Income Inequality and Housing Affordability: Evidence from Zip Codes in the United States

Shahrzad Ghourchian*

September 28, 2018

Abstract

The persistent increase in housing prices relative to household income has raised concerns about the affordability of housing in the United States. Using Internal Revenue Service (IRS) annual data and Zillow median housing price data, this paper analyzes the impact of income inequality on the housing price-to-income ratio from 2005 to 2015 for more than 12,700 zip codes. Employing various specifications, I find a positive and statistically significant relationship between the Gini coefficient and the housing affordability index. My results are robust to different methods of estimating the Gini index. Moreover, the empirical results of this study suggest that inequality has a larger impact in zip codes with higher levels of income.

Keywords: Housing market, income inequality, Gini coefficient, price-to-income ratio
JEL Codes: R21, R31, D63, D31, O18

*Corresponding Author: Department of Economics, Florida International University, Miami, FL 33199, USA; E-mail: sghou001@fiu.edu

1 Introduction

The housing market in the United States experienced a dramatic boom and bust cycle that led to a financial and economic crisis. This great recession, which started in the 2000s and ended in 2009, was the most severe economic contraction since 1947, as measured by the peak-to-trough decline in real GDP (Glick et al., 2015). Since 2011, the constant increase in the median sale price of houses in the United States and the experience derived from previous recessions raised concerns regarding the reasons behind, and the consequences of, volatility in the market. Apart from the macroeconomic impact of housing price fluctuations, an increase in housing prices will make housing unaffordable for a large number of middle and low-income households. Some poor families struggle to maintain a basic level of subsistence even if they spend a relatively low proportion of their income on housing (Chen et al., 2010).

On the other hand, income inequality has been rising continuously in the United States as well, inspiring some researchers to investigate whether there is a relationship between inequality and housing affordability. Theoretically, it is easy to argue that income inequality has an effect on housing prices. In a competitive market with a limited housing supply, housing prices will increase if the income of wealthy households increases, since they will increase the amount they are willing to pay on houses simply because they can afford it. The housing then will be less affordable for those households whose income has not changed or has decreased. Figure (1) shows the parallel trend of the housing affordability index (housing price-to-income ratio) and inequality (Gini coefficient) in the United States between 2005 to 2015. Although the positive relationship between housing affordability and income inequality over time only demonstrates correlative, and not causal, relationship between these two, it certainly motivates us to investigate this subject.

In this paper, I study the impact of income inequality on housing affordability in the United States. Using Internal Revenue Service (IRS) annual data, I calculate the Gini coefficient for each zip code from 2005 to 2015. Employing Zillow and IRS zip code level data sets, I compute the price-to-income ratio, the most widely used measure of housing affordability (Hulchanski, 1995). The price-to-income ratio, calculated using median house prices over median income, is an index of access to housing; the ratio increases as housing becomes less affordable.

Affordable housing for all citizens has been one of the main concerns of governments around the world. In the United States this subject was raised regularly, for example, by the Clinton Administration and Millennial Housing Commission in 2002. The public concerns about housing affordability arise from the fact that housing is an investment asset, and for an average family housing is the single most important component of their financial portfolio (Fairchild et al., 2014). The average household devotes roughly one-quarter of its income to housing expenditures, while poor and near-poor households commonly devote half of their incomes to housing (Quigley and Raphael, 2004), resulting in small changes in housing prices having a major impact on some households' well-being.

The recent debate over government tax cut policy focused on argument that increasing the income of wealthy individuals has an indirect "trickle-down" effect on those further down the income distribution, as explained by Matlack and Vigdor (2008); however, income increase at the high end of distribution can raise the prices of goods consumed by the poor. In reality, because of the down-payment requirements and limited affordable housing supply available to low-income households, the demand for low quality or smaller housing will increase as well, leading to an increase in all types of housing in a region. Therefore, wealthy households will gain the most advantages, not only due to income increases but also because of the capital gains associated with owning their own house, causing the inequality (wealth inequality if not income inequality) to increase even more.

