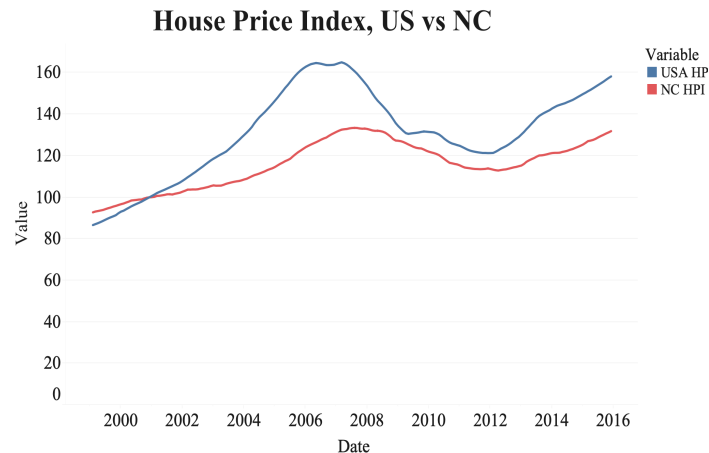


Abstract

Our group examined the housing bubble that crashed in the late 2000s. We manipulated a range of data sets to analyze the causes of the housing bubble and create interactive visualizations which could appear on a website.

US Housing Market Statistics



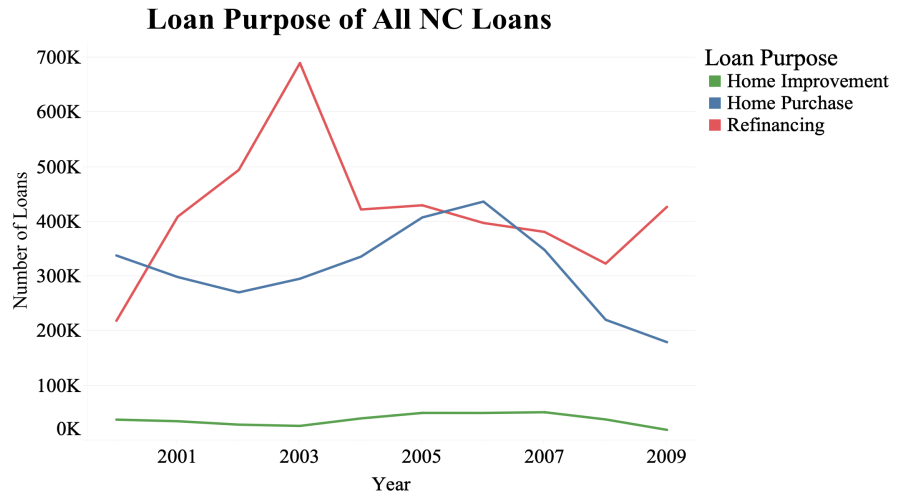
After a period of high appreciation in the years leading up to the crisis, home prices dropped starting in 2007. This decline was triggered largely by subprime and predatory lending. North Carolina shows a smaller rise than the USA index, possibly because of outliers in high home price areas.

LAR Data

Respondent ID	Lender	Applicant Income	Loan Amount	Applicant Race	Rate Spread	...
1	Bank of America	60	92	White	5.3	...
2	Country wide	46	139	Black	5.2	...
3	Wells Fargo Bank	52	72	Asian	5.1	...
4	Country wide	33	64	White	5.8	...

Above is a snapshot of our (clean) main dataset: the Loan Application Register from the Home Mortgage Disclosure Act. Provided is loan level data from all big lenders, with these variables and other important demographic/geographic factors, which aid in identifying predatory lending. Given as an example here are loans from a few of the major players in the market leading up to the housing crisis.

