Jack Coyne
Data Wrangling Final Report
05/03/2024

Analysis of The 2008 Housing Crisis in R

Introduction

In late 2007, the onset of the global financial crisis marked a pivotal moment in economic history, particularly for the U.S. housing market. Often referred to as the most severe financial crisis since the Great Depression, the turmoil began with the bursting of an 8 trillion-dollar housing bubble. The repercussions were devastating foreclosures surged, house prices plummeted, and countless financial institutions found themselves on the brink of collapse. This period of economic instability not only eroded personal wealth and consumer confidence but also had a profound impact on the broader global economy.

The primary focus of this report is to delve into the intricacies of the housing market during and around the 2008 financial crisis. It aims to explore the dynamic interactions between key economic indicators—such as mortgage rates, GDP, unemployment rates, and the Consumer Price Index (CPI)—and their influence on housing prices. This analysis is particularly relevant as the housing market is a significant component of the national economy, influencing monetary policy, investment strategies, and financial stability. Understanding how these indicators interacted during the crisis can provide valuable lessons for predicting and mitigating future economic downturns.

The motivation for this study stems from the far-reaching implications of the housing market on economic stability and individual livelihoods. Given that housing is both a basic need and a central investment for American families, fluctuations in the housing market can have widespread socio-economic consequences. By examining the factors that contributed to the housing market's instability during the crisis, this study seeks to uncover patterns and insights that could inform future policy and economic decision-making, aiming to buffer against similar crises.

In this report, I will address several key research questions that explore the relationships between economic indicators and housing market dynamics. I hypothesize that negative economic conditions, such as higher unemployment and reduced GDP, are correlated with lower housing prices and increased mortgage rates. Through a combination of correlation analysis, regression models, and time-series forecasting, this study will quantify these relationships and predict future trends. These methods will not only help in understanding

the past but also in preparing for future economic challenges, making this study essential for policymakers, economists, and stakeholders in the housing market.

Data

2.1 FHFA Single-Family Mortgage Originations

For the analysis of mortgage dynamics from 2000 to 2010, I utilized data on single-family mortgage originations provided by the Federal Housing Finance Agency (FHFA). The data included key details on mortgage originations, which were pivotal for understanding lending trends during the specified period. After obtaining the data from FHFA's records, I executed a script to clean and prepare the dataset for analysis. This process involved normalizing column names for consistency, filtering the dataset to include only the relevant years (2000 to 2010), and exporting the cleaned data to a CSV file named "cleaned_sf_mortgage_originations.csv". The script responsible for these operations is maintained within my project folder.

2.2 National Mortgage Aggregate Statistics

The analysis further incorporated national mortgage statistics, detailing quarterly data spanning the same period. This comprehensive dataset was compiled from multiple CSV files obtained from national mortgage databases. Each file corresponded to a different year and required integration into a single dataset. My script automated the reading and combination of these files, applying necessary transformations such as renaming columns for uniformity and rounding numerical values for better precision. The final dataset focused solely on national-level data, filtering out any non-relevant geographic details and was saved as "cleaned_national_mortgage_aggregate_statistics.csv".

2.3 FRED API Economic Indicators

To incorporate economic indicators into the analysis, data was fetched from the Federal Reserve Economic Data (FRED) API. This included critical metrics like GDP, unemployment rate, consumer price index, and others. The data fetched was aligned to match the quarterly format of the mortgage data, ensuring consistency across all datasets. These economic indicators were vital for analyzing the interplay between the housing market and broader economic conditions.

2.4 Data Integration and Final Merging

The final step in data preparation involved merging the cleaned datasets into a single comprehensive frame to facilitate the analysis. This included merging FRED economic

indicators with FHFA mortgage data based on the 'year' and 'quarter' columns, ensuring that each data point from the economic indicators corresponded accurately to the mortgage data. The final merged dataset, named "final_merged_fred_and_mortgage_data.csv", underwent a final cleanup process to standardize column names and formats, preparing it for the analytical phase.