This paper is not the first that investigates the impact of income inequality on housing affordability. Rodda (1994) shows a positive relationship between the income inequality and housing prices. Lamont and Stein (1999) show that in cities where homeowners are more leveraged, house prices react more sensitively to city-specific shocks, such as changes in income. The results of Quigley et al.'s (2001) estimation suggests that rather modest improvements in the affordability of rental housing or its availability can substantially reduce the incidence of homelessness in the United States. Vigdor (2002) investigates the hypothesis that an increase in the income of the wealthy causes housing affordability problems for the poor. Ortalo-Magne and Rady (2002) argue that homeownership adds to the volatility of the housing market, amplifies the dispersion of household income within a location, and raises distributional issues. They also confirm that the families who acquire the most housing gain the most from the ability to own their home. Quigley and Raphael (2004) argue that modest changes in institutional arrangements could greatly affect the affordability of homeownership, especially for young households whose incomes will increase over the life cycle of ownership. Matlack and Vigdor (2008) show that in a simple partial equilibrium model, an increase in income at the high end of the distribution can raise prices paid by those at the low end of the income distribution. Using census microdata and data on housing markets in American metropolitan areas between 1970 and 2000, they show that in markets with low-vacancy rates, increases in income at the high end of the distribution are associated with significantly higher rent per room. Gyourko et al. (2013) document large long-term differences in average housing price appreciation across metropolitan areas over the past 50 years. They show that these differences can be explained by an inelastic supply of land in some unique locations combined with an increasing number of high income households nationally. Ray et al. (2015) show that there is an affordability crisis in Los Angeles that is accentuated by income inequality. Chen et al. (2010), Zhang (2015) and Zhang et al. (2016) show that income inequality is one important factor in housing affordability in China.

I am, however, one of the first papers to study the relationship between income inequality and housing affordability in the United States. In previous studies I continuously see the importance of local economic variables. Using zip code level data sets, I am the first to capture local and regional factors affecting the housing market, as well as demographics. For instance, Abraham and Hendershott (1996) document a significant difference in time-series properties between coastal and inland cities, Capozza et al. (2004) argue that the dynamic properties of housing markets are specific to the given time and location being considered, and Hwang and Quigley (2006) argue that housing demand is a function of prices, incomes and demographic variables as well. Their studies confirm the importance of changes in regional economic conditions, income and employment on the local housing market.

In this paper, I study the impact of income inequality on housing affordability among zip codes in the United states using OLS, Fixed Effect (FE) estimations and then system GMM method, to address potential endogeneity. My results confirm that an increase in income inequality leads to an increase in the housing affordability index, meaning less affordable housing for families in the United States. Using three different estimation methods and two ways of calculating the Gini index, I show that my results are robust.

The rest of the paper is organized as follows: in Section 2, I provide a simple model that emphasizes the relationship between income inequality and housing price to income ratio; in Section 3 I discuss the data and methodology used in this research; sections 4 and 5 contain the empirical results of my analysis and robustness check; and, finally, I present my concluding remarks and policy implications in Section 6.

2 Economic Model

Following Zhang et al. (2016) and using a simple partial equilibrium model, I can show how an increase in income inequality will lead to a higher housing price-to-income ratio, i.e. less affordable housing, especially for low-income households. Without loss of generality, here are the assumptions I make:

- We have two types of households in each zip code: high income households (H) and low income households (L).
- The total number of households is standardized as a unit, with the proportion of H-type household denoted by θ.
- $0 < \theta < 1/2$, i.e., high income households are the minorities in each zip code.

- The total income of all households is denoted by Y, with the total income proportion of H-type households as γ. By definition, we have $1/2 < \gamma < 1$.

- The utility functions for H-type and L-type households take the same form, i.e. $U(x, y) = x^{\alpha}y^{1-\alpha}$, where x denotes the size of houses and y denotes all other consumption.

- The unit price of houses is denoted by p , while the unit price of other goods is normalized to a unit.

- The supply function of housing is linear in price, i.e., $S(p) = bp$, where $b > 0$.

