City-Level Dynamics of U.S House Prices and Mortgage Value , 2002-2011

Author: Brandon Smith

DATA 501 - Data Science in R

Dr. Zechariah Meunier

Willamette University – School of Computing and Information Sciences

December 5th, 2025

Introduction:

1.1 Project Overview

Currently, the U.S. is in a housing affordability crisis, with the prices for affordable homes on a downward trend year after year. Since the 1980s, the price-to-income ratio has been rising, reaching 4.01 in 2005, the highest until recently, and 4.42 in 2024 (The Outlook for U.S. Housing Supply and Affordability, 2025). However, the cause was not an aftereffect of the 2007 financial crisis, driven by subprime adjustable-rate loans given to those with low or even no FICO scores (Ceciliano, 2018). It was not even the collateralized debt obligations (CDOs) that were alone responsible for $542 billion to private financial institutions (Barnett-Hart, 2009). No, the current COVID-19 crisis has resulted in very high demand and extremely low supply, pushing prices higher under the simple laws of supply and demand (The Outlook for U.S. Housing Supply and Affordability, 2025).

The opposite of what happened in the 2007 financial crisis, where supply was high, and demand was low. Though this was not entirely due to chance, it began with the administration at the time. The Bush administration sought to create a "ownership society" and increase homeownership, specifically among minorities, by 5.5 million by 2010 (Record of Achievement - Expanding Home Ownership, 2006). To achieve this, President Bush implemented several policies. One of those policies included the American Dream Downpayment Act, which allocated $200 million to help roughly 40,000 families each year with external costs (downpayments and closing costs) (Record of Achievement - Expanding Home Ownership, 2006). He followed that up with the Zero-Downpayment Initiative, which allowed the Federal Housing Administration (FHA) to allow first-time homebuyers to purchase a home without a down payment (Record of Achievement - Expanding Home Ownership, 2006). Though he did not just create policies, he enlisted the public and private sectors in the cause, and more than a dozen organizations and businesses pledged to it, providing more than $1.1 trillion in mortgage purchases for minority homebuyers in the 2000s. (Record of Achievement - Expanding Home Ownership, 2006).

Though President Bush did not know what would happen in the coming years, collateralized debt obligations (CDOs) were a key factor in the 2007 housing bubble. CDOs contained many assets, which contained, but were not limited to, things such as "bonds, corporate debt, credit card receivables, loans, and mortgages" (Tardi, 2025). Though the specific CDOs that caused the crisis were those that contained subprime or Alt-A, adjustable-rate mortgages. These were dangerous since subprime means the person has a FICO score below 650 (Barnett-Hart, 2009). Alt-A is a classification based on their prime profile and is also the most common to default on (Kagan, 2025). The combination of these subprime and Alt-A adjustable-rate mortgages was one of the leading causes of the housing bubble, as the number of defaults was too high because homeowners were unable to pay their new, higher mortgage rates.

To understand the dynamics and conditions that lead to volatility, we use a statistical model that accounts for variation across 18 metropolitan cities. These cities include Phoenix, Los Angeles, San Diego, San Francisco, Denver, Washington D.C, Miami, Tampa, Atlanta, Chicago, Minneapolis, Charlotte, Las Vegas, New York, Cleveland, Portland, Dallas, and Seattle. Including city fixed effects, which allows for ignoring external factors that might impact the data; these factors can include zoning laws, local economic structures, and more. Allowing me to analyze the House Value in the "cities-month-SA.csv" data set as my explanatory variable. With the dependent variables being mortgage values and cities from the "nmdb-mortgage-performance-statistics-metros-quarterly.csv" file.

1.2 Research Question

RQ1: Is there a relationship between mortgage value based on delinquency or defaults and the house value that affects the time period for buyers, or is it a city fixed effect?

Analyzing this to see if there is a relationship (reject the H0). If there is a relationship, it creates a predictable signal for homebuyers and also indicates a "city fixed effect". Helping me achieve the goal of making it easier for everyone to purchase affordable housing.

RQ2: Do the cities differ in their house value and mortgage value trends over time?

Wanting to analyze whether all housing market metrics, house values, and mortgage value are uniform across cities, meaning that all cities behave the same. If this is not the case and there is no uniformity, policies can affect cities differently. For example, the American Dream Downpayment Act might positively affect Chicago, but negatively affect San Diego. Furthermore, there is a "city fixed effect," meaning that a city's policymakers are not correlated.

1.3 Hypotheses

RQ1:

- $H_0 : \beta_1 = 0$ (There is no significant relationship between house value and mortgage value)
- $H_1 : \beta_1 \neq 0$ (There is a significant relationship between house value and mortgage value)

RQ2:

- $H_0 : \beta_1 = 0$ (There is no significant difference in house value and mortgage value trends between cities)
- $H_1 : \beta_1 \neq 0$ (There is a significant difference in house value and mortgage value trends between cities)

Methods:

2.1 Statistical Test and Assumptions

Research Question 1:

Table 1: Shapiro-Wilk Normality Test Results of House Value by City

city	shapiro
AZ-Phoenix	4.22E-24
CA-Los Angeles	3.73E-37
CA-San Diego	3.56E-37
CA-San Francisco	2.08E-36
CO-Denver	2.34E-38
DC-Washington	2.84E-36
FL-Miami	6.87E-37
FL-Tampa	8.02E-28
GA-Atlanta	8.25E-39
IL-Chicago	6.28E-29
MN-Minneapolis	1.71E-32
NC-Charlotte	1.25E-38
NV-Las Vegas	1.25E-36
NY-New York	2.48E-31
OH-Cleveland	9.98E-38
OR-Portland	1.64E-30
TX-Dallas	6.05E-32
WA-Seattle	8.16E-24

Table 2: Levene's Test for Homogeneity of Variance in House Value by City

Levene's Test for Homogeneity of Variance (center = median)

	Df	F value	Pr(>F)	
group	17	735.52	< 2.2e-16	***
	33102			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

I began the diagnostic testing by first conducting a Shapiro-Wilk normality test on the house value for each city in the data group. The purpose is to predict whether the house values within each city follow a normal distribution. After reviewing the data, I can see that all cities have small p-values (Figure 1: Shapiro), which leads me to reject the null hypothesis in all cases.

Furthermore, I conducted a Levene's Test for Homogeneity of Variance to determine whether the spread, or variance, of house values was consistent across cities. I am seeing if the variances are equal across all groups (Pr(>F) > 0.05). After reviewing the table, Pr(>F) is smaller than the significance level of 0.05, indicating that the house value variance was significantly unequal across cities, violating both the assumption of normality (Table 1) and the assumption of homogeneity of variance (Table 2). With this, I cannot perform a regular parametric test such as a one-way ANOVA. I must conduct the nonparametric one-way ANOVA, which is the Kruskal-Wallis rank-sum test.

Research Question 2:

For research question 2, I modeled mortgage value (VALUE1) as a function of house value by city and examined the model diagnostics when being treated as a standard regression assumption. I performed this analysis for two specific cities, Los Angeles and San Francisco, using tests such as Residuals vs Fitted, Q-Q plots, Scale-Location, and Residuals vs Leverage. When analyzing the plots, I noticed signs of non-normality and heteroskedasticity, which led me to perform an emmeans/emtrends analysis (Appendix Figures, 3.1-3.8). I prefromed this test to focus on trend estimates rather than a regular linear regression.

2.2 Description of Both Explanatory and Response Variables

RQ1:

My first research question focuses specifically on the geographic location (GEONAME) and whether the House Value influences it. The geographic location (from the "nmdb-mortgage-performance-statistics-metros-quarterly.csv") is my Explanatory variable (X). I define the Response Variable (Y) as House Value from the "cities-month-SA.csv" file. I modeled the House Value using geographic variables and aimed to determine whether and how specific regulatory environments and regional factors systematically affected local housing price movements.