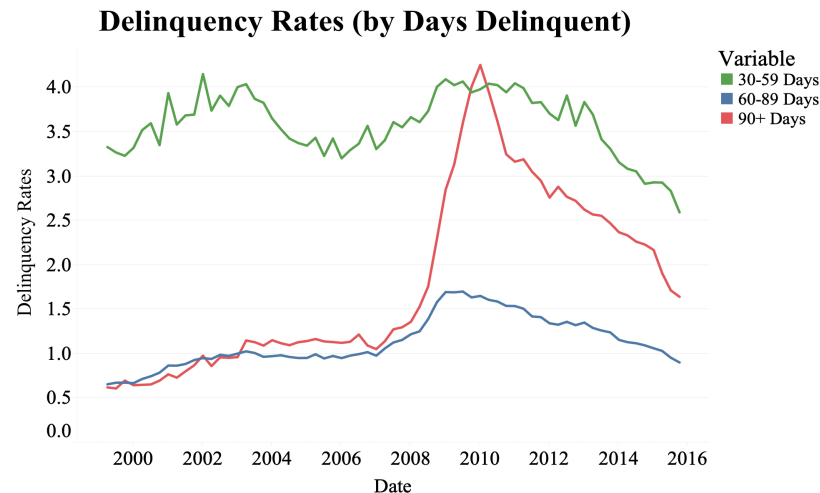
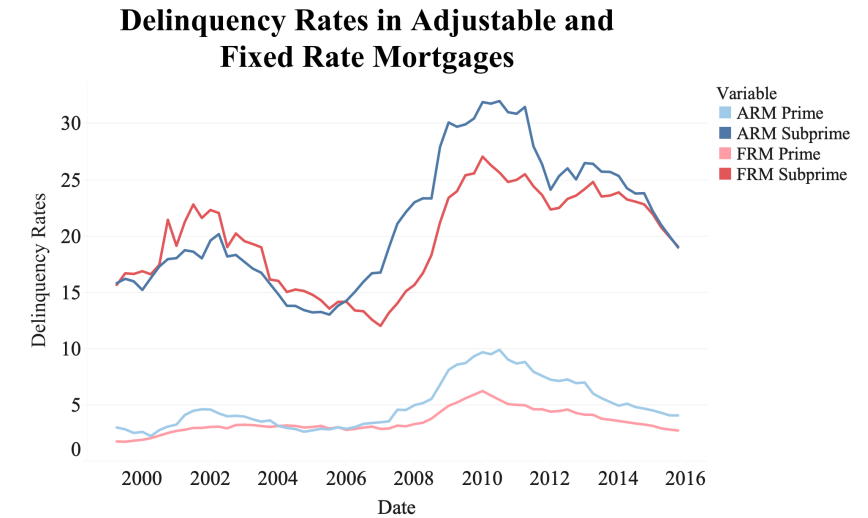
Loan Purpose



In the early 2000s, the housing market followed a predictable trend. Refinances were given to borrowers at a high rate—with housing prices continuing to rise, borrowers were reassured that they would be able to refinance to a better interest rate once the teaser rate expired on their adjustable rate mortgages. This approach worked for a long time, but it stripped equity out of homes, making it a mechanism for predatory lending. Once house prices dropped, this method stopped working, leaving borrowers with payments they could not afford.

Delinquency Rates

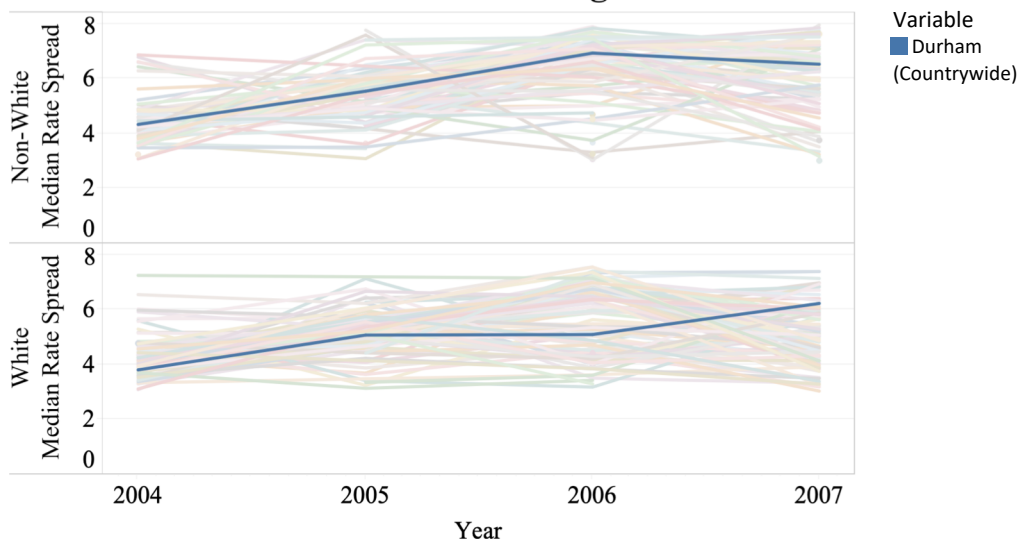
The graph to the right shows delinquency rates by mortgage type. Through the crisis, delinquency rates were consistently higher for adjustable rate over fixed rate mortgages, and subprime (identified by MBA) loans had consistently higher delinquency rates than their prime counterparts. Most rates peak in the few years after the crisis, when the recession hit the American people the hardest.



