On Using Big Financial Data for Macro-management: Explorations and Practices

Author: Ruan Jianhong, Director General of the Statistics and Analysis Department of the People's Bank of China (PBC)

Abstract: President Xi Jinping pointed out in the *Notes to the Recommendations for the 13th Five-Year Plan* that "we should coordinate the compilation of comprehensive statistics on the financial sector, and through the full-scale collection of financial sector data, strengthen financial macro-management and maintain financial stability." The meteoric rise of big data technologies in recent years has provided powerful tools for compiling and analyzing financial statistics. Bringing major improvements to the efficiency, accuracy, and relevance of the compilation and analysis of financial statistics, these technologies help enhance the framework of financial statistics. Earnestly implementing President Xi Jinping's call for compiling comprehensive financial sector statistics, the PBC extensively reviewed how developed countries wield large sets of financial data to serve macro-management objectives, and approached the task of data governance thoroughly by starting from the underlying data. Big data techniques are used to analyze macro-economic and financial situations to more effectively support macro-management and prevent and control systemic risks. The PBC has made some achievements in this endeavor, including preliminarily resolving the difficulties associated with information gathering and use.

Keywords: comprehensive financial sector statistics, big data

The rapid development of information technologies and big data algorithms in recent years has increased the momentum of economic and social development and offered powerful tools for financial statistics compilation and analysis. A new challenge now is to leverage big data technologies to improve the financial statistics framework and more effectively support macro-management. Under the framework of the *Opinions of the General Office of the State Council on Fully Advancing the Comprehensive Statistics Compilation and Analysis in the Financial Sector* (hereinafter referred to as the *Opinions*), the PBC has been fully committed to exploring and implementing ideas that boost macro-management with the help of big data in financial statistics (big financial data).

I. Big data technologies can improve the financial statistics framework

The 2008 global financial crisis turned out to be the biggest of its kind since the Great Depression. One of its contributing factors relates to financial statistics. Specifically, in this event not only did they fail to reflect financial changes in a timely manner before the crisis, partly because of the incompleteness of data, but they also failed to support an accurate judgment and estimation of risk contagion after the crisis erupted. Two major approaches have since been taken to address the issue of missing data: expanding the scope of conventional financial statistics and using big data technologies.

(I) Expanding the scope of conventional financial statistics

Expanding the scope of conventional financial statistics centers around establishing a context-appropriate statistics system for each of the areas where data have been lacking, such as shadow banking, systemically important financial institutions (SIFIs), and financial holding companies (FHCs). In this regard, China has successively created such a system for asset management products, SIFIs, and FHCs, all under the framework set out in the *Opinions*.

This approach solves the problem of missing information and is backed by proven and internationally comparable statistical methods, but is not without its faults.

First, the statistical framework is ill-adapted to new problems. Broadening the conventional statistical scope mainly remedies the information deficiencies exposed in this one particular crisis. It is not nearly as effective with regard to new issues that may surface in the future.

Second, more statistical forms make it harder to compile and analyze data and delineate the interactions between financial institutions. Greater granularity with respect to counterparties and financial instruments creates a growing number of forms in conventional financial statistics and, with it, highly complicated reporting processes. The other problem is that tabulated data tend to obscure the fine-grained relations among institutions.

Third, it is difficult to compile price and structured information. A form-based approach for gathering information on organizations' interest rates, maturities, locations, ownership type, industries, and other statistical dimensions as well as inter-dimensional data is challenging and poorly scalable.

(II) Employing big data technologies

Conventional, form-based statistical approach is beset by poor data relevance and difficulties in data gathering, analysis, and correlation. Big data technology-based financial statistical approach, in which finer-grained data are collected, addresses these shortcomings with its larger set of resulting data and problem-oriented design. Big data-powered financial statistics have the following features:

First, big data technologies complement the conventional financial statistical methods. Like the conventional approach, the big data-driven approach also focuses on financial instruments and counterparty information, which provides two benefits. The first is that the big financial data thus obtained are readily interfaced with those obtained through the conventional monetary and financial statistical method. The second is that the fine-grained data can greatly expand the range of information available for financial statistical analysis. Big financial data consist of massive amounts of standardized data and have high information density, thus better meeting the needs of macro-management.

Second, big data technologies enable more scalable and accurate financial statistics. With standardized information on financial instruments and counterparties, the statistical recording methods for most financial transactions will converge. For example, information

about the credit instruments of different financial institutions can be compiled and analyzed by similar methods. This means that financial statistics are more scalable and the data gap can be bridged more efficiently. Furthermore, fine-grained data make it easier to detect outliers and correct erroneous values promptly, resulting in higher-quality statistics.

Third, big data technologies help produce financial statistics that more faithfully reflect the reality, and therefore lead to more accurately targeted policies. Big data-powered, fine-grained financial statistical information can reflect financial transactions more accurately. The data thus gathered will support not only macro-level aggregate analysis but also micro-level structural analysis, thereby greatly improving the accuracy of structural policies.

