

## Sources of Foreclosure Data

### INTRODUCTION

Continued weakness and home price declines in some regional economies, tighter underwriting standards and the recent seizure of the nonconforming securitization market (or the private-label securitization market) have all contributed to the increased spate of foreclosures in the United States over the past couple of years.

From the media to investors to government agencies, increasing attention has been given to this trend, and several different foreclosure data sources are often cited to provide insights into its size and impact. Clouding some of the public and policy discussions of foreclosure trends is the fact that many of the definitions and data collection methodologies used by these different sources are very different, as are the business models that drive them. And while it is true that, regardless of which source of data is quoted, the general trend shows that foreclosures are rising, the magnitude and extent of the increase in reported foreclosures varies considerably.

This DataNote provides an overview of the main sources of foreclosure data receiving recent media and analyst attention.<sup>1</sup> The table in the appendix summarizes the basic facts that were gathered.

### METHODOLOGY FOR VARIOUS FORECLOSURE MEASURES

#### DATAQUICK

DataQuick is a public records information service that has compiled data from county assessor and recorder offices since 1979. DataQuick's data currently cover an estimated 80 percent to 95 percent of property sales, refinancing, and foreclosure activity nationwide. Data scrubbing is a high priority in its process in or-

der to maximize the accuracy of the information. The company provides only raw, non-seasonally adjusted data.

DataQuick provides data on the number of properties in the foreclosure process and the number of foreclosed properties for the four regions and for selected counties.

Since DataQuick is a provider of real estate information solutions, which includes property data and analytics, one of the strengths of the data is that users can conduct analysis at very granular geographic levels, including at the zip code, city, and county levels. The source also provides information on sales prices and resale trends. Based on its foreclosure news releases, DataQuick is very strongly placed in California and has detailed monthly news releases on foreclosures in the state. While DataQuick covers various other states in terms of sales data, they do not release foreclosure statistics on these other states.

#### FORECLOSURES.COM

ForeclosureS.com is an established foreclosure website with a database of over 5.5 million property listings. The company collects foreclosure data from publicly recorded documents in each state. The company bases its analysis on the number of formal notices filed against a property during the foreclosure process. This can include notice of default, notice of foreclosure auction, and trustee's deeds/REOs (real estate owned by lender through foreclosure). The company does not have full coverage of many geographic areas because many counties do not have accessible data, but is confident that the data are representative of the markets covered. Newly collected records are compared with an internal property database and the updated records are offered on a daily basis for a fee to clients.

ForeclosureS.com specializes in foreclosed property listings and providing investment advice, and has been in business for over twenty years. Clients use the data to help them in the process of making property purchase or investment decisions. Since property listings are their specialty and they collect property specific data from public records, the company's data cover very detailed geographies and are updated monthly. ForeclosureS.com publishes monthly press releases covering national foreclosure trends.

<sup>1</sup> The MBA worked with all the organizations cited to obtain basic information on their data and collection methods. We would like to thank the following for their assistance: John Karevoll, DataQuick; Mark Carrington, Mark Fleming, Sam Khater, Pete Kreiser, and Bob Visini, First American CoreLogic/Loan Performance; Alexis McGee, ForeclosureS.com; Celia Chen, Moody's Economy.com; and Rick Sharga, RealtyTrac.

## LOAN PERFORMANCE

LoanPerformance, now integrated into First American CoreLogic, is a data cooperative for the mortgage industry, used by secondary market investors and mortgage originators and servicers. Members of the data cooperative (i.e. data licensees) submit loan-level data from their institutions to the company. LoanPerformance then augments those data with third party data sources to complete the data in the database. For example, issuers, broker/dealers and servicers, who are licensees, contribute their securitized loan data to the Securities Database. LoanPerformance supplements the data with other third party sources, such as bond trustees, master servicers, etc., for the Securities Database. Currently, the servicing database covers approximately 80 percent of active first-lien mortgages, while the non-agency securities database covers over 95 percent of active securitized mortgages representing over \$2 trillion in loan-level data.

In terms of quality control, the data are checked against publicly available information and various regulatory documents, where applicable. The company also uses a comprehensive set of proprietary automated checks and business rules in its processing systems. For some of the databases, each data contributor must validate the accuracy of the contributed data on a monthly basis.

