

Financial Research Advisory Committee Meeting

February 25, 2016

Discussion Topic: Data Quality, Data Scope, and Data Accessibility Programs

Data Quality

The financial crisis demonstrated to regulators and industry that standardizing data to promote systemwide analysis is necessary for effective oversight of the financial system.

Data standards do not assure quality, but without standards, comparing, aggregating, and analyzing data for financial stability analysis are nearly impossible. Data quality is particularly important to the OFR and FSOC. Although the financial data needed for analysis may be available in some form, they might not be complete, accurate, and timely for effectively identifying and analyzing vulnerabilities in the U.S. financial system. That's why the OFR has a mandate to standardize the types and formats of data we report and collect, and to assist FSOC member agencies with the development and use of data standards.

Unlike other industries, the financial services industry has been slow to define and implement consistent data standards and formats to support its risk management and reporting, as well as systemic risk analysis by government regulators. Solving this "collective action problem" and speeding development of data standards requires concerted effort from the public sector. Much of our work with our FSOC counterparts and international authorities involves developing and promoting data standards so financial data can be easily shared and integrated, while protecting confidentiality to support broad analysis of financial system risks.

The data quality program has five components:

1. Develop and publish a financial instrument reference database — an authoritative source for precise, common definitions, and for descriptive data, known as metadata.
2. Continue to identify and implement data standards for markets and critical instruments, such as derivatives, repos, and mortgages, by collaborating with domestic and international authorities and industry standard-setting bodies.
3. Enhance internal data stewardship function to create and maintain complete, accurate, and timely datasets required to analyze threats.
4. Work toward requiring the use of the legal entity identifier (LEI) and appropriate instrument standards for all reporting entities to improve data quality, and promote their use on a consistent, global basis.
5. Add information to the LEI system to identify relationships among entities, such as ownership and subsidiaries.

Data Scope

The OFR has a mandate to collect from any financial company the data necessary to assess the extent to which a financial activity or financial market poses a threat to U.S. financial stability. Financial markets innovate and evolve, so assessing data gaps and prioritizing efforts to fill them will always be a moving target. Through this program, we will use our ongoing research and collaboration with our regulatory colleagues to identify key questions and the data needed to answer them.

To measure financial activity for financial stability monitoring, data scope must be both comprehensive and detailed. It must be comprehensive to analyze sources and uses of funds, the behavior of borrowers and lenders, and risks wherever they arise. It must be detailed (at times to transaction-level) because assessing vulnerabilities involves measuring risk throughout the distribution of outcomes, not just at the mean or median.

The means by which we undertake data collection is important. Before building permanent data collections, we will engage with industry — the sources for financial data — and conduct pilot projects when necessary, to ensure that the data collected are defined precisely and meet specific data-quality criteria.

We use data inventories or catalogs to compare the needs with the available data and to prevent the duplication of existing data collection efforts. We will require the use of the LEI and other data standards, as they become available, in all OFR data collections.

Working with our FSOC counterparts, we are preparing for new, permanent collections of bilateral repo data and securities lending data. These datasets will provide a clearer view of the functioning, risks, and vulnerabilities in two significant markets that are essential for securities financing, price discovery, and market making. We have been careful to align these efforts with those of other agencies.

The data scope program has three components:

1. Assess and prioritize data needs for critical risk topics using mapping techniques and metadata catalogs.
2. Develop and implement pilot projects to ensure data are fit for the intended purpose.
3. Develop and implement rules for our permanent data collections that require the use of appropriate data standards and appropriate security protocols.

Data Accessibility

During the financial crisis, the inability to access or share certain data prevented market participants and regulators from fully understanding the size and scope of risks throughout the financial system. The need to share data more easily in a timely fashion, either for monitoring before an event or forensic analysis afterward, remains today. The Dodd-Frank Act requires the OFR to collect data on behalf of FSOC, provide data to FSOC and its member agencies, and maintain data security and confidentiality. Further, to increase transparency and facilitate research on the financial system, the OFR will provide data to financial industry participants and

the public, appropriately and after consultation with member agencies. Thus, secure and appropriate data sharing is a key priority.

One key component of this program is to promote the use of catalogs of metadata — data about the data. Metadata catalogs inform parties about the data without having to first grant access to the actual datasets. We will build on the existing Interagency Data Inventory (posted on the OFR website) that catalogs basic information collected by FSOC member agencies. Will also help the agencies link their metadata catalogs or create such catalogs if they do not exist.

Linked metadata catalogs can provide information useful in crafting agreements to share data with other regulators in the U.S. and globally and, as previously noted, in deciding how to fill data gaps. In addition, linked catalogs help in applying consistent safeguards and controls to assure security and confidentiality. We plan to create and promote a set of best practices for data sharing among and between regulators and the public.

The data accessibility program has three components:

1. Move beyond the existing Interagency Data Inventory and help FSOC member agencies link their metadata catalogs or create such catalogs.
2. Create and promote a set of best practices and techniques for data sharing among regulators and between regulators and the public.
3. Promote the standardization and efficient development of memorandums of understanding that set the terms for timely and appropriate access to nonpublic data for the regulatory community.

Questions for discussion

1. How should the data scope, quality, and accessibility components be prioritized?
2. What is the best model for partnering with stakeholders and engaging with industry?
3. The OFR has initiated development of a reference data dictionary and is focusing on swaps, repurchase agreements, and securities lending data. What other asset classes should we prioritize?
4. What approaches should we consider as we plan to build, link, and maintain metadata repositories? What type of governance models should we consider?
5. What best practices should we consider for collecting data, given that the need for data has become increasingly granular and extensive, to meet our mission while minimizing the reporting burden? Are pilot projects the best way to explore undertaking a collection? What are some possible public/private partnerships that could accelerate the process?