RQ2:

For the second research question, I swapped the explanatory and response variables from the first question. The goal is to analyze the impact of a housing price shift on mortgage value . House Value, my Explanatory Variable (X), from the "cities-month-SA.csv" file. I hypothesized that changes in House Value values can predict mortgage value . VALUE1 (mortgage_value), my Performance Index (Response Variable Y), is sourced from "nmdb-mortgage-performance-statistics-metros-quarterly.csv" and it identified by the SERIESID; summarize mortgage performance for each metro (e.g, delinqueny or default rates) in standard index form. This analysis allowed me to test whether changes in house values were significant predictors of changes in mortgage value (delinquency or default) within those same metropolitan areas.

2.3 R-Packages Used

I used five different R libraries to provide diagnostic tests, data manipulation, etc. I used the tidyverse suite of packages to streamline my manipulation and visualization, taking advantage of their shared data repositories and design to work together. I used the car package, which allowed me to perform additional calculations on the fitted models, such as testing for homogeneity of variance. Similarly, I used the lmtest package to diagnose specific regression relationships and ensure my model assumptions met the requirements. To look into the finer details, I used emmeans to obtain the marginal means for specific factors, which are necessary for examining contrasts among them. Finally, I employed the broom package to convert the statistical analysis objects into tibbles for readability and reporting.

Results:

3.1 Summary Statistics and Figure 1

Table 3: Kruskal-Wallis rank sum test results for House Value by City

```
        Kruskal-Wallis rank sum test

data:  house_value by
city
Kruskal-Wallis chi-squared = 20061, df = 17, p-value < 2.2e-16
```

The Kruskal-Wallis test I performed addresses the null hypothesis that the distribution of house values is the same across all cities. The results, I observed showed the Kruskal-Wallis chi-squared value was 20,061 on 17 degrees of freedom, with that the p-value was less than 2.2e – 16, which is a significant difference to the a. Therefore, I reject the null hypothesis (H_0), meaning that there is a significant result that at least one city's house value distribution is statistically differnt from the others (Table 3). Given the results, I plan to use the emmeans (estimated marginal means) package to conduct pairwise comparisons and identify which specific cities differ.

Figure 1: House Value by City

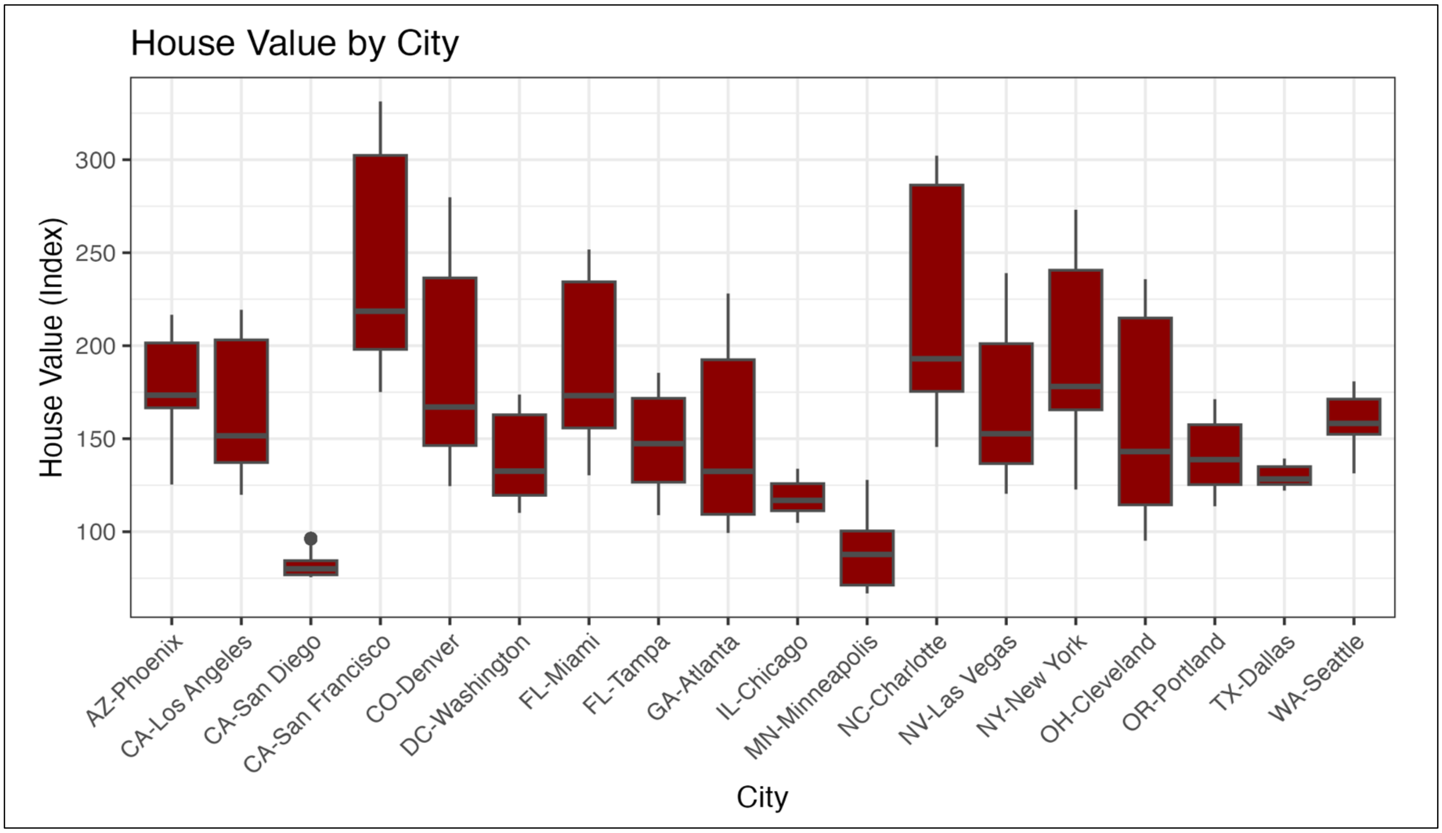


Table 4: Summary Statistics of House Values by City (Median and IQR)					
city	median_house	sd_house	q1	q3	iqr
AZ-Phoenix	173.397	24.673247	166.61	201.499	34.889
CA-Los Angeles	151.55	33.11046	137.24	203.116	65.876
CA-San Diego	80.065	5.587468	76.837	84.452	7.615
CA-San Francisco	218.575	50.689857	198.024	302.301	104.277
CO-Denver	167.046	51.672082	146.345	236.415	90.07
DC-Washington	132.581	21.901078	119.592	162.852	43.26
FL-Miami	173.151	39.521534	155.784	234.3	78.516
FL-Tampa	147.334	23.994399	126.621	171.736	45.115
GA-Atlanta	132.466	44.661935	109.382	192.428	83.046
IL-Chicago	116.838	8.813881	111.283	125.877	14.594
MN-Minneapolis	87.784	18.239204	71.296	100.378	29.082
NC-Charlotte	192.96	52.379163	175.44	286.343	110.903
NV-Las Vegas	152.669	37.865211	136.627	201.115	64.488
NY-New York	178.068	45.631493	165.546	240.583	75.037
OH-Cleveland	143.077	50.909715	114.452	214.883	100.431
OR-Portland	138.762	17.716164	125.317	157.528	32.211
TX-Dallas	128.344	5.126243	125.398	134.97	9.572
WA-Seattle	158.221	12.429811	152.379	171.288	18.909

3.2 Inferential results for RQ1

I analyzed the median and interquartile range (IQR) to account for the non-normal distribution I previously identified. After reviewing, I concluded that there is a massive difference in central tendency across the cities. Specifically, San Francisco has the highest median house value at 218.575. Cities like Charlotte (192.96), New York (178.068), and Phoenix (173.397). Furthermore, I can find cities like San Diego (80.065) and Minneapolis (87.784) with the lowest median house values. The baseline for a house varies significantly nationally and by state.

Beyond this analysis, when I look at both the Standard Deviation and the IQR. I understand volatility within each market location. When I compare San Francisco (the city with the highest median house value), I see it is incredibly volatile, with an IQR of 104.277. Meaning that the difference between quartile one and quartile three is 104.277. However, this is not even the highest difference, with Charlotte having an IQR of 110.903, quartile one at 136.627, and quartile three at 286.343. These points point to a highly diverse market with many house sales at both low and high prices. This is not the same for all cities; San Diego and Dallas have the lowest IQR, with San Diego at 7.615 and Dallas at 9.572. This shows that some cities are tightly clustered, and others vary widely.

