Setting Standards for Microdata[1]

When industry and regulators work together, all can benefit from improved data quality

By Cornelius Crowley
Deputy Director and Chief Data Officer
U.S. Office of Financial Research

Introduction

Financial market regulatory authorities today collect and analyze more microdata than ever before, prompted by the lessons of the 2007-09 global financial crisis. At the same time, financial companies benefit from investing in and using microdata to improve internal risk management and reporting. The experiences of the U.S. Office of Financial Research (OFR) in working with regulatory authorities and with industry participants underline the need for the public and private sectors to collaborate to get the most out of these data.

Microdata have been defined as data describing "individual entities, transactions or instruments."[2] Collecting data at this micro level permits a fine-grained view of a financial system or market. The benefits of granularity accrue from collecting information about important details — such as "name," "quantity," and "maturity date" — for each of the individual entities, transactions, instruments, or products in a dataset. Data can then be linked and integrated for more effective macroprudential supervision, that is, for monitoring of risks across the financial system. Regulators and the financial industry have made some progress toward this since the 2008 crisis, when data — even when available — were not sufficiently granular or detailed to support key decisions. The crisis highlighted the need for identification standards such as the legal entity identifier (LEI) and for clearer standards for regulatory data collections. After eight years, though, continued collaboration still is needed.

In principle, financial companies and their regulators should use the same underlying data for risk management. Companies use these microdata for recordkeeping and risk management. The data also could be used for regulatory reporting. If the microdata were well-structured and

---

[1] This document was prepared for the Eighth ECB Conference on Statistics, July 2016. Views and opinions are those of the author and do not necessarily represent official positions or policy of the Office of Financial Research or the U.S. Treasury Department.

[2] See Bruno Tissot, "Closing information gaps at the global level — what micro data can bring," Bank for International Settlements, 2015 (available at https://www.bis.org/ifc/events/ws_micro_macro/tissot_paper.pdf, accessed June 30, 2016).

standardized, regulators could link and integrate them to create well-defined aggregated systemic data.

Companies say that when regulators request aggregated data that draw on underlying microdata that those companies do not already collect or create, compliance becomes expensive. To address this, regulators can as much as possible base data collections on information gathered by companies in their normal course of business. That should reduce the compliance burden on firms. The OFR recently discussed this and other principles for regulatory data collections in a Viewpoint Series paper reviewing best practices for such collections.[3]

To promote mutual benefits, the OFR has emphasized industry and regulatory cooperation to address issues related to the scope, quality, and accessibility of financial data.

To fully benefit from microdata, companies need to improve their back-office data practices and systems so they can more easily produce, link, and integrate information. Company data often suffer from lack of structure and standardization. For instance, many companies have not consolidated their data systems after mergers and acquisitions. Efforts to improve data practices have been defined on a global level by the Basel Committee on Banking Supervision's "Principles for effective risk data aggregation and risk reporting," known as BCBS 239. Evidence, though, shows the industry has been slow to meet those principles.

This discussion will focus largely on development, application, and use of data standards to link and integrate data, and thus improve data quality. That is critical for the most effective use of microdata by both industry and regulators. Data standards provide the common language that, among other things, enables: (1) unique identifiers and classifiers of financial data and (2) elements that provide the descriptive data for entities, positions, and transactions.

Research into financial stability issues requires the ability to assemble and view the complex structures formed by legal entities, the financial instruments they hold, and the transactions that give rise to their positions. For example, knowledge of these structures is critically important to risk managers evaluating exposure to market shocks by entities holding portfolios of financial instruments. These structures also reveal the interconnections among entities, not just inside a single complex financial conglomerate, but across the many conglomerates and networks formed through legal agreements that create ownership and control relationships.

