idc/groups/public/@newsroom/docum
The role of a leader here is to identify and draw attention to key issues, keep other parties motivated, learn from mistakes, work with key stakeholders and counterparts at home and abroad, and use the force of law needed to produce uniformity of adoption once a useable set of standards and data elements have been identified. In that regard, as noted earlier, the Financial Stability Board has provided leadership through its Data Gaps Initiative to coordinate efforts to improve the scope and quality of global financial data.\textsuperscript{38}

Another key element is willingness by all to listen, learn and compromise. As the experience in this domain reflects, going it on one’s own doesn’t work. Uniformity is key, and can only be achieved when there is broad buy-in and a well-designed standard. Similarly, the more regulators are willing to work together to produce standards that can serve multiple aims, the greater the potential efficiency gains in terms of reduced reporting requirements. These small compromises can be hard won but over time can enable a far more productive dialogue by getting disparate regulators and market participants on the same page.

Industry can also play an important leadership role in helping to develop usable standards and to work with regulators to help them understand how data can help them achieve specified aims. As the flurry of post-crisis reforms recede but the volume of financial activity increases, and fragilities remain, now is a ripe moment for progress.

b. Fintech

The current excitement surrounding fintech may seem to make standardization all the more banal. Why is there a need for standardization and all of the associated upfront costs when innovations such as Distributed Ledger Technology (DLT), which reliably stores information about transactions, and artificial intelligence, might obviate the need for such infrastructure?

We believe that just the opposite is closer to the truth. Technology demands precision and interoperability, and standards help provide both.\textsuperscript{39} DLT and machine learning, for example, demand quality data inputs.\textsuperscript{40} Industry-wide technology

\textsuperscript{38} Financial Stability Board, “Addressing Data Gaps,” online: <http://www.fsb.org/what-we-do/policy-development/additional-policy-areas/addressing-data-gaps/>

\textsuperscript{39} It is worth noting here that both standards and a governance structure to standardize data across a firm are needed.

standards are also critical to avoid repeating past mistakes, with siloed solutions that can't be reconciled.

History also shows that financial-system vulnerabilities often grow just outside the regulated perimeter. With the financial system evolving at an ever more rapid rate, and the potential for new entrants to gain market share quickly, the threat of unexpected systemic disruptions cannot be ignored. New modes of gathering and analyzing disparate bits of data may yield real gains, but without quality data inputs, they could also risk create new information-related vulnerabilities. When the limits of these developments come to the fore, as they did with securitization a decade ago, high-quality, credible information will be critical to minimizing the disruption that results.

V. Data Standardization as Microcosm

The core claim here is that data standardization is important and should be prioritized. But the essay also has an important second lesson. The firms and countries that have long been the financial leaders are facing increasing threats for that dominance. Fintech and globalization are among the forces that make it possible for power dynamics to change quickly, even if the top players seem inevitable right up until their downfall. That well-designed standardization has the potential to enhance the operations of firms in addition to promoting resilience suggests that delays in implementation may well signal bigger problems. Firms and regulators can get trapped by inertia. Legacy assets within firms may go underutilized or pose risks that are not justified at the entity-level because of a failure to accept the short-term costs necessary to realize long-term gains. Turf wars and more subtle forms of provincialism among regulators can also impede widespread adoption of standards. To remain dominant in today's dynamic financial system requires that countries and firms be ahead—or at least not too far behind—in implementing value-creating reforms. When that is not happening, and the slow pace and mixed quality of data standardization suggests it is not, more monumental change may well be lurking in the shadows.

The core take away is this: Details matter. Stylized accounts of finance suggest regulation is only needed to address market failures, and those failures are most apparent when interests conflict. Although there are positive externalities from standardization, there is also a lot of overlap between the types of reforms that will benefit industry and those that will promote public aims. The slow progress on data standardization suggests that frictions too banal to merit serious academic inquiry can have first-order effects on progress, or lack thereof. A willingness to grapple with these details, institutional and substantive, is key to producing financial regulatory schemes that work as well in practice as they do in theory.

Conclusion

We close the way we opened: Information is the lifeblood of finance. New technologies are only accentuating its importance and making urgent the need to govern and manage it. As finance continues to evolve, the capacity of the financial system to achieve the promise of enhanced efficiencies will depend in part on its
capacity to gather, analyze and make productive use of available information. Information is also critical to promoting the resilience and health of the financial system. Data standards that are well designed, broadly understood, and widely used are a critical component to these undertakings. The process of designing standards forces communication and lays the groundwork for those standards to evolve when and as needed.

As the pace of finance continues to reach ever more dizzying speeds, the value of high-quality information and the threats posed by information gaps continue to grow. Given the myriad frictions that stand in the way of optimal policy, leadership, creativity, and a willingness to look to the future, and work across firm, industry, and national bounds is critical to success. Some progress has been made already. More is needed, and we think possible. Only vision and leadership are the missing ingredients.
Works Cited


Committee on Payments and Market Infrastructures & Board of the International Organization of Securities Commission. 2018. “Harmonisation of Critical OTC Derivatives Data Elements (other than UTI and UPI).”
https://www.bis.org/cpmi/publ/d175.pdf.


Data Coalition. n.d. :The Financial Transparency Act.”


https://cdn.ymaws.com/edmcouncil.org/resource/collection/16D6DC67-430E-4F75-9E07-08B1EC228091/FIBO_Primer_v0.2.pdf


FSB & IMF. 2009. The Financial Crisis and Information Gaps, Report to the G-20 Finance Ministers and Central Bank Governors:


Global Legal Entity Identifier Foundation. 2018. “Regulatory Use of the LEI.”


https://www2.usgs.gov/datamanagement/plan/datastandards.php
The European Corporate Governance Institute has been established to improve corporate governance through fostering independent scientific research and related activities.

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.
Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute’s Web-site (www.ecgi.global/content/working-papers) or SSRN:

|------------------------|----------------------------------------|

www.ecgi.global/content/working-papers