3.3 Relationship plots and Figure 2

Figure 2: Comparison Plot of San Francisco vs Los Angeles House Value by Mortgage Value

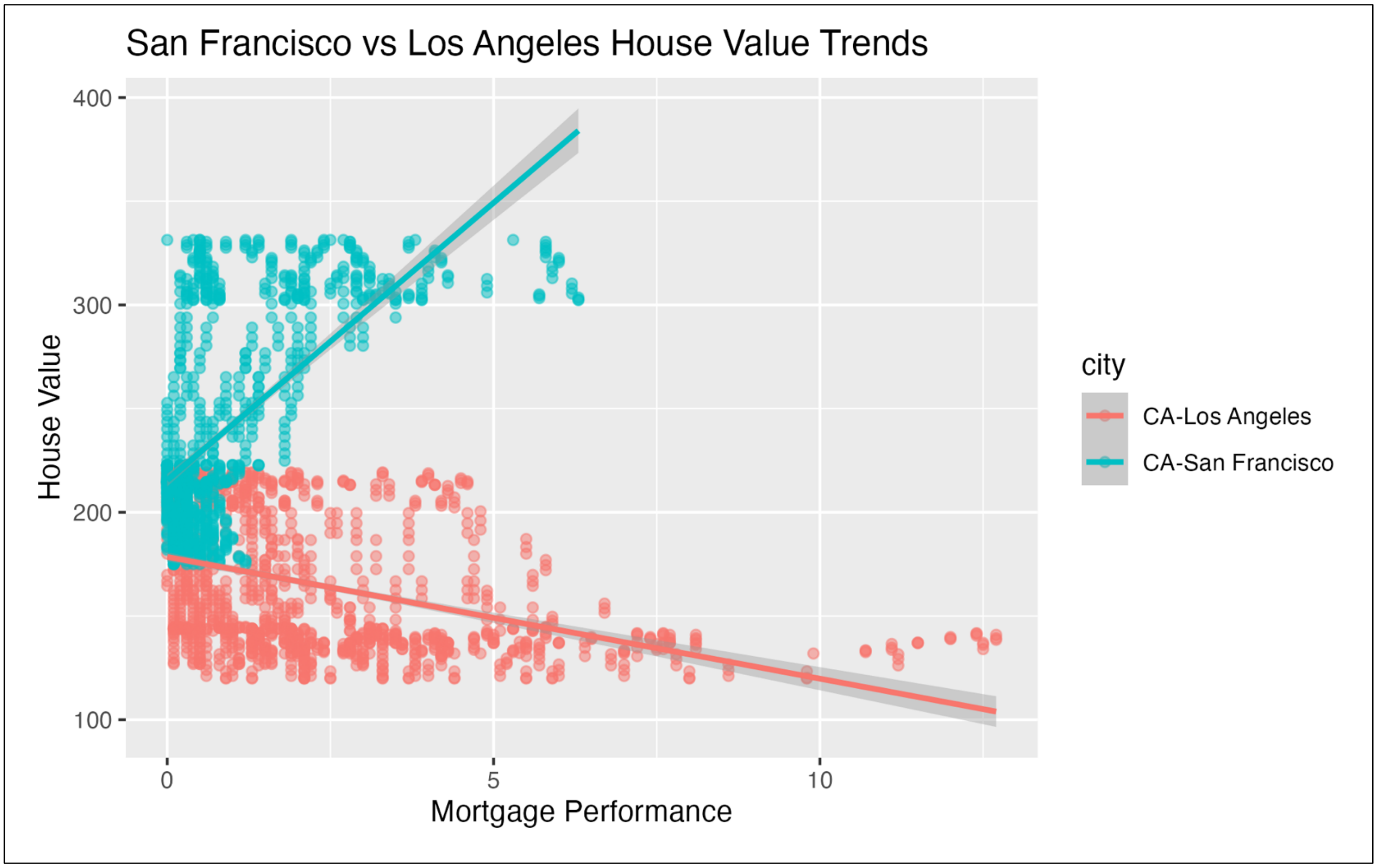


Figure 3 : Mortgage Value by House Value subgroupped by Geographical Location

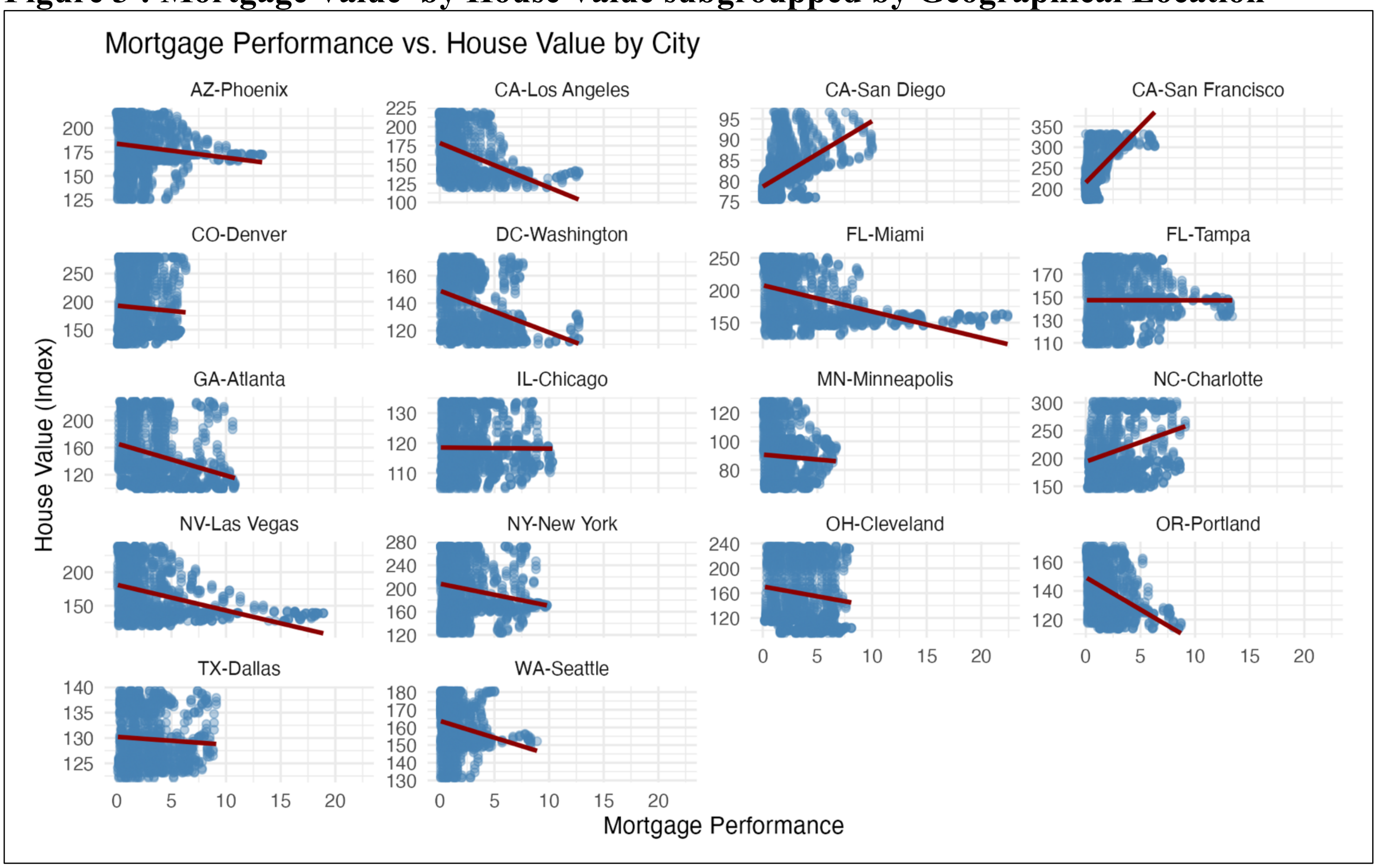


Table 5: Estimated Trends in Mortgage Value by City

City	Trend Estimate	SE	df	Lower CL	Upper CL
AZ-Phoenix	-1.4581	0.335	24471	-2.114	-0.802
CA-Los Angeles	-5.8718	0.376	24471	-6.608	-5.135
CA-San Diego	1.5758	0.459	24471	0.677	2.475
CA-San Francisco	26.7542	0.817	24471	25.153	28.355
CO-Denver	-1.8996	0.573	24471	-3.023	-0.776
DC-Washington	-3.0374	0.356	24471	-3.735	-2.34
FL-Miami	-4.0559	0.192	24471	-4.432	-3.68
FL-Tampa	-0.0129	0.31	24471	-0.621	0.595
GA-Atlanta	-4.7086	0.334	24471	-5.364	-4.054
IL-Chicago	-0.0377	0.353	24471	-0.729	0.653
MN-Minneapolis	-0.6823	0.585	24471	-1.829	0.465
NC-Charlotte	6.9644	0.412	24471	6.157	7.772
NV-Las Vegas	-3.8346	0.237	24471	-4.299	-3.371
NY-New York	-3.7989	0.37	24471	-4.524	-3.074
OH-Cleveland	-3.1681	0.441	24471	-4.032	-2.304
OR-Portland	-4.4882	0.556	24471	-5.577	-3.399
TX-Dallas	-0.1548	0.427	24471	-0.992	0.683
WA-Seattle	-1.9079	0.612	24471	-3.107	-0.709

3.4 Inferential results for RQ2

After rejecting the null hypothesis in the Kruskal-Wallis test, I use the emmeans package to estimate trends and confidence intervals for mortgage values across cities. Analysis of the Trend Estimate column revealed significant volatility in how these values behaved across cities. There were two versions of Trend Estimate: positive trends (+) or negative trends (-). The three cities with the worst negative trend scores are Los Angeles (-5.8718), Atlanta (-4.7086), and Portland (-4.4882). On the opposite end of the spectrum, I see that cities like San Francisco (26.7542), Charlotte (6.9644), and San Diego (1.5758) exhibit positive trends that exceed 1.

Furthermore, when I look into both the Lower and Upper Confidence Intervals, I have identified that cities like Tampa, Chicago, Minneapolis, and Dallas have intervals that cross 0 (e.g, Dallas ranged from -0.992 to 0.683). The trends in these specific cities were not statistically significant, which contrasts with those in the other cities, which were more robust and unique (e.g., San Francisco ranged from 25.153 to 28.355).

I examined the multi-city scatter plots and found that most cities show a negative trend (Figure 3). This includes cities such as Los Angeles, Phoenix, and New York, which have some of the highest median house values (Table 4: median_house). Though there are cities that are outliers and show positive trends, the three are San Francisco, San Diego, and Charlotte. Finally, some cities, such as Chicago, show no significant relationship between the two variables.

I then isolated the California market, specifically Los Angeles and San Francisco (Figure 2). I can see there is a dramatic contrast in their trajectories: San Francisco showed a steep positive slope, with house values skyrocketing within a narrow range of mortgage values. This is the opposite of what Los Angeles displays: a slow, negative trend. Los Angeles's data points are more spread horizontally and exhibit greater performance variance, indicating greater stagnation and declining values. Continuing this, I performed a Linear Regression on both cities, finding that their b coefficients are significantly different: San Francisco's is 26.75, and Los Angeles's is -5.872. Further disproving the claim that house values and mortgage values differ across cities and rejecting the null hypothesis. It also showed me how much of the model is explained by using house value. For San Fransico, the model shows an R^2 of roughly 0.41 (Appendix: Figure 4.2). Los Angeles's model explained only 17% of it, giving it an R^2 of roughly 17 (Appendix: Figure 4.1). This suggests that houses values play more of a factor (whether values are high or low) when determing mortgage value.

The analysis showed that mortgage value is not uniform when predicting values across all geographic locations. This means there is a negative trend: the most common approach excludes outliers like San Francisco, which has the most economic factors. I determined that even geographically close cities could operate as entirely different economic locations. This is shown in San Francisco, where it decoupled from the trends I saw.

Discussion :

Discussion 1:

RQ1:

I examined whether there is a relationship between mortgage rates based on delinquency and default, and the House Value, which affects the time to buy, or whether it is a city fixed effect. Based on our analysis and results, I can reject the null hypothesis, meaning there is statistical relationship between mortgage rates and house value that affects the time to buy. This implies that a "city fixed effect" is driving market behavior. This also there is no single statistical signl that tells buyer when to purchase their home. I can further show this with Los Angeles, where lower property values are associated with higher delinquency metrics for those properties. When I compare this to San Diego, which is still in the same state as Los Angeles, it shows a positive trend, indicating divergence in specific markets that causes fluctuations in mortgage value metrics.

