Investors and Housing Affordability*†

Carlos Garriga‡, Pedro Gete§, Athena Tsouderou¶

February 2021

Abstract

This paper studies the impact of a new and growing class of institutional investors on U.S. housing affordability. We find that investors' purchases increase the price-to-income ratio, especially in the bottom price-tier, the entry point for first-time buyers. Investors cause a short-run reduction in the vacancy rate and a medium-run positive response of construction, which is stronger for multi-unit buildings. The supply response mitigates the effect on affordability, although not enough to reverse the effect. The transmission channels depend on the housing supply elasticity. In highly elastic areas investors affect rents more than prices, whereas in highly inelastic areas investors have the opposite effect.

Keywords: Institutional investors, housing affordability, house prices, supply elasticity, propensity to invest.

*We thank Itzhak Ben-David, Morris Davis, Anthony DeFusco, David Echeverry, Andra Ghent, Jonathan Halket, Lu Han, Nina Karnaukh, Finn Kydland, Jose Maria Liberti, David Ling, Christos Andreas Makridis, Charles Nathanson, Michael Reher, Stephen L. Ross, Martin Schneider, and participants at Durham, ECB, Hebrew University of Jerusalem, NEOMA, Notre Dame, IE, Ohio State, and 2018 UEA, 2019 HULM, 2019 AREUEA International, 2019 AEFIN Finance Forum, 2019 SED, 2020 Econometric Society and 2020 AREUEA Conferences.

†The views expressed herein do not necessarily reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System. The results and opinions are those of the authors and do not reflect the position of Zillow Group. Research reported in this paper was partially funded by the Spanish Ministry of Economy and Competitiveness (MCIU), State Research Agency (AEI) and European Regional Development Fund (ERDF) Grant No. PGC2018-101745-A-I00.

‡Federal Reserve Bank of St. Louis. carlos.garriga@stls.frb.org.

§IE Business School, IE University. pedro.gete@ie.edu.

¶IE Business School, IE University. athenatsouderou@gmail.com.

1 Introduction

Housing affordability is one of the most pressing challenges for most cities in the world. In the United States in 2019, the median house price exceeded three times the median household's annual income in 55% of the MSAs. Moreover, in 42% of the MSAs in 2019 the rent-to-income ratio exceeded the threshold of rental affordability of 27%. As Figure 1 shows, the recent lack of housing affordability resembles the situation during the large housing boom of the 2000s.

This paper studies the link between the recent worsening of affordability and the emergence of a new class of investors in real estate markets, namely, *institutional investors.* Increasing demand from investors can move house prices directly, but it can also impact the rental market by affecting the supply of tenant-occupied housing. Investors' purchases can also influence the characteristics of the newly constructed units, for example favoring small or multi-family units instead of more traditional single-family. Investors can shape the composition of the stock of residential housing and can alter homeownership rates (Halket, Nesheim, and Oswald 2020).

There is a growing literature studying housing investors, like Agarwal et al. (2019), Cvijanovic and Spaenjers (2020), DeFusco et al. (2018), Favilukis and Van Nieuwerburgh (2021) and others that we discuss below. Our contribution is to show that post-financial crisis institutional investors affected housing prices and affordability. The impact of investors' purchases differs across market segments based on price-tiers, building sizes, and locations with different supply restrictions. We connect the asset pricing perspective with the traditional view of affordability from urban economics that focuses on geographical supply constraints (Glaeser and Gyourko 2018; Molloy, Nathanson, and Paciorek 2020).

We define institutional investors as legal entities who purchase multiple housing units under the name of an LLC, LP, Trust, REIT, etc. Figure 2 shows a new fact: MSAs that experienced the largest increase in the price-to-income ratio post-crisis also had the largest market share of housing purchases by institutional investors.[1] Establishing a causal connection is not direct. For example a standard OLS regression of house price growth on investors' market share would be biased downwards if investors were attracted to areas where prices collapsed following the crisis. To overcome this challenge we use an instrumental variable approach with a rich dataset covering the U.S. MSAs from 2000 to 2017.

We use a "Bartik inspired instrument" that takes advantage of the Quantitative Easing (QE) programs that the Fed implemented during the financial crisis. QE resulted in a sharp decline

[1] Figure A1 shows the positive correlation of the raw data in a scatter plot, also denoting the population of each MSA.

in the returns of safe assets, which encouraged risk-taking behavior by investors. Martínez-Miera and Repullo (2017) and Rodnyansky and Darmouni (2017) document that QE triggered a search for yield. While QE was a national shock, different regions reacted differently based on the pre-shock propensity for investments among the local high-income population. High-earning, sophisticated residents directed more capital to the housing market, through new or existing legal entities.

We capture such pre-shock local propensity to invest with the share of the top earners' business income over total income in each MSA in 2007. In the panel data analysis our instrument is this share interacted with the growth rate in the rate on certificate of deposits (CDs). This is like the housing net worth channel of Mian and Sufi (2014) that exposes certain areas to larger macro effects from declines in housing prices due to their housing leverage. In our case, we expose investment-prone areas to the QE national shock.

It is straightforward to show that our instrument is strongly correlated with the geographical presence of institutional housing investors post-crisis. The new housing investors post-crisis are mainly small and local, and create legal entities to buy houses throughout the U.S., while the large private equity investors' housing purchases are geographically concentrated in "superstar cities" (Garriga, Gete, and Tsouderou 2020). Thus, the instrument satisfies the relevance condition.

We perform several tests to rigorously assess the exclusion restriction. That is, conditional on observables, the share of the top earners' business income is uncorrelated with factors that determine house prices: (1) We provide extensive evidence that in the pre-QE period the instrument does not predict changes in housing prices or new construction. Areas with the highest or lowest levels of the instrument exhibit parallel pre-trends. Placebo tests confirm that the instrument only captures post-crisis shocks to investment in the housing markets. The assumption that there were no pre-existing differences in the growth of housing prices and that the common unconventional monetary policy shock caused the changes in price growth seems plausible (Goldsmith-Pinkham, Sorkin, and Swift 2020). (2) We provide evidence that it is very hard to predict the investment attitude of an MSA, as analogously it is difficult to explain which cities become hubs for entrepreneurship.[2] That is, even if factors such as the tax regime, natural amenities, or the structure of the population have some forecasting power, most of the cross-sectional geographical variation in entrepreneurship, and in the share of business income, is unexplained. Moreover, since the instrument is determined two years prior to the period of study (2009 to 2017) this reduces risk of reverse causality. (3) We run a number

[2]For example when trying to predict cross-regional differences in entrepreneurship the largest R-squared in Davidsson (1991) is 25%, and in Rocha and Sternberg (2005) is 23%.

of specifications to show that the results are not driven by shifts in the composition of labor demand of MSAs during the post-crisis period. (4) We control thoroughly for an array of fixed effects and of local activity trends (income, population flows, unemployment, GDP, wages and labor force participation), so it is unlikely that the error term reflects common movers of both investors and housing market variables. (5) Finally, we confirm the robustness of the results to alternative measures of the investors' presence and different geographical units.

Our first result is that institutional investors increase housing prices and worsen affordability. Between 2009 to 2017, one standard deviation higher purchases by institutional investors leads to 1.46 percentage points higher housing price growth for the median house. Moreover, we find that prices grow significantly faster than income. Further investigation indicates that these results do not come from the variation driven by the so-called "superstar cities" as discussed by Gyourko, Mayer, and Sinai (2013), or by the purchases of so-called "Wall Street Landlords" often discussed in the popular press.[3] The analysis that excludes the top cities by large investors, often correlated with superstar cities, still arrives to the same conclusions.

Housing markets are often segmented, hence, the effect of house purchases might not have the same pass-through across market tiers, consistent with evidence from Armesto and Garriga (2009), and Piazzesi, Schneider, and Stroebel (2020). We find that the market segment that is more sensitive to purchases by institutional investors is the bottom price-tier. In the bottom price segment, one standard deviation higher purchases causes 2.29 percentage points higher housing price growth. As first-time buyers tend to purchase housing from the bottom price tier, it is apparent that investors had large negative effects on affordability especially for this group.

Since real estate developers and the construction sector need to anticipate well in advance future housing demand, it is important to explore the effects of these purchases in the timing and the composition of the housing supply. Our analysis indicates that during 2009-2017 institutional investors increase the overall supply of housing by 5.1% on average, for every one percentage point increase in the share of investors. One percentage point increase in the share of investors increases the number of new construction permits for single-unit buildings by 4.8% on average, and for buildings of 5 or more units by 16.4% on average. There are clear compositional effects in the characteristics of the newly constructed stock of residential housing.

Moreover, it is relevant to separate the short-run effects where the housing supply is more inelastic from the long-run where housing supply can respond. To address this question, the second part of the paper quantifies the dynamic effects of institutional investors on housing affordability, price growth, and the supply response employing the projection method developed

[3]See the Wall Street Journal (2017), or the ACCE Institute Report (Abood 2018).

by Jordà (2005).

The effect of purchases by institutional investors varies over time. Much of the cross-sectional results are driven by a powerful short-run response of price increases that weakens over time, as in the medium-run housing supply, measured in terms of the number of construction permits, responds. More specifically, the dynamic analysis illustrates how investors' purchases reduce the vacancy rate in the short-run, and generate a medium-run response of construction. These equilibrium responses slow down the price growth, however they do not reverse the effects of investors on worsening affordability.

The effects on price-to-income and price-to-rent ratio differ once we split the sample of MSAs by the housing supply elasticity. In highly elastic areas, investors affect rents more than prices, whereas in areas that are highly inelastic investors have the opposite effect. As a result, in MSAs with low housing supply elasticity, the short-run fluctuations in prices and worsening of affordability are much larger than in MSAs with high supply elasticity, but the opposite effect happens via rents.

These differences in the dynamic response of the different measures of affordability (price-to-income or rent-to-income ratios) have important consequences for the design of policies. Officials in several cities have enacted or are discussing policies to block institutional investors. For example, New York and California, where presence of institutional investors has reached unprecedented highs, recently approved statewide rent controls (Business Insider 2019). Amsterdam has discussed banning institutional investors from purchasing and renting properties (Bloomberg 2018), Berlin is considering expropriating large, private, profit-seeking landlords (The Wall Street Journal 2019), and Spain recently imposed measures to penalize institutional investors (Bloomberg 2019). The implications from our analysis is that such policies need to take into account different market segments, the local supply elasticities and the composition of new supply. For example, our analysis shows that investors caused minimal price increases in MSAs where there are loose supply restrictions.

The rest of the paper is organized as follows: Section 2 outlines the theory and summarizes the existing literature. Section 3 describes the data. Section 4 presents the cross-sectional analysis. Section 5 presents the dynamic analysis. Section 6 assesses the validity of the instrument and the robustness of the results. Section 7 concludes.

2 Theory and Related Literature

The finance literature views housing as an asset that provides services. The nature or identity of the buyers can influence the valuation of the asset. Several papers have focused on *short-term investors* (known as flippers), like for example Agarwal et al. (2019); Albanesi, De Giorgi, and Nosal (2017); Bayer, Mangum, and Roberts (2021) and Ben-David (2011). Another part of the literature has explored the contribution of deep pocket *foreign and out-of-town investors*, like Chinco and Mayer (2016); Cvijanovic and Spaenjers (2020); Davids and Georg (2020) or Favilukis and Van Nieuwerburgh (2021).

