The Impact of the Dynamic of Housing Market on Income Inequality in the United States: An Empirical Analysis of the Burst-Boom Housing Cycle From 2010 to 2014

Kaichong (Matt) Zhang

Department of Economics

ECON 381 Introduction to Econometrics

Professor Amy Damon

May 4th, 2020

Abstract

This paper investigates the causality that the dynamic of the housing market causes the increase in income inequality during the burst-boom housing cycle from 2010 to 2014 in US. I used the county level panel data, and employ Pooled Ordinary Least Square, Fixed Effect, and Instrumental Variable specifications to conduct empirical analysis and capture the causality. The empirical evidence suggests that the causality within each county over time is economically and statistically significant, but as I expand the picture to include both cross time and cross county dimensions, the macro-level economic factor indicated by GDP per capita starts to kick in and played a determinant role. In short, building on the empirical result, I conclude that the housing market is explaining a proportion of income inequality caused by the economic development. Future works can be conducted by taking into account of more contemporary political and economic shock within the period of investigation.

Keywords: Housing market, Income inequality, Housing cycle

I. Introduction:

Starting from 2000, US house price doubled in 18 years where the median house price skyrocketed from $160,000 to $340,000 and the average house price bounced from $200,000 to $400,000 according to the Federal Reserve (2020). However, as the value of real estate asset increased and housing investments became more lucrative, economic disparity measured by Gini index escalated almost 3%. Interestingly, the booming housing market before 2008 was coincident with approximately 2% increase in Gini index; the 2008 financial crisis primarily caused by the subprime mortgage that quickly cooled down the housing market was also accompanied with about 2% drop in Gini index; and the subsequent recovery of housing market was followed by almost 3% rise in Gini index. A report published by Conference Board and the Demand Institute confirmed that since 2000, the value of housing the wealthiest 10% accounted for rose from 52% to 90% while the housing wealth held by the bottom 40% just increased from 8% to 13%. From 1998 to 2018, the 20 years of coincidence of the trends of US housing cycle and that of Gini index implied that there are at least some correlations between housing market and income inequality in US (Cohen et al. 2012).

The influence of housing market on people's wealth has been explored by various scholars. Benjamin, Chinloy, and Jud (2004) used the data from 1952 to 2002 that included several housing cycles to show that people had relatively higher marginal propensity to consume on the real estate assets than on the financial assets because house purchasing brought both consumption and investment through the refinancing against housing. In other words, the investment in housing market was an effect method of wealth accumulation, and, for minority groups and low-income households, wealth accumulation through non-housing investments was minor or even negative (Boehm and Schlottmann, 2008a). Accordingly, while the rich had much more opportunities to take advantage of the lucrative housing market where the poor were mostly excluded, the issue of booming housing market has been a crucial factor of social inequality by challenging the traditional welfare system for wealth redistribution in US, namely the tax system and the social security system (Gulbrandsen and Sandlie, 2015).

On the other hand, the influence can be the other way around. Niku Määttänen and Marko Terviö (2013) revealed that the rising income inequality during the US boom-burst housing cycle from 1998 to 2007 negatively affected the average house price in 6 metropolitan area in the way that the low-income households would drag down the overall ability to consume and therefore the house price. But the rising inequality would also drive up the demand and therefore the cost for low quality and inexpensive houses (Quigley, Raphael, and Smolensky, 2001), overwhelmed the low-income population, and eventually drove them out of the city. Therefore, house prices would climb up again (Baranoff, 2016).

In the more specific framework of one-way causality from housing market to income inequality, earlier scholars had acquired some empirical results by researching the effectiveness of housing public policy. Boehm's and Schlottmann's (2008a) analysis on policies to increase homeownership opportunity revealed that the wealth accumulation generally appeared to be positively affected by homeownership and the access to the housing market could affect the income inequality. This reasoning was essential to my analysis but their micro-focused methodology using individual level survey data was likely to suffer from selection and sample bias and their models forgot to take into account of the effect of housing cycle. In short, the lack of comprehensive consideration would eventually threaten the internal validity of the research.

More recent scholars' analyses centered around the 2008 financial and housing crisis. Jung Hyun Choi and Richard K. Green (2017) provided a comprehensive and insightful analysis on the causal relationship between house price and income inequality in both strong and weak housing market. However, their regression models, although controlling for housing cycle and taking into account nominal and real income inequality, missed three important variables: racial segregation, education, and parental heritage were proven to be important theoretically and statistically by several scholars.

First, racial segregation was a phenomenon in US where nonwhites were historically on average less affluent than whites due to the discrimination. Redlining is a prominent example that was described as the refusal to make mortgage loans in certain areas regardless of the creditworthiness of the applicants (Holmes & Horvitz, 1994). The percentage of redlining was tested to be positively correlated with the percentage of minority groups and therefore the income inequality. Sean F. Reardon's and Kendra Bischoff's (2011) empirical research also showed that while the housing market was largely based on the affordability, some nonwhites with similar level of wealth generally had less residential options in the housing market compared to the white households. To some extent, the limitation of accessibility to housing market restricted their wealth accumulation.

Second, education and skill formation, which is widely considered as a determinant factor of future employability and income, is also correlated with the booming housing market. In the boom period, less skilled labor would be less willing to pursue more formal and higher level education, attracting the talents to the positions that could make "easy money," like using financial leverage for real estate speculation (Laeven and Popov, 2016). In the burst period, these unskilled "talents" would be more likely to lose their job and their mortgage debt stemming from the booming period would increase the effective cost of schooling, which made them even less likely to be reeducated and employed.

In addition, the boom-burst cycle of the housing market affected not only the current generation, but also the future one. Personal wealth inherited from parents could substantially enlarge both intragenerational and intergenerational wealth inequality. Housing wealth accumulated by parents could effectively push their offspring into housing market and the rising house price would then amplify the parental financial support (Gulbrandsen and Sandlie, 2015). As house price kept rising, the rich's wealth accumulation through house ownership (renting and arbitrage) provided them substantial income while the low-income household who remained struggling about sheltering not only could not enjoy such generous return, but also suffered from the rising cost of living due to the rising rental (Boehm and Schlottmann, 2002b). That being said, there is a threshold naturally set in the booming housing market that limit the poor to access for considerable economic return. On the other hand, Boehm's and Schlottmann's (2002b) research showed that children from low income household, especially those without house ownership, tended to be less educated, be less virtuous, and have more social problem, which would hinder their way to success. Overall, if parents cannot even afford their sheltering, it is difficult for them to take care their children's education and promote their children's wealth accumulation in the future.

In this paper, I analyzed the causal relationship between the dynamic of housing market and income inequality in the burst-boom cycle of the housing from 2010 to 2014. It mainly serves as an extension of Luc Laeven's and Alexander Popov's pieces (2016) that showed the negative impact of the boom-burst cycle in the early 2000s on education, employability and income inequality. Building on Laeven's and Propov's analysis, I also synthesized Reardon's and Bischoff's idea of racial distribution and Boehm's and Schlottmann's idea of parental heritage into the discussion. I tried to avoid the trap of omitted variables bias and sample bias that occurred in the Boehm's, Schlottmann's, Choi's, and Green's pieces. The remaining sections are organized as follows: Section II will explain the existing theories and build up empirical models for the relationship between housing market and income inequality; Section III will introduce the county-level data from United States Census Bureau and Zillow and discuss some bivariate visualizations; Section IV will employ the Fixed Effect and supplemented by Instrumental Variable specification and discuss the regressions results; and Section V will conclude the paper and provide some policy implications.

II. Theories and Empirical Models

Economic Theory

I will start with the assumption that the real estate property is a normal goods coming with considerable economic return in the way that helps buyers to save tax, exert leverage, and earn passive rental (Frankel, 2019). Comparing S&P 500 total return with Vanguard Real Estate ETF total return, I found that the housing market in many times had better performance. which implied that having access to the housing market investment was effective for wealth accumulation (Dewilde and Lancee, 2012).

Under this assumption, we would have Mattlack's and Vigdor's (2006) simple partial equilibrium describing the demand side of the housing market: as both the market demand and house price increased since the rich started to invest more in the housing market, the poor would either have less access to the housing market or devote a higher percentage of income on the housing consumption that could potentially curtail the expenditure on necessities or other investments. on the supply side of the housing market, the story is similar; as the employment in the housing market, mainly in the construction industry, was typically low income population who rented instead of owned the house, the housing boom, although generating a positive impact on wage by increasing demand and wage for construction labors, would hurt their real income through higher rentals and higher house price (Choi and Green, 2017). This is because the increase in wage could hardly catch up the increase in house price, or sometimes even the increase in rental. In short, the heat of housing market would bring more benefit for the rich and much less for the poor, and therefore would widen the wealth inequality.