RQ2:

I determined that the cities' house values and mortgage value trends over time differ significantly, so I rejected the null hypothesis and concluded a statistically significant difference in house values and mortgage value across the cities. Of the 18 geographical locations, the majority followed a negative trend, in which their city's mortgage value metrics correlated with lower house values. These cities include Portland, Washington, D.C., and others. Three cities show positive trends: San Francisco, San Diego, and Charlotte. When analyzing, I concluded that Los Angeles' data spread is more horizontal and more scattered, unlike San Francisco, which is more linear and contains more clusters (Figure 2). The relationship between mortgage value and house value varies by local economic environment. This means there is no "one size fits all" model for this.

Discussion 2:

The Bush administration sought to expand homeownership, aiming to create a "ownership society" (Bush, 2004; The White House, 2002). The goal was to "increase the number of minority homeowners by 5.5 million families" (Bush, 2004; The White House, 2002), with a deadline by the end of the decade. This was followed by President Bush asking the private sector for help, which was met with dozens of companies and organizations supporting the movement. In total, more than $1.1 trillion in mortgage purchases for minority homebuyers happened throughout the rest of the decade (Bush, 2004; The White House, 2002). Though in 2007, there was the mortgage market crash, which caused a decrease in homeownership and more. This led the Obama administration to create more stable, preventive programs that, for the time being, pivoted away from the Bush policies. (U.S. Department of the Treasury, 2016).

The Federal Housing Administration (FHA) was the primary agency responsible for stabilizing the housing market after the housing bubble. They provided liquidity when private capital (banks) withdrew (Griffith, 2012). However, this did not stop the local government from enacting policies that helped their citizens during this crisis. I can see this in the analysis of the comparison between San Francisco and Los Angeles, with their unique correlations. Proving that local policies have more impactful decisions than the federal government when it comes to dictating house values and mortgage value over time.

Lastly, my rejection of a "time to buy" rule is further proved after the analysis of Los Angeles and San Francisco. There is no central U.S. Housing market that signals or alerts everyone that it is time to buy. Local factors, including policies, zoning laws, and the cost of living, determine house values. Also, the proximity of cities is shown not to affect house values or mortgage value . I can see this when comparing Los Angeles and San Francisco. San Francisco had a positive relationship, and most points stayed within the best-fit line, showing that not only are their house values higher, but their mortgage value is also better than that of Los Angeles (Figure 2).

Discussion 3:

Some limitations that hindered this study's performance include the fact that the overall analysis was based solely on 18 geographical locations. This makes the analysis inapplicable to rural or smaller markets with local economic density. Secondly, cities are not separated into distinct neighborhoods within each city, making it seem that cities with negative trends, like Washington, D.C., have lower home values, when they might be correlated. Finally, I understand that my linear regression test did not account for external factors, such as unknown variables, which could be the proper drivers of the trend visualized.

Discussion 4:

For both research questions, I rejected the null hypothesis. This was done through assumption testing, and the results were presented in figures and tables. This indicates a relationship between mortgage value and house value, suggesting a city-fixed effect and that there is not one U.S Housing Market. This is similar to research question two, showing that there is a significant difference in house values and mortgage value across cities. This was backed by our linear regression of Los Angeles (negative slope and stagnant) and San Francisco (positive slope and outliers), which shows that, even though the two cities are close in proximity and reside in the same state, they differ in slope and stability. The local economy has a greater influence on house values and mortgage value. This creates a strong argument that federal factors influence both house values and mortgage value, but the most impactful ones are local legislation. This gives us something to look for, something that Obama's Administration passed to allow Americans, regardless of minority or non-minority status, to purchase a house without worrying about the risk of high default rates due to a shift in mortgage preformance.