LoanPerformance is one of the few industry sources for loan-level data. The access to loan-level data enables analysts to compile a broad array of statistics that include aspects such as foreclosure statistics by credit score, loan-to-value ratio, mortgage origination year, ARM reset date, etc., as well as various granular geographic categories. While still extremely broad, LoanPerformance's data coverage is limited to only those companies with which it has a relationship and covers only the securitized portion of the market.

## FIRST AMERICAN CORELOGIC

First American CoreLogic (FACL) is a data and analytics company that collects property address level data from public records at county recorder's offices, courthouse filings, tax assessors, sheriff's offices, newspaper filings, proprietary sources and selected vendors on a daily, weekly or monthly basis for the number of new and outstanding unique notices of default (pre-foreclosure notices), as well as from notices of trustee sales. It captures the data using document images, electronic access or field collection. In the majority of counties, it compiles data directly from the recorded document images.

Depending on the jurisdiction and the laws governing foreclosure, there are different procedures and public filings requirements, but they fall into three distinct categories similar to what other sources collect: notice of entry into foreclosure (which may require a single or multiple notices); the actual foreclosure filing; and the exit foreclosure sale. FACL also collects tax assessment, land values and appraisal data on over 60 million active outstanding mortgage loans in over 3,000 counties in the US.

With extremely broad and detailed coverage, including over 140 million properties covering 99 percent of the US population, 3.5 billion document images, mortgage servicing data, mortgage-backed securities and asset-backed securities loan-level data, and user-submitted data from appraisers and real estate agents, the FACL data are one of the main industry sources for data on REOs and other foreclosure measures. These data are also available at most geographical levels.

## MOODY'S ECONOMY.COM

Moody's Economy.com is an analytics and consulting firm and its foreclosure data are sampled data from the Equifax National Consumer Database, which is a credit bureau database and contains information collected from over 60,000 contributors. The credit bureau has data on tradelines, collections, public records and demographics across a broad range of industries, with updates on a daily basis. New information added to the database goes through a rigorous quality assurance process, including automated and manual reviews, to ensure the accuracy. Every month, Equifax selects a 5 percent random sample from its National Consumer Database, which amounts to approximately 14 million files. The selection always occurs at month end so that in-month reporting fluctuations are kept to a minimum. Data are scaled for the sampling so the data are representative of the entire national mortgage markets. Regular seasonal variation is removed from the data using standard X-12 programming. Equifax aggregates this data to zip codes before providing it to Moody's Economy.com.

Moody's Economy.com uses the data for its CreditForecast.com subscribers to analyze delinquency and foreclosure trends. The measures include delinquency rates, foreclosure rates and the foreclosure counts as a percent of households in the US. This data set is rich in regional detail and provides a consistent look at credit quality conditions across regions and product types.

## MORTGAGE BANKERS ASSOCIATION

The Mortgage Bankers Association (MBA) is a trade association representing the real estate finance industry. Its National Delinquency Survey (NDS) is a quarterly survey of over 120 servicers who submit delinquency and foreclosure counts. There are two foreclosures categories: foreclosure starts and loans currently in the foreclosure process. Survey participants self-classify loans in foreclosure since the definition depends on investor and state requirements. In general, a loan is classified as being in foreclosure when foreclosure proceedings have been initiated by an attorney.

The foreclosure starts measure represents loans that entered the foreclosure process during the reporting quarter and is a flow measure so it is seasonally adjusted. For loans that are in the foreclosure process, that is a measure of all loans at any point in the foreclosure process at the end of the reporting quarter. A number of factors influence this measure, for example, state foreclosure statutes that establish how long it takes for a loan to work its way through the foreclosure process. Thus, two states with the same rates of loans going into foreclosure can have a

different number of loans in foreclosure at the end of the quarter. This is a stock measure and is not seasonally adjusted.

The MBA calculates the delinquency and foreclosure rates on the number of loans serviced and not the dollar value. For example, the foreclosure starts rate is the number of loans for which foreclosure proceedings have been initiated during a specific quarter as a percentage of all loans serviced that quarter. This survey does not collect information on the dollar value of loans serviced. The US and regional information is available in both seasonally adjusted and non-seasonally adjusted forms. Only non-seasonally adjusted state data are available.