A growing literature starts to look into *institutional investors*. For example, Garriga, Gete, and Tsouderou (2020) document that the new institutional investors are mainly focused on the search for yield. Mills, Molloy, and Zarutskie (2019) document the purchases of single-family homes by several large firms securitizing these investments in capital markets. Lambie-Hanson, Li, and Slonkosky (2019) use as identification strategy the First Look program to study the effects of large institutional investors on local house prices. Ganduri, Xiao, and Xiao (2019) find that the bulk purchases of distressed single-family homes by large institutional investors have a positive spillover on nearby home values. Graham (2019) finds that during the latest housing bust investors substitute for falling homeowner demand, lessening the declines in housing prices. Brunson (2020) studies the impact of institutional investors in Charlotte, NC, and Allen et al. (2018) in Miami-Dade County, FL. Wu, Xiao, and Xiao (2020) find that while institutional landlords extract greater surplus from renters, they also improve the quality of rental services.

The standard view of affordability comes from the urban economics literature that views housing as a localized consumption good. It suggests that affordability problems are related to geographical areas with constraints in the production of this good. Gyourko, Mayer, and Sinai (2013) suggest that the inelastic supply of land along with an increasing number of high-income households leads to persistent high house prices in large MSAs and crowds out lower-income households. Molloy, Nathanson, and Paciorek (2020) develop a dynamic model that predicts that supply constraints should have a larger effect on house prices than rents.[4] This literature abstracts from the contribution of the type of housing investors on housing affordability.

This paper combines the approaches from urban economics and asset pricing. We determine the causal effect of investors' purchases on house prices, rents and affordability over different time horizons. We separate areas where housing supply can react to prices, from areas with strict supply restrictions.

[4]For a summary of new research on housing affordability from the urban economics perspective see Ben-Shahar, Gabriel, and Oliner (2020).

3 Data

Data on investors in the U.S. housing market come from the Zillow Transaction and Assessment Dataset (ZTRAX, Zillow 2017).[5] The database covers all ownership transfers as recorded by the counties' deeds. We focus on ownership transfers of residential properties, including multi-family and single-family, from January 1st, 2000 to December 31st, 2017. Our final sample, from which we construct the investors' purchases variable, consists of about 85 million transactions.

We follow a rigorous methodology to classify institutional investors. First, we distinguish between individual and non-individual buyers based on the buyer name. Second, we filter out buyers that are relocation companies, NGOs, construction companies, national or regional authorities, mortgage lenders, GSEs, and the state taking ownership of foreclosed properties. Third, our variable of investors' presence is the share of the dollar value of purchases by investors over the dollar value of all purchases, that is, by investors and households.[6]

Our instrument uses zip code data of individual tax returns from the Statistics of Income of the Internal Revenue Services (IRS). It is the average share of business income over total income of high earners (annual adjusted gross income above $100K) in each MSA in 2007. We weight by the total income of high-earners to aggregate to the MSA level. To construct the panel version of the instrument, we use the average one-year certificate of deposits (CD) rate from Bankrate, a consumer financial services company.

We use the Zillow Home Value indices for all homes, top-tier homes and bottom-tier homes at the MSA level. The bottom-tier segment of the market is the bottom third of the housing price distribution in each MSA. The bottom-tier price is the median price of the segment, that is, the bottom 17th percentile of the prices of the total market within an MSA.[7] Housing rents come from the Zillow Rent Index for all homes. We collect the number of new construction permits from the Census Bureau's annual Building Permits Survey. Finally, population comes from the Census, the unemployment rate from the Bureau of Labor Statistics, and income from the Statistics of Income of the IRS and Zillow. We calculate the 17th, 50th and 83rd percentiles of individual income from the IRS to get the price-to-income ratio for the corresponding tiers.

[5] We include a detailed description of the data sources in the Appendix *A*.

[6] The number of purchases would underestimate presence in the apartments market. For example the number of purchases would equate a purchase of one condominium to the purchase of one apartment building of 100 apartments. For robustness checks we use alternative measures of the presence of investors based on the number of properties or the number of units purchased.

[7] In a symmetrical way, the top-tier segment of the market is the top third of the price distribution in each MSA, and the top-tier price is the top 83rd percentile of prices within an MSA.

Table 1, Panel A summarizes the key statistics of the cross-sectional sample. There are 332 MSAs with the full set of average housing variables and investors' market share for the years 2009-2017, control variables beginning in 2000, and tax-returns for the year 2007. On average, investors purchase 12.37% of the market annually, over 2009-2017. Prices for a median house grow on average by 0.47 annually in real terms, while for an individual with median income, the price-to-income ratio is 4.76 on average. Table 1, Panel B summarizes the key panel variables, at the MSA-year level, we use in the dynamic analysis.

4 Investors and Affordability in the Cross-Section

Our goal is to study the effects of institutional investors on housing affordability. Formally, we run the following cross-sectional regressions:

$$y_{m,09-17} = \beta_0 + \beta_1 Inst_{m,09-17} + \gamma C_m + \alpha_s + u_m, \tag{1}$$

where $y_{m,09-17}$ denotes the relevant housing variables for a given MSA indexed by m and for the period 2009-2017. The housing variables we study are the average annual real housing price growth rate, the price-to-income ratio for different price and income percentiles, and the log number of construction permits for different kind of houses. $Inst_{m,09-17}$ is the average share of institutional investors' dollar value of purchases over the total purchases in MSA m over the same period. The term C_m summarizes the MSA-specific controls: population growth, income growth, unemployment rate change and real housing price growth over the periods 2000-2006 and 2006-2007. We also control for the log number of building permits in 2007, to account for new supply. The term α_s includes state dummies to account for the time-invariant state-specific influences.

A direct OLS estimation of specification (1) is likely to be biased downwards. This is because the estimates might capture "reverse causality" as investors target MSAs where prices fell more after the crisis and were slow to pick up. To overcome this problem, we use an instrument for the investors' market share of purchases.

4.1 The instrumental variable: Propensity to invest

We use an instrumental variable that allows us to exploit variation in the geographical presence of investors and that is plausibly exogenous to other drivers of housing markets. This instrument is the average share of value of business income over total income of the top earners in an MSA for the year 2007. Top earners are residents that file total income larger than $100K in their tax returns, and they are the ones most likely to have the scope to invest in housing.[8]

The economic rational for this instrument is that it measures the local exposure to the sharp drop in returns of safe assets caused by the unconventional monetary policy shock during the financial crisis. The QE programs reduced the supply of safe assets in the market. The federal funds rate and the returns on certificates of deposits and other safe assets fell close to zero. This national shock triggered a search for yield (Martínez-Miera and Repullo 2017; Rodnyansky and Darmouni 2017; Daniel, Garlappi, and Xiao 2018; Campbell and Sigalov 2020). Areas with high-earning, knowledgeable, risk-seeking investors experienced higher investment in the local housing markets. Consistent with this theory, De Stefani (2020) documents that the investment attitude towards housing increased significantly among the wealthy U.S. population following the financial crisis.

Crucially for the validity of our identification, the share of business income is uncorrelated with factors that drive housing markets, conditional on our multiple controls. The literature on entrepreneurship finds that it is very hard to explain geographical differences in entrepreneurship (Davidsson 1991; Rocha and Sternberg 2005; Bosma and Kelley 2019). Section 6 contains multiple tests that all suggest that the instrument satisfies the exclusion restriction. That is, given our various controls for observable factors, exposure to top earners' business income in 2007 is uncorrelated with other drivers of housing markets over 2009-2017.

Table 2 assesses the relevance of the instrument, showing the results of the first stage of the 2-stage least squares (2SLS) regression based on specification (1). After controlling for the relevant MSA-level controls, and state dummies, the instrument is significantly correlated with the investors' purchases. The Wald F statistic of 19.4, reported in Table 3, allows to reject that the instrument is weak.

4.2 Cross-sectional results

Table 3 summarizes the effects of institutional investors on housing price growth and on the price-to-income ratio, by price and income tier, over the period 2009-2017. The first column reports the OLS estimation of (1) for the median house price and median income. The smaller coefficient of the OLS estimation is consistent with the expected downward bias of the OLS, since the prices were falling significantly up to 2012, and investors were likely to select areas

[8] As robustness tests, we have also constructed instruments using the average share of business income in the MSA and different moments of the distribution. The results, not reported here, hold for different versions of the instrumental variable.

were prices collapsed. From the summary statistics (Table 1, Panel A) the average growth in real housing prices between 2009 and 2017 was 0.47%. Our results show that the effect of investors was to prevent even larger drops in housing prices and eventually recover the positive growth.

Looking at the standardized estimates,[9] Table 3 shows that one standard deviation higher purchases by institutional investors (7.78% from Table 1) causes 0.827 standard deviations, or 1.46 percentage points, higher housing price growth for the median house.[10] However, the largest effects are on the bottom price tier of the market. For this bottom tier, one standard deviation higher purchases causes 0.909 standard deviations, or 2.29 percentage points, higher housing price growth.[11]

Turning to price-to-income ratio, Table 3 also shows positive effects, that are also larger for the bottom price tier of the market. For example, from Table 1 we know that the average price-to-income ratio in the bottom tier of the market is 8.48. If we add the change caused by one standard deviation higher purchases by investors then the ratio becomes around 20.[12]

Our findings are robust to excluding superstar cities. These are the areas that are heavily affected by purchases coming from Wall Street investors. Table A1 replicates the analysis from Table 3 for two different samples: first excluding the top 10 cities, and second excluding the top 20 cities based on large investors' purchases. The results from Table A1 show that all effects become larger as we remove the top cities. For example, one percentage point increase in the share of investors' purchases increases bottom-tier price-to-income by 1.538 in the full sample, 1.540 in the sample without the top 10 superstar cities, and 1.728 in the sample without the top 20 superstar cities.

To check the robustness of the results to the geographical unit, we perform the same analysis with counties instead of MSAs. Table A2 shows that the results remain unchanged when we use counties.

Table 4 summarizes the effects of institutional investors on new construction over the period 2009-2017. The first column reports the IV estimation of (1) for the log number of construction permits for all houses. One percentage point increase in the share of investors increases the number of new construction permits by 5.2% on average ($e^{0.051} - 1$). By analyzing separately

[9] The standardized estimates use the standardized share of investors and standardized dependent variables, for easier comparison and derivation of the economic significance of the results.We restrict the sample of the standardized variables to the MSAs for which we have Zillow housing prices for all price tiers, to facilitate comparison.

[10] That is, 0.827 from Table 3 multiplied by 1.77 from Table 1.

[11] That is, 0.909 from Table 3 multiplied by 2.52 from Table 1.