References:

1. Barnett-Hart, A. K. (2009). *M-RCBG associate working paper series | no. 4*. https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/Barnett-Hart_2009.pdf
2. Brookings Institution. (2017, April 5). *The Federal Housing Administration can do more with more*. https://www.brookings.edu/articles/the-federal-housing-administration-can-do-more-with-more/
3. Bush, G. W. (2004, August). *Record of Achievement – Expanding Home Ownership (Chapter 7)*. The White House. https://georgewbush-whitehouse.archives.gov/infocus/achievement/chap7.html
4. Bush, G. W. (2004, September 2). *Increasing affordable housing and expanding homeownership* [Press release]. The White House. https://georgewbush-whitehouse.archives.gov/news/releases/2004/09/20040902-5.html
5. Ceciliano, L. (2018, September 19). *What Really Caused the Great Recession? – Institute for Research on Labor and Employment*. Institute for Research on Labor and Employment. https://irle.berkeley.edu/publications/irle-policy-brief/what-really-caused-the-great-recession/
6. Griffith, J. (2012, October 11). *The Federal Housing Administration saved the housing market*. Center for American Progress. https://www.americanprogress.org/article/the-federal-housing-administration-saved-the-housing-market/
7. Tardi, C. (2025). *Understanding Collateralized Debt Obligations (CDOs) and Their Impact*. Investopedia. https://www.investopedia.com/terms/c/cdo.asp
8. *The Outlook for US Housing Supply and Affordability*. (2025, October 21). Goldmansachs.com. https://www.goldmansachs.com/insights/articles/the-outlook-for-us-housing-supply-and-affordability
9. The White House, Office of the Press Secretary. (2002, June 17). *Fact sheet: President Bush calls for expanding opportunities to homeownership*. The White House, George W. Bush Archives. https://georgewbush-whitehouse.archives.gov/news/releases/2002/06/20020617.html
10. U.S. Department of Housing and Urban Development. (2003). *Budget summary for Massachusetts, Fiscal Year 2003*. https://archives.hud.gov/budget/fy03/fy03massachusetts.pdf
11. U.S. Department of Housing and Urban Development. (2003). [Press release 03-140]. U.S. Department of Housing and Urban Development Archives. https://archives.hud.gov/news/2003/pr03-140.cfm
12. U.S. Department of Housing and Urban Development. (2015, September). *The Obama Administration's efforts to stabilize the housing market and help American homeowners (National Scorecard)*. https://home.treasury.gov/system/files/initiatives/financial-stability/reports/Documents/huddoc.pdf#
13. U.S. Department of Housing and Urban Development. (2017, January 5). *Housing as a platform for opportunity: A memo to the American people*. Cabinet exit memo. https://obamawhitehouse.archives.gov/administration/cabinet/exit-memos/department-housing-urban-development
14. U.S. Department of Housing and Urban Development, & U.S. Department of the Treasury. (2016, December). *The Obama administration's efforts to stabilize the housing market and help American homeowners (National Scorecard, December 2016)*. HUD User Archives. https://archives.huduser.gov/scorecard/pdf/Dec2016-scorecard-2016.pdf

Appendix:

Figure 3.1: Residuals vs Fitted for San Francisco

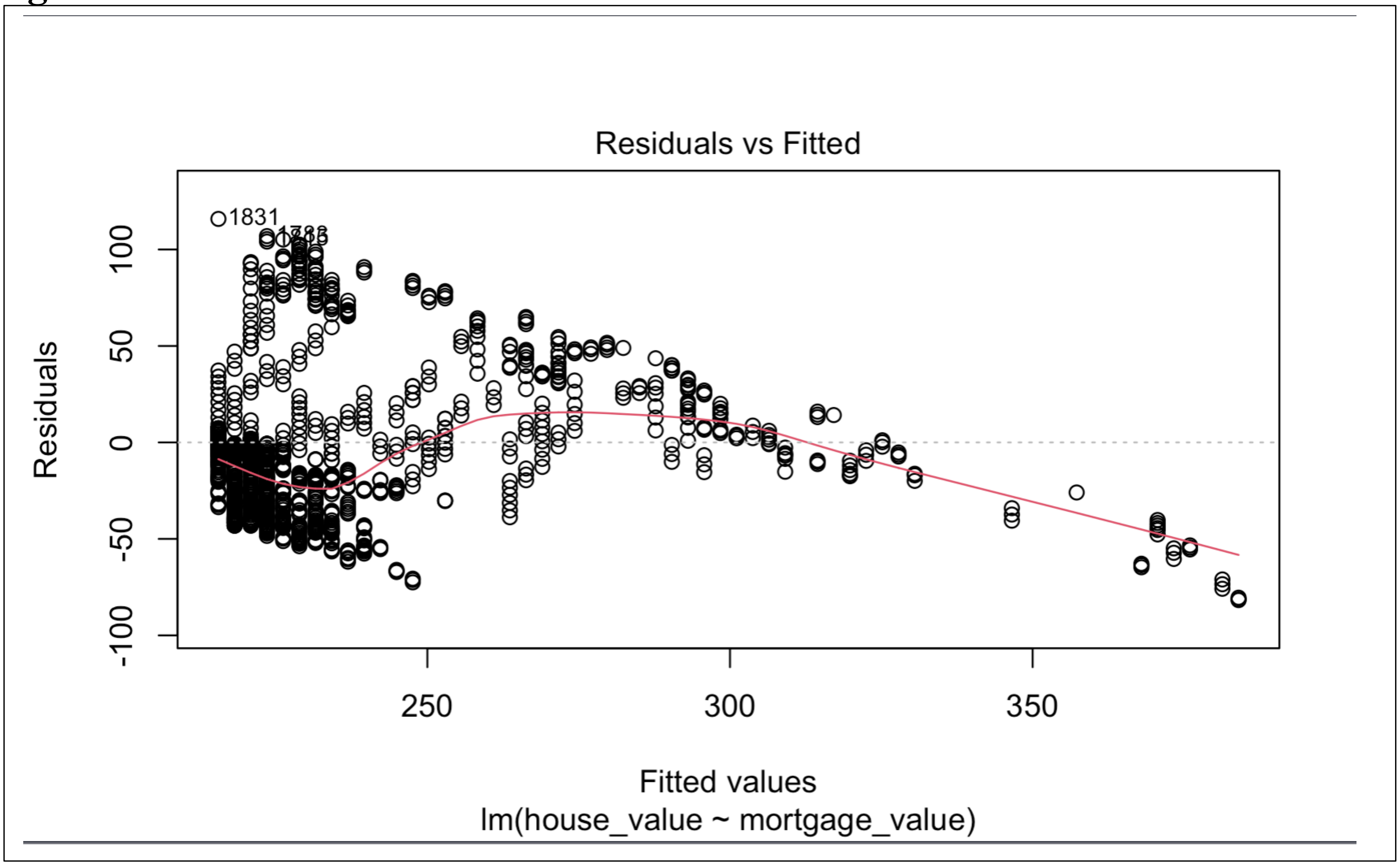


Figure 3.2: Residuals vs Leverage for San Francisco

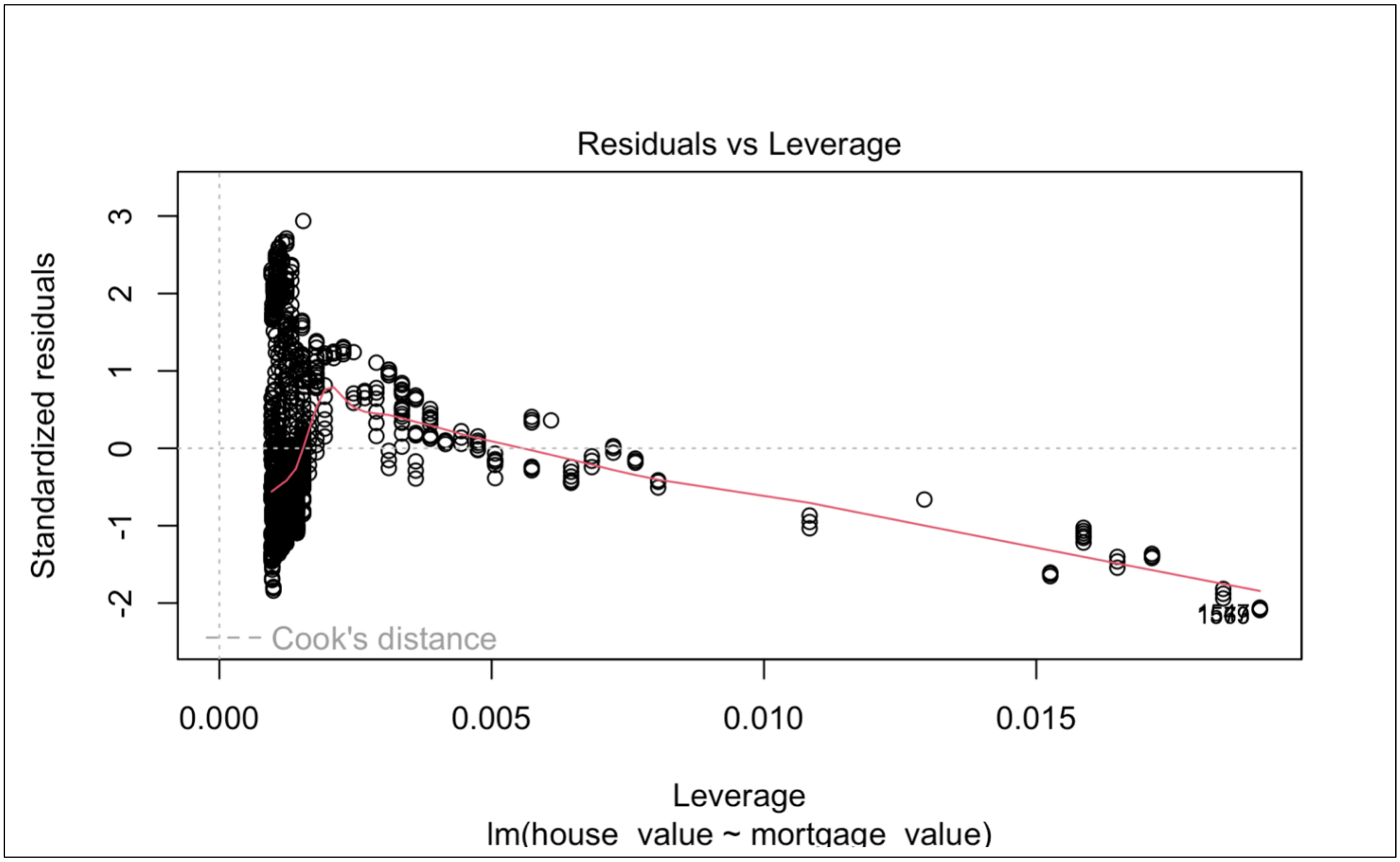


Figure 3.3: Q-Q Residuals for San Francisco

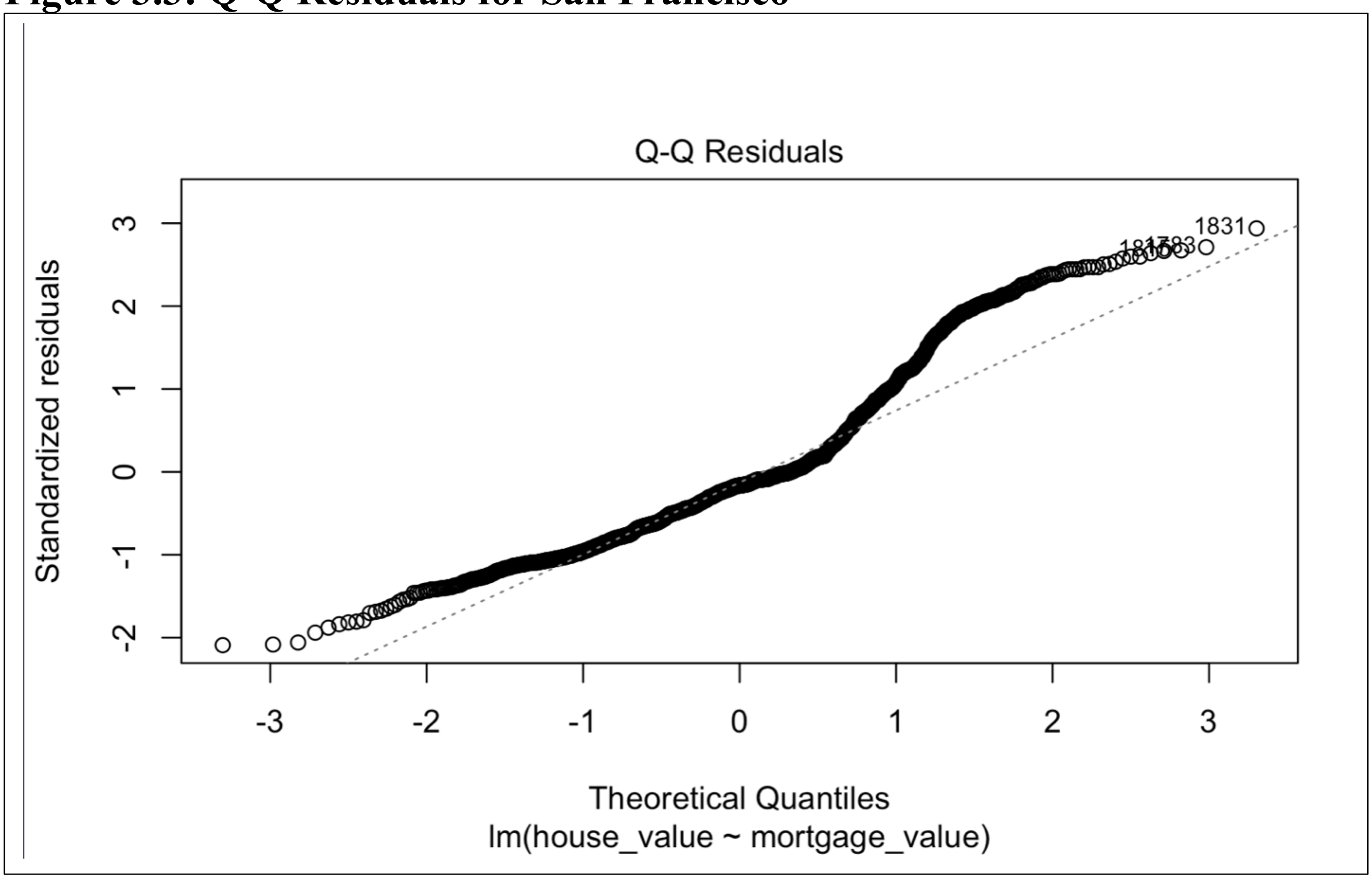


Figure 3.4: Scale-Location for San Francisco

Figure 3.5: Residuals vs Fitted for San Francisco

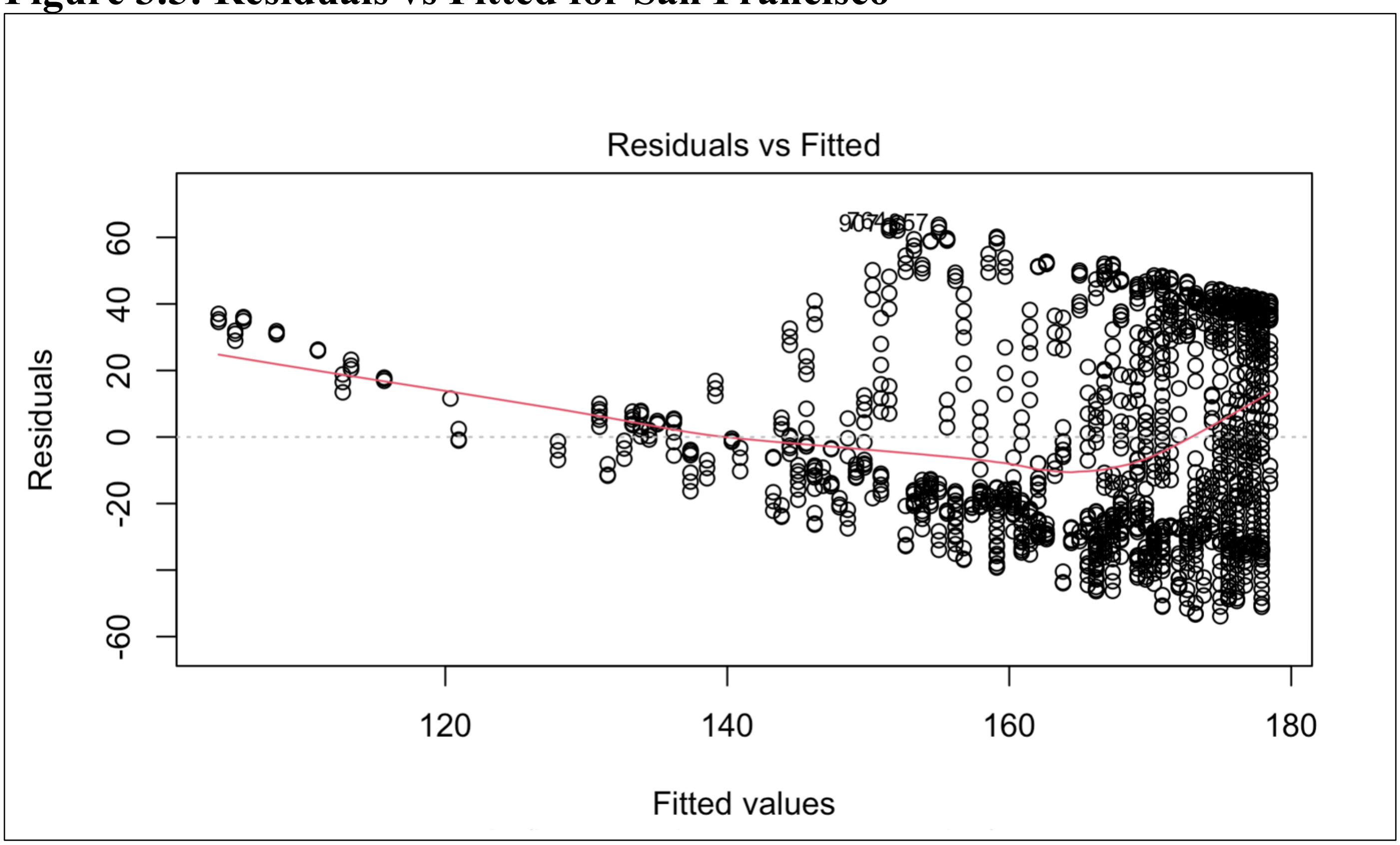


Figure 3.6: Q-Q Residuals for Los Angeles

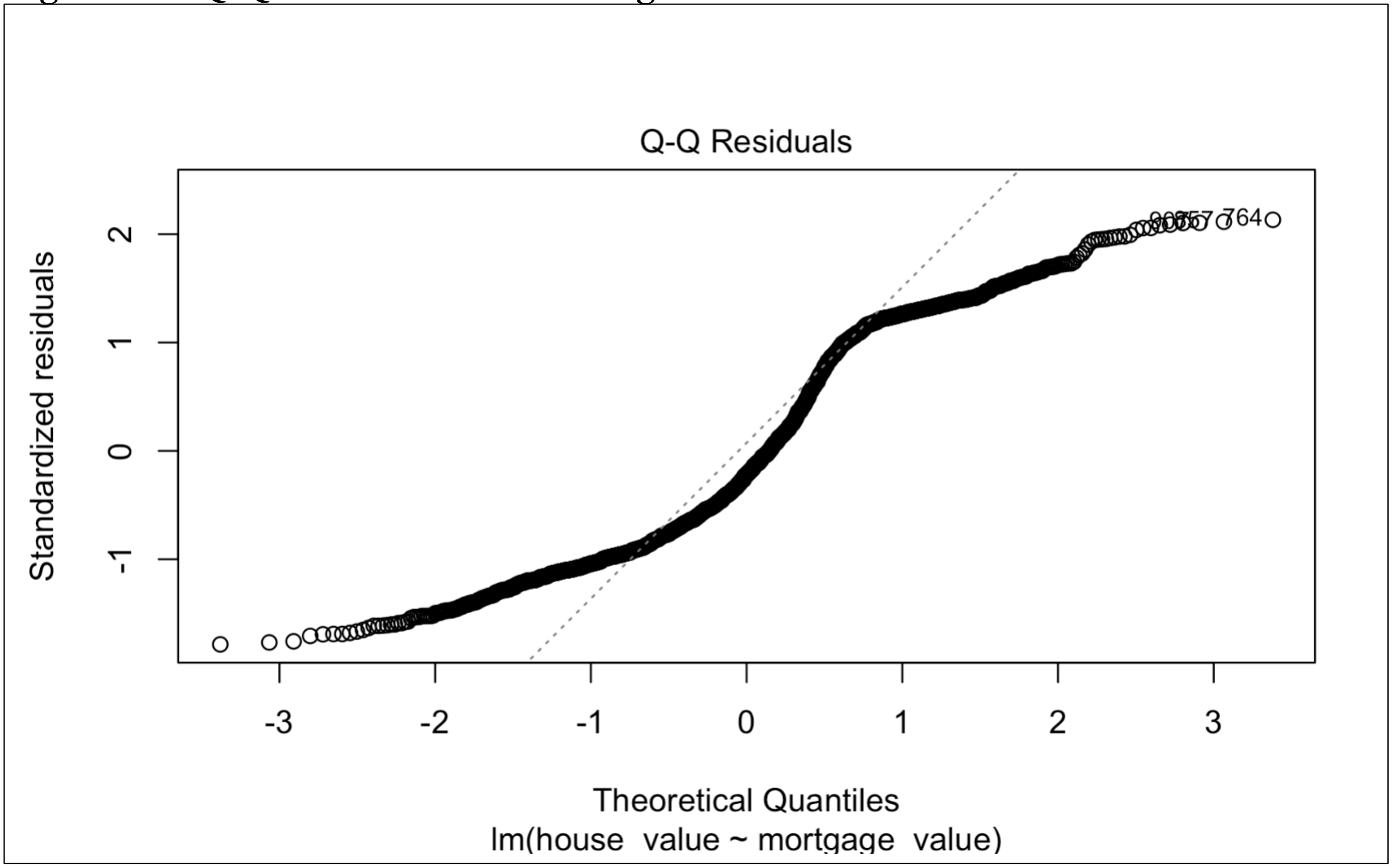


Figure 3.7: Scale-Location for Los Angeles

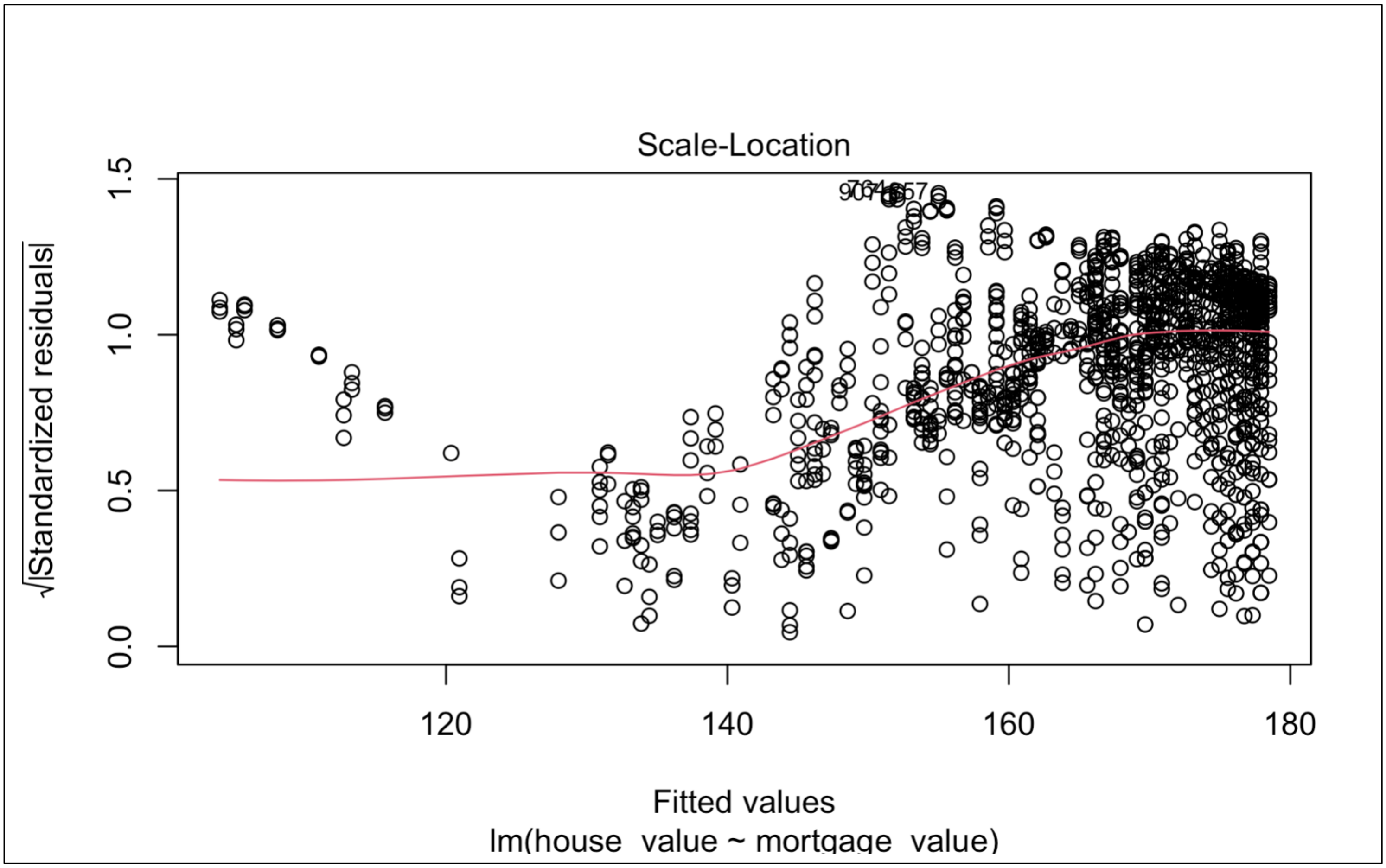
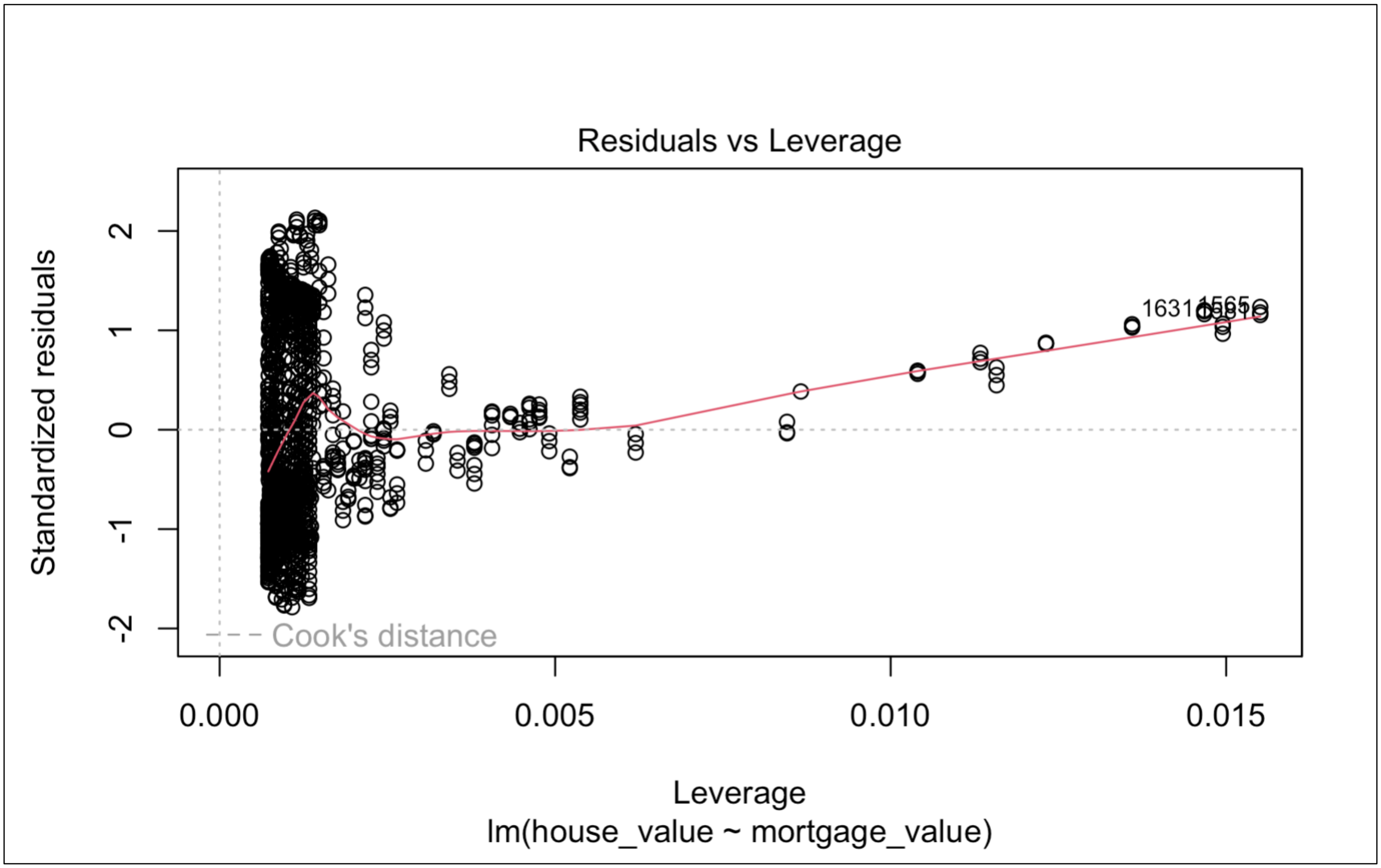


Figure 3.8: Residuals vs Leverage for Los Angeles



4.1 Linear Regression Summary of Los Angeles

```
Call:
lm(formula = house_value ~ mortgage_value, data = filter(combined,