After simple calculations, I have the Gini coefficient, G, equals $\gamma-\theta$. Then, solving the utility-maximization problem of the household gives us the housing demands of H-type households and L-type households as:

$$x_H = \frac{\alpha\gamma Y}{p\theta} \tag{1}$$

and

$$x_L = \frac{\alpha(1-\gamma)Y}{p(1-\theta)} \tag{2}$$

and therefore the demand for housing in each zip code is:

$$D(p) = \theta x_H + (1-\theta)x_L = \frac{\alpha Y}{p} \tag{3}$$

In equilibrium, I have the housing price p as:

$$p = \sqrt{\frac{\alpha}{bY}} \tag{4}$$

and aggregate demand for housing as:

$$X = \sqrt{\alpha bY} \tag{5}$$

The housing price to income ratio, R, is then:

$$R = \sqrt{\frac{\alpha}{bY}}\Big(1 + \frac{G}{1-\gamma}\Big) \tag{6}$$

which is median housing price over median household income. Looking at equation (6) we see that an increase in income inequality (Gini coefficient, G) will cause the housing affordability index (R) to increase (Zhang et al 2016). Intuitively, when inequality increases, so does median income, since I assume that some households become wealthier while the income of others stays the same[1], which as argued previously, leads to an increase in prices for all types of housing, causing an increase in the median housing price. However, since the majority of the population consist of middle or low-income households, changes in the median housing price will be much bigger than changes in median income, causing the housing price-to-income ratio to rise.

3 Data and Methodology

In this paper, I use Internal Revenue Service (IRS) data to calculate the frequently used inequality index, the Gini coefficient, from 2005 to 2015. This zip code level annual data is drawn from the number of returns and adjusted gross income (before taxes), based on administrative records (individual income tax returns) from the Internal Revenue Service's Individual Master File (IMF) system[2]. Since these data is based on individual income tax returns filed with the IRS, I believe self-reported measurement error is minimized.

The published IRS data, i.e., Individual Income Tax Statistics (SOI), is in group form. I rely upon the studies of Cowell and Mehta (1982), Cowell (1995) and Frank (2009), to construct a compromise Gini coefficient. Accordingly, the lower limit of the Gini coefficient can be derived based on the assumption that all individuals in a group receive exactly the

[1] or even decreases.

[2] SOI Tax Stats - Individual Income Tax Statistics ZIP code documentation guide

mean income of the group:

$$G_L = \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{n_i n_j}{n\mu} |\mu_i - \mu_j| \tag{7}$$

where n is the number of individuals, μ is mean income, and subscripts i and j denote within-group values. The upper limit of the Gini can be derived based on the assumption that individuals within the group receive income equal to either the lower or the upper bound of the group interval:

$$G_U = G_L + \sum_{i=1}^{k} \frac{n_i^2 (a_{i+1} - \mu_i)(\mu_i - a_i)}{n^2 \mu (a_{i+1} - a_i)} \tag{8}$$

The compromise Gini coefficient proposed by Cowell and Mehta (1982) is then simply $2/3G_U + 1/3G_L$ (Frank 2009).

The data used for median housing prices was gathered from the Zillow data set. Zillow Home Value Index (ZHVI) is a seasonally adjusted measure of the median estimated home value across a given region (here different zip codes) and housing types[3]. I use ZHVI and median income (calculated using same IRS zip code-level data) to estimate my dependent variable, housing price to income ratio. This ratio has been used widely in the literature as a measure of housing market situation; a high housing price-to-income ratio is an indication of the housing market status, sometimes even a measure of a housing bubble (Green and Malpezzi, 2003; Jensen, 1998; Girouard et al., 2006). A threshold for this measure is often employed to judge whether housing bubbles exist or not, although the choice of threshold is under debate and varies across different contexts[4] (Zhang et al, 2016).

To estimate the impact of inequality on housing affordability, I use OLS and FE methods. FE estimation was chosen to eliminate omitted variable bias. However, I am also facing endogeneity. To address this issue, I add the system GMM developed by Arellano and Bover

[3]https://www.zillow.com/research/data/

[4]Renaud (1991) claimed that the housing price to income ratio in a healthy housing market should have a value between 2 and 6, whereas a higher value may reflect housing bubbles.

(1995) and Blundell and Bond (1998) to my analysis.

The model, as elaborated in previous literature, is as simple as:

$$R_{it} = \alpha + \beta G_{it} + X_{it} + u_{it} \tag{9}$$

where R_{it} is housing price to income ratio of the zip code i at time t, G_{it} is the constructed compromise Gini coefficient and X_{it} is a set of control variables that I believe may impact housing affordability, including share of minorities (African American households), proportion of members of the male gender and population (all at zip code level), and GDP per capita (at the state level). u_{it} is the error term including the city fixed effect. Zip code level demographic data was gathered from the 2010 U.S. Census and the state level GDP per capita was gathered from the Bureau of Economic Analysis (BEA). β is my coefficient of interest.