Choi and Green (2017), on the other hand, provided a more comprehensive analysis of the impact of housing cycle on income inequality: while low-income population was usually more uncompetitive in the job market and more likely to work in the industries that were more sensitive to the local economy, like construction, food, and accommodation industry, their nominal incomes were more vulnerable under the downturn of housing market. But shifting the gear to the real income, Choi and Green indicated that as the low-income population had to spend a much higher percentage of income on cost of living and as many of them rented rather than owned the house and their rent would fluctuate, their real income suffered more from the volatility of housing market.[1]

Empirical Models

Building upon the theories above, we could derive the dynamic of housing heat driven by increasing housing demand and housing price would cause the income inequality by rising the debt of the low-income population and preventing them accessing to the lucrative housing market. Within this relationship, there were two omitted variables that were time-variant: education level (more educated regions are usually more populated urban area with higher house price and greater income inequality) and parental heritage (the higher the wealth of the middle class, the smaller the income inequality gap). In order to cancel out the effect of time-invariant omitted variables, like the culture or tradition of the area, on the regression, I used panel data and the county level fixed effect (entity-demeaned) model to capture the causality within county over time. I also added the year as dummy variable to further explore the cross time and cross county empirical result. The fixed effect regression was written as

$$Y_{it} = \beta_0 + \beta_1 H_{it} + \beta_2 E_{it} + \beta_3 F_{it} + u_{it} \qquad (1)$$

where i denoted the county and t denoted year of selection; Y_{it} refered to the income inequality; H_{it} refered to the measurements of the dynamic of housing market; E_{it} and F_{it} were the two control variables: education level and parental heritage; β_i indicated the change in income inequality given one unit change of each explanatory variable; and u_{it} was the error term of

[1] The increase in nominal income could not offset the increase in cost of living caused by the dynamic of housing market, so the real income would drop.

the regression. This fixed effect model could capture the relatively unbiased estimated coefficients and causality as we minimized the effect of both time-variant and time-invariant omitted variables that might distort the relationship. But it was noteworthy that the model was not perfect in the way that it can't control for all the time-variant variables, which might lead to biased estimated coefficients.

In order to reexamine my results from fixed effect specification, I employed the instrumental variable model with the racial distribution ((Z_S) that measured the "redlining status[2]" and county level rental return(Z_R) that estimated the both the housing return and the cost of living of low-income population as instrumental variables. They all affected income inequality through housing market. Specifically, I used the Two Stage Least Square (TSLS): while the first stage (equation 2) utilized two instrumental variables (Z_P and Z_R) plus the two controls (E and F) I mentioned above to predict the dynamic of housing market ($\widehat{H}$), the second stage (equation 3) regressed the income equality (Y) on the unbiased predicted values of the variable of interest ($\widehat{H}$) with the two controls (E and F). This Two Stage Least Square (TSLS) could avoid the omitted variable bias and the simultaneous causality with the help of the instrumental variables that met both instrument relevance ($corr(Z_i, H) \neq 0$) and instrument exogeneity ($corr(Z_i, u) = 0$) conditions. Therefore, the second regression could capture the unbiased estimated coefficients and causality between the dynamic of housing market and the income inequality.

$$H = \pi_0 + \pi_1 Z_R + \pi_2 Z_S + \pi_3 E + \pi_4 F + v \qquad (2)$$

$$Y = \beta_0 + \beta_1 \widehat{H} + \beta_2 E + \beta_3 F + u \qquad (3)$$

Still, I expected to have measurement error and response bias where the data might be error as the data entry was mistook and as respondents did not report the true numbers. I worried

[2] I have mentioned in the Introduction section that there is a positive relationship between the percentage of minority groups and the proportion of redlining.

about the strength of instrumental variables and the efficacy of control variables that strongly affected the empirical results. But more importantly, I concerned the effectiveness and preciseness of the measurements due to the limited forms of data. Less matched measurements would directly affect the conclusion I drew, and this problem would be further discussed in the empirical regression section.

III. Data Description

In order to investigate how does the dynamic of housing market affect the income inequality in the burst- boom cycle, I used the county level panel data from 2010 to 2014, which, based on the average sales price of houses sold in US published by Federal Reserve, was the beginning part of the burst-boom period. Ideally, a perfectly unbiased individual panel data with large random sample from random control trials across US can provide a precise causality. However, such data does not exist, and county level data that collected through strict randomization and reflect the general picture of regional income inequality is the best alternative I can get. The data were gathered from United States Census Bureau and Zillow[3], whose credibility and scope of business can provide the reliable and representative data. After the data cleaning and reshaping, the data included 925 observations, which consisted five consecutive years data in 185 counties.

Based on the empirical models, income inequality was measured by Gini index while the dynamic of the housing market, the independent variable of interest, was mainly measured by average monthly volume of sales and annual median house price. The housing market boom stimulated by the rising demand would bring up both sales and price simultaneously, and the market burst would bring them down. Because the absolute values of the two explanatory

[3] Zillow is an online real estate database company with highest market share where over 80% of homes in US have been viewed on.

variables were so large that would make the estimated coefficients very small numerically, I converted their units into thousands for convenience.

The summary statistics was shown in Table 1, which presented the roughly normal distribution of Gini index and right-skewed distribution of median house price and average monthly volume of sales. This implied that the housing market during the 2010-2014 burst-boom period started to emerge from the shadow of the 2008 Great Recession and some counties were already ready to boom.

Delving into the bivariate correlation, according to Figure 1 and 2, the scatterplots between the median price and Gini index and that between the average monthly volume of sales and Gini index presented a positive relationship. More specifically, in Figure 1, the five points on the top right were all New York County in five years, implying that the urban area with higher median house price would normally have high income inequality. However, the correlations described here were too simple to be unbiased due to the omitted factors.

For the two controls, education level was measured by the percentage of population with high school or higher degree and that with bachelor or higher degree and parental heritage was measured by the median family income. According to Figure 3 and 4, the scatterplots revealed a positive relationship between percentage of bachelor or above degree and Gini index and a negative relationship between percentage of high school or above degree and Gini index. This seemed to be counterintuitive. But using the income pyramid could provide some valuable explanations with the assumption better education leads to higher income: on one hand, as more people were getting much higher pay given their bachelor or above degree, the income disparity would be greater; on the other hand, as more people were getting only high school or college degree, the middle class would expand and, to some extent, might even contribute to the decreasing inequality because part of the middle class were rising from the lower class. Figure 5 that showed the negative relationship between Gini index and percentage of population with only high school or college degree[4] confirmed this argument.

Figure 6 revealed a weak negative relationship between Gini index and median family income[5], which was also consistent with the original hypothesis through the lens of income pyramid in the way that as the middle class had higher wages, the gap between their wealth and that of the upper class would shrink. In other words, while the inequality was driven by the elites who own the vast majority of wealth, the rising income of the middle class would reduce the relative proportion of wealth held by the rich and therefore relieve the income inequality.

For the two instrumental variables, racial distribution or the redlining status was measured by the percentage of nonwhite population and the total rent was measured by the county level aggregate gross rent. Figure 7 and Figure 8 that showed a positive relationship between Gini index and percentage of nonwhite and a positive relationship between Gini index and aggregate gross rent were consistent with the previous hypothesis. The higher percentage of minority groups in the region was correlated with more serious redlining status and potentially increasing income inequality. The high rent, on the other hand, would raise the cost of living of the poor and would induce more housing market investment by the rich and therefore drive up the income inequality. This process could be, to certain extent, described as the money transfer from the poor to the rich through the hot and lucrative housing market.

However, it was unlikely to get the full picture of the relationship between the dynamic of housing market and the income inequality merely based on the scatterplot-based bivariate analysis and simple descriptive statistics. Therefore, I pushed further by including multivariate regressions and various statistical models to capture the causality.

IV. Regression Analysis

[4] The result is calculated by using the percentage of population with high school degree or above minus that with bachelor or above degree.

[5] The best fit line is mainly dragged by the outliers.

Pool Ordinary Least Square

I started with the Pooled OLS approach to see if there was a correlation between the housing market and the Gini index. It was noteworthy that because the two measurements for housing market, sales and median price, were highly collinear, I included them separately in the following multivariate models, and the following analysis would be based on 95% confidence level. The two univariate models in Table 2 showed that, using 95% confidence level, the relationship was statistically significant but only economically significant if we measured the dynamic of the housing market by sales instead of median price. Specifically, one-thousand-dollar change in annual house sales was correlated with 0.91275% change in Gini index while the same change in median house price only had trivial effect. However, the model was too simple to provide insightful information due to the strong omitted variable bias where each county had different conditions that would potentially affect the Gini index.