[12] That is, adding 8.48 from Table 1 plus the product of 2.108 from Table 3 and 5.48 from Table 1.

the permits for buildings of different number of units, we find that the share of investors leads to an increase in permits of 4.8% ($e^{0.047} - 1$) for single-unit houses. The effect of investors on construction of building of 5 or more units is more than double, at an average increase of 16.4% ($e^{0.152} - 1$). Table $A3$ confirms that when we take out of our sample the superstar cities, the results for new construction hold.[13]

5 Dynamic Real Effects of Investors

This section studies how the response of housing prices and quantities to the institutional investors' purchases changes over time. We follow Jordà (2005) and estimate sequential regressions of the dependent variable shifted forward.[14] That is, we estimate:

$$y_{m,t+i} = \beta_0 + \beta_1^{(i)} Inst_{m,t-1} + \beta_2 y_{m,t-1} + \gamma C_{m,t-1} + \alpha_m + b_t + u_{m,t}, \qquad (2)$$

where t indexes years and m MSAs, and $y_{m,t}$ denotes the housing variables: real housing price growth rate from year $t-1$ to year t, for top, mid and bottom tier houses, the price-to-income and rent-to-income ratios, the price-to-rent ratio, and the log number of new construction permits.

$Inst_{m,t-1}$ is the institutional investors' share of dollar value of purchases over the total market value for the year $t-1$ in MSA m. $C_{m,t-1}$ are the time-varying MSA-specific controls that include the population growth rate, the median income growth rate, and the unemployment rate change.[15] The location fixed effects α_m capture the time-invariant MSA-specific influences, and the time fixed effects b_t account for the time-varying factors common to all MSAs, like national mortgage rates. We include a lagged dependent variable $y_{m,t-1}$ to allow the growth response to be temporary.

The estimate of interest is the vector of $\{\beta_1^{(i)}\}$, where $i = 0, 1, ..., 6$ is the time horizon of the response, that is, the number of years after the investors' purchases. Each $\beta_1^{(i)}$ corresponds to the effect of investors' share of purchases at horizon i. When $i = 0$, this gives the usual panel specification. We estimate (2) for the full panel data from 2009 to 2017. In the estimation

[13] Along this line, comparing the effects of investors on the single-family market to the effects on the multi-family market, we estimated the effects separately for single-family prices and multi-family prices. However, we do not find any significant differences between the two.

[14] Favara and Imbs (2015) also apply this method to study house prices, and Mian, Sufi, and Verner (2017) to study GDP growth.

[15] Controlling for contemporaneous income and population growth, and unemployment rate change doesn't change the results.

we cluster standard errors by MSA to allow for within-MSA correlation throughout the sample period.[16]

For the analysis, we employ the panel version of the instrumental variable defined as the 2007 local exposure to top earners' business income interacted with the certificate of deposits (CD) interest rate growth. The idea is that QE triggered a national shock to the CD rate, which is equal for all locations and it is not driven by local factors. The exposure of each location to the national shock is unrelated to local factors affecting the housing markets, as we assess in Section 6. The exposure is also predetermined, fixed in 2007, which minimizes the possibility of reverse causality. Thus, this instrument captures which MSAs are more likely to have housing investors after the QE policies. The rational for our panel instrument is analogous to the housing net worth channel of Mian and Sufi (2014) that exposes certain areas to larger macro effects from declines in housing prices due to their housing leverage. In our case, we expose investment-prone areas to the QE shock. Table A4 shows that the relevance condition is satisfied. Section 6 assesses the exogeneity condition, which is derived from the cross-sectional dimension.

The dynamic effect of purchases by institutional investors varies over time as illustrated in Figures 3 to 7 that display the results from the instrumental variable estimation of specification (2).[17] Initially, the purchases of institutional investors have positive effects on price and rent growth. However, the positive dynamics on house price growth become zero in the third year, and in the fourth for rent growth. This means, that while the growth of prices and rents in the short-term is increasing from one year to the next, in the medium-term this acceleration stops and the annual growth is decreasing. The cumulative responses in Figure 3 confirm that after three years from the investors' purchase shock the price and rent growth slow down, although on average the effects remain positive. Figure 4 shows an average result across MSAs for the effects of investors' purchases on price-to-income and rent-to-income ratios.

The larger short-run effect in the panel regressions, relative to the cross-section results, is due to the lack of response of the housing supply. Notice that the response of new construction, in Figure 5, measured by building permits, is a hump shape that peaks after two or three years and remains positive for several years. In other words, investors generate price increases that motivate a strong response from housing supply that slows down the affordability effects. Another way to look at this supply reaction is to look at vacancies in Figure 5. In the short-term vacancies decrease as investors purchases meet an inelastic supply of housing. Vacancies

[16] The results remain unchanged when we alternatively allow for Newey-West standard errors that allow for heteroskedasticity and within-MSA serial autocorrelation of the error term.

[17] Tables $A5$ and $A6$ have the results of the estimation.

increase as new constructions arrives to the market.[18]

Consistent with the supply response, the effects on price-to-income and price-to-rent ratio differ once we split the sample of MSAs by the housing supply elasticity. The average effects (from Figure 4) look very different based on the housing supply elasticity of the area, as we show in Figures 6 and 7. In highly inelastic areas, the short-run fluctuations in prices and worsening of price affordability are much larger than in MSAs with high supply elasticity. In other words, in these areas with low supply elasticity, investors drive prices and don't seem to move rents in the short-run. As a result the price-to-rent ratio increases, the price-to-income ratio also increases, and the rent-to-income ratio is constant. In areas with high supply elasticity the opposite effect is true. The price-to-rent ratio decreases in the short run and most of the effect on affordability comes from rents and not prices.

6 Validity of the Instrument

In this section we assess the validity of the instrumental variable and the robustness of the previous results. We examine at length the exclusion restriction. Section 4.1 already showed that the instrumental variable is relevant as it is strongly correlated with the investors' share of purchases. Figure $A2$ provides visual support of the strong correlation between the instrument and the share of investors' purchases over 2009-2017, while Figure $A3$ supports visually the relevance condition for the panel version of the instrument.

Our instrumental variable in the cross-section is the share of income reported as business income in 2007 by high-earner residents. In the panel and dynamic analyses, this share measures the exposure of each MSA to the national shock of the sudden drop in interest rates. Using a Bartik-like instrument is equivalent to using the local shares as instruments, hence the exclusion restriction should be interpreted in terms of the shares (Goldsmith-Pinkham, Sorkin, and Swift 2020).[19]

The identification concern for our instrumental variable is whether the exposure of the local investors to the drop of interest rates is correlated with changes in housing prices that come through channels other than property purchases. In their seminal paper, Goldsmith-Pinkham, Sorkin, and Swift (2020) set out the strategies to test for the validity of Bartik-like instruments,

[18] Ben-David, Towbin, and Weber (2019) argue that one way to identify housing booms is to look at the response of vacancies for owner-occupied and rental houses.

[19] While a typical shift-share instrument utilizes the inner product of local exposure (e.g. industry shares) and local growth rates, we can think of our instrument as using only one relevant weight - the business income share of the top tier of the income distribution - times the national shock of the drop in CD rates.

which we employ in our setting.[20]

We do the following exercises: (1) parallel pre-trends and placebo tests; (2) extensive local economy controls; (3) controls for shifts in the composition of labor demand; (4) exploration of predictors of business income and inspection of correlation with standard drivers of housing markets. Finally, we show the robustness of the results to alternative specifications and definitions of investor purchases.

6.1 Parallel pre-trends

The use of a Bartik-like instrument and the availability of pre-period trends, make our empirical strategy analogous to difference-in-differences. In a difference-in-differences setting the MSAs with the largest exposure to business income of top earners in 2007 is the treated group, and the MSAs with the smallest exposure is the control group. The year 2008 is the "treatment" year, when the Fed implemented the first wave of unconventional monetary policy, which led to a large drop in interest rates and caused an increase in investors, especially in the MSAs with higher investment attitude.

Figure 8 plots the annual log number of building permits and the annual real price growth of bottom-tier homes for MSAs ranking in the top and bottom 25% of exposure to top earners' business income in 2007. Figure 8 shows that prior to the shock there are parallel dynamics in housing construction and prices between the high and low exposure groups. The divergence starts post-2008. That is, in the period when QE does not exist and there are no incentives to have investors into housing markets, the MSAs behave similarly. We only see differences during and after the QE period when the MSAs more exposed to potential investors see those investors move to the housing market in search for yields.[21] The parallel pre-trends suggest that the instrument is driving construction and prices only in the post-crisis period. In other words, the instrument is not capturing other factors that could make housing prices to have permanently different dynamics across locations. Our empirical design satisfies the parallel pre-trends, an important assessment towards the plausibility of the exogeneity assumption

[20] "Because the shares are typically equilibrium objects and likely co-determined with the level of the outcome of interest, it can be hard to assume that the shares are uncorrelated with the levels of the outcome. But this assumption is not necessary for the empirical strategy to be valid. Instead, the strategy asks whether differential exposure to common shocks leads to differential changes in the outcome. (...) Hence, the empirical strategy can be valid even if the shares are correlated with the levels of the outcomes" (Goldsmith-Pinkham, Sorkin, and Swift 2020, p.2588).

[21] Consistent with the dynamic results, Figure 8 shows that the largest positive effect in the housing price growth due to investors happens in the first three years, while the response of construction is positive throughout the post-crisis period.

(Goldsmith-Pinkham, Sorkin, and Swift 2020).

In a similar analysis, we run a placebo test with the pre-crisis period 2000-2006 when QE was not operating and thus institutional investors were not actively seeking for yield (Figure *A*4). The scatterplots control for the same variables as specification (1). The MSAs are binned by percentiles so that each point represents around 15 MSAs. The bottom panel of the figure demonstrates strong positive correlation between the instrument and housing price growth over 2009–2017. This correlation is absent in the pre-crisis placebo sample that is in the top panel. This evidence suggests that the instrument is not contaminated by pre-crisis price growth.

To confirm the message from Figure *A*4, we conduct various placebo tests over the 2000–2006, 2001–2006, and 2000–2005 periods in Table 5.[22] We ask if, when using a specification analogous to (1), the exposure to the top earners' business income can explain housing price growth over any of these periods. The placebo point estimates are insignificant across periods. That is, the instrument is only capturing post-crisis positive shocks in housing investment. None of the factors operating pre-QE period are correlated with the instrument.

Table *A*7 contains the results of placebo tests for the panel analysis, for pre-crisis periods. Figure *A*5 plots a placebo experiment linking the instrument to prices, and Figure *A*6 to new construction. The instrument does not contribute to changes in prices or number of construction permits in time periods pre-crisis. Overall the above tests make us more comfortable that the instrument is unrelated to drivers of changes in housing markets that operate through different channels, other than the exposure to the national shock in interest rates.

6.2 Contemporaneous local economy controls

To rule out the possibility that local economic conditions drive the results, Table 6 reestimates the baseline specification controlling for a range of variables that capture contemporaneous local economic activity: average annual unemployment rate change, labor force participation growth, real GDP per capita growth, and median hourly wage per capita growth from 2009 to 2017. It is not clear that these variables are good controls, since they can be part of the transmission channel of the effect of investors. Nevertheless, Table 6 displays results very similar to Table 3. Importantly, the estimated coefficients are in close range (plus-minus 8%) of the baseline coefficient of 0.234 from Table 3. A large change in the coefficient would hint at omitted variables biasing the estimation. Our results alleviate concerns of omitted variable

[22]The selection of placebo periods is restricted by a lower bound of the year 2000, since this is when our investors' data begin. The upper bound is 2006, since we want to avoid an overlap and potential co-determination of the investors' share and our instrumental variable that is constructed using 2007 data.

bias. These results suggest that the local economic activity and the institutional investors are both important for housing price growth, but investors also affect housing markets even when keeping local economic activity constant.

