

Commercial data sources

Philip Papaelias¹

The purpose of this paper is to describe how the Statistics Function of the Reserve Bank of New York (FRBNY) uses commercial securities data in its analysis of security-by-security data. Commercial securities data sources are primarily used in the quality assurance process for the data received from the reporters and to supplement reported data with information not provided directly by the reporters for publication purposes. This paper describes the benefits and drawbacks of using commercial securities data.

Use of commercial data sources

The FRBNY uses commercial data sources primarily in the quality assurance for the data received from the reporters on the security-by-security reports and to obtain data that are used in conjunction with reported data for publication and advanced analysis. No information acquired from any commercial data sources is published directly. However, data such as coupon rate and dividend rates are used in calculations, the results of which are published.

Due to cost and resource restrictions, the number of commercial resources used by the FRBNY is limited. As a result of these limitations, the sources used are selected based on their ability to provide comprehensive coverage of various types of instruments (eg debt, equity and asset-backed securities). In addition, to reduce cost, the FRBNY only receives commercial data for securities that are reported by the security-by-security data reporters.

Commercial data sources do not primarily exist to meet FRBNY data collection needs, which are a very small percentage of the vendors' overall business. Therefore, these data are often not aligned with the data definitions used for FRBNY analysis or publications needs. Since it is not cost-effective for vendors to alter the data to meet specific data collection needs, the FRBNY selects from data fields that already exist in the vendors' databases that are designed for their primary customers.

Drawbacks of vendor data

Commercial data sources are a valuable resource that enables the FRBNY to validate and analyse the data received from reporters of the security-by-security data. However, there are limitations to most commercial data sources that make it impractical to use these data in place of the reported data. The main drawback is that commercial data sources cannot provide information on securities that are reported with internally generated security identifiers (ie IDs that are assigned by the reporting entities). Almost every major US reporter reports some holdings using internally generated security identifiers. On average, approximately 3% of foreign securities and 2% of US securities reported on the United

¹ Supervising Reports Analyst, Securities Reports Division, Federal Reserve Bank of New York, philip.papaelias@ny.frb.org.
The views expressed should be interpreted as those of the author and do not necessarily reflect views of the Federal Reserve Bank of New York.

States' security-by-security data are reported with internally generated security identifiers. In 2004, the value of foreign securities reported with these identifiers was less than 2% of the total reported data and comprised predominately of private placements.

The values of foreign securities reported with these identifiers are increasing as a result of investments in offshore limited partnership investments. Since 2003, the market value of these reported securities increased approximately 300%. The detailed attributes, such as security type, issue and maturity date (if debt), currency of denomination and market value, associated with these securities would be lost from our dataset if we did not ask our reporters to provide it. This would be a significant loss to users of these data since they would not get an understanding of how large the markets for these securities are.

A solution to this issue is for reporters to provide all the detailed attributes for securities reported with internally generated security identifiers while reporting only specific fields such as type of holder and quantity for securities with security identifiers issued by a recognised numbering agency (eg ISIN, CUSIP). This assumes that vendor data would be able to provide the fields not reported. However, US reporters have indicated that this solution would significantly increase reporting costs because they would have to identify which securities require complete information rather than providing the information they have for all securities.

Another drawback of vendor data is that short-term debt information is typically not available once the issue has matured. Data are provided by US reporters approximately 60 days after the as-of date. This prevents the FRBNY from obtaining detailed attributes about short-term instruments that have matured. This is also an example of where the reporter would provide limited data because the short-term instrument has a valid CUSIP or ISIN, but FRBNY would not have access to the attributes data.

Comparison of information across reporters vs vendor data

As a result of these drawbacks, the FRBNY has found the best method of analysing reported data is to compare individual security attributes across different reporters. As reporters have gained familiarity and experience with compiling security-by-security reports, the quality of the data they provide has increased. This increased quality enables the FRBNY to analyse the attributes of individual securities reported by different reporters to determine the correct value. If a majority of the reporters report the same attributes for a security, then a high level of confidence is gained that the data have been reported correctly. However, if there is a discrepancy between different reporters for an individual security or if there is only one reporter of a particular security, then commercial data sources assist in identifying the proper attributes of the security.

In order for this comparison to be effective, the reporters must be trained extensively in the methods of proper reporting. The FRBNY holds seminars for reporters in which reporting issues are discussed, as well as one-on-one sessions with representatives from individual reporting institutions to discuss specific reporting issues. In addition to the seminars and one-on-one meetings, we have asked reporting institutions with significant reporting issues to provide action plans that describe the steps they will implement to address these systemic issues. Once the action plans are submitted, follow-up discussions are held in order to ensure that the action plans are being implemented. These measures have led to significantly increased data quality and have enabled us to rely less on vendor data.

Conclusion

Commercial securities data are a very useful tool for analysing security-by-security data. However, a drawback such as high cost and the lack of information on non-registered securities and short-term debt that has already matured prevents us from using these sources as a primary source of information.