Table (1) shows a summary of the statistical data I use in this study. My data set covers more than 12,700 zip codes of the United States from 2005 to 2015. The upper part of the table shows the average of the variables across all zip codes and their corresponding standard deviation. As indicated in this table, housing price-to-income ratio and Gini coefficient are 8.84 and 0.46 on average across all zip codes with a standard deviation of 5.4 and 0.07, respectively. The average of the share of minorities (African Americans) and proportion of members of the male gender in households is 20 and 49 respectively, with standard deviations of 21 and 2 respectively, showing the vast differences especially in the share of different races between regions of this country. Percentages of family members with a bachelor's degree has an average of 29 with standard deviation of 14. These data set covers zip codes with average population of 20821 per square mile. The last column shows the state level GDP is 48904 dollars per capita on average with a standard deviation of 9226.

This table clearly indicates the significant differences across zip codes, confirming again that housing is a local market, hence a comprehensive zip code analysis has more advantages over country or state level studies, as are commonly used in other investigations. A detailed data set, such as mine, will capture the importance of regional factors affecting the fluctuations in this market. Examples of previous studies that verify the significance of local factor are Del Negro and Otrok (2007) who argue that historically, movements in housing prices are mainly driven by local factors, rather than variations in national factors, or Fratantoni and Schuh (2003) who explain that housing is determined in local markets and heavily dependent upon regional factors. The bottom part of table (1), which shows the correlation between these variables, confirms a positive correlation between the Gini coefficient and housing price to income ratio.

4 Empirical Results

In this section, I present the regression results using Equation (9). Table (2) shows the results of regression specifications examining the impact of income inequality on housing affordability, which is measured using the logarithm of the Gini Coefficient so that these regressions examine the effect of variation in income inequality operating both through variation in their own and others' income (Matlack and Vigdor, 2008). I use 3 methods of estimation: OLS, FE and system GMM. All variables have a positive sign, as expected, especially for the coefficient of interest, the Gini coefficient, showing that inequality is positively related with housing affordability, regardless of the method used. OLS results without and with control variables are shown respectively in the first two columns of table (2). The coefficient of the Gini coefficient in the second column suggests that a ten percent increase in the Gini coefficient leads to 0.89 increase in housing price to income ratio. This estimation is statistically significant at the 0.1 percent level.

In the next two columns, (3) and (4), I turn to FE estimation to control for unobserved zip code and time heterogeneity. The results using fixed effect show a negative coefficient of -0.04, statistically significant at the 0.1 percent level. However, controlling for the year fixed effect in column (4), the point estimate of the Gini coefficient becomes positive again with the magnitude of 0.02, which is smaller than the OLS estimate, and still significant. The difference between the estimates with and without control variables implies that the potential endogenous bias may be severe, unlike what Zhang et al. (2016) argue.

To address the reverse causality problem, I turn to system GMM estimation results, shown in the last column. The Gini coefficient here shows that an increase in inequality by ten percent, will lead to an increase in housing price to rent ratio index by 0.75, statistically significant at the 0.1 percent level. The points estimate the Gini coefficient using all methods are consistent with my hypothesis that an income inequality measured by Gini coefficient is significantly and positively related to housing affordability. My results in table (2) indicates that an increase in inequality leads to an increase in housing affordability index, meaning less affordable housing in the United States[5].

5 Robustness Check

5.1 Income Inequality Measured with Salaries and Wages

There has been a concern regarding an endogeneity problem in previous studies, arguing that housing prices may affect income inequality as well. The reverse causality problem comes from the fact that there is a capital gain associated with housing assets, especially when housing prices are rapidly increasing[6]. Moreover, owners in this market may benefit from rental income. The endogeneity problem will cause my OLS and FE estimation to be biased. However, I included the system GMM in my specifications to address in this issue and my regression results in table 2 show different number as the Gini coefficient among various methods, suggesting that the potential omitted variable bias might be severe enough to alter my conclusions qualitatively Unlike what we observe in previous literature (for instance Zhang et al., 2016).