The NDS data cover about 45 million first-lien mortgages on 1-4 unit residential properties. The subprime sample of the survey is around 5.5 million loans. The NDS covers approximately 85 percent of all first-lien mortgage loans outstanding in the market. For each geographic classification, delinquency and foreclosure rates are available by loan type (prime, subprime, FHA and VA) and for fixed rate and adjustable rate mortgages.

The NDS foreclosure data are collected directly from servicers of the mortgages and have broad market coverage. While the company-level data submissions are not audited, each quarter's results are compared with past submissions and any significant changes are verified with participants before they are included in the final statistics. It is important to note that the data only include loans currently in the foreclosure process or if foreclosure proceedings had been initiated in a specific quarter, and once a property has been foreclosed on, it is excluded from the data set. The data are available at the national, regional, and state level.

## REALTYTRAC

RealtyTrac is a real estate resource website that collects and processes home sales and foreclosure data. The company assists consumers, investors and real estate professionals in locating, evaluating, buying and selling properties. The foreclosure data collection counts all filings made on a specific property as it goes through the foreclosure process. It captures information on properties that have default notices, auction sale notices, and bank repossession notices filed on them. RealtyTrac's geographic coverage extends to over 1.3 million properties from around 2,200 counties across the country.

RealtyTrac publishes both listings of foreclosure filings in all covered geographies as well as statistical releases on the number of foreclosure filings by state and by county. Measures include the absolute number of foreclosure filings and an estimate of how many foreclosure filings there are per household in the nation, state, or county. If more than one foreclosure filing on the same property occurs during the same month (as it goes through a different stage of the process), only the most recent filing is counted for that month. However, if the most recent filing is in a separate month, that will be considered two separate filings over those two months. On a quarterly basis and again on an annual basis, RealtyTrac rolls up the data by unique property so that the same property is only counted once in each summation period. For the household ratios, RealtyTrac uses the Census housing unit data, which includes some apartment units.

## CONCLUSION

Definitions are fundamental to any data. When definitions are different, it is extremely difficult to compare data elements, let alone account for data that come from different sources and collected by different methods. Each source of foreclosure data discussed has its respective strengths and weaknesses. It is thus important to understand the difference in definitions and data collection, the limitations of each data source, and the fact that due to these differences, not all foreclosure data are comparable. There may not be a single "best" measure of foreclosures in the United States, but the context in which the data is used — whether compiling local foreclosure listings for individual investors or analyzing aggregated market foreclosure trends for policy development — can determine which sources of data are most suitable (and which are not).

### About Research DataNotes

.....  
**Author:**

Joel Kan, Associate Director, Single Family Research,  
Mortgage Bankers Association

Any questions or for more information, please contact  
Joel at [jkan@mortgagebankers.org](mailto:jkan@mortgagebankers.org).

Research DataNotes are a series produced by members of MBA's research and economics group designed to explain and explore technical aspects of the real estate finance industry.

## General Characteristics of Foreclosure Data Sources

	Release Schedule	Market Coverage	Covered Geographies	Data Collection Channel	Published Measures
<b>Data Quick</b>	Monthly	80-95% of residential sales activity	Zip, city, county	Public records	Foreclosure filings
<b>ForeclosureS.com</b>	Monthly	5.5 million property listings	Zip, city, county	Public records	Foreclosure filings
<b>LoanPerformance</b>	Monthly	80% of first-liens, 95% of non-agency securities	Zip, city, county	Securities issuers, broker/dealers, servicers	Foreclosure and delinquency rates
<b>First American CoreLogic</b>	Monthly	140 million properties, 99% of US population	Zip, city, county	Public records, proprietary sources, other selected vendors	Foreclosure filings
<b>Moody's Economy.Com</b>	Monthly	5% random sample, sample size of 14 million records	Zip, city, county	Credit bureau records	Foreclosure and delinquency rates
<b>Mortgage Bankers Association</b>	Quarterly	85% of loans outstanding	State, region, nation	Mortgage servicers	Foreclosure and delinquency rates
<b>RealtyTrac</b>	Monthly	1.3 million properties, 2,200 counties nationwide	Zip, city, county	Public records	Foreclosure filings