6.3 Controls for spacial spillovers

While we include several controls for economic conditions, a remaining concern is that the instrument is likely to be correlated with the industrial composition of the local labor market, and therefore related to shifts in the composition of labor demand during the post-crisis period.[23] To address this concern we reestimate the baseline specification controlling for changes in employment in the largest industry sectors within the MSAs (Table $A8$). The changes are accounted for, starting from the base year of the instrumental variable, that is, from the annual change from 2007 to 2008, up to the annual change from 2016 to 2017. Employment changes in some industries, such as Real Estate, Rental and Leasing could be considered bad controls, as they might be part of the transmission channel of the investment effect on prices. Even with this prudent analysis, after controlling for employment growth of up to ten industries, the estimated effect of investors holds, and it is close to the baseline effect. Having the estimated coefficient within the range of the previous estimations in Table 6 provide extra confidence that our controls for observables capture the influential factors. It is unlikely that the instrument is correlated with any remaining unobserved drivers of housing prices.

Moreover, Table $A9$ reestimates the dynamic results accounting for the lagged annual shifts in the composition of labor demand. The dynamic patterns of housing price growth remain unchanged when we include the employment growth controls for the largest industries in the MSAs. The shifts in the composition of labor demand during the post-crisis period do not seem to be driving the results.

6.4 Unpredictable instrumental variable

Ideally, for the identification to be valid, we would have that the cross-MSA differences in the share of business income is random. The parallel pre-trends, we documented earlier, show that the share of business income was not related to the housing market dynamics before 2008. In addition to the previous result, we show that it is very difficult to predict the share of

[23]For example, Monte, Redding, and Rossi-Hansberg (2018) study the importance of spatial spillovers due to local labor demand shocks through changes in commuting patterns.

business income, or the investment/entrepreneurship attitude of an MSA, based on a variety of factors that the literature found to be linked to those attitudes. In the introduction we discuss papers showing that most of the cross-regional differences in investment attitude are as good as random. We confirm this result in Table $A10$. We regress the share of the top earners' business income in each MSA in 2007 on several factors that may explain investment or entrepreneurship activity. These factors are demographic (median age and share of immigrants), regulatory (tax rate for high earners), geographical (natural amenity index) and the ranking of MSAs in the ease of doing business. While some of these factors are correlated with the top earners' business income, their explanatory power is low. The demographic and regulatory factors explain 11% of the variation in the top earners' business income share, as we see by the R-squared of the first column of Table $A10$. Including the geographical factor the R-squared becomes 22%.

Moreover, in Table $A11$ we study whether the standard drivers of the housing market are correlated with the instrument, given our controls. We regress the local share of top earners' business income on the pre-crisis trends of homeownership and median age within each MSA. To better gauge the magnitude of these partial correlations, the table normalizes all variables to have a mean of zero and a variance of one. This allows us to assess both the magnitude and statistical significance of any correlations. Importantly, there is no relevant correlation between the common drivers of housing variables and the MSA share of top earners' business income.

Thus, the failure to predict the instrument indicates that large part of the variation in this share is random, unrelated to other drivers of housing markets. Moreover, given the expansive set of controls we include in all our specifications, the exclusion restriction for the instrument seems satisfied.

6.5 Robustness to other specifications

We check that the core results survive to changes in the specifications. For example, Figure $A7$ plots the estimated impulse responses for the top and bottom price tiers in the panel case. Consistent with the cross-sectional evidence, we find that investors have larger effects on the bottom-tier of the market.

We redo the analysis only for the single-family segment of the housing market. Table $A12$ shows that the response of prices to investors is exactly as statistically significant in the single-family segment as in the total market.[24] The lower panel of Table $A12$ shows the results of the analysis for single-unit properties, which are again as statistically significant as the results for

[24] Ninety percent of the properties in the Zillow Home Value Index are single-family and the rest are condominiums and cooperatives.

the total market.

We use additional controls in all our models, to control for total demand for housing or demand for housing by institutional investors. These controls are the total dollar value of purchases in the market or the total dollar value of purchases by investors. Controlling for either of these levels of demand does not change any of the results.[25] Our baseline controls (population, income, unemployment, MSA and year fixed effects) already capture a large part of the variation in housing demand.

Table *A*13 shows that our results are robust to using alternative measures of investors' share based on number of purchases and number of units.

7 Conclusions

The explosive growth of investors in residential housing markets after the 2008 Global Financial Crisis has been central to many affordability debates. Cities around the world are designing policies to deal with these new investors. By analyzing a large database covering the whole U.S., this paper showed that the response of price-to-income and rent-to-income ratios to the investors' purchases is positive and economically significant. Investors drove most of the recovery in housing prices, especially in low-tier housing, and housing affordability worsened. Especially affected were the single-family homes at the bottom of the price distribution. These are usually starter homes that otherwise would be purchased by young households.

The presence of investors triggered an equilibrium response of supply. One to three years after the investors' purchases, there was a substantial positive effect on new building permits, especially in the multi-unit segment. This equilibrium effect weakened the growth of price-to-income and rent-to-income ratios. After five to six years the price-to-income and rent-to-income ratio response became zero. In the medium term the investors helped to lessen the effects on worsening affordability, however they were far from reversing the substantial price and rent increases they caused.

Investors caused minimal price increases in MSAs where there are loose supply restrictions. In those areas investors affected rents more than prices and worsened rent affordability. On the other hand, the price increases caused by investors were particularly large in areas where there are strict supply restrictions, even after taking into account increases in new buildings. Thus, all together the paper suggests that the institutional investors affected differently the

[25] We do not report the tables of these results, as they are similar to the previous results.

price and rent affordability, depending on the supply restrictions of each area. Overall, our results suggest that investors had a significant role in worsening housing affordability, and even the medium-term supply response was not enough to reverse the effects.

References

Abood, M.: 2018, Wall Street landlords turn American Dream into a nightmare, pp. 1–50. ACCE Institute: Los Angeles, CA, USA.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S. and Evanoff, D.: 2019, Mitigating investor losses due to mortgage defaults: Lessons from the Global Financial Crisis. In Franklin Allen, Ester Faia, Michalis Haliassos, Katja Langenbucher, eds., Capital Markets Union and Beyond, MIT Press.

Albanesi, S., De Giorgi, G. and Nosal, J.: 2017, Credit growth and the financial crisis: A new narrative. NBER Working Paper No. 23740.

Allen, M. T., Rutherford, J., Rutherford, R. and Yavas, A.: 2018, Impact of investors in distressed housing markets, *The Journal of Real Estate Finance and Economics* **56**(4), 622–652.

Armesto, M. T. and Garriga, C.: 2009, Examining the housing crisis by home price tier, *Economic Synopses* **2009**(34), 1–2. Federal Reserve Bank of St. Louis.

Bayer, P., Mangum, K. and Roberts, J. W.: 2021, Speculative fever: Investor contagion in the housing bubble, *American Economic Review* **111**(2), 609–51.

Ben-David, I.: 2011, Financial constraints and inflated home prices during the real estate boom, *American Economic Journal: Applied Economics* **3**(3), 55–87.

Ben-David, I., Towbin, P. and Weber, S.: 2019, Expectations during the U.S. housing boom: Inferring beliefs from vacant homes. NBER Working Paper No. 25702.

Ben-Shahar, D., Gabriel, S. and Oliner, S. D.: 2020, New research on housing affordability, *Regional Science and Urban Economics* **80**, 1–4.

Berger, D., Guerrieri, V., Lorenzoni, G. and Vavra, J.: 2018, House prices and consumer spending, *The Review of Economic Studies* **85**(3), 1502–1542.

Bernstein, A., Gustafson, M. T. and Lewis, R.: 2019, Disaster on the horizon: The price effect of sea level rise, *Journal of Financial Economics* **134**(2), 253–272.

Bloomberg: 2018, The big problem with investing in Amsterdam's hot housing market. https://www.bloomberg.com/news/articles/2018-08-08/the-big-problem-with-investing-in-amsterdam-s-hot-housing-market.

Bloomberg: 2019, Spain is latest battleground for global affordable housing. https://www.bloomberg.com/news/articles/2019-06-19/spain-is-latest-battleground-in-global-affordable-housing-fight.

Bosma, N. and Kelley, D.: 2019, Global Entrepreneurship Monitor 2018/2019 Global Report. Global Entrepreneurship Research Association.

Brunson, S.: 2020, Wall Street, Main Street, your street: How investors impact the single-family housing market.

Business Insider: 2019, California becomes the third state nationwide to pass a rent control bill to address its affordable housing crisis. https://www.businessinsider.com/california-pass-rent-control-bill-to-address-affordable-housing-crisis-2019-9?IR=T.

Campbell, J. Y. and Sigalov, R.: 2020, Portfolio choice with sustainable spending: A model of reaching for yield. NBER Working Paper No. 27025.

Chinco, A. and Mayer, C.: 2016, Misinformed speculators and mispricing in the housing market, *The Review of Financial Studies* **29**(2), 486–522.

Cvijanovic, D. and Spaenjers, C.: 2020, 'We'll always have Paris': Out-of-country buyers in the housing market, *Management Science* . Forthcoming.

Daniel, K., Garlappi, L. and Xiao, K.: 2018, Monetary policy and reaching for income. NBER Working Paper No. 25344.

Davids, A. and Georg, C. P.: 2019, The cape of good homes: Exchange rate depreciations, foreign demand and house prices. SSRN Working Paper No. 3349815.

Davidsson, P.: 1991, Continued entrepreneurship: Ability, need, and opportunity as determinants of small firm growth, *Journal of Business Venturing* **6**(6), 405–429.

De Stefani, A.: 2020, House price history, biased expectations and credit cycles: The role of housing investors, *Real Estate Economics* . Forthcoming.

DeFusco, A., Ding, W., Ferreira, F. and Gyourko, J.: 2018, The role of price spillovers in the American housing boom, *Journal of Urban Economics* **108**, 72–84.

Favara, G. and Imbs, J.: 2015, Credit supply and the price of housing, *The American Economic Review* **105**(3), 958–992.

Favilukis, J. and Van Nieuwerburgh, S.: 2021, Out-of-town home buyers and city welfare, *Journal of Finance* . Forthcoming.

Ganduri, R., Xiao, S. C. and Xiao, S. W.: 2019, Tracing the source of liquidity for distressed housing markets. SSRN Working Paper No. 3324008.

Garriga, C., Gete, P. and Tsouderou, A.: 2020, Search for yield in housing markets. (Working Paper).

Garriga, C., Manuelli, R. and Peralta-Alva, A.: 2019, A macroeconomic model of price swings in the housing market, *American Economic Review* **109**(6), 2036–72.