Shown is the delinquency rate (by number of days delinquent) through the crisis. The large spike comes from loans that were more than 90 days delinquent, which rose by over 3% during the crisis.

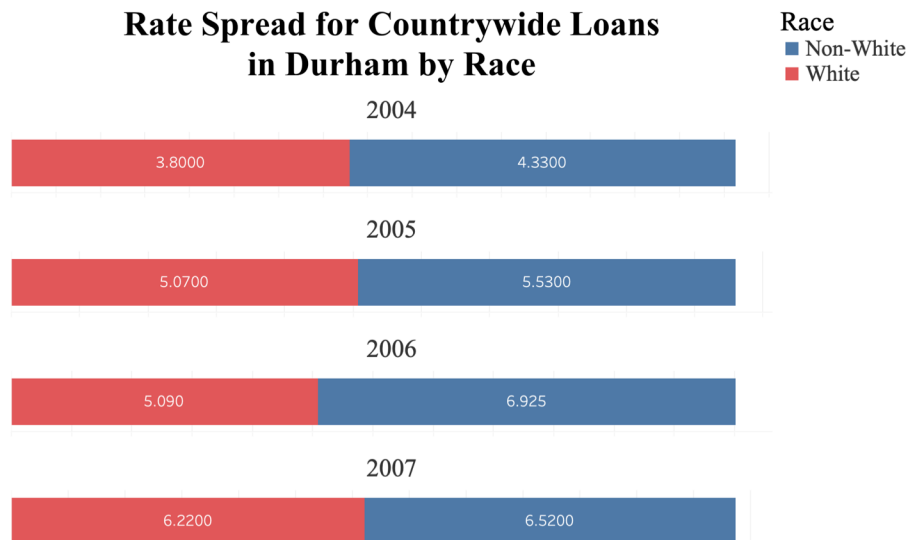
By-Loan Analysis

White vs. Non-White Rate Spread in Countrywide Home Purchase/Refinancing Loans



Shown is the difference in rate spread (a proxy for identifying subprime loans) by race in Durham. We selected the county for the diversity and similar loan to income ratios by race—meaning there is no systematic difference in loans undertaken by borrowers with the same income. We see consistent rises in the rate spread for both, meaning Countrywide made riskier subprime loans as time went on.

Rate Spread for Countrywide Loans in Durham by Race



Even more significant is the digression of the lines by race. The rate spread is higher, (in 2006, 1.835 points higher) in the Non-White than White category. Despite only a slight difference in some years, any difference provides a window into potential racial targeting by Countrywide Financial.

Conclusion/Next Steps

The Bass Connections team at Duke will continue this work in the fall, collecting testimonies from those who were affected by predatory lending practices in North Carolina in an effort to gain a holistic view of the housing crisis, and the recession that followed. From there, an interactive website will be created with the goal of providing the general public with a complete view of the causes and happenings of the crisis, coupled with applets that allow for interactive data exploration.

Acknowledgements

The Data+ team would like to thank Kate Coulter, Lee Reiners, Emma Rasiel, Keith Ernst, Jim Parrott, Michael Stegman, Joseph Smith, and Edward Balleisen for their contributions to this project, as well as Moody's Analytics for providing a vast amount of data for analysis.