Based on the literatures I mentioned above, education (Laeven and Popov, 2016) and parental heritage (P. Boehm's and Alan M. Schlottmann 2002) were two elements that play an important role in the relationship between housing market and income inequality. Therefore, I included them as control variables into the model. The coefficients and the t-values of two partial models in Table 2 revealed that by controlling these two factors, the relationship between the dynamic of housing market and income inequality got weaker both statistically and economically. In detail, although the volume of house sales was still statistically significant, the median house price became insignificant and both variables of interest now contributed much less to the change in Gini index. In the last two full models in Table 2, I included all the variables as controls to check the robustness and the results had similar coefficients but both measurements of housing market were now statistically significant. Overall, the Pooled OLS approach provided a sense that there should be at least some correlation between the housing market and the income inequality, but biases in this approach restricted further analysis.

Fixed Effect specification

In order to further eliminate the bias from county-level time-invariant omitted variable, like the culture, ideology, or history of the counties, I adopt the entities demeaned fixed effect model. The two entity demeaned models in Table 3 showed that the average monthly volume of sales was both economically and statistically significant while the median price was insignificant in both cases. Although not completely and confidently, we could infer some degree of causality where the hot housing market could cause the income inequality, because the main time-variant omitted variables were included in the controls and all the time-invariant ones were “swept away” by the fixed effect specification. The F-test that jointly tested the significance of each variable with the p- value less than 0.05 also confirmed the effectiveness of the model. The estimated coefficients in the entity demeaned models in Table 3 revealed that one thousand increase in the annual sales from average caused 0.70896% increase in Gini index from average, and such change was substantial because since 2000, US had never experienced a change in Gini index greater than 7% within any 3-year period.

The control variables that measured the education level were significant economically and statistically; specifically, 1 percent increase in percentage of population with high school or above degree and that with bachelor or above degree would cause around 0.1% increase in Gin index. This was consistent with the Laeven’s and Popov’s theory that the education would affect income inequality.

This causality was further examined by adding time fixed effect model. The two simple time fixed effect models in Table 4 with much less significant variables of interests and controls but very significant year dummies indicated that the causality between the heat of housing market and the Gini index was insignificant cross county and cross time. From here, some macro level economic trends should kick in as the current variables could not capture the cross time and cross county variations. I included the county level GDP per capita[6] as an indicator of county level macro economy in the controls. The two full time fixed effect models showed that the GDP per capita was both economically and statistically significant in the way that the one thousand dollars change in GDP per capita was at least correlated with approximately 0.5% change in Gini index, which was also substantial compared to the annual change of US Gini index. Such results were robust as I added and dropped other control variables.

Undoubtedly, the four time fixed effect models provided a different perspective compared to the entity demeaned model. Specifically, the latter showed the causality within each county while the former one argued that the macro level economy took a more determinant role in the relationship between housing market and income inequality by, for example, affecting the investment confidence and emotion and shocking the values of assets held by the rich. A prominent example was the economic crisis in 2008 that was accompanied with the housing crisis where consumers and investors, usually the middle and upper class, experienced the wealth shocks and chose to hold instead of spend money, which, to some extent, would relive or even reverse the increasing income inequality. This explanation was well supported the in Figure 9 through 11 where GDP per capita was strongly correlated with the both income inequality and the dynamic of housing market and Table 5 where GDP per capita as well as median house price remained significant in the entity demeaned model. So, the main takeaway here is that the macro-economy would generate income inequality (Gallo 2002) and the dynamic of housing market could explain part of it.[7]

Instrumental Variable Specification

Finally, I implemented the instrumental variable model to check if the conclusion still held.

---

[6] The county level GDP per capita was gathered from US Census Bureau and was included in the attached summary statistics

[7] The macro-level economic indicator cannot be the instrumental variable because it is correlated with the error term.

The two instrumental variable models in Table 6 inferred a similar conclusion. In the first stage of the least square regression, we can see that rent was both statistically significant and jointly significant with other controls while the percentage of nonwhite was insignificant. To improve the model, I dropped the percentage of nonwhites and the new first stage regression presented the total gross rent as a strong instrumental variable, because it is has substantial coefficients and t-value and the F-test was greater than 10. In the second stage, both average monthly sales and median house price were also economically and statistically significant. Statistically, the average monthly volume of sales caused 0.60785% increase in Gini index while the median house price caused 0.1758% increase in Gini index. The causality could be suggested by excluding the effects of omitted variables on regressors and minimizing the measurement error, selection bias and simultaneous causality via instrumental variable mode. Also, the result here was consistent with the one from the entity-demeaned fixed effect model, which implied that the causality robustly survived through different specifications.

It was noteworthy that the ineffectiveness of the percentage of nonwhite as instrumental variable can be explained by the fact that the redlining behaviors were not well captured by the percentage of nonwhites or had trivial impacts on income inequality through housing market as the racial equality was further accentuated in recent days.

V. Conclusion

The empirical result from the three statistical models suggested that there is a causal relationship between the dynamic of housing market and income inequality within each county in the burst-boom period from 2010 to 2014. However, as I expanded the picture to cover the cross year and cross county variation, the rise and fall of macro-economy became dominant by affecting both housing market and income inequality. More specifically, I found that the dynamic of housing market helped to explain part of the uneven distribution of income from economic development. Overall, the effect of the housing market on income inequality was statistically and economically significant and the conclusion robustly survive through both fixed effect and instrumental variable specifications, especially when I used the average monthly volume of sales as measurement of the heat of housing market.

The empirical evidence can serve as a support for policies that promote more comprehensive home ownership to relieve the rising income inequality during the burst-boom housing period. More specifically, the government can try to intervene the housing market by increasing the amount of affordable housing or low-rent housing. As the cost of living occupies a large proportion of expenditure of the low-income population, increasing housing affordability can effectively increase their opportunities for education and investments.

However, it is noteworthy that my empirical research was not perfect. First, the controls might be not complete enough to minimize bias on the coefficient of the variables of interests measured by average monthly volume of sales and median house price. Second, the lack of strong measurements of parental heritage might lead to the biased estimated coefficients. Therefore, I suggest future research focusing on the relationship between housing market and income inequality to expand the spectrum of control variables and improve the quality of instrumental variables. Future research on the related field of study, on the other hand, can also delve into different periods of the housing cycle and take into account of the political and economic shocks occurred during that periods.

Reference

Allen F., Barth J. R., & Yago G. (2013). Restructuring the U.S. Housing Market. Brookings Institution Press. https://www.jstor.org/stable/10.7864/j.ctt6wpcrp.4.

.

Baranoff O. (2016). Housing Affordability and Income Inequality: The Impact of Demographic Characteristics on Housing Prices in San Francisco. John Hopkins University Economics Department. https://econ.jhu.edu/wp-content/uploads/sites/27/2016/07/Olga_Baranoff_senior_thesis.pdf.

Becker, G. S. & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. Journal of Political Economy, Vol. 87, No. 6 (Dec., 1979), pp. 1153-1189. https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_g uide/apa_sample_paper.html.

Benjamin, J. D., Chinloy, P., & Jud, G. D. (2004). Why Do Households Concentrate Their Wealth in Housing? The Journal of Real Estate Research, Vol. 26, No. 4 (2004), pp. 329-344. https://www.jstor.org/stable/10.2307/24887967.

Boehm, T. P. & Schlottmann, A. (2008b). Wealth Accumulation and Homeownership: Evidence for Low-Income Households. Cityscape, Vol. 10, No. 2, Homeownership Experience of Low-Income and Minority Households (2008), pp. 225-256. https://www.jstor.org/stable/20868659.

Boehm, T. P. & Schlottmann, A. (2002a). Housing and Wealth Accumulation: Intergenerational Impacts. In Retsinas, Nicolas P. Editor & Belsky, Eric S. Editor (Eds.), Low-Income Homeownership: Examining the Unexamined Goal (pp. 407-426). Brookings Institution Press, Harvard University Joint Center for Housing Studies.

Burke J. (2019). Zillow vs. Realtor vs. Redfin. Investopedia. https://www.investopedia.com/articles/personal-finance/060516/zillow-realtor-redfin-which-best-real-estate-website-z-nws.asp.

Case K. E., Quigley J. M., & Shiller R. J. (2001). COMPARING WEALTH EFFECTS:
THE STOCK MARKET VERSUS THE HOUSING MARKET. NATIONAL BUREAU OF ECONOMIC RESEARCH Working Paper No. 8606. http://www.nber.org/papers/w8606.