II. How developed countries apply big financial data to macro-management

Because big financial data help improve the financial statistical framework and address data deficiencies, financial regulators in developed countries are employing big data technologies to enhance financial statistics, in particular leveraging big financial data to monitor the risks of financial interconnectedness and the performance of the real economy.

1. The Deutsche Bundesbank: large-credit registration database. The enhanced money market statistical reporting (MMSR) program spearheaded by the Deutsche Bundesbank, the German central bank, was officially launched in July 2016. Covering all financial institutions in Germany whose total assets exceed EUR1 billion, MMSR monitors the connected transactions in secured, unsecured, foreign exchange, and overnight assets between them, especially those between banks. Data are reported on a transaction-by-transaction basis under MMSR, allowing the Deutsche Bundesbank to monitor intra-financial system information such as the price of investment and financing transactions, counterparty concentration, and maturities at a day-by-day granularity. It then uses these data to analyze the degree of interconnectedness between the investment and financing activities of various types of financial institutions, in assets inside and outside Germany and in different asset classes, and provide high-frequency data to facilitate timely market risk warning.

2. The European Central Bank: AnaCredit for loan-by-loan information. In 2011, the European Central Bank initiated project AnaCredit (Analytical Credit Dataset), a new module for monetary and financial statistics, which was officially put into operation in 2018. AnaCredit expands and refines the monetary and financial statistics that focused on financial instruments and counterparties. More significantly, it offers a financial-statistics data model for the era of big data. AnaCredit's data model encompasses information about the financial instrument, collateral, and counterparty of every credit transaction, organized into 88 data attributes. With AnaCredit, the European Central Bank is able to analyze the destination and risks of a loan and gain clarity into the connected transactions between financial institutions, thus assessing the degree of interconnectedness within the financial system and the systemic stability.

3. U.S. Federal Reserve: data on real-time counterparties' assets and liabilities. Since 2015, the Federal Reserve has been collecting data on real-time counterparties' assets and liabilities. This counterparty-based statistics system covers the changes in the on-balance sheet items and off-balance sheet assets of global systemically important banks in the U.S. All

data collected are granular, containing information such as the countries of the domestic and overseas counterparties, types of instrument, currencies, and maturity structure. These data are then aggregated and published on a quarterly basis. The data on real-time counterparties' assets and liabilities assist the Federal Reserve in dynamically monitoring the risk exposure of the global systemically important banks.

4. The Federal Reserve and the U.S. Bureau of Economic Analysis (BEA): using big financial data to monitor consumer spending. The Federal Reserve and the BEA are applying big data technologies to each credit card and debit card transaction to monitor consumer spending. Daily monitoring data are produced three days after a transaction. The information gathered is closely related to that published and contains abundant structural information to support policy-making.

III. China's explorations and practices in applying big financial data to macro-management

Big data technologies can significantly improve the efficiency, accuracy, and relevance of financial statistics compilation and analysis and better satisfy the demand of monetary and financial stability policies. China has taken a demand-oriented approach in exploring and implementing big financial data applications in financial statistics programs.

(I) Explorations in applying big financial data to macro-management

The PBC has been an early explorer in using big financial data to address the information deficiency in economic aggregate reports. Its efforts are mainly as follows:

First, gathering data on wealth management and trust to serve the monitoring of cross-sector financial products. To comprehensively evaluate the impact of financial institutions' wealth management and trust businesses on monetary policy transmission and financial stability, in 2010 the PBC started to collect data on a per-transaction and per-product basis. This initiative has become a "testing site" for big financial data and enabled the PBC to keep a pulse on the cross-sector financial products in China. Building on this explorative effort, in 2018, China's four financial regulators—the PBC, the China Banking and Insurance Regulatory Commission, the China Securities Regulatory Commission, and the State Administration of Foreign Exchange—jointly established a transaction-by-transaction reporting system for asset management products offered by financial institutions, which fully covers the RMB80 trillion aforementioned products. The system tracks each product throughout its life cycle, from source to use and from release to termination, including complex asset management products.

Second, establishing a standardized deposit and loan statistics system to serve interest rate monitoring. As the PBC is proactively promoting the market-based reform of interest rates in recent years, interest rate monitoring is particularly important for monetary policy decisions and financial risk prevention. Accordingly, the PBC started to compile standardized sampling statistics on deposits and loans in 2012. The program gathers the per-transaction details—product type, customer type, maturity, credit line, interest rate and asset quality among others—of nearly 100 million deposit and loan transactions every month from over

5,000 top-level sub-branches (i.e., the highest-level sub-branches) of more than 500 legal-person financial institutions nationwide. With the combination of the standardized deposit and loan statistics and economic aggregate statistics, a deposit and loan statistics system which provides insights into both aggregate and structure, both quantity and price, and both stock and flow has taken shape.