However, the standards could and should be more effectively used, by companies for their own data management, and by regulators in their collections, to improve and ensure data quality and the effective use of microdata. As Richard Berner, Director of the OFR, said in a speech earlier

[3] See Office of Financial Research, "Developing Best Practices for Regulatory Data Collections," OFR Viewpoint 16-01, May 10, 2016 (available at https://financialresearch.gov/viewpoint-papers/files/OFRvp-2016-01_Best-Practices-Data-Collection.pdf, accessed June 8, 2016).

this year, "The use of data standards to assure data quality is far from ubiquitous, and some data collections are proceeding without appropriate standards. Swap data reporting is still fraught."[4]

Created in response to the global financial crisis, the OFR's statutory purposes include "standardizing the types and formats of data reported and collected" on behalf of the Financial Stability Oversight Council (FSOC) and "assisting [FSOC]-member agencies in determining the types and formats of data … to be collected."[5] As part of its mission, the OFR has consistently called for universal adoption of the LEI, which uniquely identifies the parties to each financial transaction. It is a critical — but not sufficient — standard for aggregating risks across entities and for understanding how those risks can spread and threaten financial stability. In addition, standards are needed to uniquely identify transactions and products. For derivatives market reporting, the OFR actively participates with both the private sector (the International Organization for Standardization (ISO)) and the public sector to develop global standards for the unique transaction identifier (UTI), the unique product identifier (UPI), and other critical data elements.

This paper will begin with an overview of the OFR and its data mission, including its focus on data standards and best practices for data collection. Then, it will discuss the OFR's experience by looking at three case studies that shed light on benefits from public-private collaboration on data standards: (1) LEI, (2) standards for swap data repositories, and (3) recent pilot data collections on securities financing. It will conclude with an examination of lessons learned and discussion of next steps.

The OFR and its data mission

The Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), the U.S. financial system reform law adopted in reaction to the global financial crisis, established the OFR. The law gave the OFR the authority to collect and standardize financial data, to perform essential research, and to develop new tools for measuring and monitoring risk in the financial system.[6] By working to improve the scope, quality, and accessibility of financial data, the OFR

---

[4] See Richard Berner, director, Office of Financial Research, speech at the Financial Data Summit: Data Standards Mean Business!, hosted by the Data Transparency Coalition, March 29, 2016, Washington, D.C. (available at https://financialresearch.gov/public-appearances/2016/03/30/financial-data-summit/, accessed June 7, 2016).

[5] See U.S. Congress, *Dodd-Frank Wall Street Reform and Consumer Protection Act*, Public Law 111-203, Washington: U.S. Government Printing Office, July 21, 2010, § 153(a)(1),(7) (available at https://www.gpo.gov/fdsys/pkg/PLAW-111publ203/html/PLAW-111publ203.htm, accessed June 8, 2016).

[6] See Dodd-Frank Act; also see Office of Financial Research, *2012 Annual Report*, Washington: Office of Financial Research, July 12, 2012, p. iii (available at https://financialresearch.gov/annual-reports/files/office-of-financial-research-annual-report-2012.pdf, accessed June 8, 2016).

aims to expand its capacity and that of its stakeholders to spot and track vulnerabilities in the financial system.[7]

To OFR, improved **scope** means data must be comprehensive for a broad view across the financial system. As well, it must be granular, to help identify emerging activities and potential sources of systemic risk. Improved **quality** means data must be high quality to inform good policy decision-making and good risk management. Although data standards do not guarantee quality, they are a key to achieving uniqueness, accuracy, consistency, and completeness. Improved data **accessibility** means data must be accessible to key stakeholders for analysis of risks across the system. Sharing data appropriately with and among regulators and companies helps to align interests, reduce burden, support risk management, and facilitate cross-market analysis. Scope, quality, and accessibility are essential and interdependent. For example, the application of standards is necessary, but not sufficient, for quality and supports the effective sharing of financial data, because standards enable precise conversations about the data to be shared.[8]

Without standards, comparing, aggregating, and analyzing the data essential for financial stability analysis becomes effectively impossible. The financial crisis demonstrated to regulators and industry that standardizing data collected from financial services companies is necessary for effective oversight of the financial system and its parts.

The financial services industry, unlike many other industries, has been slow to agree on consistent data standards and formats required for high-quality data. Solving this collective action problem and speeding development of data standards requires concerted effort.