To address the reverse causality problem as a robustness check, I analyze the impact of

[5]However, the magnitude of this impact depends on the estimation method.
[6]Zhang et al. 2016.

inequality using the Gini coefficient measured only with salaries and wages reported to the IRS. As we see in the previous studies, although salaries and wages may still not be entirely independent of the housing prices, they are less likely to be affected and less likely to have measurement errors, compared to reported total income.

Tables (3) and (4) represents the summary statistics of the data and same regression results, only this time I am using a Gini coefficient measured with salaries and wages. In table (3), the correlation between income inequality and housing price to rent ratio is a bit smaller, but still positive. Using salaries and wages, the coefficient of interest in my analysis shown in table (4) has (mostly) a positive sign and is statistically significant at the 0.1 percent level. However, the coefficient reported in columns (5) of table (4) is smaller than corresponding column in table (2), suggesting that after including city and time fixed effect and addressing endogeneity problem, the impact of inequality on housing affordability is smaller when Gini index is measured using salaries and wages. My main results, however, stays the same; with an increase in income inequality, housing affordability index rises, meaning less affordable housing for households.

5.2 Impact of Income Inequality on Zip Codes with Different Levels of Income

Now I turn to the impact of inequality on housing affordability in different zip codes. One might expect zip codes with higher levels of income to experience higher levels of housing prices as income inequality increases; wealthier households bid higher on houses in their zip code because they can afford paying higher prices, whereas low income households may struggle even for low quality and cheap types of housing. However, as I mentioned before, Lamont and Stein (1999) argue that in cities where homeowners are more leveraged housing prices react more sensitively to city-specific shocks, such as changes in income.

Table (5) shows the regression results of my analysis on the relationship between income Gini coefficient (using total income) and the housing price-to-income ratio at different percentiles. Following the literature, considering the endogeneity problem and the robustness of my results among different methods used in the previous sections, I used the fixed effect method here. The coefficients in the Gini index in this table shows that a ten percent increase in income inequality has a larger negative effect on housing affordability for zip codes with households with higher levels of income, compared to low and middle-income households, which is as we expected and in contrary with Lamont and Stein's argument. Table (6) shows the results of same regression analysis, this time using salaries and wages to measure Gini coefficient. The results are the same, suggesting that in wealthier zip codes, the housing price-to-income ratio is more affected by the inequality. My results also confirm the research of Abraham and Hendershott (1996) that documents a significant difference between coastal and inland cities.

6 Conclusion and Policy Implications

Housing prices and income inequality have been rising rapidly since 2011 in the United States. Some scholars and policy makers are now concerned that prices are too high relative to median household income, causing affordability problems for many families especially the poor. This paper estimates the impact of income inequality on housing affordability. The relationship between these two variables can be explained theoretically. Wealthier households bid higher on houses when their income increases, while middle and low-income households struggle for low quality housing, leading prices to rise for all types of housing.

Employing zip code level data from 2005 to 2015, this paper empirically studies the impact of inequality on housing affordability in the United States. My indicator of income inequality is the Gini coefficient, and for housing affordability I use the housing price-to-income ratio. The analysis in this study yields three main findings. First, using IRS data on income and Zillow data on median housing prices, this paper argues that an increase in income inequality (Gini coefficient) is associated with an increase in the housing affordability index. Using OLS, FE and system GMM methods, I provide results that suggest that the

Gini coefficient is positively and significantly related to the housing price to income ratio. The consistency of the results across various specifications indicates a robust relationship.

Second, to address the endogeneity problem, I use data on salaries and wages (published by the IRS) to calculate the Gini coefficient. The results present consistent patterns; a higher Gini coefficient is associated with a higher housing price-to-income ratio. According to the literature, in highly-leveraged cities, the reaction of housing prices to a change in income is greater[7]. I test these findings in this paper; in zip codes with higher levels of income, we observe more substantial impact of inequality on housing affordability compared to zip codes with middle and low-income households. Furthermore, the empirical results in my paper serve as evidences against "trickle-down" theory. My findings confirm previous studies and reveal that a rise in wealthy households' income leads to an increase in product prices faced by low-income families and thus makes the objective well-being of the poor worse (Zhang, 2015).

In sum, my analysis proves a positive relationship between inequality and housing affordability. Given the rapid rise in housing prices and inequality in many countries, this is a crucial policy related subject. The importance of income redistribution policies is, therefore, clearly the crux of my argument. Higher prices lead to higher rents, thus forcing the poor families to spend large fractions of their income on shelter (Quigley and Raphael, 2004). Once they have covered their housing costs, there will be less income available for saving or other consumption (Matlack and Vigdor, 2008). To maintain a healthy developed economy, governments needs to adopt redistribution policies to alleviate income inequality, as it was also suggested by Zhang et al. (2016) for China.