Cohen J. P., Coughlin C. C., & Lopez D. A. (2012). The Boom and Bust of U.S. Housing Prices from Various Geographic Perspectives. Federal Reserve Bank of St. Louis REVIEW. https://files.stlouisfed.org/files/htdocs/publications/review/12/09/341-368Cohen.pdf.

Choi J. H. & Green R. K. (2017). House Price Shock and Changes in Inequality across Cities. Association for Public Policy Analysis and Management. https://appam.confex.com/appam/2017/webprogram/Paper22961.html.

Dewilde, C and Lancee, B (March 2012). INCOME INEQUALITY AND ACCESS TO HOUSING IN EUROPE. Gini Growing inequalities' impact. https://www.bramlancee.eu/wp-content/uploads/2017/03/DewildeLancee.pdf.

Dwyer, R. E. (2008). Cohort Succession in the US Housing Market: New Houses, the Baby Boom, and Income Stratification. Population Research and Policy Review, Vol. 27, No. 2 (April 2008), pp. 161-181. https://www.jstor.org/stable/41217943.

Federal Reserve Bank of St. Louis Economic Research (2016). GINI Index for the United States [U.S. Census Bureau and U.S. Department of Housing and Urban Development]. Federal Reserve Bank of St. Louis Economic Research. https://fred.stlouisfed.org/series/SIPOVGINIUSA#0.

Federal Reserve Bank of St. Louis Economic Research (2016). Gross domestic product per capita [U.S. Census Bureau and U.S. Department of Housing and Urban Development]. Federal Reserve Bank of St. Louis Economic Research. https://fred.stlouisfed.org/series/A939RC0A052NBEA.

Federal Reserve Bank of St. Louis Economic Research (2019). Median Sales Price of Houses Sold for the United States [U.S. Census Bureau and U.S. Department of Housing and Urban Development]. Federal Reserve Bank of St. Louis Economic Research. https://fred.stlouisfed.org/series/MSPUS.

Frankel M. (2020). Real Estate vs. Stocks: Which Has the Better Historical Returns? Millionacres. https://www.fool.com/millionacres/real-estate-investing/articles/real-estate-vs-stocks-which-has-better-historical-returns/.

Gallo C. (2002). ECONOMIC GROWTH AND INCOME INEQUALITY: THEORETICAL BACKGROUND AND EMPIRICAL EVIDENCE. University College London The Bartlett Development Planning Unit. https://www.ucl.ac.uk/bartlett/development/sites/bartlett/files/migrated-files/WP119_0.pdf.

Gholipour, H. F., Nguyen, J., & Farzanegan, M/ R. (2016, November 21st). Higher property prices linked to income inequality: study. The Conversation. https://theconversation.com/higher-property-prices-linked-to-income-inequality-study-68664

Gulbrandsen, L and Sandlie, H C (2002). Housing Market and Family Relations in a Welfare State. Critical Housing Analysis. Vol 2, Issue 1, (2015), pp.74-81. http://housing-critical.com/home-page-1/housing-market-and-family-relations-in-a-welfar.

Holmes A. and Horvitz P. (1994). Mortgage Redlining: Race, Risk, and Demand. The Journal of Finance, Vol. 49, No. 1 (Mar., 1994), pp. 81-99. https://www.jstor.org/stable/2329136.

Laeven, L. & Popov, A. (2016). A Lost Generation? Education Decisions and Employment Outcomes during the US Housing Boom-Bust Cycle of the 2000s. The American Economic Review, Vol. 106, No. 5. https://www.jstor.org/stable/43861096.

Määttänen N. & Terviö M. (2013). Income Distribution and Housing Prices: An Assignment Model Approach. Journal of Economic Theory Volume 151 (May 2014) Pages 381-410. https://www.sciencedirect.com/science/article/pii/S0022053114000040?via%3Dihub.

Matlack J. L. & Vigdor J. L. (2006). DO RISING TIDES LIFT ALL PRICES? INCOME

INEQUALITY AND HOUSING AFFORDABILITY. NATIONAL BUREAU OF ECONOMIC RESEARCH Working Paper No. 12331. http://www.nber.org/papers/w12331.

Moench, E. & Ng, S. (2011). A hierarchical factor analysis of U.S. housing market dynamics. The Econometrics Journal, Vol. 14, No. 1 (2011), pp. C1-C24. https://www.jstor.org/stable/23127214.

Reardon, S F. and Bischoff, K (2011). Income Inequality and Income Segregation. American Journal of Sociology. Vol. 116, No. 4 (January 2011), pp. 1092-1153. https://www.jstor.org/stable/10.1086/657114?seq=1#metadata_info_tab_contents.

The Links Between Income Inequality, Housing Markets, and Homelessness in California. (2011, September). Public Policy Institute of California. Retrieved April, 2020, https://www.ppic.org/content/pubs/rb/RB_1001JQRB.pdf.

Tilly, C. (2006). The Economic Environment of Housing: Income Inequality and Insecurity. In Bratt, Rachel G., Stone, Michael E., and Hartman, Chester (Eds), A Right to Housing: Foundation for a New Social Agenda (pp. 20-37). Temple University Press.

Paiella M. (2009). THE STOCK MARKET, HOUSING AND CONSUMER SPENDING: A SURVEY OF THE EVIDENCE ON WEALTH EFFECTS. Journal of Economic Surveys Volume23, Issue5 Pages 947-973. https://onlinelibrary.wiley.com/doi/10.1111/j.1467-6419.2009.00595.x.

Upadhyaya K. P., Dhakal D., & Mixon F. G. (2017). Housing prices, stock prices and the US economy. Routledge Taylor and Francis online APPLIED ECONOMICS VOL. 49, NO. 59, 5916–5922. https://doi.org/10.1080/00036846.2017.1361009.

Valentini, P., Ippoliti, L., & Lara F. (2013). MODELING US HOUSING PRICES BY SPATIAL DYNAMIC STRUCTURAL EQUATION MODELS. The Annals of Applied Statistics, Vol. 7, No. 2 (June 2013), pp. 763-798. https://www.jstor.org/stable/23566412.

Wachter S. (2014). THE MARKET STRUCTURE OF SECURITISATION AND THE US HOUSING BUBBLE. National Institute Economic Review, No. 230 (November 2014), pp. R34-R44. https://www.jstor.org/stable/44396494.