Embracing standards offers real benefits for the industry. For instance, at a 2011 conference sponsored by the OFR and FSOC, Karla McKenna, a top Citibank executive who has long been active in international standards-setting efforts, laid out some industry uses for the LEI. Among them were facilitating aggregation and analysis of clients' transactions and interactions with counterparties across their banking relationships, as well as making key business processes more efficient. Those key processes include managing business relationships, such as "Know Your Customer"; identifying parties involved in financial transactions; managing counterparty and concentration risk; improving "Straight Through Processing" of transactions electronically without manual involvement; and simplifying reconciliation activities.[9]

Regulators internationally recognize that because markets are global, well-defined data standards must also be global for the public and private sectors to see the benefits. Thus, in 2015, along

---

[7] See Office of Financial Research, *2015 Annual Report*, Washington: Office of Financial Research, Jan. 27, 2016, p. i (available at https://financialresearch.gov/annual-reports/files/office-of-financial-research-annual-report-2015.pdf, accessed June 8, 2016).

[8] See Berner.

[9] See Karla McKenna, "Data Standards for Financial Services; Focus on Legal Entity Identifier," Dec. 1, 2011, at Conference on The Macroprudential Toolkit: Measurement and Analysis (available at https://www.treasury.gov/initiatives/wsr/Documents/OFR-%20McKenna.pdf, accessed June 6, 2016).

with the Bank of England and the European Central Bank, the OFR sponsored two workshops on "Setting Global Standards for Granular Data."[10] These workshops were designed to bring together participants to discuss global standards for the increasingly granular data central banks and financial regulators now have at their disposal. A third workshop is planned for early 2017.

In carrying out its mission, the OFR has conducted two pilot data collections on securities financing, as well as working with FSOC member agencies and international regulatory authorities to analyze and improve standards for derivatives data. In doing so, it has accumulated experience in the best practices for such collections. The OFR outlined those best practices in detail in a Viewpoint Series paper published in May 2016.[11] That paper suggested that regulators designing any data collection focus on "meticulous planning, data transmission controls to avoid introducing errors, and validation techniques to ensure the data are fit for purpose." It also said that regulators should understand the business activities and processes involved in producing the data to work most effectively with market participants that report data. Doing so is key to getting good data while minimizing regulatory burden.

The industry also has a responsibility. To take full advantage of the promise of microdata, the financial industry must improve its information technology practices and thus the quality of the microdata it produces. The outline for doing so was laid out in 2013 by the Basel Committee on Banking Supervision in BCBS 239, its risk data aggregation and reporting principles.[12] The principles call for banks to develop the ability to aggregate risk exposures and identify concentrations quickly and accurately at the bank group level, across business lines, and between legal entities. That would allow data aggregation both for regulatory stress testing and for internal uses.

For the largest banks, those designated global systemically important banks (G-SIBs), the principles went into effect in January 2016. But as of December 2015, the Basel Committee expected most of the G-SIBs to miss the deadline.[13] The committee called for an independent evaluation of progress, but no evaluation has yet been published. However, a private-sector

---

[10] See Bank of England, "Setting Global Standards for Granular Data: Joint Bank of England, European Central Bank and US Office of Financial Research Workshop, 15-16 January 2015 – Programme" (available at http://www.bankofengland.co.uk/research/Pages/conferences/programme0115.aspx, accessed June 6, 2016); also see Office of Financial Research, "Second Joint Workshop by the Bank of England, European Central Bank, and the OFR, 'Setting Global Standards for Granular Data,'" Oct. 6, 2015, (available at https://financialresearch.gov/conferences/2015/10/06/setting-global-standards-for-granular-data/, accessed June 28, 2016).

[11] See Office of Financial Research, "Developing Best Practices for Regulatory Data Collections."

[12] See Basel Committee on Banking Supervision, "Principles for effective risk data aggregation and risk reporting," BCBS 239, Bank for International Settlements, January 2013 (available at http://www.bis.org/publ/bcbs239.pdf, accessed June 8, 2016).