Glaeser, E. L. and Gyourko, J.: 2018, The economic implications of housing supply, *Journal of Economic Perspectives* **32**(1), 3–30.

Goldsmith-Pinkham, P., Sorkin, I. and Swift, H.: 2020, Bartik instruments: What, when, why, and how, *American Economic Review* **110**(8), 2586–2624.

Graham, J.: 2019, House prices, investors, and credit in the Great Housing Bust.

Guren, A. M., McKay, A., Nakamura, E. and Steinsson, J.: 2020, Housing wealth effects: The long view, *The Review of Economic Studies* .

Gyourko, J., Mayer, C. and Sinai, T.: 2013, Superstar cities, *American Economic Journal: Economic Policy* **5**(4), 167–99.

Halket, J., Nesheim, L. and Oswald, F.: 2020, The housing stock, housing prices, and user costs: The roles of location, structure, and unobserved quality, *International Economic Review* **61**(4), 1777–1814.

Jordà, Ò.: 2005, Estimation and inference of impulse responses by local projections, *American Economic Review* **95**(1), 161–182.

Kleibergen, F. and R. Paap: 2006, Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* pp. 97–126.

Lambie-Hanson, Lauren, W. L. and Slonkosky, M.: 2019, Leaving households behind: Institutional investors and the U.S. housing recovery. Working Paper 19-1. Federal Reserve Bank of Philadelphia, PA, USA.

Martinez-Miera, D. and Repullo, R.: 2017, Search for yield, *Econometrica* **85**(2), 351–378.

Mian, A. and Sufi, A.: 2014, What explains the 2007–2009 drop in employment?, *Econometrica* **82**(6), 2197–2223.

Mian, A., Sufi, A. and Verner, E.: 2017, Household debt and business cycles worldwide, *The Quarterly Journal of Economics* **132**(4), 1755–1817.

Mills, J., Molloy, R. and Zarutskie, R.: 2019, Large-scale buy-to-rent investors in the single-family housing market: The emergence of a new asset class, *Real Estate Economics* **47**(2), 399–430.

Molloy, R., Nathanson, C. and Paciorek, A.: 2020, Housing supply and affordability: Evidence from rents, housing consumption and household location. FEDS Working Paper No. 2020-044.

Monte, F., Redding, S. J. and Rossi-Hansberg, E.: 2018, Commuting, migration, and local employment elasticities, *American Economic Review* **108**(12), 3855–90.

Piazzesi, M., Schneider, M. and Stroebel, J.: 2020, Segmented housing search, *American Economic Review* **110**(3), 720–59.

Rocha, H. O. and Sternberg, R.: 2005, Entrepreneurship: The role of clusters theoretical perspectives and empirical evidence from Germany, *Small Business Economics* **24**(3), 267–292.

Rodnyansky, A. and Darmouni, O. M.: 2017, The effects of quantitative easing on bank lending behavior, *The Review of Financial Studies* **30**(11), 3858–3887.

Saiz, A.: 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* **125**(3), 1253–1296.

Stroebel, J.: 2016, Asymmetric information about collateral values, *The Journal of Finance* **71**(3), 1071–1112.

The Wall Street Journal: 2017, Meet your new landlord: Wall Street. https://www.wsj.com/articles/meet-your-new-landlord-wall-street-1500647417.

The Wall Street Journal: 2019, In Berlin, a radical proposal to combat rising rents: Expropriate big landlords. https://www.wsj.com/articles/in-berlin-a-radical-proposal-to-combat-rising-rents-expropriate-big-landlords-11554202800.

Wu, J., Xiao, S. C. and Xiao, S. W.: 2020, Do Wall Street landlords undermine renters' welfare? SSRN Working Paper No. 3410281.

Zillow: 2017, "ZTRAX: Zillow Transaction and Assessor Dataset, 2017-Q4". Zillow Group, Inc. http://www.zillow.com/ztrax/.

Figures

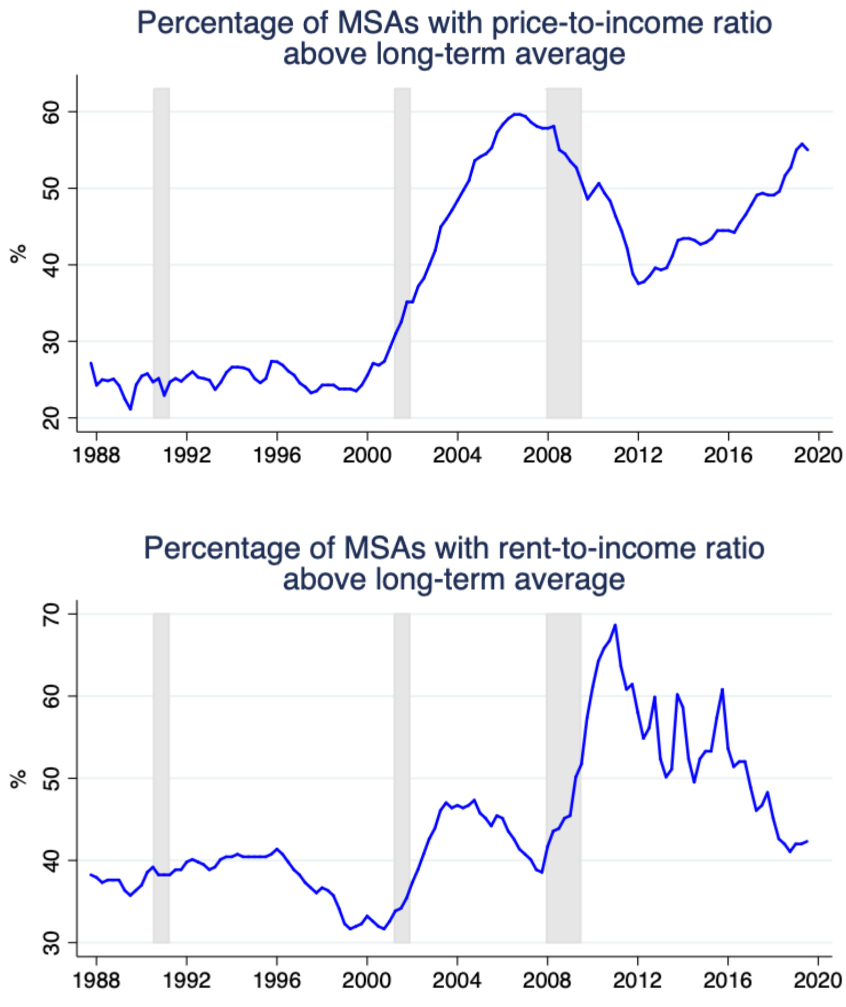


Figure 1. Housing affordability in the U.S. relative to long-term average. The top figure plots the share of MSAs where price-to-income ratio is above three, which has been the national average in the U.S. over 1987-2019. That is, the share of MSAs where the median housing price is higher than three times the median annual household income. The bottom figure plots the share of MSAs where rent-to-income ratio is above 0.27, the national average over 1987-2019. That is, the share of MSAs where the median annual housing rent is higher than 27% of the median annual household income. The gray areas illustrate the U.S. Recessions. The price-to-income and rent-to-income ratios come from Zillow.

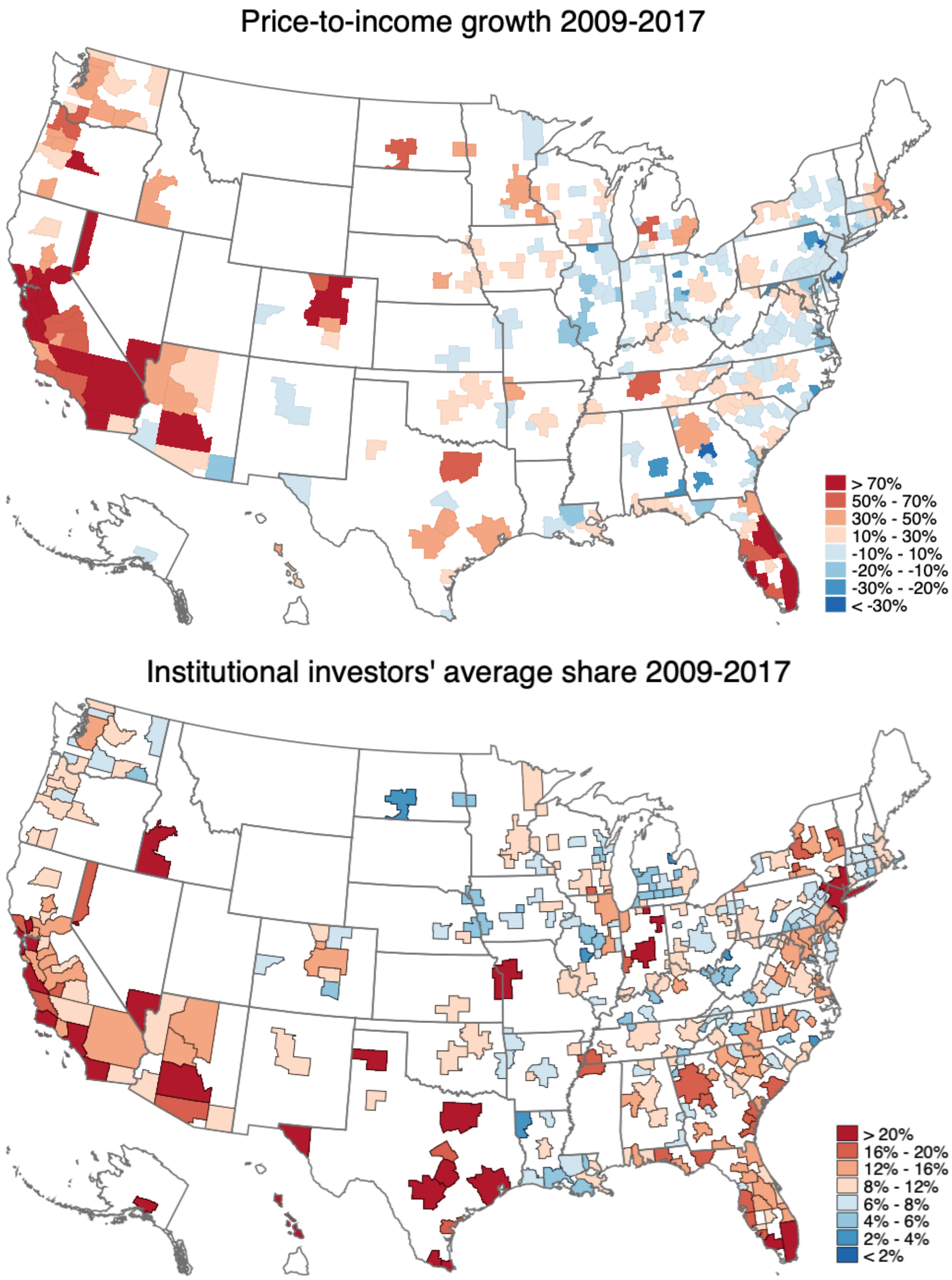


Figure 2. Affordability and institutional investors in the U.S. The top map shows the percentage growth of price-to-income ratio from 2009 to 2017 in each MSA for bottom-tier houses. The bottom map shows the average market share of dollar purchases by institutional investors from 2009 to 2017 in each MSA. Figure A1 shows the correlation of the raw data in a scatter plot.