Appendix A: Figures and Graphs

Figure 1: Relationship between Gini index and median house price

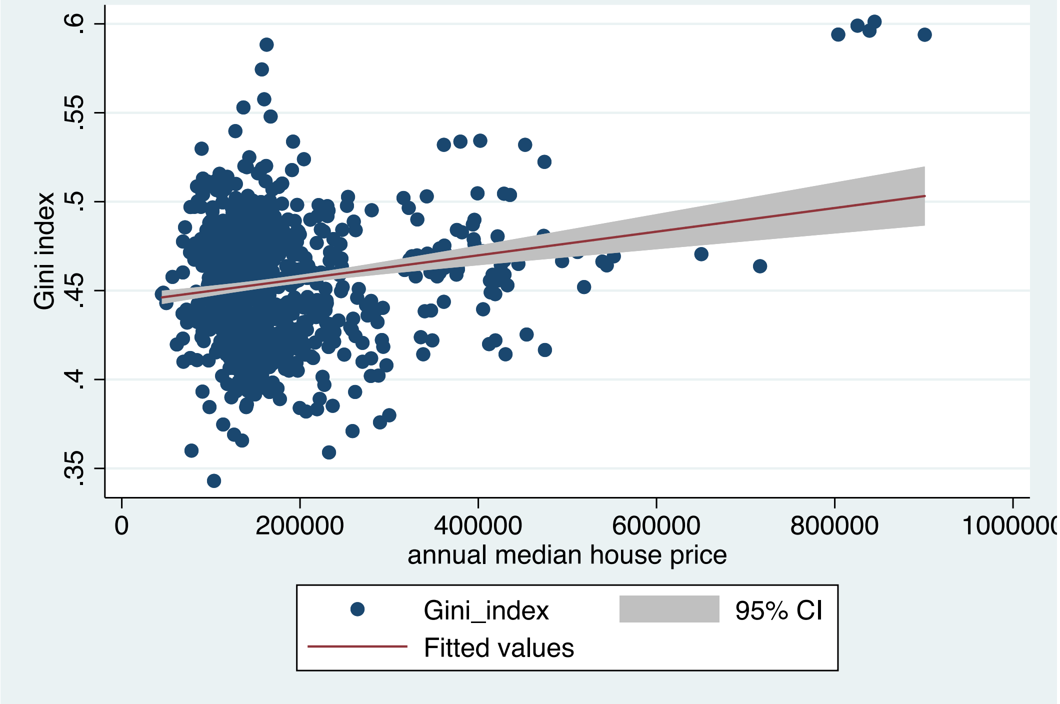
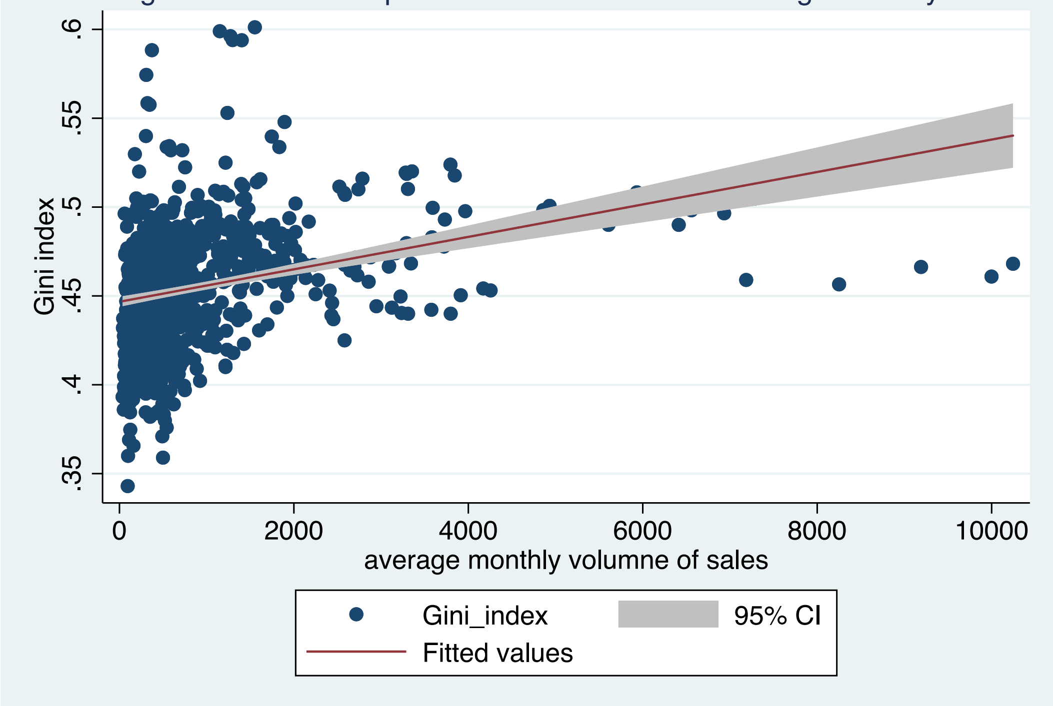


Figure 2: Relationship between Gini index and average monthly sales



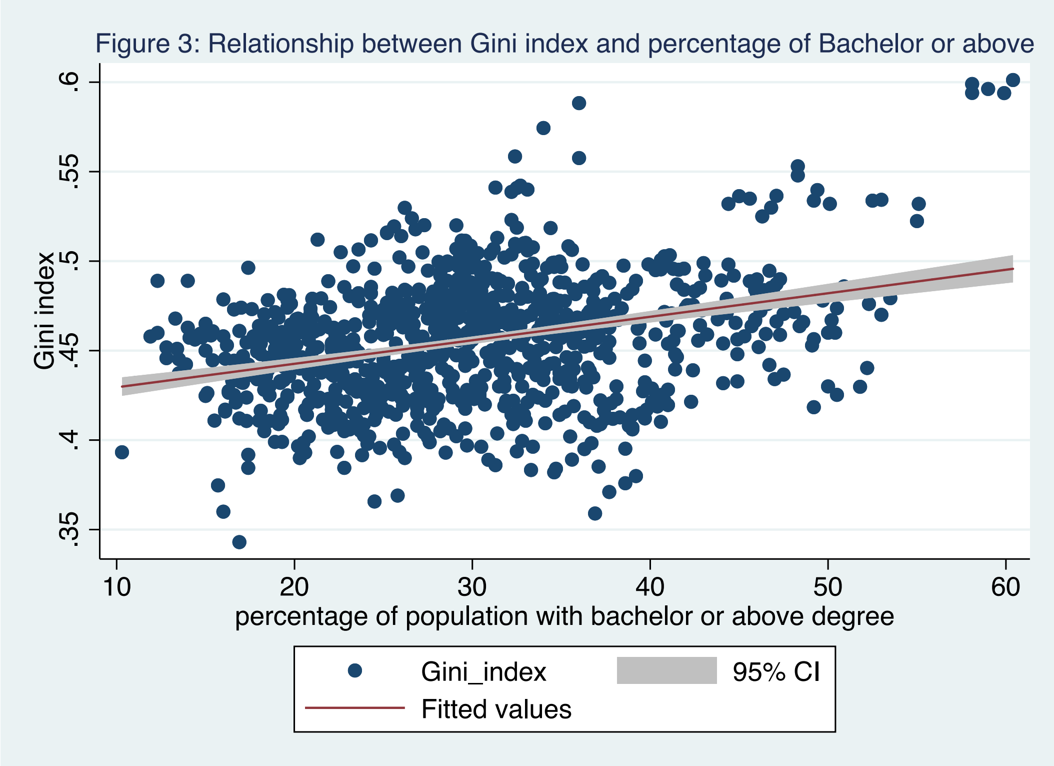


Figure 3: Relationship between Gini index and percentage of Bachelor or above
Gini index
.35
.4
.45
.5
.55
.6
10
20
30
40
50
60
percentage of population with bachelor or above degree
Gini_index
95% CI
Fitted values

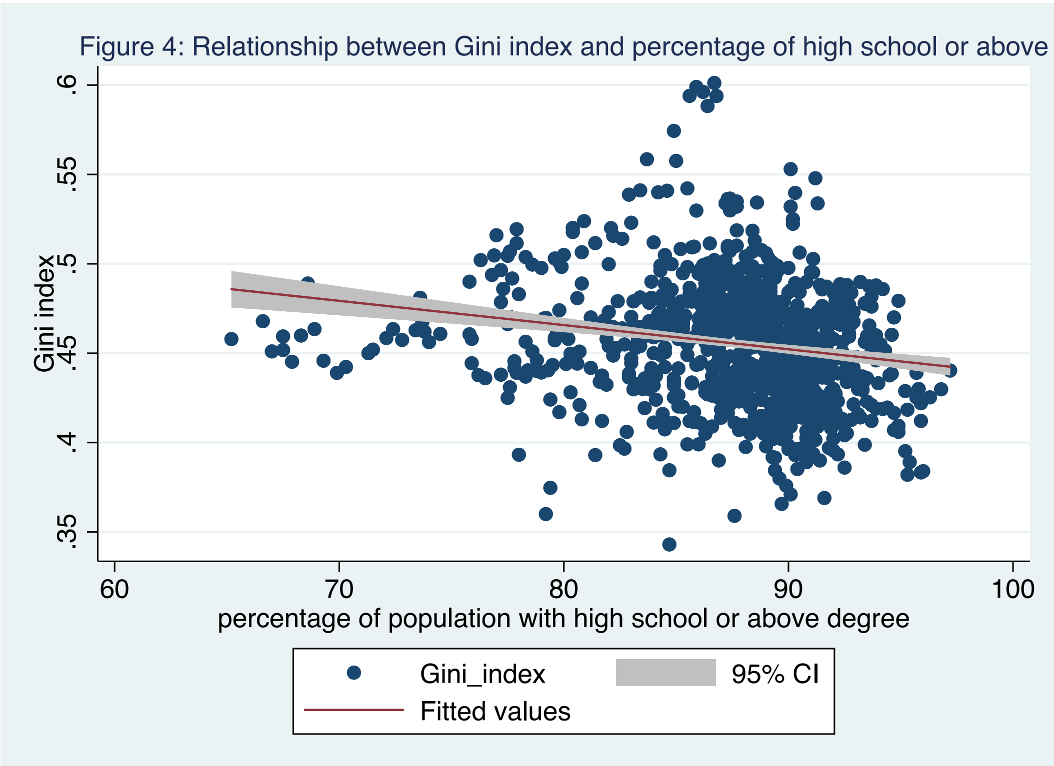


Figure 4: Relationship between Gini index and percentage of high school or above
Gini index
.35
.4
.45
.5
.55
.6
60
70
80
90
100
percentage of population with high school or above degree
Gini_index
95% CI
Fitted values

Figure 5: Relationship between Gini index and percentage of high school and college