[13] See Basel Committee on Banking Supervision, "Progress in adopting the principles for effective risk data aggregation and risk reporting," third report, Bank for International Settlements, December 2015 (available at http://www.bis.org/bcbs/publ/d348.pdf, accessed June 8, 2016).

consulting firm, Chartis, reported in January 2016 that no G-SIBs were compliant in mid-2015, and that just 5 percent would be compliant within a year. "It is obvious that there is no chance the BCBS 239 deadline will be met," Chartis said.[14]

Case studies: The OFR experience with data standards

The OFR actively participates in U.S. and global efforts to improve data quality through the development of financial data standards. Three case studies shed light on the importance of public-private partnerships to these efforts.

Legal Entity Identifier (LEI)

The universal adoption of the LEI has been a top OFR priority from the first, and the Office has worked with the private sector and with both U.S. and global regulators to encourage adoption. The development of the system illustrates the value of shared goals and cooperation between the public and private sectors. Government regulators and private financial firms recognized the tremendous benefit that the LEI could bring, such as supporting management of counterparty risk. Industry provided the technical expertise. Officials solved the collective action problems that were hurdles to development by establishing a process and encouraging market participants to adopt and use the identifier. Progress on adoption has been most significant in the derivatives industry, where the financial crisis revealed serious difficulties in managing counterparty risk.

However, because adoption is not yet universal, the system is not as helpful as it could be. That is why OFR advocates that regulators mandate the use of the LEI in regulatory reporting.

The LEI is a 20-character alphanumeric code developed by the International Organization for Standardization and described in the ISO 17442 standard. The code is assigned to legal entities that are counterparties to financial transactions. It is frequently described as similar to a bar code for each entity. In its current format, a publicly accessible LEI database provides "business card" information such as legal name, address, country of formation, legal form, and business registry information.[15] A next step, to be phased in beginning in late 2016, will add information about the parents and subsidiaries of each entity.[16]

---

[14] See Chartis, "Risk Data Aggregation & Reporting Solutions 2016," January 2016 (available at http://www.chartis-research.com/research/reports/risk-data-aggregation-reporting-solutions-2016, accessed June 6, 2016).

[15] See Global Legal Entity Identifier Foundation, "About the LEI," undated (available at https://www.gleif.org/en/lei-focus/about-the-lei, accessed June 6, 2016).

[16] See LEI Regulatory Oversight Committee, "Collecting data on direct and ultimate parents of legal entities in the Global LEI System — Phase 1," March 10, 2016 (available at http://www.leiroc.org/publications/gls/lou_20161003-1.pdf, accessed June 3, 2016).

As of mid-June 2016, almost 450,000 LEIs had been issued. About a quarter of them were in the United States; most of the rest were in European Union nations.[17] While that is an impressive number, it is just a tiny fraction of the hundreds of millions of estimated eligible entities.[18] (However, many of those eligible entities are likely to be smaller firms.)

The LEI is a fundamental standard that supports optimum use of microdata. By uniquely identifying the parties to each transaction, it provides a way for market participants and regulators to know exactly who is involved in each deal. This identification permits transparency into counterparty risk and improves firm-level risk management. Such insight was missing at the time of the Lehman Brothers collapse. Firms had difficulty assessing their exposures. And at that time of unprecedented financial system stress, regulators and policymakers found themselves trying to aggregate data from systems that had proprietary naming conventions for both counterparties and instruments.[19]

As the LEI becomes more widely used, it is expected to cut costs and improve risk management. For firms, these savings will come largely from operational efficiencies such as reducing transaction failures; lowering data reconciliation, cleaning, and aggregation costs; and cutting regulatory reporting costs. The LEI will provide long-term benefits to companies by identifying their counterparties and customers, and improving internal risk management. It is premature to estimate how much the financial industry will save by adopting the LEI, but industry estimates range from $300 million to $10 billion.[20]

Although there have been many attempts over the years to establish a system to identify the parties to financial transactions, the roots of the current global LEI system can be traced back to the crisis aftermath. In 2011, the Cannes summit of G20 leaders directed the Financial Stability Board (FSB) to establish the LEI system. In 2012, FSB reported back to the G20 with a proposed structure, and the G20 leaders approved it.