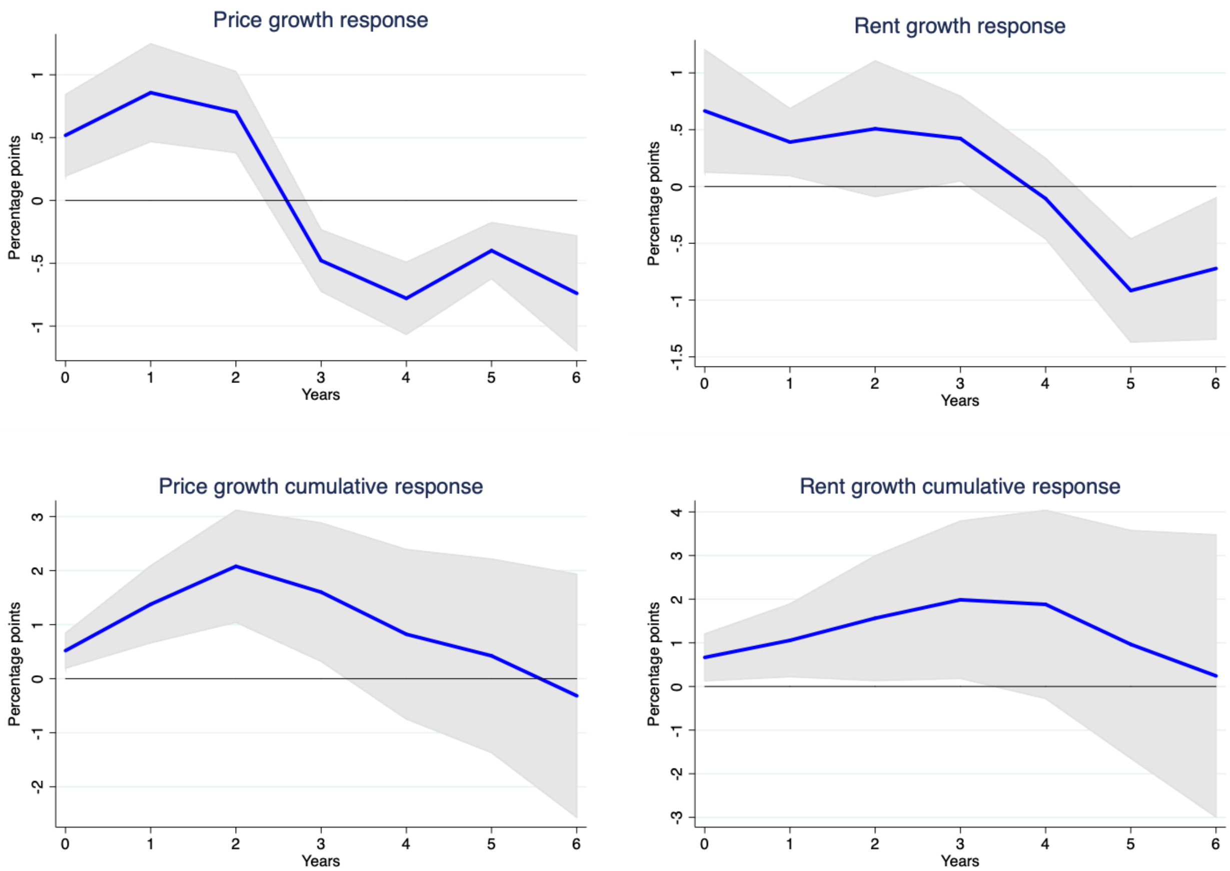


Figure 3. Dynamics of housing prices and rents after investors' purchases. The top figures plot the estimates from sequential regressions of the price growth and rent growth on the instrumented past investors' share of purchases. The bottom figures plot the cumulative effects, calculated as the cumulative sum of the previous coefficients. Prices and rents are adjusted for inflation. Section 5 contains the methodology that follows Jordà (2005). We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.

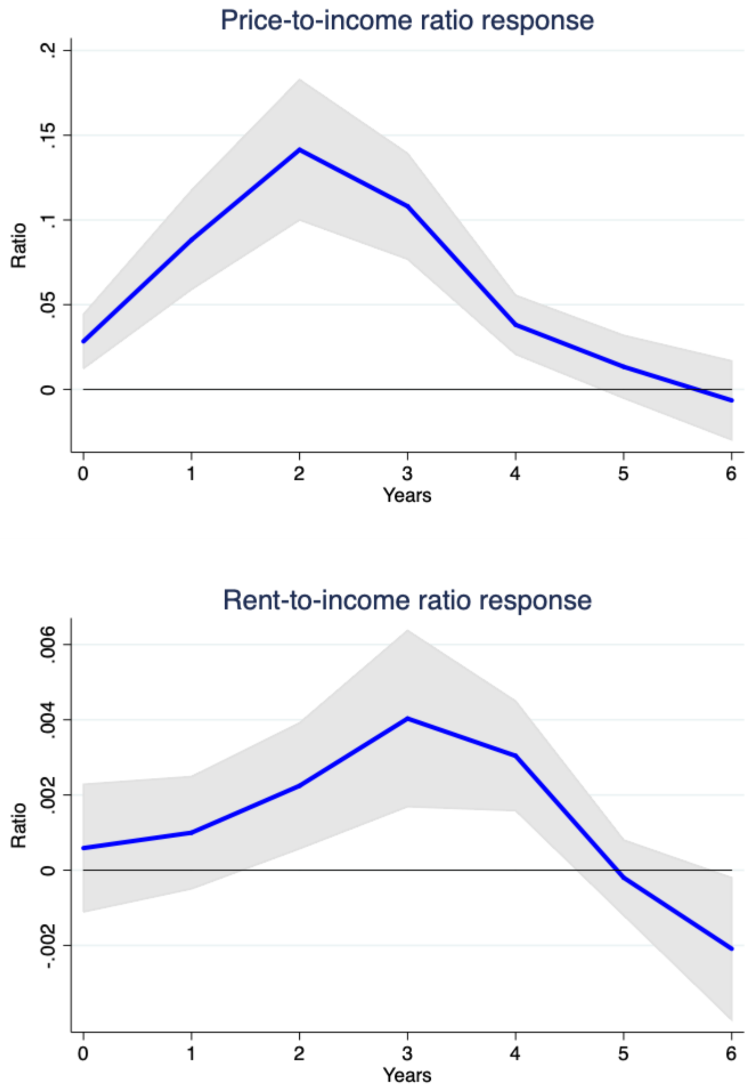


Figure 4. Dynamics of housing affordability after investors' purchases. The figures plot the estimates from sequential regressions of (a) the price-to-income ratio, and (b) the rent-to-income ratio on the instrumented past investors' share of purchases. The price-to-income ratio is the median housing price over the median annual household income in an MSA. The rent-to-income ratio is the median annual housing rent over the median annual household income in an MSA. We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.

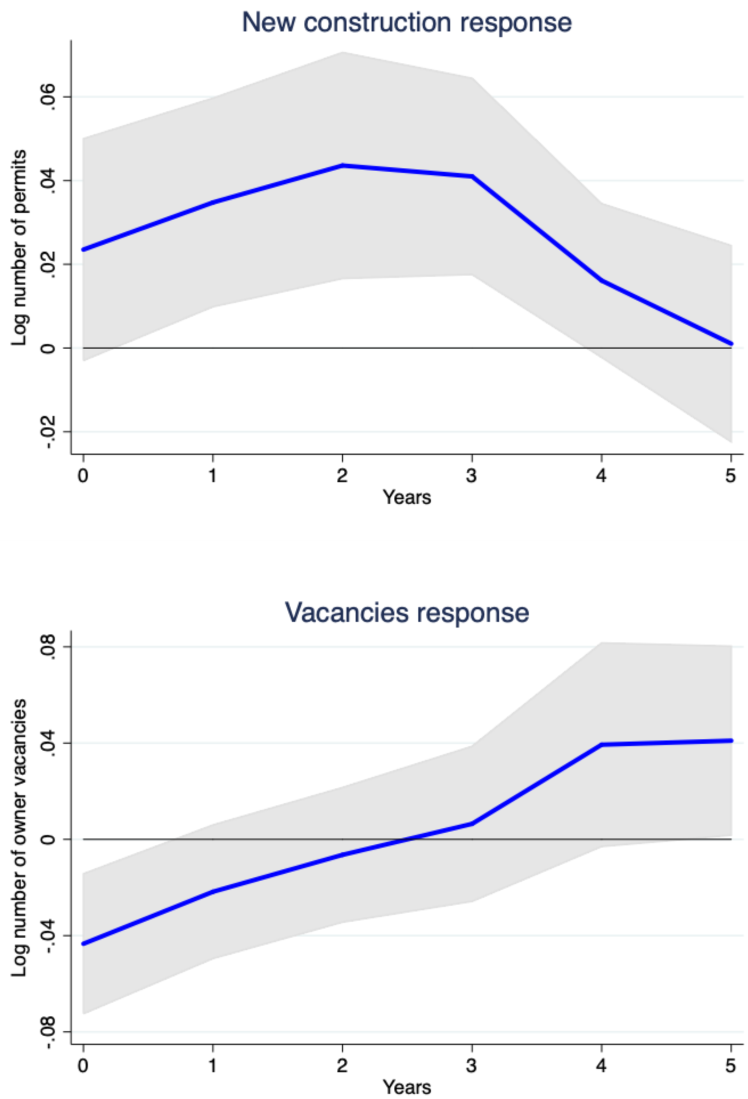


Figure 5. Dynamics of housing supply after investors' purchases. The figure plots the estimates from sequential regressions of (a) the log number of building permits, and (b) the log number of homeowner vacant units on the instrumented past investors' share of purchases. We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.

