REFINITIV™

DATA IS JUST
THE BEGINNING™

DATA-FIRST THINKING

THE MORTGAGE MARKET'S DATA CHALLENGE

EXECUTIVE SUMMARY

Produced in collaboration with Greenwich Associates, this paper – the second in a series of four reports that will analyze current trends and challenges in the fixed income space – takes an in-depth look at the vast mortgage securities market and its associated data challenge.

> Using sophisticated analytics, investors can now dig into the individual loans within any MBS and see exactly what they're buying.

US mortgage debt currently stands in excess of $12 trillion, with the majority of this debt ($9.4 trillion in the second quarter of 2019) held by institutional investors via mortgage-backed securities (MBS), collateralized mortgage obligations (CMO) and agency (i.e., FHLMC, FNMA, GNMA) bonds.

These institutional investors and mortgage dealers are tasked with developing accurate risk models to help them make decisions relating to price, yield and risk, and other criteria. Any robust and meaningful analysis of course relies on quality data but given that most of the mortgages involved in these complex computations were initiated before financial markets embraced digitization, there remains a vast quantity of mortgage data that is neither standardized nor usable.

Furthermore, even as the breadth and depth of data available in the mortgage market continues to grow, incoming data may be 'dirty', disconnected and non-standardized. Leading-edge technology and emerging solutions can harness and make sense of this data, delivering insightful analysis and empowering professionals in the complex world of mortgage finance and related investments.

In this paper we dive into these new technologies and emerging solutions, to uncover how mortgage market professionals can solve the data challenge.

Adam Quinones
Global Head of Mortgages and ABS
Refinitiv

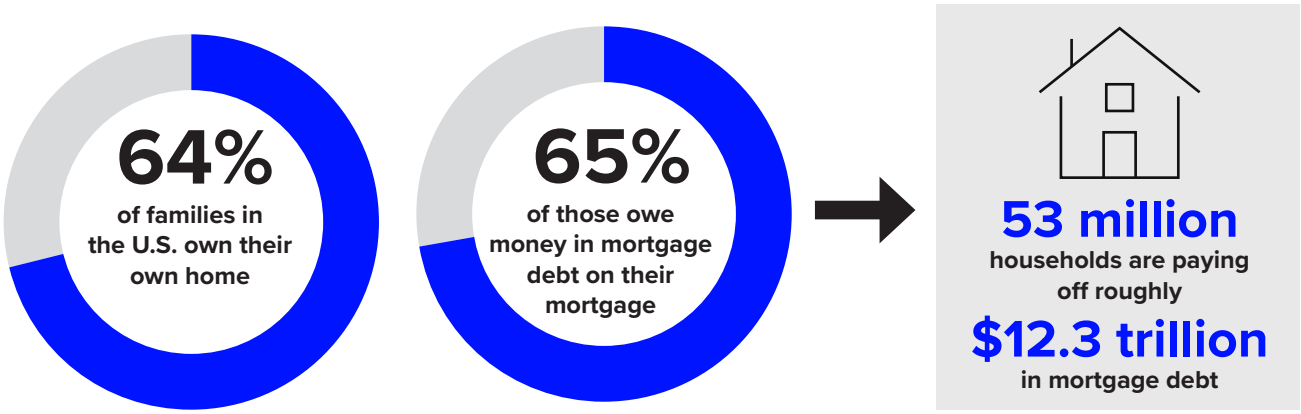
$9.4 trillion

in mortgage debt is held by institutional investors via mortgage-backed securities, CMOs and agency bonds

INTRODUCTION

Mortgages are a very familiar type of loan, simply because so many people have them. The latest census data shows that nearly 64% of families in the U.S. own their own home, with 65% of those owing at least some money on that mortgage[1]. That equates to roughly 53 million households that are in the process of paying off a combined total of roughly $12.3 trillion in mortgage debt[2].