That summer, FSB set up the Private Sector Preparatory Group, an international group of nearly 300 private-sector representatives who offered opinions on the system in workshops, conference calls, and an online collaboration space. Regulators credit participants with "detailed and

---

[17] See Global Legal Entity Identifier Foundation, "LEI Statistics," undated (available at https://www.gleif.org/en/lei-data/global-lei-index/lei-statistics, accessed June 17, 2016).

[18] See LEI Regulatory Oversight Committee, "The Global LEI System and regulatory uses of the LEI," Nov. 5, 2015, p. 22 (available at http://www.leiroc.org/publications/gls/lou_20151105-1.pdf, accessed June 6, 2016).

[19] See Office of Financial Research, *2012 Annual Report*, Washington: Office of Financial Research, July 12, 2012, pp. 107-119 (available at https://financialresearch.gov/annual-reports/files/office-of-financial-research-annual-report-2012.pdf, accessed June 8, 2016).

[20] See Office of Financial Research, "Legal Entity Identifier — Frequently Asked Questions," undated (available at https://financialresearch.gov/data/legal-entity-identifier-faqs/, accessed June 6, 2016).

substantive" work on operational issues and on the decision to take an accounting consolidation approach to developing relationship data.[21]

Private sector participation remains key. Structurally, the Global Legal Entity Identifier Foundation (GLEIF), the international body created to support implementation of the LEI system, relies on public-private partnership. The 15-member GLEIF Board of Directors consists mostly of private-sector representatives, including bankers, academics, and technology company executives. The LEI Regulatory Oversight Committee, which coordinates and oversees the LEI system, includes public authorities from more than 40 countries. The OFR's Chief Counsel was the committee's first chairman and remains a member of its executive committee.

The financial industry has been broadly supportive of the LEI.[22] According to a financial industry trade group, the Securities Industry and Financial Markets Association, "The lack of a standard identification system for financial counterparties makes it difficult for financial firms to develop a consistent and integrated view of their exposures."[23]

One thing that's notable about LEI data: The quality appears high. The registering entity provides the data. The issuing organization verifies them against other sources. The information is reviewed and revalidated annually by the entity and the issuer. GLEIF reports monthly on data quality; in recent months, the global quality score has generally been between 98 percent and 99 percent. Some LEI issuers have scored 100 percent.[24]

It is also notable, however, that LEI adoption has been uneven. Rules requiring or encouraging use are in place in at least 40 jurisdictions globally.[25] Adoption has been widespread in sectors where regulators require it, as opposed to just encouraging it; in the United States, those include derivatives transaction reporting and home mortgage lending, where the Consumer Financial Protection Bureau is requiring many lenders to provide their LEIs when reporting loans.[26] Adoption is spotty at best in sectors where the LEI is not required.

---

[21] See LEI Regulatory Oversight Committee, "1st progress note on the Global LEI Initiative," March 8, 2013 (available at http://www.leiroc.org/publications/gls/roc_20130308.pdf, accessed June 6, 2016).

[22] See International Organization for Standardization, "ISO financial services standard wins industry support six months ahead of publication," July 25, 2011 (available at http://www.iso.org/iso/home/news_index/news_archive/news.htm?refid=Ref1449, accessed June 6, 2016).

[23] See Securities Industry and Financial Markets Association, "Legal Entity Identifier Resource Center; Overview," undated (available at http://www.sifma.org/issues/operations-and-technology/legal-entity-identifier/overview/, accessed June 6, 2016).

[24] Global Legal Entity Identifier Foundation, "GLEIF Data Quality Management," undated (available at https://www.gleif.org/en/lei-data/gleif-data-quality-management, accessed June 6, 2016).

[25] See LEI Regulatory Oversight Committee, Nov. 5, 2015, p. 10.

[26] See LEI Regulatory Oversight Committee, Nov. 5, 2015, Annex 1, pp. 13-17; also see Consumer Financial Protection Bureau, "Home Mortgage Disclosure (Regulation C)," final rule and official interpretations, Oct. 15, 2015 (available at http://files.consumerfinance.gov/f/201510_cfpb_final-rule_home-mortgage-disclosure_regulation-c.pdf, accessed June 6, 2016).