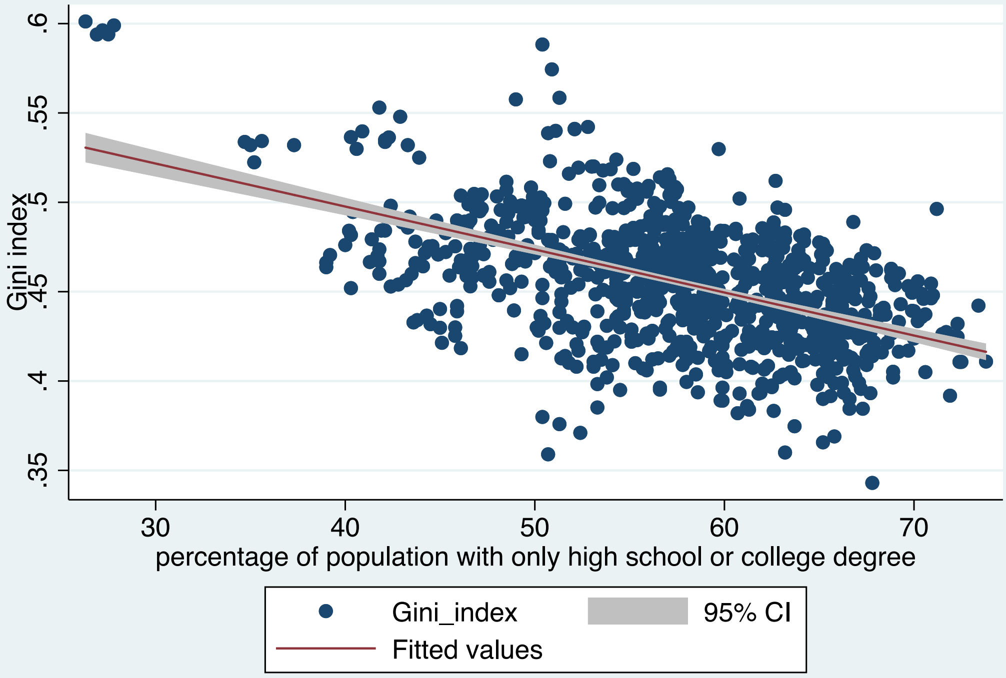


Figure 6: Relationship between Gini index and median family income

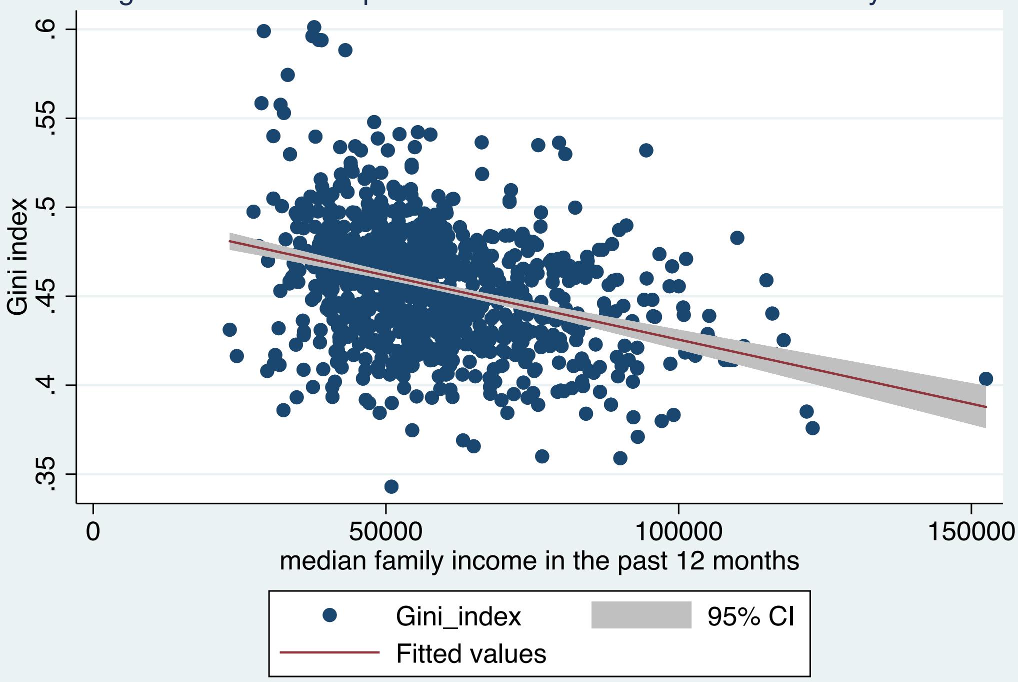


Figure 7: Relationship between Gini index and percentage of nonwhite

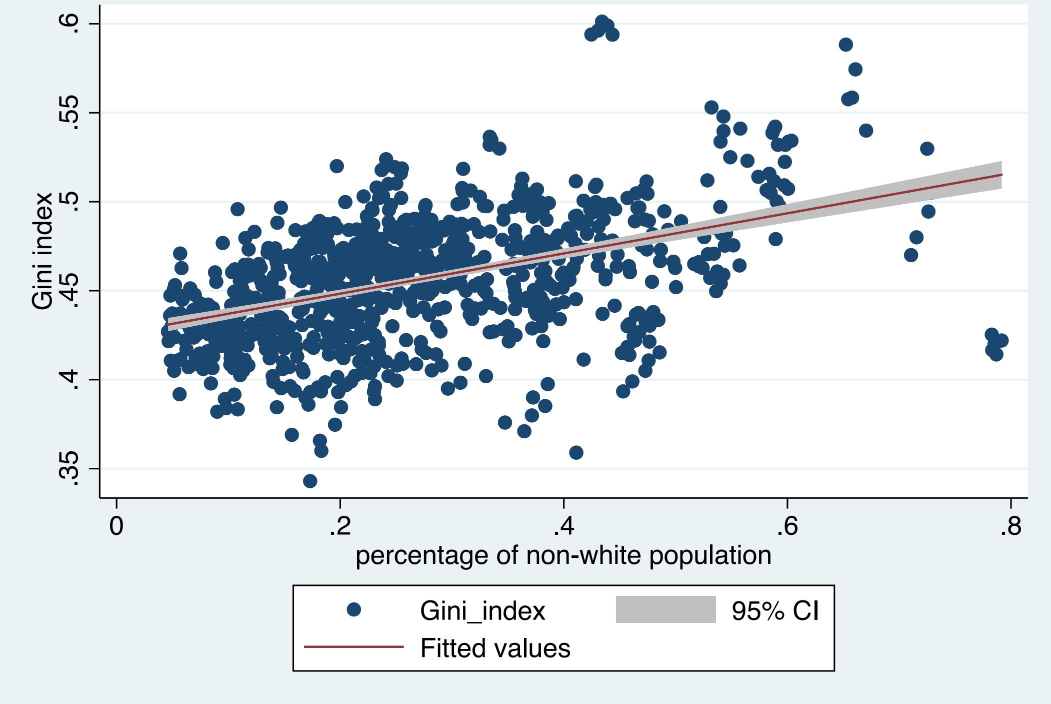


Figure 8: Relationship between Gini index and total rent