U.S. HOUSEHOLDS WITH MORTGAGE DEBT



Source: Statista 2018

[1] https://www.statista.com/statistics/183635/number-of-households-in-the-us

[2] https://www.federalreserve.gov/data/mortoutstand/current.htm

MORTGAGE SECURITIES—A VAST MARKET

Data is the lifeblood of mortgage finance, as it is with most robust markets around the world. Banks, of course, want to know your credit score, credit history and expectations for income before lending you a few hundred thousand dollars for 30 years. But that is only a tiny tip of the very large iceberg. The majority of mortgage debt—$9.4 trillion in the second quarter of 2019—is ultimately held by institutional investors via mortgage-backed securities (MBS), collateralized mortgage obligations (CMO) and agency bonds (i.e., FHLMC, FNMA, GNMA). That $9.4 trillion outstanding generated $7.8 trillion in trading in July 2019[3], making it the second largest market in the world, topped only by the U.S. Treasury market, at $15.6 trillion outstanding.

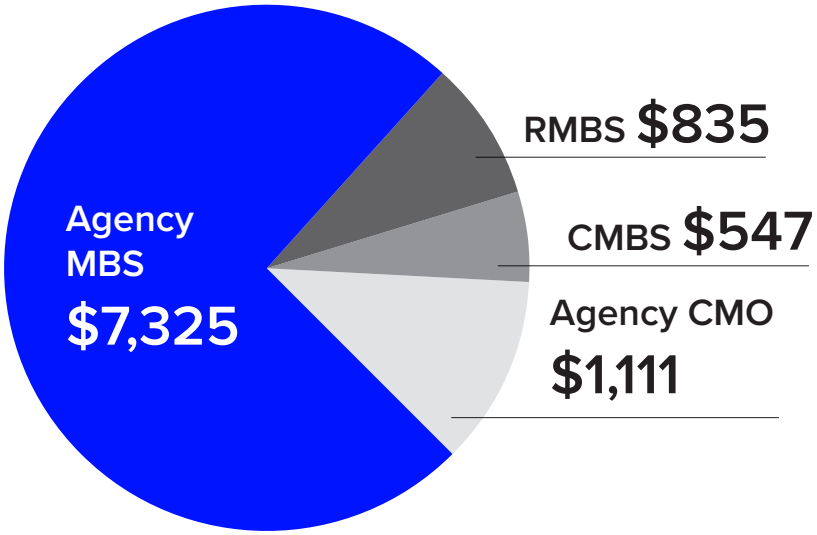
Institutional investors also care about your ability to repay your mortgage (and the likelihood that you'll pay it off before the term is up). But as these mortgages get packaged up into securities along with thousands of other loans, investors no longer have the advantage of getting to know the individuals behind each loan personally.

Instead, they need to work with complex risk models to help them determine the right price, yield, prepayment risks, and default risks, among many other critical metrics. The largest institutional mortgage dealers, such as J.P. Morgan, Wells Fargo and Bank of America, spend huge amounts of money and time to improve these models and to employ experts who can create and read them better than anyone else.

But even the best mortgage traders and analysts depend on the availability, quality and standardization of mortgage-related data. While this might sound like no big deal, given how digitized financial markets information has become over the past few decades, remember that nearly all of those mortgages were put into place in a community bank with borrowers signing a stack of paper a few inches thick. Transforming that stack of paper into usable data—for over 60 million mortgages—is anything but trivial.

A lack of due diligence on these mortgage documents played a huge part in the housing crisis of 2008. "Garbage in" clearly created "garbage out." Even risk-averse investors who thought they had done the necessary analysis of their investment came to learn that the data they were analyzing was at best lacking and at worst, downright wrong. This left many feeling that mortgage investing isn't good for the market and the economy—which is not true. However, greater data quality and data analysis is crucial to avoiding another market meltdown and to growing the market beyond its current state. But improving both of those areas remains the challenge for the market.