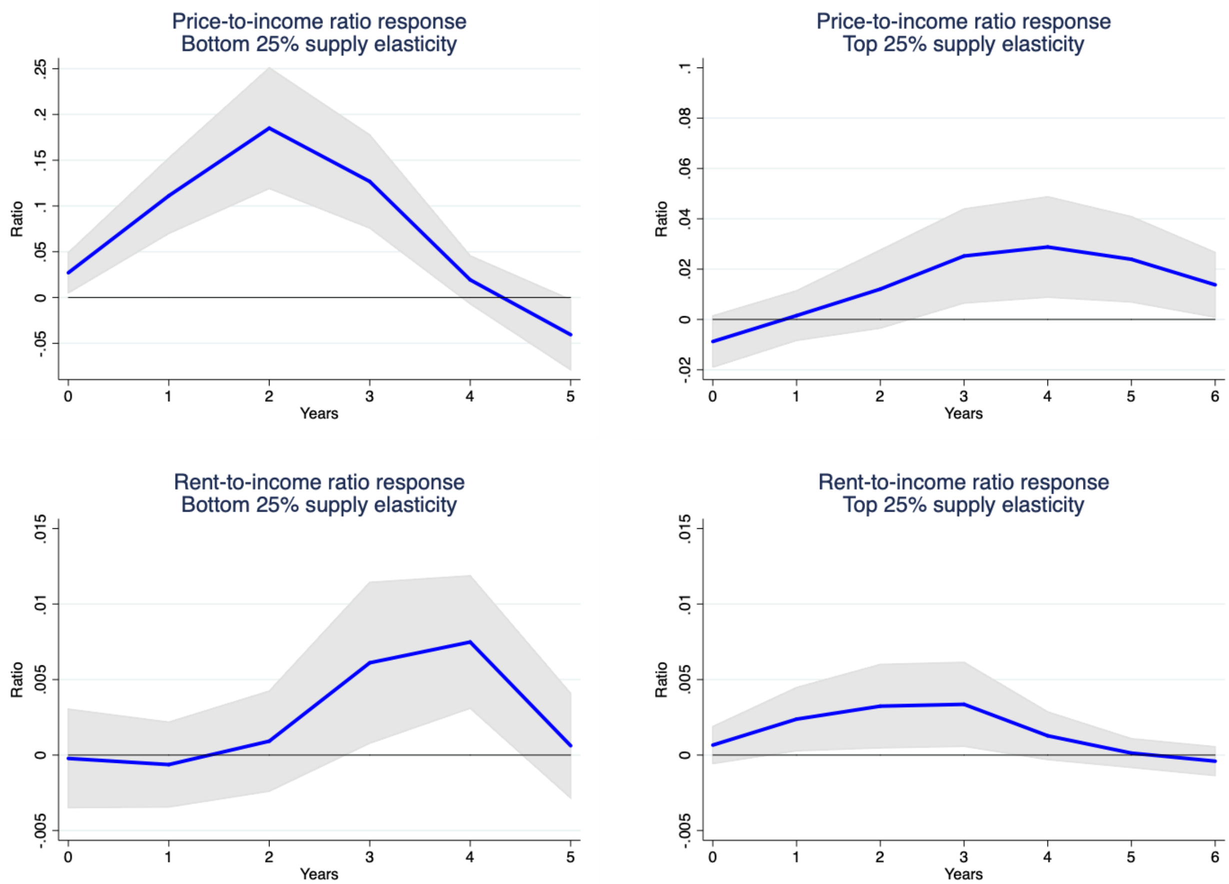


Figure 6. Dynamics of housing affordability after investors' purchases and supply elasticity. The figure plots the estimates from sequential regressions of (a) the price-to-income ratio, and (b) the rent-to-income ratio for MSAs at the bottom and top quartiles of the supply elasticity distribution. The housing supply elasticity comes from Saiz (2010). The bottom quartile of the supply elasticity is 1.56, and the top quartile is 2.89 in our sample. The bottom quartile has an average price-to-income ratio of 4.2 and rent-to-income ratio of 0.33, over 2009-2017. The top quartile has an average price-to-income ratio of 2.5 and rent-to-income ratio of 0.27, over the same period. We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.

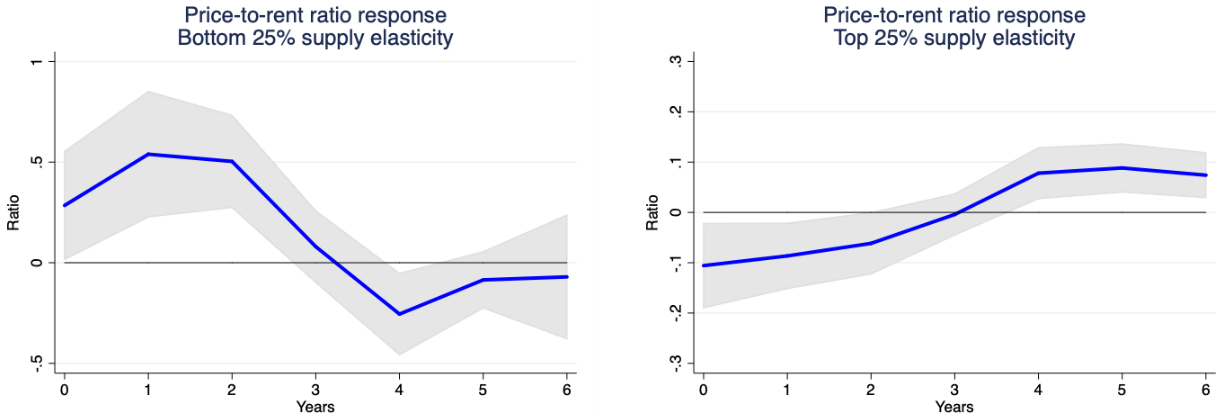
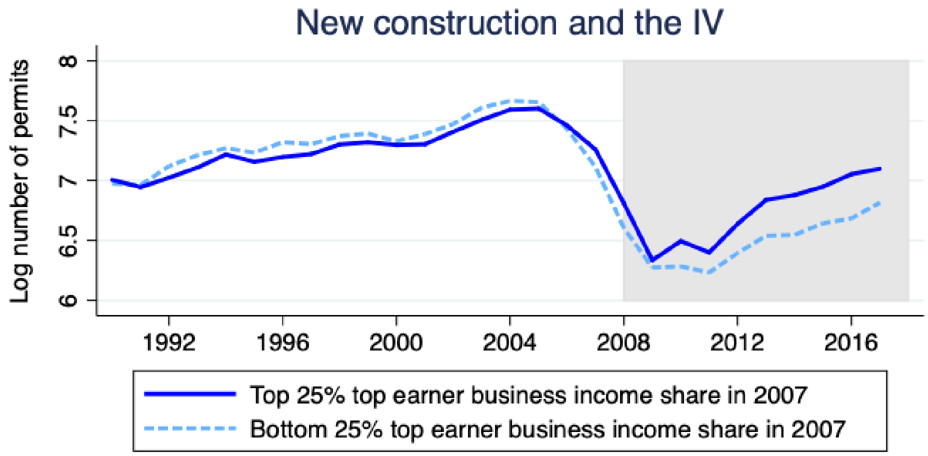


Figure 7. Dynamics of price-to-rent ratio after investors' purchases and supply elasticity. The left figure plots the estimates from sequential regressions of the price-to-rent ratio on the instrumented past investors' share of purchases, for MSAs in the bottom 25th percentile of supply elasticity. The right figure redoes the previous plot but for MSAs in the top 25th percentile of supply elasticity. The bottom quartile value of supply elasticity is 1.56, and the top quartile value is 2.89. The bottom supply elasticity quartile has an average price-to-rent ratio of 13.0, whereas the top quartile has an average price-to-rent ratio of 9.4, over 2009-2017. We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.



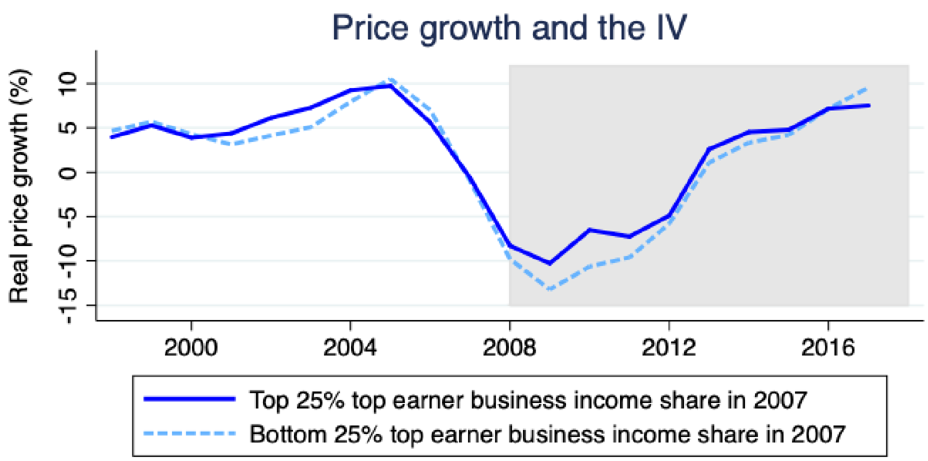


Figure 8. Parallel trends. The top figure plots the time series of the log number of new building permits, and the bottom figure the time series of the bottom-tier real price growth for MSAs in our sample ranking in the top and bottom 25% of exposure to our instrumental variable: the 2007 top earners' average share of business income over total income in an MSA. The gray shaded area shows the period from the introduction of the Fed's QE in 2008 onwards.

Tables

Table 1. Summary statistics

Panel A - MSA level

	Obs	Mean	SD	Min	Max
Investors' share of purchases (%)	332	12.37	7.78	3.10	41.26
Top tier price growth (%)	328	0.43	1.55	-4.26	6.45
Mid-tier price growth (%)	332	0.47	1.77	-5.15	5.96
Bottom tier price growth (%)	296	0.17	2.52	-8.97	7.04
Top tier price-to-income ratio	328	3.09	1.22	1.40	9.50
Mid tier price-to-income ratio	332	4.76	2.40	1.50	16.98
Bottom tier price-to-income ratio	296	8.48	5.48	1.05	38.68
Log number of building permits all properties	332	6.49	1.27	2.33	10.33
Log number of building permits single-unit	332	6.44	1.28	2.24	10.31
Log number of building permits 2–4 units	330	2.25	1.23	0	6.73
Log number of building permits 5+ units	327	2.41	1.27	0	6.43
Top earner business income share$_{2007}$(%)	332	2.77	0.94	1.03	9.09

Panel B - Panel data

Investors' share of purchases (%)	2,997	11.50	8.40	0.65	75.95
Top tier price growth (%)	2,853	0.46	5.61	-24.92	28.41
Mid tier price growth (%)	2,901	0.47	6.67	-25.51	36.47
Bottom tier price growth (%)	2,610	0.13	9.87	-53.03	34.09
Rent growth (%)	2,583	0.52	6.12	-35.07	49.65
Price-to-income ratio of median household	2,849	3.24	1.27	1.12	9.97
Rent-to-income ratio of median household	2,583	0.29	0.05	0.14	0.61
Log number of building permits all properties	2,997	6.46	1.36	1.10	10.58
Log number of homeowner vacancies	2,554	7.57	1.13	3.14	10.96
Lagged population growth (%)	2,994	0.71	0.90	-4.45	7.99
Lagged median household income growth (%)	2,853	1.41	2.61	-7.98	11.01
Lagged unemployment rate change (%)	2,997	0.04	1.56	-4.54	9.29
Top earner business income share$_{07}$(%)$\times$CD rate growth$_{t-1}$	2,997	-0.57	0.76	-4.98	1.58

The top panel presents summary statistics of the key variables at MSA level, and the bottom panel at MSA-year level, in 2009-2017. Prices and rents are inflation adjusted to reflect 2012 dollars. Detailed description of the variables and data sources is in Appendix A.

Table 2. First stage: Investors' share and the instrumental variable

	Investors' share of purchases$_{m,09-17}$
Top earner business income share$_{m,07}$	1.441***
	(0.327)
MSA-level controls	Yes
State dummies	Yes
R-squared	0.689
Observations	332

Heteroskedasticity robust standard errors are in parentheses. The controls are the population growth, income growth, unemployment rate change and real housing price growth over the periods 2000-2006 and 2006-2007, and the log number of building permits in 2007. Each observation is an MSA. ***significant at the 1% level.

