Big Regulation's Big Data

Banks are accustomed to thinking about Big Data in terms of marketing but regulatory compliance nowadays also requires some very big data. BY RICHIK SARKAR AND BENJAMIN C. GALEA

May 8, 2013 | 0 Comments

In March 2012, the Obama Administration announced its Big Data Research and Development Initiative. The program involves a total of over $200 million in new commitments from several departments and agencies to improve the federal government's ability to make meaning out of the astoundingly large amount of data available to it. Notably absent from the list of participants was any department or agency responsible for regulating the financial services industry. It would be a mistake, however, to conclude that these regulators are not serious about big data.

Congress passed the Dodd-Frank Wall Street Reform and Consumer Protection Act to create a framework to ensure the stability of the nation's financial system and to prevent harm to consumers. Financial services firms operate in one of the world's most data-rich industries, and the effectiveness of Dodd-Frank will depend on regulators' ability to use information that they gather from markets and firms to improve the nature of their supervisory efforts. The three primary federal entities that Dodd-Frank created – the Financial Stability Oversight Council (FSOC), the Office of Financial Research (OFR) and the Consumer Financial Protection Bureau (CFPB) – all recognize big data's potential and have begun to implement regulatory strategies based on the collection and analysis of new data in new ways. Although these efforts are still largely in the development stage, the three agencies are headed toward a new era of data-based regulation.

Understanding the Dodd-Frank regime's big data goals, capabilities, and challenges with regard to systemic stability and consumer protection will not only improve market participants' ability to anticipate their compliance obligations but also offers firms the opportunity to optimize their own big data initiatives.

Agency Watch

At the current time, the two agencies for bankers to watch from a big data perspective are the OFR and CFPB. The mandate of the OFR is to serve as the FSOC's research and analysis desk for the task of collecting data from market participants, standardizing that data to facilitate analysis, translating that analysis into research on financial stability and risk management and using its experience in data analytics to assist the FSOC and its member agencies on data collection and analysis issues.

Considering the overwhelming number of applications for data analytics in the regulation of financial markets, the OFR's first data initiative has been relatively modest. In November 2010, the OFR began promoting the adoption of a global standard for identifying the parties to financial transactions. This standard, known as the Legal Entity Identifier (LEI), aims to add a layer of reference data to financial markets to make firms' trading activities, as intermediaries and as principals, more transparent.

To accomplish its various consumer protection goals, the CFPB meanwhile must monitor the behavior of firms and consumers within the financial services arena. The CFPB has the statutory authority to compel firms to report on their activities and practices, or respond to the agency's specific questions, regarding the consumer financial products and services that they offer. Despite this power, CFPB's primary constituents have been the chief source of its data to date. In July 2011, the agency started to collect consumers' grievances about their credit cards through a portal on its website. To submit complaints, consumers populate a form that allows participants some flexibility to tell their stories, but also imposes structure on the submissions through the use of drop-down menus, yes/no selections and check boxes. Each completed form creates a record in the CFPB's consumer complaint database.

After five months focusing on credit cards, the CFPB expanded its data collection to mortgage-related complaints, depository bank accounts and services, private student loans and other consumer loans. The CFPB now also takes complaints on credit reporting and money transfers and plans to gather information on payday loans and debt collection in the near future.

At one level, the CFPB's relational complaint database is nothing new, but the way that the CFPB has extended its application demonstrates the agency's commitment to exploring the possibilities of big data. In June 2012, the CFPB made its credit card complaint data available online and, on March 28, 2013, posted its entire database. Through a platform developed in conjunction with Socrata, a Seattle-based company offering open-data solutions to governments, users can apply filters to the database to focus on subsets of fields and create a wide variety of charts to illustrate the relationships among them.

Despite the CFPB's commitment to open government and transparency, all of this public disclosure does risk unauthorized access to consumers' personally identifiable information. Although the CFPB's policy is to strip the publicly available records of such information, including the narratives that the customer and the firm provide, the security of the CFPB's dataset is an issue. On March 28, 2013, the Federal Reserve's Office of the Inspector General released the results of its audit of the CFPB's information security controls under the Federal Information Security Management Act of 2002 and found that the agency could improve its system in nine areas, which the CFPB acknowledged and pledged to address.

The forays that the OFR and the CFPB have made into big data demonstrate several best practices that financial services companies would do well to keep in mind:

Focus on human capital first. The OFR recognized its need for a big data platform after surveying viable models for measuring systemic risk and considering the data requirements of their application. Similarly, banks should consider the objectives that they intend to achieve with data analytics and the models that exist or that they can build to shed light on relevant trends. This creative exercise requires that firms bring individuals who are able to define the business problems and their potential solutions together with those who are able to design datasets and processing tasks to address them.

Develop sound data platforms. As the OFR's LEI initiative demonstrates, standardizing the format of data is one way to increase the potential of data analytics, but the exponentially expanding amount of data available from a variety of sources will demand a capable data platform. To provide a foundation for big data, firms should invest in a platform that incorporates low-cost, high-volume storage, structured and unstructured data analytic capacity, visualization and discovery tools and security monitoring capabilities.

Apply systemic risk measures proactively. The stress tests that Dodd-Frank mandates are not the only measures of systemic risk available to the OFR. Just as the Office considers the stability of a financial system by feeding data into a variety of models, firms should also adopt this practice, which the OFR has facilitated by making the source code for its risk analytics available to the public. As a proactive measure, firms should use internal and market data to run these models and consider the significance of the results not only from a business perspective, but also as their regulators would see them.

Append publicly available datasets to internal data. The CFPB understands the value that its consumer complaints database may create when firms combine its data with their internal customer service records. Even if consumer-reported data may be unreliable, it nonetheless offers firms the opportunity to isolate and address potential customer service issues and gain a sense of the likelihood of a CFPB examination.

Shop for datasets smartly. In its efforts to learn about the payday lending market through data analytics, the CFPB has provided firms with a path to the responsible acquisition of datasets from third parties. As firms appreciate the bid data's benefits, there will be an increase in demand for external datasets, and firms should conduct due diligence on their data providers to ensure that the data that they receive was not collected in violation of any contractual privacy policies or statutes. The value of a dataset depends on the quality of the data that it contains and the manner in which it was obtained.

Mr. Sarkar and Mr. Galea are attorneys with [Ulmer & Berne LLP](), a full-service law firm with offices in Chicago, Cincinnati, Cleveland and Columbus. They can be reached at [rsarkar@ulmer.com]() and [bgalea@ulmer.com]() respectively.