For the final analysis, the following variables from the merged dataset were utilized:

Variable Name	Type	Source	Description
year	Integer	Both	The year of the data record, identifying the timeframe of the data.
quarter	Integer	Both	The quarter of the year for the data record.
mortgagerate	Numeric	FRED	The average mortgage interest rate per quarter.
realgdp	Numeric	FRED	Real Gross Domestic Product, indicating economic growth.
unemploymentrate	Numeric	FRED	The unemployment rate, a key economic performance indicator.
consumerpriceindex	Numeric	FRED	Measures inflation as experienced by consumers in the economy.
federalfundsrate	Numeric	FRED	The interest rate at which depository institutions lend funds.
delinquencyrate	Numeric	FRED	Percentage of loans where payments are overdue.
homepriceindex	Numeric	FRED	An index that measures the price changes of residential housing.
mean_loans_average_loan_amount	Numeric	FHFA	Average amount of loans processed in the period.
mean_loans_average_property_value	Numeric	FHFA	Average property value associated with the loans.
mean_loans_average_vantagescore	Numeric	FHFA	Average credit score of borrowers during the quarter.

Table 1 Data Dictionary

This data dictionary provides an overview of each variable used in the analysis, ensuring clarity and transparency in the processing and examination of the data. Each variable was chosen based on its relevance to the research questions posed in the study, focusing on

the economic impact on mortgage dynamics, market conditions, crisis impact, and predictive analysis of future trends.

https://www.fhfa.gov/DataTools/Downloads/Pages/National-Mortgage-Database-Aggregate-Data.aspx
https://www.fhfa.gov/DataTools/Downloads/Pages/Current-Market-Data.aspx
https://fred.stlouisfed.org/docs/api/fred/

Analysis

3.1 Economic Impact on Mortgage Dynamics

In my exploration of how various economic indicators impact mortgage rates, a detailed correlation analysis was initially conducted. The correlation matrix, pivotal in this analysis, highlights the intrinsic relationships between several economic parameters and mortgage rates. For instance, a notable negative correlation of -0.65 between Real GDP and mortgage rates suggests that economic growth often coincides with reduced mortgage rates. This trend might be due to lower interest rates during periods of economic expansion, which enhance borrowing and investment activities.

Similarly, the unemployment rate exhibits a robust negative correlation of -0.80 with mortgage rates. Typically, higher unemployment rates are associated with economic downturns, which can prompt lower interest rates as a countermeasure to stimulate economic activity. Conversely, the Consumer Price Index (CPI), with a correlation of -0.72, indicates that as inflation increases, mortgage rates tend to decline, perhaps reflecting central bank policies aimed at controlling inflation through lower interest rates.

The Federal Funds Rate displayed a positive correlation of 0.79 with mortgage rates. This relationship underscores the direct impact of central bank interest rate policies on mortgage rates, where increases in the federal funds rate, an instrument to curb inflation or cool an overheating economy, directly translate to higher mortgage rates.

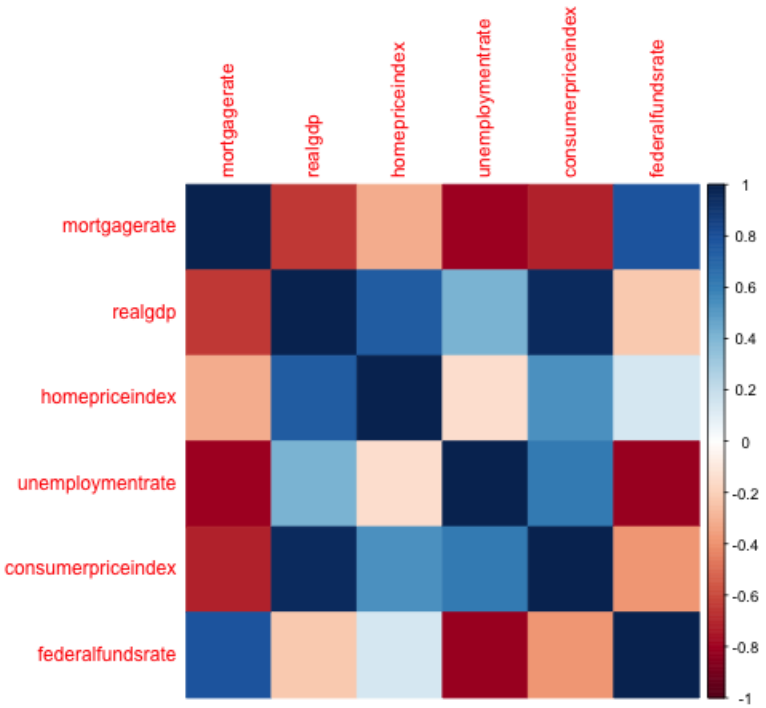


Figure 1 Correlation Matrix

The regression analysis further quantified these effects, showing that each unit increase in the unemployment rate decreases mortgage rates by approximately 0.262 percentage points, holding other factors constant. This model, which accounted for 90.07% of the variance in mortgage rates (Adjusted R-squared = 0.8905), emphasizes the significant impact of economic conditions on lending rates. Additionally, CPI's positive coefficient (0.065) indicates that as consumer prices increase, reflecting inflation, mortgage rates tend to rise as well, likely due to the inflationary pressures prompting lenders to demand higher returns.