[7]Lamont and Stein, 1999.

References

Arellano, Manuel and Olympia Bover, "Another look at the instrumental variable estimation of error-components models," *Journal of econometrics*, 1995, *68* (1), 29–51.

Blundell, Richard and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of econometrics*, 1998, *87* (1), 115–143.

Capozza, Dennis R, Patric H Hendershott, and Charlotte Mack, "An anatomy of price dynamics in illiquid markets: analysis and evidence from local housing markets," *Real Estate Economics*, 2004, *32* (1), 1–32.

Chen, Jie, Qianjin Hao, and Mark Stephens, "Assessing housing affordability in post-reform China: a case study of Shanghai," *Housing Studies*, 2010, *25* (6), 877–901.

Cowell, Frank, "Measurement of inequality," 1998.

Cowell, Frank A and Fatemeh Mehta, "The estimation and interpolation of inequality measures," *The Review of Economic Studies*, 1982, *49* (2), 273–290.

Fairchild, Joseph, Jun Ma, and Shu Wu, "Understanding housing market volatility," *Journal of Money, Credit and Banking*, 2015, *47* (7), 1309–1337.

Frank, Mark W, "Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures," *Economic Inquiry*, 2009, *47* (1), 55–68.

Fratantoni, Michael and Scott Schuh, "Monetary policy, housing, and heterogeneous regional markets," *Journal of Money, Credit, and Banking*, 2003, *35* (4), 557–589.

Girouard, Nathalie, Mike Kennedy, Paul van den Noord, Christophe André et al., "Recent House Price Developments: The Role of Fundamentals," Technical Report, OECD Publishing 2006.

Glick, Reuven, Kevin J Lansing, Daniel Molitor et al., "What's different about the latest housing boom?," *FRBSF Economic Letter*, 2015, *34*, 1–5.

Green, Richard K and Stephen Malpezzi, *A primer on US housing markets and housing policy*, The Urban Insitute, 2003.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai, "Superstar cities," *American Economic Journal: Economic Policy*, 2013, *5* (4), 167–99.

Hendershott, Patric H and Jesse M Abraham, "Patterns and determinants of metropolitan house prices, 1977-91," Technical Report, National Bureau of Economic Research 1992.

Hulchanski, J David, "The concept of housing affordability: Six contemporary uses of the housing expenditure-to-income ratio," *Housing studies*, 1995, *10* (4), 471–491.

Hwang, Min and John M Quigley, "Economic fundamentals in local housing markets: evidence from US metropolitan regions," *Journal of regional science*, 2006, *46* (3), 425–453.

Jensen, M, "Affordability Indicators," *Encyclopedia of Housing. Thousand Oaks: Sage*, 1998, pp. 11–12.

Lamont, Owen, Jeremy C Stein et al., "Leverage and House-Price Dynamics in US Cities," *RAND Journal of Economics*, 1999, *30* (3), 498–514.

Matlack, Janna L and Jacob L Vigdor, "Do rising tides lift all prices? Income inequality and housing affordability," *Journal of Housing Economics*, 2008, *17* (3), 212–224.

Negro, Marco Del and Christopher Otrok, "99 Luftballons: Monetary policy and the house price boom across US states," *Journal of Monetary Economics*, 2007, *54* (7), 1962–1985.

Ortalo-Magne, Francois and Sven Rady, "Homeownership: Low household mobility, volatile housing prices, high income dispersion," 2002.

Quigley, John M and Steven Raphael, "Is housing unaffordable? Why isn't it more affordable?," *Journal of Economic Perspectives*, 2004, *18* (1), 191–214.

_ , _ , and Eugene Smolensky, "Homeless in America, homeless in California," *Review of Economics and Statistics*, 2001, *83* (1), 37–51.

Ray, Rosalie, Paul Ong, and Silvia Jimenez, "Impacts of the Widening Divide."

Rodda, David T, "Rich man, poor renter: A study of the relationship between the income distribution and low-cost rental housing.," 1995.