(II) Practices in enhancing macro-management with big financial data

Based on sufficient experiments and explorations, the PBC has recently proactively applied big financial data to macro-management within the framework set out in the *Opinions*.

First, the PBC has launched the National Basic Financial Database. The Database is a major pillar for big financial data. The hardware and software infrastructures for this database have been put in place to create an industry-leading, big data-based intelligent cloud architecture. The centralized financial data collecting system and the intelligent platform for big data analysis have been deployed. Featuring a petabyte-level storage capacity, the database is currently connected to all PBC branches and over 4,600 financial institutions for intelligent and integrated collection and usage of data.

Second, the PBC has implemented a statistical system for basic financial data. In order to address problems of inconsistent data standards, as well as difficulties in data collection and usage, the PBC has explored solutions from their fundamental causes, and redoubled efforts on studies and formulation of comprehensive statistical standards for the financial sector. In July 2020, the PBC introduced the statistical system for basic financial data. Reflecting the transition to the big data era, the statistics are standardized, transaction-by-transaction data covering deposits, loans, interbank products, bonds, equities, special purpose vehicles, and other types of financial instruments. Involving over 600 informational dimensions, these data can better meet the information needs of macro-management. The first round of collecting basic financial data was completed in September 2020, which will be followed by further data collection in the future.

Third, the PBC has built and improved entity information databases for effective data correlation. Based on existing in-house data and sophisticated external databases, the PBC has developed two entity information databases, namely the financial institution entity database, covering over 4,600 legal-person financial institutions, and the enterprise entity database, covering over 20 million enterprises. From a micro perspective, the entity data set a solid base for data correlation, as the correlation could facilitate the monitoring of equity chain, guarantee chain, capital chain and other complex relations and the profiling of financial institutions, markets, counterparties, and activities, so as to break down complexities in the financial system, reflect interconnection between financial institutions, and measure the risk contagion of financial markets.

Fourth, the PBC has leveraged basic data to strengthen studies on financial risk contagion for forestalling and controlling systemic risks. Financial risks are mainly transmitted through various forms of business correlations. By leveraging the national basic financial data, the PBC has systematically sorted interbank relations between financial institutions and credit relations between banks and enterprises, established the webs of

interbank relations and shared credit relations. The PBC has conducted in-depth studies on the correlation, contagion, hazard, and prevention of financial risks as well as financial fragility from the perspective of market mechanisms, and created models to simulate the process of financial risk contagion.

Fifth, the PBC has strengthened the analytical and application capabilities based on big data to facilitate macro-management. The PBC seeks to gradually realize coverage of all financial institutions, infrastructures, and activities, which, combined with existing monitoring indicators such as the macro leverage ratio, wealth management products, real-sector financing costs, and loan maturity, can strengthen the effectiveness of macro policies. Per-transaction basic financial data help display the distribution of loans, bonds, and other financial instruments among various enterprises. An accurate depiction of the allocation of financial resources would help to optimize the resource allocation and deepen financial reforms. Big data technology allows the PBC to provide targeted data support for the financial sector in serving the real economy. The problem-orientated technology strengthens the capability of the financial sector to serve the real economy by facilitating more direct and targeted measures.

(III) Future plan for big financial data

First, improving the National Basic Financial Database in line with high standards and positioning it as a key financial statistics infrastructure. Under the principle of "unified platform, centralized management, coordinated deployment, and integrated application," the PBC plans to harness the latest technologies to build the National Basic Financial Database into a national financial cloud that hosts comprehensive, unified, and connected data with advanced, domestically supporting and controllable technologies.

Second, effectively implementing the system-centric financial statistics and data centralization plans. Plans outline the timetable, roadmap, and task lists necessary to successful financial statistics compilation and analysis. Accordingly, the PBC is committed to promoting high standard financial statistics in accordance with the development plans for financial statistics compilation and analysis as well as for the National Basic Financial Database, so as to supply the underlying data needed for financial policy-making and evaluation of policy effects.

Third, building an efficient and dedicated laboratory to promote the analysis and use of basic financial data. This basic financial data lab will be a problem-driven and technology-empowered program for end-to-end data analysis covering business requirements, statistical standards, data collection, algorithm implementation, and data presentation. The PBC will make full use of the national basic financial data to perform its duties, with the overarching goals of supporting the real economy, preventing and controlling financial risks, and promoting financial reforms.

Fourth, expediting work on financial data governance and improving the quality of financial statistics. The PBC seeks both quantitative and more importantly, qualitative improvement on financial statistics. Efficient interconnection and deep utilization of the data require better financial data governance and innovative statistical analysis approaches, in

addition to cloud computing, big data, and other advanced technologies.

Fifth, improving data sharing to fully realize the value of financial statistics. The PBC will use modern data-sharing technologies to upgrade from the sharing of mere statistics to the sharing of underlying indicators, statistics systems and statistics definitions, and to make data sharing among regulatory authorities easier and more efficient.