Figure 2 Impact of GDP and CPI on Mortgage Rates

The plot of GDP and CPI against mortgage rates visualizes these relationships, depicting how economic growth and inflation correlate with changes in mortgage rates. This visual representation helps stakeholders understand the economic levers affecting mortgage finance.

3.2 Market Conditions and Borrower Profiles

Focusing on how market conditions influence borrower behavior, regression analysis on the average loan amounts relative to economic indicators revealed that higher consumer prices (CPI) are associated with larger loan amounts, with a coefficient of 3.640 indicating a substantial increase in loan amounts per unit increase in CPI. This result, statistically significant ($p < 0.001$), suggests that as prices rise, so do the costs of properties, consequently increasing the loan amounts borrowers need to procure. Conversely, the negative coefficient for the federal funds rate (-3.596) indicates that higher interest rates reduce loan amounts, likely due to the increased cost of borrowing.

A t-test was conducted to evaluate the impact of the 2008 financial crisis on mortgage rates, comparing pre-crisis and post-crisis periods. The results were statistically significant, with the pre-crisis period averaging higher mortgage rates (6.48%) compared to the post-crisis period (5.26%), with a p-value of less than 0.001. This significant difference

underscores the profound impact of the financial crisis on mortgage rates, reflecting the aggressive rate cuts and monetary easing policies implemented to stimulate the economy.

The relationship between average loan amounts and property values was illustrated through a scatter plot, where different levels of CPI are represented by varying colors.

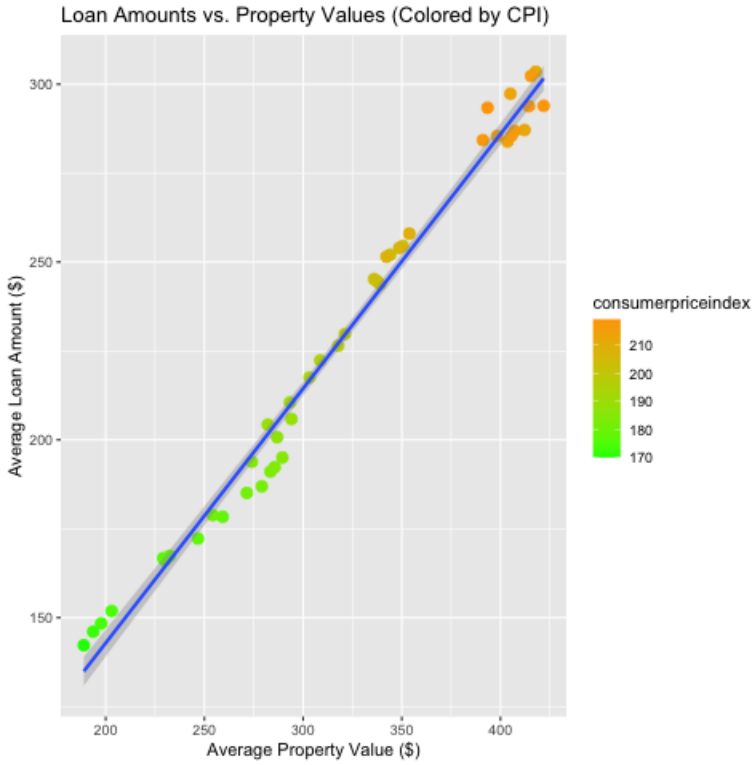


Figure 3 Loan Amounts vs. Property Values (Colored by CPI)

This visualization effectively conveys how economic conditions, reflected through CPI levels, influence both property values and the amount borrowed. The trend line in the plot further helps to clarify the positive correlation between property values and loan amounts, affirming that as property values increase, so do the amounts that borrowers are willing to finance.

This analytical section provides a clearer understanding of how macroeconomic factors affect borrower behavior and market conditions in the mortgage industry. By detailing the specific influences of economic indicators on mortgage characteristics, this analysis aids stakeholders in making informed decisions in the housing and finance sectors.