Vigdor, Jacob L, "Does Gentrification Harm the Poor?," *Brookings-Wharton Papers on Urban Affairs*, 2002, *2002* (1), 133–182.

Zhang, Chuanchuan, "Income inequality and access to housing: Evidence from China," *China Economic Review*, 2015, *36*, 261–271.

_ , Shen Jia, and Rudai Yang, "Housing affordability and housing vacancy in China: The role of income inequality," *Journal of Housing Economics*, 2016, *33*, 4–14.

Tables

	HPIR	Gini Index	Minorities (%)	Male (%)	Bachelor's Degree (%)	Population	GDP
Mean	8.843613	.4682712	19.99664	49.1919	29.35671	20821.54	48904.52
SD	5.401261	.0709575	21.07048	2.029753	14.41628	15821.32	9226.227
HPIR	1.0000						
Gini Index	0.3257	1.0000					
Minorities	0.0723	-0.2282	1.0000				
Male	0.0167	-0.0225	-0.0721	1.0000			
Bachelor's Degree	0.4819	0.5127	-0.2114	-0.1415	1.0000		
Population	0.1467	-0.0812	0.4595	-0.1489	0.0380	1.0000	
GDP	0.3882	0.0893	0.0452	-0.0135	0.2672	0.0204	1.0000

Table 1: Summary Statistics of data, 2005 – 2015. All variables are zip code level data (except for GDP that is State level). Gini Index in this tables is calculated using Adjusted Gross Income (AGI). The upper part of this table shows mean and standard deviation and the bottom part shows correlation between these variables.
Source: IRS, Zillow data, Census data 2010, FRED.

	Dependent variable: Housing Price-to-Income Ratio				
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	FE	FE	GMM
Gini Index	0.114***	0.0891***	-0.0455***	0.0201***	0.0751***
	(109.28)	(78.42)	(-53.34)	(12.87)	(104.53)
Minorities (%)		0.0399***	-	-	0.0337***
		(57.78)	(.)	(.)	(57.47)
Male (%)		0.258***	-	-	-0.0669***
		(41.69)	(.)	(.)	(-8.67)
Bachelor's degree (%)		0.123***	-	-	0.0236***
		(112.12)	(.)	(.)	(39.81)
Population		0.323***	-	-	0.776***
		(36.17)	(.)	(.)	(85.79)
GDP		1.579***	3.412***	2.227***	0.617***
		(112.56)	(55.51)	(48.81)	(82.63)
Lag of the ratio					0.795***
					(533.24)
Year fixed effect	No	No	No	Yes	Yes
N	111026	111026	111026	111026	100496
R^2	0.097	0.427	0.119	0.560	-
adj. R^2	0.097	0.427	0.119	0.560	-

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: Regression results of panel data analyses of the impact of inequality (calculated using Adjusted Gross Income, AGI) on housing price-to income ratio. Robust standard errors clustered at the city level are in parentheses. For the system GMM estimation in the last column, P-value of the Arellano-Bond test suggests that instruments are valid.

	HPIR	Gini Index	Minorities (%)	Male (%)	Bachelor's Degree (%)	Population	GDP
Mean	8.226779	.4206804	20.26871	49.18923	29.56019	21283.32	48920.66
SD	4.89593	.0562135	21.01639	2.085061	14.05455	15773.48	9101.493
HPIR	1.0000						
Gini Index	0.2705	1.0000					
Minorities	0.0874	-0.3088	1.0000				
Male	0.0320	-0.0386	-0.0732	1.0000			
Bachelor's Degree	0.4450	0.5728	-0.2084	-0.1164	1.0000		
Population	0.1467	-0.1052	0.4619	-0.1498	0.0115	1.0000	
GDP	0.1373	0.1333	0.0412	-0.0101	0.2632	0.0127	1.0000

Table 3: Summary Statistics of data, 2005 – 2015. All variables are zip code level data (except for GDP that is State level). Gini Index in this tables is calculated using salaries and wages.
Source: IRS, Zillow data, Census data 2010, FRED.