That's why the OFR has been so vocal in calling for universal required adoption. Additionally, regulators repeatedly hear from industry that even though the LEI is beneficial, adopting it has costs — just a few hundred dollars a year for registration and maintenance of an identifier, but orders of magnitude more to improve in-house data systems. Without a regulatory requirement, corporations do not put those costs at the top of their budget priorities, thus depriving firms of the longer-term benefits of improved risk management.

Swap data repositories

The experience of U.S. regulators with swap data repositories (SDRs) highlights how a lack of universally agreed-upon standards between the public and private sectors hurts data quality and usefulness. The OFR is working with U.S. and overseas regulators to address the issue.

Like the LEI initiative, the effort to improve data about derivatives and swaps grew from the aftermath of the financial crisis. At the 2009 G20 summit in Pittsburgh, leaders called for improved oversight of derivatives, especially those traded over the counter (OTC). In the United States, the Dodd-Frank Act gave the Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC) oversight over the derivatives market. The law also required that data related to swap transactions be reported to SDRs. Swap data are critical to understand exposures and connections across the financial system.[27]

The rule establishing SDRs did not rely upon existing industry data standards or set data definitions and structures — even at a basic technical level, such as the allowable characters in a field. There are currently four U.S. SDRs, all privately operated. They each collect similar information, but designate it and aggregate it differently. They are not required to validate or clean data they receive.

To test the quality of swap data, the OFR analyzed data made public in October 2015 by the four U.S. SDRs on their websites for credit default swaps and interest rate swaps.[28] Researchers assessed which values in certain fields were null or missing. That is one factor affecting overall data quality.

Problems were widespread. The researchers found the collateralization field empty or null more than 40 percent of the time in data posted by certain repositories. Collateralization is important in financial stability analysis for measuring counterparty risk, demand for collateral for other types of transactions, and overall market liquidity. A number of other fields were routinely blank, making it difficult to analyze swap market volumes. Some fields displayed null values even when other fields with related information had data, implying the first set of fields should be populated.

---

[27] See Office of Financial Research, "Data Quality in Swap Data Repositories," undated (available at https://financialresearch.gov/data/data-quality-in-swap-data-repositories/, accessed June 6, 2016).

[28] See Office of Financial Research, 2015 Financial Stability Report, Washington: Office of Financial Research, Dec. 15, 2015, pp. 84-85, (available at financialresearch.gov/financial-stability-reports/files/OFR_2015-Financial-Stability-Report_12-15-2015.pdf, accessed June 6, 2016).

Each of the four repositories posts real-time transaction data on its website, which third-party vendors integrate to create consolidated U.S. datasets and data platforms. The vendors clean and normalize the data because each repository structures and names data elements differently. Each vendor makes its own assumptions and has its own methodology for aggregating and representing the data.

Because repositories report data differently, discrepancies may be compounded. For example, market participants reported three-month LIBOR, a commonly traded floating rate in interest rate swaps, in different ways. The data field for some trades used "USD-LIBOR-BBA 3M" while others had "USD.LIBOR.3M.BBA." As a result, it is hard to identify specific trade characteristics, such as underlying asset or rate.

The OFR is working closely with the CFTC to improve data standards in swap data reporting to assure quality and utility. Under a 2014 memorandum of understanding, the OFR has been assisting the CFTC on a project to enhance the quality, types, and formats of data collected from SDRs. In December 2015, the CFTC requested comments from the public about specifications for some swap data, and is considering how to proceed.[29]

Complicating the data quality issue further, derivatives reporting requirements differ across borders, a problem in a global market. According to the Financial Stability Board, aggregation of data across the various national repositories is necessary to ensure that authorities can obtain a comprehensive global view of the OTC derivatives market and activity. That will require more widespread standardization of data and other advances.[30]

International efforts are under way to harmonize derivatives data reporting, and the OFR is actively participating in those, bringing to bear its experience working with the CFTC and on other data standardization efforts. In 2014, the Committee on Payments and Market Infrastructures (CPMI) and the International Organization of Securities Commissions (IOSCO), jointly known as CPMI-IOSCO, established a working group to address the harmonization of the data elements essential for the reporting and aggregation of derivatives data.