3.3 Crisis Impact Analysis

The regression model analyzing the impact of the 2008 financial crisis on delinquency rates and home price indices revealed that economic downturns, characterized by rising unemployment and CPI, significantly increase delinquency rates, while also negatively impacting home prices. The coefficients from the model indicated that unemployment rate increases are strongly associated with higher delinquency rates (coefficient = 1.601, $p < 0.001$), while rises in Real GDP slightly buffer against declines in home prices (coefficient = 0.074, $p < 0.001$).

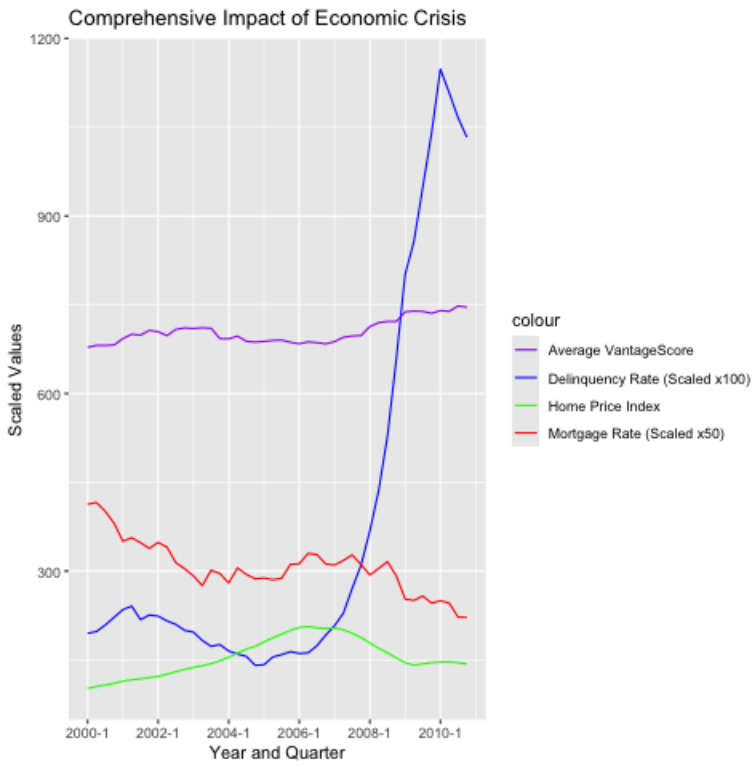


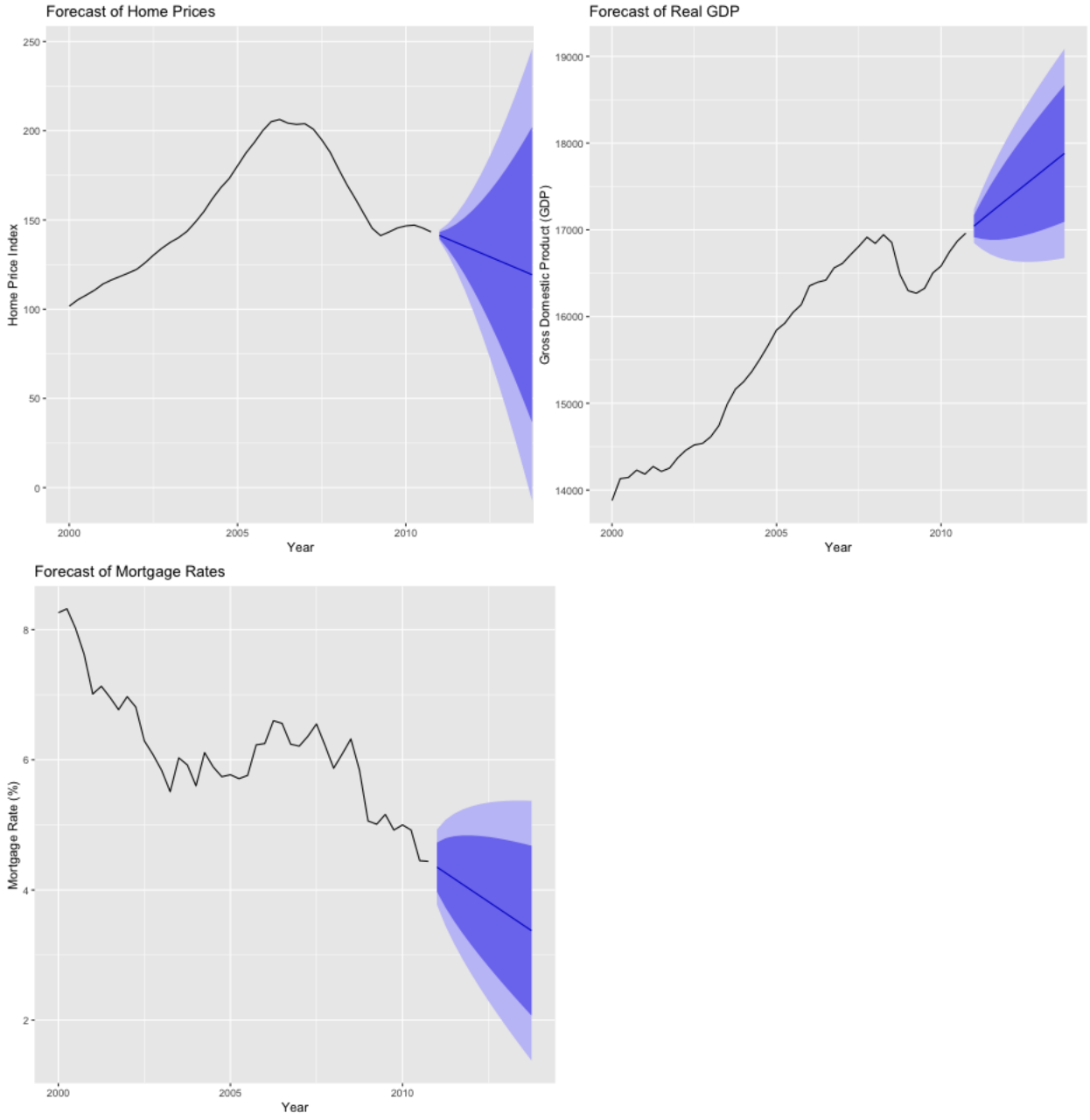
Figure 4 Impact of Economic Crisis

This plot presents the trends over time, illustrating how economic factors like GDP, unemployment, and consumer prices impact delinquency rates and home values through the crisis period, offering a dynamic view of the economic landscape's influence on the mortgage market.

3.4 Forecasting and Predictive Analysis

Using time series analysis, forecasts for mortgage rates, home prices, and Real GDP were generated to predict future economic conditions. The ARIMA model forecasts suggest a continued decline in mortgage rates and home prices over the next three years, reflecting a

potentially easing monetary policy environment or continued recovery from previous economic shocks. The forecast for GDP indicates modest growth, suggesting a gradual economic recovery.



Figures 5, 6, 7, Forecasts of Home Prices, Real GDP, and Mortgage Rates

These visual forecasts allow stakeholders to gauge future economic conditions, assisting in strategic planning and decision-making by providing a predictive outlook on mortgage rates, property values, and overall economic activity.

This section of the report synthesizes complex economic data into actionable insights, using both statistical analyses and visual tools to illustrate the profound impacts of economic variables on the mortgage market. By providing a detailed and nuanced understanding of these dynamics, the report aids stakeholders in navigating the complexities of the financial landscape.

Conclusion