	Dependent variable: Housing Price-to-Income Ratio				
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	FE	FE	GMM
Gini Index	0.0956***	0.0939***	-0.146***	0.0172***	0.0488***
	(82.48)	(70.67)	(-113.11)	(7.73)	(25.36)
Minorities (%)		0.0438***	-	-	0.0579***
		(62.36)	(.)	(.)	(60.77)
Male (%)		0.243***	-	-	0.402***
		(40.82)	(.)	(.)	(30.90)
Bachelor's degree (%)		0.0997***	-	-	0.0671***
		(87.59)	(.)	(.)	(34.05)
Population		0.276***	-	-	0.949***
		(31.14)	(.)	(.)	(61.61)
GDP		1.366***	3.016***	2.239***	0.873***
		(97.10)	(54.00)	(52.04)	(67.02)
Lag of the ratio					0.741***
					(277.22)
Year fixed effect	No	No	No	Yes	Yes
N	96745	96745	96745	96745	77496
R^2	0.066	0.400	0.331	0.552	-
adj. R^2	0.066	0.400	0.331	0.552	-

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: Regression results of panel data analyses of the impact of inequality (calculated using salaries and wages) on housing price-to income ratio. Robust standard errors clustered at the city level are in parentheses. For the system GMM estimation in the last column, P-value of the Arellano-Bond test suggests that instruments are valid.

	Dependent variable: Housing Price-to-Income Ratio		
	(1)	(2)	(3)
	$< 10^{\text{th}}$ percentile	$25^{\text{th}} >$ and $<75^{\text{th}}$ percentile	$> 90^{\text{th}}$ percentile
Gini Index	-0.00118	0.0301***	0.0906***
	(-0.24)	(12.81)	(6.16)
Fixed effect	Yes	Yes	Yes
N	11103	55510	11103
R^2	0.481	0.506	0.579
adj. R^2	0.481	0.506	0.578

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5: Regression results of panel data analyses of the impact of inequality (calculated using Adjusted Gross Income, AGI) on housing price-to income ratio for zip codes in different percentiles of income. The results in this table were estimated using FE model (column 4 of table 2).

	Dependent variable: Housing Price-to-Income Ratio		
	(1)	(2)	(3)
	$< 10^{\text{th}}$ percentile	$25^{\text{th}} >$ and $<75^{\text{th}}$ percentile	$> 90^{\text{th}}$ percentile
Gini Index	-0.0314***	0.0109**	0.128***
	(-4.42)	(3.27)	(8.41)
Fixed effect	Yes	Yes	Yes
N	9675	48371	9675
R^2	0.449	0.483	0.661
adj. R^2	0.449	0.483	0.661

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 6: Regression results of panel data analyses of the impact of inequality (calculated using salaries and wages) on housing price-to income ratio for zip codes in different percentiles of income. The results in this table were estimated using FE model (column 4 of table 2).

Figures

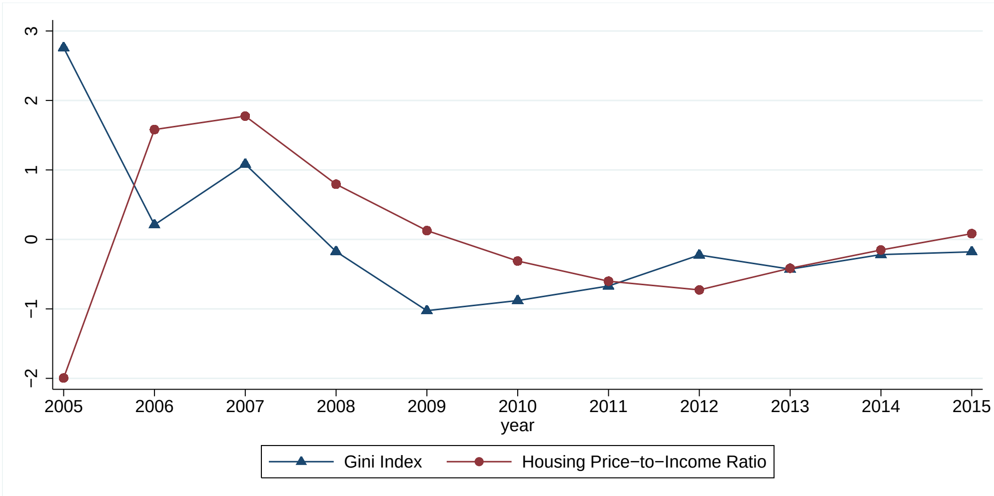

Figure 1: Dynamics of Gini coefficient and housing price to income ratio over 2005 to 2015