The CPMI-IOSCO effort addresses the need for a unique transaction identifier (UTI), unique product identifier (UPI), and additional critical data identified as "Other Data Elements." This

---

[29] See Commodity Futures Trading Commission, "Draft Technical Specifications for Certain Swap Data Elements," Dec. 22, 2015 (available at http://www.cftc.gov/idc/groups/public/@newsroom/documents/file/specificationsswapdata122215.pdf , accessed June 8, 2016); also see Commodity Futures Trading Commission, "CFTC Staff Extends Request for Comment Period on Draft Technical Specifications for Certain Swap Data Elements," press release, Feb. 18, 2016 (available at http://www.cftc.gov/PressRoom/PressReleases/pr7329-16, accessed June 8, 2016).

[30] See Financial Stability Board, "Feasibility study on approaches to aggregate OTC derivatives data," Sept. 19, 2014, pp. 1-3 (available at http://www.fsb.org/wp-content/uploads/r_140919.pdf, accessed June 17, 2016).

effort remains a work in progress that has produced several consultative reports and regularly seeks input from the industry.[31]

The UTI sounds simple: Each transaction gets an identifier, and all parties use that identifier. Exchanges use such identifiers routinely. But OTC derivatives trades do not, which results in the each counterparty identifying the same transaction in a different way. Thus, one trade may be double counted in aggregated data for jurisdictions that have dual reporting requirements. Setting up a UTI seems straightforward at first glance, but choosing among various systems and setting one global standard is not.[32]

The UPI is even more complex. It involves classifying each type of derivative, based on collections of the instrument's data elements and their contents. Classification for business purposes is generally organized around subsets of descriptive data, such as those that refer to the asset class of the underlying instrument, the product type (forward/future, option, or swap), the tenor, interest rates where appropriate, delivery specifications, etc. (UPI is different from an instrument identification, i.e., a specific interest rate swap between two counterparties. It's the difference between the classification of "common stock" — a product — and a share of "Apple Inc." — an instrument.) Regulators need a globally accepted UPI to consistently aggregate risk across related products and markets. Firms need it to better understand their own exposures.

Those "other data elements"—information about things such as the value of a trade—each also require a definition, naming convention, standard, format, list of allowable values, and cross-references. There were about 60 of them when the CPMI-IOSCO effort began in 2014. [33] The list continues to evolve. As that demonstrates, data requirements change, and thus standard-setting processes require continuous monitoring and investment. Business and government have long recognized this need in other fields. For instance, the ISO develops voluntary market-relevant standards for fields as diverse as manufacturing and health care. The ISO developed and maintains the technical standard that is the basis of the LEI system.

---

[31] See Bank for International Settlements, Committee on Payments and Market Infrastructures Board of the International Organization of Securities Commissions (CPMI-IOSCO), "Consultative report: Harmonisation of the Unique Product Identifier," December 2015 (available at http://www.bis.org/cpmi/publ/d141.pdf, accessed June 8, 2016); also see CPMI-IOSCO, "Consultative report: Harmonisation of the Unique Transaction Identifier," August 2015 (available at http://www.bis.org/cpmi/publ/d131.pdf, accessed June 8, 2016); and also see CPMI-IOSCO, "Consultative report Harmonisation of key OTC derivatives data elements (other than UTI and UPI) – first batch," September 2015 (available at http://www.bis.org/cpmi/publ/d132.pdf, accessed June 8, 2016).

[32] See International Swaps and Derivatives Association, "Unique Trade Identifier (UTI): Generation, Communication and Matching," updated July 20, 2015 (available at http://www2.isda.org/functional-areas/technology-infrastructure/data-and-reporting/identifiers/uti-usi/, accessed June 6, 2016).

[33] See CPMI-IOSCO, "Consultative report: Harmonisation of key OTC derivatives data elements (other than UTI and UPI) – first batch."