This project provided a detailed examination of the economic factors influencing mortgage dynamics and housing market conditions during and after the 2008 housing crisis. By integrating and analyzing data on mortgage rates, home prices, GDP, unemployment rates, and other relevant economic indicators, we gained critical insights into how this financial upheaval affected the U.S. housing market.

The analysis revealed significant relationships between economic downturns—marked by rises in unemployment and drops in GDP—and the deterioration of housing market health, evidenced by falling home prices and rising mortgage delinquencies. The regression analyses quantified these effects, highlighting the sensitivity of mortgage rates and housing stability to broader economic shifts. The period comparison using t-tests provided statistical proof of the lasting impacts of the crisis, showing a significant decrease in mortgage rates post-crisis as a response to federal intervention and economic stimulus measures.

However, the study's scope was limited to historical data, which may not fully predict future market behaviors under different crisis conditions. Furthermore, while the analysis incorporated several key economic indicators, other potential influences such as consumer confidence indices and international economic trends were not examined, which could provide further insights into the housing market's susceptibility to global economic shifts.

Future research could build on these findings by examining more recent economic crises, such as the COVID-19 pandemic, to understand similarities and differences in impact compared to the 2008 crisis. Additionally, exploring the effects of specific government interventions or policy changes on the housing market could offer valuable lessons for mitigating future financial crises.

In conclusion, this project underscores the complex interplay between economic indicators and the housing market, providing valuable lessons for economists, policymakers, and investors aiming to navigate or mitigate future financial downturns. The

insights garnered here reinforce the importance of vigilant economic monitoring and proactive policy interventions to sustain housing market stability during economic shocks.