U.S. MORTGAGE-RELATED SECURITIES OUTSTANDING (BILLIONS)



Source: Federal Reserve, HUD, FHFA, Fannie Mae, Freddie Mac, Ginnie Mae

[3] DTCC

Greater data quality and data analysis is crucial to avoiding another market meltdown and to growing the market beyond its current state.

MACHINE LEARNING AND BIG DATA

Greater data quality and data analysis, and their associated obstacles, have made the mortgage market, which had its beginnings in the 1930s, a prime case for the use of machine learning and big data. The data comes in varying frequencies, in large volumes, is often dirty, and delivered in varying formats. While traditional relational databases can handle mortgage data in its simplest form—think principal balances and interest rates in structured formats—more complex data technology is needed to store and perform the needed analytics on millions of mortgages and mortgage securities. Refinitiv, for instance, estimates that capturing the entirety of the mortgage market and its related securities would require 7 petabytes of data—nearly triple the amount estimated the human brain is able to store.

The need for more advanced data technology is particularly critical for institutional mortgage investors who can't afford to wait until the next morning to understand the risk in their portfolio or the current value of what could be dozens of different mortgage-backed securities.

And as the analytics move beyond relatively standard agency securities into multi-issuer pools and complex CMOs—where diverse use-cases requiring complex ad-hoc queries and dynamic calculations are required—the challenge (and importance of the results) becomes even greater.

But running overnight batch jobs is still common in the world of mortgage investing, a practice that is only now starting to come to an end, as the ability to access the necessary data and analytics in the cloud is allowing near real-time calculations and results in any dimension.

BETTER DATA AND ANALYTICS IMPROVE TRANSPARENCY

That being the case, it is up to the capital markets community to improve the ways in which they capture and utilize that raw data to enhance transparency in the mortgage market. While mortgages and mortgage-related investments were the primary cause of the financial crisis, things have certainly improved in the decade since. It is now possible to dig into the individual loans within any given mortgage-backed security, enabling investors to deeply examine exactly what they are buying, rather than simply accepting the word of their banks and credit-rating agencies.

Furthermore, nearly two-thirds of MBS are now traded electronically, according to Greenwich Associates data, a huge accomplishment for a complex over-the-counter market. And as history has demonstrated, more data leads to more electronic trading, which leads to more data—a virtuous cycle.

Industry-agreed standards have also helped, as agency mortgage pools have gone from generic groups of mortgages to more specific groupings (often referred to as "spec pools"), allowing investors to target certain loan characteristics such as geography, FICO score or specific originators.

This and additional clarity around MBS construction helps market participants analyze the risks of holding and trading these instruments and feel more confident in the results.

Yet, there is more progress to be made. Market participants must not feel beholden to existing models for measuring risk, but be open to new solutions based on new data and new analytics technology. Improving the raw data and the industry's trust in it will also be critical, as the industry works to incorporate machine-learning technology that can digest mortgage documents, making the information within more accessible to all. And as mortgage-related products get more complex, as they inevitably will, it is critical that analytics that can generate results faster are made available to both issuers and investors to ensure the true risks and appropriate prices are well understood. Empowering lenders, loan servicers and investors with higher quality, more easily accessible data will ultimately mitigate risk and create a more robust mortgage market for all.

As history has demonstrated, more data leads to more electronic trading, which leads to more data – a virtuous cycle.

Refinitiv is one of the world's largest providers of financial markets data and infrastructure, serving over 40,000 institutions in over 190 countries. We are new market pioneers with 167 years of confidence. We provide leading data and insights, trading platforms, and open data and technology platforms that connect a thriving global financial markets community – driving performance in trading, investment, wealth management, regulatory compliance, market data management, enterprise risk and fighting financial crime.

Learn how Refinitiv can help you solve the challenges posed by expanding and fragmented mortgage data, with Advanced Mortgage Analytics. Visit: **refinitiv.com/mortgages**

REFINITIV™

DATA IS JUST
THE BEGINNING™

RE1033433/11-19