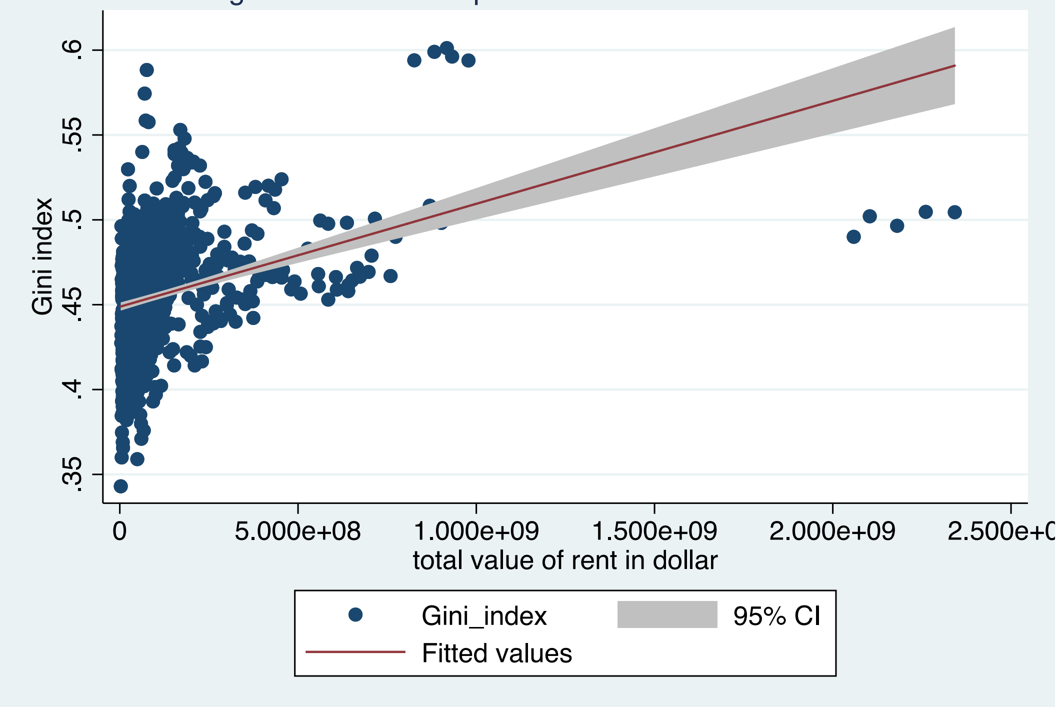


Figure 9: Relationship between Gini index and GDP per capita

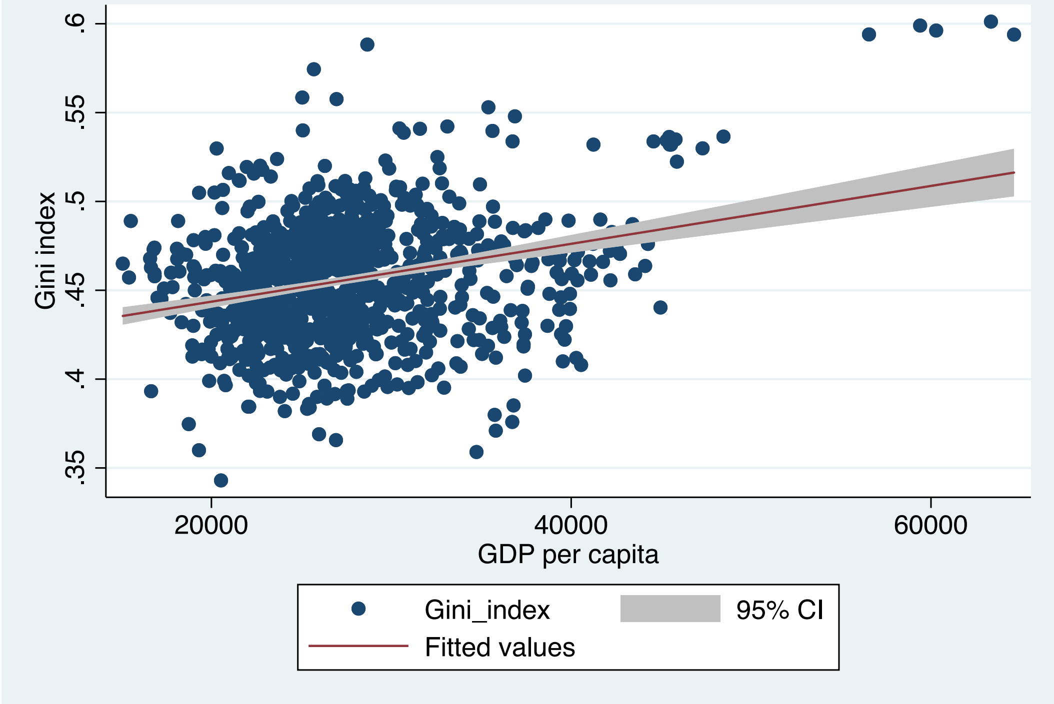


Figure 10: Relationship between GDP per capita and average monthly sales

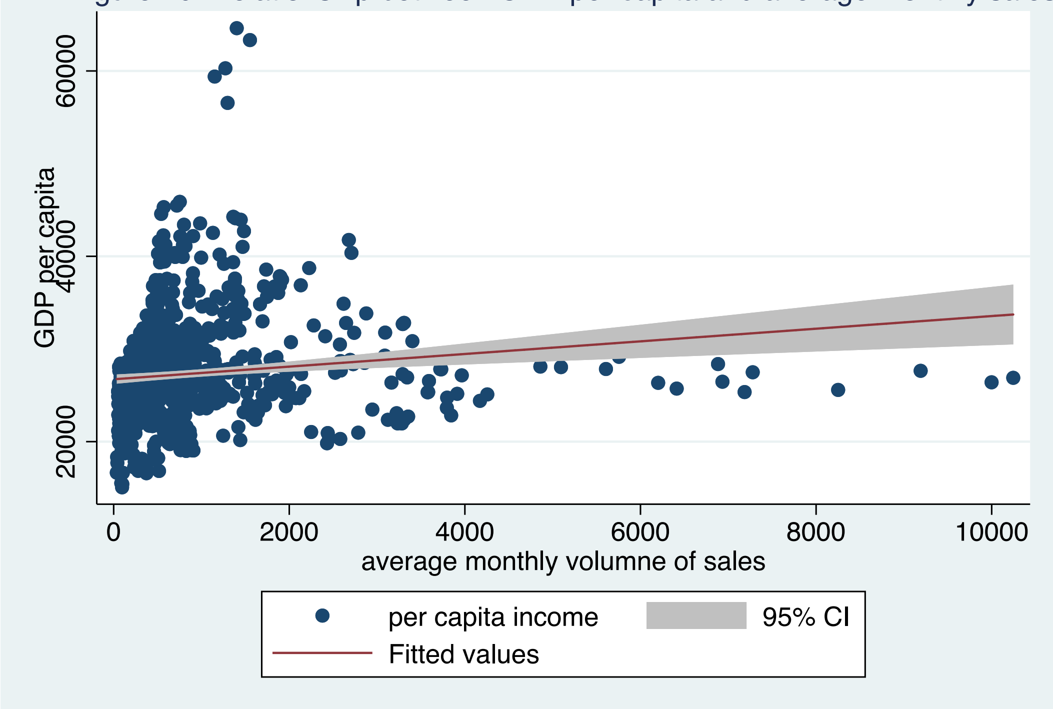
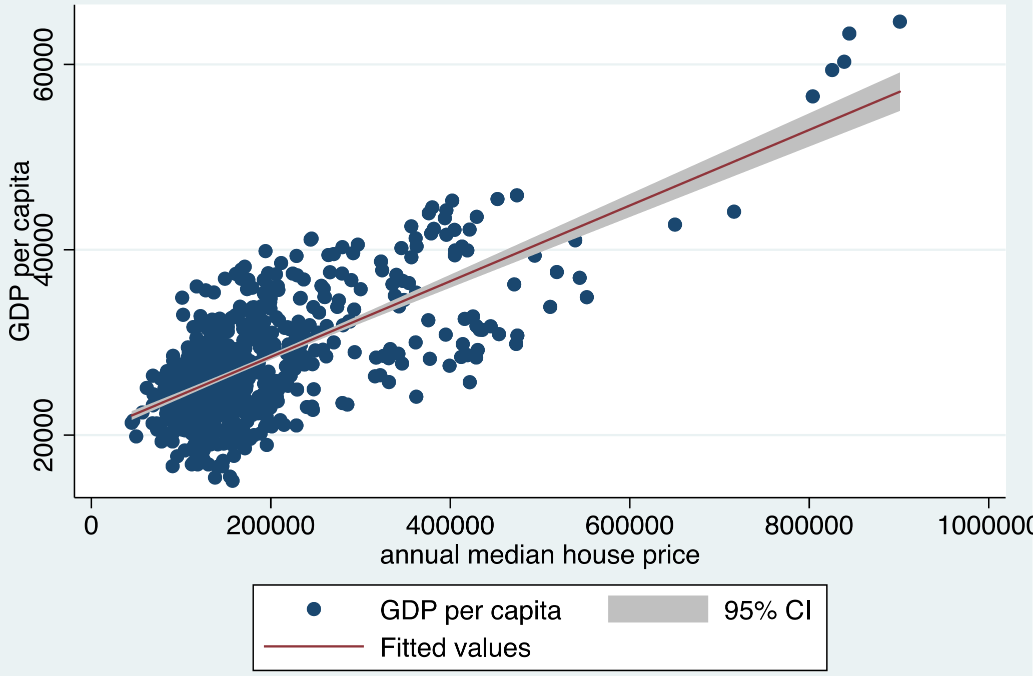