    city == "CA-Los Angeles"))

Residuals:
    Min      1Q  Median      3Q     Max
-53.882 -27.010  -6.538  31.512  64.344

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    178.4931     1.1263  158.48   <2e-16 ***
mortgage_value  -5.8718     0.3505  -16.75   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 30.19 on 1378 degrees of freedom
  (460 observations deleted due to missingness)
Multiple R-squared:  0.1692,    Adjusted R-squared:  0.1686
F-statistic: 280.7 on 1 and 1378 DF,  p-value: < 2.2e-16


Call:
lm(formula = house_value ~ mortgage_value, data = filter(combined,
    city == "CA-Los Angeles"))

Coefficients:
   (Intercept)  mortgage_value
       178.493          -5.872
```

4.2 Linear Regression Summary of San Francisco

```
Call:
lm(formula = house_value ~ mortgage_value, data = filter(combined,
    city == "CA-San Francisco"))

Residuals:
    Min      1Q  Median      3Q     Max
-81.724 -28.181  -6.377  18.071 115.854

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    215.4733     1.5491  139.09   <2e-16 ***
mortgage_value  26.7542     0.9956   26.87   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 39.47 on 1045 degrees of freedom
  (793 observations deleted due to missingness)
Multiple R-squared:  0.4086,    Adjusted R-squared:  0.4081
F-statistic: 722.1 on 1 and 1045 DF,  p-value: < 2.2e-16


Call:
lm(formula = house_value ~ mortgage_value, data = filter(combined,
    city == "CA-San Francisco"))

Coefficients:
   (Intercept)  mortgage_value
        215.47           26.75
```