Securities financing data collection pilot projects

The OFR has recently completed pilot data collections for two important segments of the securities financing market, bilateral repurchase agreements (repos) and securities lending. These collections were performed in partnership with the Federal Reserve and the SEC to fill data gaps. The government agencies worked closely with industry survey participants to define the data to be collected. For planned future collections, the OFR will leverage industry standard definitions and will collaborate with the industry. As the regulators look to fill data gaps, they will also look to fill data standards gaps.

In their analysis of the first pilot collection, bilateral repos, OFR researchers identified specific challenges. The limited scope of the pilot, which collected data from nine firms on a voluntary basis, meant it did not provide a full picture of the market — a challenge to be expected in any pilot project. But two other challenges emerged related to microdata, data quality, and benefits to industry: lack of data standards and separate data systems.[34] These are relevant to any microdata-based regime for financial system data. As the researchers explained:

Lack of data standards: Given the voluntary nature of the pilot, the collection attempted to leverage each firm's internal reporting systems, rather than impose external requirements. The lack of use of data standards undermined the quality of the data received. In particular, a lack of standardized counterparty information, such as use of the LEI, limited the ability to analyze market interconnectedness because the same firm had different names in dealers' reporting systems. Repo market participants are not required to use LEIs in regulatory reporting, although many filing forms recommend or allow LEIs. Without LEIs and associated mapping to specific industry sectors, identification of counterparties is a substantial challenge. Although the OFR requested industry sector information for counterparties for the pilot, inconsistent sector mapping by participating firms resulted in low-quality counterparty data. In future collections, having respondents submit their counterparties' LEIs would greatly increase the value of the data.

Separate data systems: The reporting systems of participating firms evolved to meet their business needs and were not designed to support reporting of granular data at the enterprise level. For example, data elements specific to a trade, such as principal amount and rate, are kept in one trading system, while counterparty data might be kept in a separate back-office system. This problem illustrates the challenges firms face in using their existing reporting systems to respond to requests for ad hoc collections of trade-level data. Permanent collections might create incentives for firms to invest in automating the reporting of the data, reducing manual interventions, and making the process more efficient.

The OFR and financial regulators are working together to develop a permanent granular data collection for the bilateral repo market, building on the lessons learned from the pilot. They also

---

[34] See Viktoria Baklanova, Cecilia Caglio, Marco Cipriani, and Adam Copeland, "The U.S. Bilateral Repo Market: Lessons from a New Survey," Office of Financial Research, Jan. 13, 2016 (available at https://financialresearch.gov/briefs/files/OFRbr-2016-01_US-Bilateral-Repo-Market-Lessons-from-Survey.pdf, accessed June 8, 2016).

participate in international working groups to harmonize reporting definitions, concepts, and requirements.

Conclusion

The experiences of the OFR underline that standards developed through public-private partnerships best allow regulators and industry to mutually benefit from the increasing demand for microdata. Those experiences also underline that global regulators and the industry still have much to accomplish, eight years after the financial crisis.

From its work on the LEI, the OFR has seen the value of shared goals and cooperation between the public and private sectors in developing and popularizing data standards. In contrast, the regulatory experience with SDRs shows how a lack of agreed-upon standards between the public and private sectors hurts data quality and usability. Based on those lessons and others, the OFR will work closely with industry and build on industry standards as it develops its permanent data collections for the securities financing market.

The OFR has also seen that even though industry participants praise standards, without regulatory mandates, they may not adopt them. An incentives mismatch remains. Firms have demonstrated that they will not spend money on data-related issues to keep their own houses in order without significant public-sector involvement, as the LEI and SDR experiences show. Neither group can solve this mismatch separately, presenting a challenge that must be addressed if the industry and public sector are to use the same underlying microdata to improve risk management and reporting.

Resolving that challenge requires that regulators continue outreach to the industry as well as participation in multinational standards-setting initiatives. It also requires that industry participants collaborate in joint standards development efforts, then adopt and use those standards. Development and adoption of standards obviously is neither fast nor easy. The result, though, should be improved data quality and lower cost for both regulators and industry, with reduced reporting burden for the industry.

###