Figure 11: Relationship between GDP per capita and median house price



Appendix B: Tables

Table 1: Summary Statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Gini index	925.00	0.46	0.03	0.34	0.60
average monthly volume of sales	828.00	900.18	1,173.37	36.08	10,245.50
annual median house price	810.00	176,023.77	102,770.13	44,966.67	900,983.31
percentage of population with high school or above degree	925.00	87.39	4.95	65.20	97.20
percentage of population with bachelor or above degree	925.00	29.96	8.87	10.30	60.40
median family income in the past 12 months	925.00	58,250.01	16,976.51	23,324.00	152,503.00
aggregate gross rent in dollar	925.00	114422385.84	208082927.41	2,857,700.00	2.34e+09
GDP per capita	925.00	27,461.56	6,093.25	15,074.00	64,618.00
percentage of non-white population	925.00	0.27	0.14	0.05	0.79

Note: index is roughly normal and both average monthly sales and median house price are right skewed

Table 2: Pooled Ordinary Least Square Models

VARIABLES	(1) MODEL bivariate_1	(2) MODEL bivariate_2	(3) MODEL partial_1	(4) MODEL partial_2	(5) MODEL full_1	(6) MODEL full_2
average monthly volume of sales of in thousands dollar	0.00913***		0.00433***		0.00310***	
	(0.00126)		(0.000785)		(0.000976)	
percentage of population with high school or above degree			-0.00245***	-0.00265***	-0.00185***	-0.00208***
			(0.000204)	(0.000220)	(0.000205)	(0.000221)
percentage of population with bachelor or above degree			0.00251***	0.00255***	0.00146***	0.00144***
			(0.000132)	(0.000145)	(0.000173)	(0.000176)
percentage of non-white population					0.0430***	0.0420***
					(0.00710)	(0.00737)
median family income in thousand dollars			-0.000940***	-0.000959***	-0.000978***	-0.000904***
			(6.38e-05)	(7.37e-05)	(6.29e-05)	(6.97e-05)
aggregate gross rent in million dollars					5.08e-06	2.38e-05***
					(4.74e-06)	(4.52e-06)
GDP per capita in thousand dollars					0.00131***	0.00186***
					(0.000239)	(0.000299)
median house price in thousands dollar		6.66e-05***		1.64e-05		-5.09e-05***
		(1.92e-05)		(1.20e-05)		(1.35e-05)
Constant	0.447***	0.443***	0.644***	0.663***	0.580***	0.591***
	(0.00158)	(0.00337)	(0.0159)	(0.0164)	(0.0162)	(0.0166)
Observations	828	810	828	810	828	810
R-squared	0.095	0.039	0.513	0.490	0.549	0.547

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: two bivariate models show significant correlations; two partial models show much less economically significant estimated coefficients of both variables of interest; two full models suggest that the variables of interest are statistically significant again

Table 3: Fixed Effect models

VARIABLES	(1) MODEL entity_demeaned_1	(2) MODEL entity_demeaned_2
average monthly volume of sales of in thousands dollar	0.00709***	
	(0.00253)	
percentage of population with high school or above degree	0.00103**	0.00104**
	(0.000452)	(0.000458)
percentage of population with bachelor or above degree	0.00106***	0.00120***
	(0.000381)	(0.000381)
median family income in thousand dollars	1.22e-05	4.27e-06
	(4.28e-05)	(4.38e-05)
median house price in thousands dollar		2.65e-05
		(2.13e-05)
Constant	0.326***	0.322***
	(0.0361)	(0.0366)
Observations	828	810
R-squared	0.061	0.050
Number of Countyname	169	169
Country FE	YES	YES
corr	-0.148	-0.244
F_f	18.62	19.14

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: average monthly sales is both statistically and economically significant, implying the causality between the dynamic of housing market and income inequality

Table 4: Fixed Effect models

VARIABLES	(1) MODEL time_simple_1	(2) MODEL time_full_1	(3) MODEL time_simple_2	(4) MODEL time_full_2
average monthly volume of sales of in thousands dollar	0.000885	-0.00221		
	(0.00270)	(0.00251)		
percentage of population with high school or above degree	0.000309	-1.19e-05	0.000355	5.80e-05
	(0.000456)	(0.000423)	(0.000459)	(0.000424)
percentage of population with bachelor or above degree	0.000252	-0.00109***	0.000279	-0.00108***
	(0.000402)	(0.000393)	(0.000405)	(0.000395)
median family income in thousand dollars	-1.04e-05	-7.73e-05**	-1.90e-05	-8.22e-05**
	(4.15e-05)	(3.89e-05)	(4.23e-05)	(3.95e-05)
GDP per capita in thousand dollars		0.00538***		0.00543***
		(0.000510)		(0.000513)
year of the data = 2011	0.00574***	0.00299**	0.00572***	0.00255**
	(0.00123)	(0.00116)	(0.00127)	(0.00121)
year of the data = 2012	0.00458***	8.95e-05	0.00424***	-0.000709
	(0.00134)	(0.00131)	(0.00135)	(0.00133)
year of the data = 2013	0.0104***	0.00233	0.0106***	0.00227
	(0.00147)	(0.00156)	(0.00144)	(0.00155)
year of the data = 2014	0.00803***	-0.00254	0.00842***	-0.00201
	(0.00158)	(0.00177)	(0.00165)	(0.00182)
median house price in thousands dollar			-1.96e-05	-5.44e-05**
			(2.48e-05)	(2.31e-05)
Constant	0.414***	0.348***	0.414***	0.348***
	(0.0388)	(0.0364)	(0.0394)	(0.0369)
Observations	828	828	810	810
R-squared	0.135	0.261	0.132	0.262
Number of Countyname	169	169	169	169
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: GDP per capita is significant in the time fixed effect models, implying that the dynamic of housing market failed to affect the income inequality when we consider both cross time and cross county variations

Table 5: Fixed Effect models with GDP per capita

VARIABLES	(1) MODEL sales	(2) MODEL median_price
average monthly volume of sales of in thousands dollar	-0.00259	
	(0.00244)	
percentage of population with high school or above degree	-0.000106	-6.15e-05
	(0.000420)	(0.000420)
percentage of population with bachelor or above degree	-0.00126***	-0.00125***
	(0.000396)	(0.000394)
median family income in thousand dollars	-8.11e-05**	-8.44e-05**
	(3.96e-05)	(4.00e-05)
GDP per capita in thousand dollars	0.00503***	0.00536***
	(0.000421)	(0.000430)
median house price in thousands dollar		-7.30e-05***
		(2.07e-05)
Constant	0.371***	0.369***
	(0.0330)	(0.0331)
Observations	828	810
R-squared	0.230	0.236
Number of Countyname	169	169
Country FE	YES	YES
corr	-0.393	-0.306
F_f	22.83	23.81

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: GDP per capita is also significant in the entity demeaned model, implying that the macro level factors are also affecting income inequality

Table 6: Instrumental Variable models

VARIABLES	(1) MODEL iv_1_FirstStage	(2) MODEL iv_1_FirstStage_rentonly	(3) MODEL iv_1_SecondStage	(4) MODEL iv_1_FirstStage	(5) MODEL iv_1_FirstStage_rentonly	(6) MODEL iv_2
aggregate gross rent in million dollars	0.00420***	0.00420***		0.131***	0.130***	
	(0.000138)	(0.000138)		(0.0131)	(0.0129)	
percentage of non-white population	-0.0233	-0.0233		-4.995		0.0381***
	(0.216)	(0.216)		(21.02)		(0.00794)
percentage of population with high school or above degree	0.00405	0.00405	-0.00230***	-5.167***	-5.110***	-0.000882**
	(0.00725)	(0.00725)	(0.000216)	(0.688)	(0.644)	(0.000377)
percentage of population with bachelor or above degree	-0.00620	-0.00620	0.00244***	6.040***	6.005***	0.00104***
	(0.00418)	(0.00418)	(0.000122)	(0.399)	(0.370)	(0.000313)
median family income in thousand dollars	-0.000807	-0.000807	-0.000941***	1.989***	1.993***	-0.00126***
	(0.00166)	(0.00166)	(5.29e-05)	(0.158)	(0.157)	(9.80e-05)
average monthly volume of sales of in thousands dollar			0.00608***			
			(0.00105)			
median house price in thousands dollar						0.000176***
						(3.79e-05)
Constant	0.278	0.278	0.632***	313.9***	308.5***	0.534***
	(0.584)	(0.584)	(0.0172)	(55.28)	(50.41)	(0.0272)
Observations	828	828	828	810	810	810
R-squared	0.587	0.587	0.510	0.523	0.523	0.386
Prob>F	0.0000	0.0000		0.0000	0.0000	
F	233.5	233.5		176.1	220.3	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: the instrumental variable models show that that rent is a valid instrumental variable and reconfirm the impact of the dynamic of housing market on income inequality