Table 3. Housing price growth and affordability by price tier

	Price growth$_{m,09-17}$			
	Mid Tier		Bottom Tier	Top Tier
Investors' share$_{m,09-17}$	0.034**	0.243***	0.302***	0.180**
	(0.015)	(0.083)	(0.100)	(0.070)
Estimation	OLS	IV	IV	IV
1st stage F-test excluded instruments		19.430	19.457	19.969
Underidentification test p-value		0.000	0.000	0.000
Observations	332	332	296	328
			Standardized	
Investors' share$_{m,09-17}$		0.827***	0.909***	0.768**
		(0.313)	(0.302)	(0.338)
Observations		293	293	293
		Price-to-income ratio$_{m,09-17}$		
Investors' share$_{m,09-17}$	0.108***	0.539***	1.538***	0.304***
	(0.021)	(0.161)	(0.375)	(0.096)
Estimation	OLS	IV	IV	IV
Observations	332	332	296	328
			Standardized	
Investors' share$_{m,09-17}$		1.679***	2.108***	1.884***
		(0.490)	(0.509)	(0.585)
Observations		293	293	293

Heteroskedasticity robust standard errors are in parentheses. Prices are inflation adjusted. Bottom tier refers to the 17th percentile, and top tier to the 83rd percentile of the housing prices and individual income in each MSA. The standardized results show the estimated effects of the standardized independent variable on the standardized dependent variables, for the sample of MSAs for which we have price series for all price tiers. All models include state dummies and MSA-level controls: population growth, income growth, unemployment rate change and real housing price growth over the periods 2000-2006 and 2006-2007, and the log number of building permits in 2007. Table 2 contains the first stage of the IV regression. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in MSA m in the year 2007. The weak identification F statistic is the Kleibergen and Paap Wald F statistic. The underidentification test is from Kleibergen and Paap (2006). Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level.

Table 4. Housing construction by property type

	Log number of permits$_{m,09-17}$			
	All	Single-unit	2-4 units	5+ units
Investors' share$_{m,09-17}$	0.051***	0.047**	0.107*	0.152***
	(0.020)	(0.019)	(0.055)	(0.046)
Estimation	IV	IV	IV	IV
1st stage F-test excluded instruments	19.430	19.430	19.707	19.453
Underidentification test p-value	0.000	0.000	0.000	0.000
Observations	332	332	330	327

Heteroskedasticity robust standard errors are in parentheses. Single-unit refers to permits for the construction of single-unit properties, 2-4 units refers to permits for the construction of properties that have between 2 and 4 units, and 5+ units refers to permits for the construction of properties of 5 units or more. All models include state dummies, MSA-level controls and the instrumental variable as in Table 3. The weak identification F statistic is the Kleibergen and Paap Wald F statistic. The underidentification test is from Kleibergen and Paap (2006). Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table 5. Placebo: Housing price growth and investors' share pre-crisis

	Price growth$_{m,[t_1,t_2]}$		
$[t_1, t_2]$	2000-2006	2001-2006	2000-2005
Investors' share of purchases$_{m,[t_1,t_2]}$	0.027	0.870	-0.036
	(0.807)	(1.680)	(2.238)
Estimation	IV	IV	IV
MSA-level controls	Yes	Yes	Yes
State dummies	Yes	Yes	Yes
Observations	307	303	306

Heteroskedasticity robust standard errors are in parentheses. Prices are inflation adjusted. The controls are the population growth, income growth, unemployment rate change and real housing price growth over the periods 1991-1997 and 1997-1998, and the log number of construction unit permits in 1998. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in MSA m in the year 2007. Each observation is an MSA.

Table 6. Estimation including additional local economic drivers

	Price growth$_{m,09-17}$			
Investors' share of purchases$_{m,09-17}$	0.259***	0.252***	0.223***	0.247***
	(0.084)	(0.085)	(0.083)	(0.084)
Unemployment rate change$_{m,09-17}$	-3.072**			
	(1.268)			
Labor force participation growth$_{m,09-17}$		-0.006		
		(0.208)		
Real per cap. GDP growth$_{m,09-17}$			0.204	
			(0.126)	
Per cap. wage growth$_{m,09-17}$				0.001
				(0.189)
First stage F-test of excluded instruments	20.747	21.058	19.276	21.231
Underidentification test p-value	0.000	0.000	0.000	0.000
Observations	332	331	332	332

Heteroskedasticity robust standard errors are in parentheses. Unemployment rate change$_{m,09-17}$ denotes the average unemployment rate change in MSA m over 2009-2017. Labor force participation growth$_{m,09-17}$, real per capita GDP growth$_{m,09-17}$ and per capita wage growth$_{m,09-17}$ denote the average annual growth rate of those variables in MSA m over 2009-2017. Prices are inflation adjusted. The specifications include MSA-level controls, state dummies and the instrumental variable as in Table 3. The underidentification test is that of Kleibergen and Paap (2006) and the F statistic is the Kleibergen and Paap Wald F statistic. Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level.

Online Appendix (NOT FOR PUBLICATION)

A Detailed Description of Database

In this appendix we describe our data sources, how we cleaned the data, and the key variables used in our analysis.

Investors' purchases

The investors' data come from the Zillow Transaction and Assessment Dataset (ZTRAX), a large raw database of U.S. deeds data. The transactions database of ZTRAX contains all property ownership transfers that are documented in the County deeds. Each record contains the date of the transfer, the address of the property, the type of the property, the sale price, and the names of the buyer and seller. We keep transactions between January 1^{st}, 2000 and December 31^{st}, 2017. We restrict the data to ownership transfers, dropping observations that refer exclusively to mortgages or foreclosures.[26] We drop transactions with deed type "Life Estate", since this is not an immediate transfer of ownership. We also drop transactions that had "Cancellation" in the deed type. We restrict the data to residential property transfers based on the ZTRAX property land use standard codes, which include both single-family and multi-family properties. Table A14 contains the classification of the property land use standard codes in single-family and multi-family from ZTRAX. This amounts to 139 million transactions nationally. We then drop transactions with purchase price missing or smaller than $10,000, a common practice with deeds data (Bernstein, Gustafson and Lewis 2019; Stroebel 2016). This leaves 85 million transactions.

With the previous cleaning criterion, most of the transactions are dropped in the non-disclosure states. These states or counties do not require that the sale price is submitted to the county office. Specifically, all transactions are dropped in five non-disclosure states: Mississippi, Missouri, Montana, Utah and Wyoming. We keep in our data seven non-disclosure states, with a total of 28 MSAs, in which some of the transactions record sales price. We drop from our final dataset MSA-years that have fewer than 200 transactions, to avoid large outlier values, due to very few observations. The final dataset contains the following MSAs in non-disclosure states: Anchorage, Alaska; Boise City, Idaho; Alexandria, Baton Rouge, Hammond, Houma-Thibodaux, Lafayette, Lake Charles, Monroe, New Orleans-Metairie and Shreveport-Bossier City, Louisiana; Kansas City and Wichita, Kansas; Albuquerque, New Mexico; Bis-

[26] The mortgage and foreclosure deeds have a separate corresponding deed for the ownership transfer.

marck and Fargo, North Dakota; Amarillo, Austin-Round Rock, Brownsville-Harlingen, Corpus Christi, Dallas-Plano-Irving, El Paso, Fort Worth-Arlington, Houston-The Woodlands-Sugar Land, Killeen-Temple, Lubbock, McAllen-Edinburg-Mission and San Antonio-New Braunfels, Texas. Additional results, not reported here, contain our baseline cross-sectional and dynamic analyses, dropping completely all non-disclosure MSAs. The results of both analyses hold with the same significance and even stronger results for the relevance tests for our instrumental variable.

To identify *institutional investors*, we first use the ZTRAX classification of buyer names into individual and non-individual names. The non-individual names frequently end with the words "LLC", "LP", "INC", "TRUST", "CORPORATION", "PARTNERS", but they also contain entity names without the description in the end of the name.[27] Thorough inspection of the data confirms that the classification by ZTRAX of individual and non-individual names is as expected, with very minimal (human) errors. Our institutional investors' identifier contains the deeds where the buyer has a non-individual name. From these names we filter out names of relocation companies, non profit organizations, construction companies, national and regional authorities, banks, Ginnie Mae, Fannie Mae, Freddie Mac and other mortgage loan companies and credit unions, homeowner associations, hospitals, universities (not when is university housing), churches, airports, and the state, names of the county, city and municipality. To identify relocation companies, non profit organizations and construction companies we use public data of lists of the top relocation companies, non profit organizations and construction companies in the U.S. We also manually check the names of the 200 largest non-individual buyers in each state using online search engines to classify them in the right category, and iterate this procedure several times to ensure the largest buyers are correctly classified.

To further increase the accuracy of the largest institutional investors' classification we collect from industry reports and news reports the names of the top 20 institutional investors in the single-family rental market. For example Amherst Capital's 2018 market commentary report[28] provides a comprehensive list of the top 20 single-family rental institutions and the number of homes owned based on their calculations. We also collect the names of the residential real estate companies that belong to the S&P 500 Real Estate Index, most of which are apartment REITs. We then search for the names of these top investors and their subsidiaries in the ZTRAX database and ensure they are classified as institutional investors. We use public SEC filings and other business websites to track down the names of the subsidiaries of these large

[27] For example "Invitation Homes" and "Invitation Homes LP" are both included as non-individual names.

[28] Amherst Capital report is retrieved from https://www.amherstcapital.com/documents/20649/22737/Amherst+Capital+Market+Commentary+-+April+2018+vF/f06bd51a-44c7-4f8f-87e3-ca8d795bf42a Last visited: 03-05-2019.

investors. This procedure results in calculating the exact holdings of the top single-family and multi-family institutional investors.

We calculate the market share of investors as the *dollar value of institutional investors' purchases* divided by the dollar value of all purchases. Using the dollar value, accounts correctly for purchases of buildings with multiple units.

Alternatively, we use the *number of units*, instead of the dollar value. The number of units is coded by ZTRAX, in the tax assessment dataset, which we merge with the transactions dataset, using the RowID unique identifier. We use the property type code (PropertyLandUseStndCode) to fill in the missing number of units. Specifically, we fill in number of units 2 if number of units is missing and the property type is duplex or multifamily dwelling (generic any combination 2+). We fill in number of units 3 for triplex, 4 for quadruplex, and 5 for apartment building (5+ units) or court apartment (5+ units). We fill in number of units 100 for apartment building (100+ units). With this criterion, when the number of units is missing we assign the lower bound of the number of units to the property, inferred by the qualitative description. For the rest of the multi-family property types and all the types we classify as single-family in Table A14 that do not specify number of units, we assign 1 unit. We double-check with the sales price and confirm that these refer to single-unit purchases.

Finally, we use the crosswalk file from Census Bureau to match the County FIPS codes in ZTRAX to the Census Bureau MSA's 2017 core based statistical area (CBSA) code. For submetro areas of the largest MSAs, we use the CBSA division code. In total we match 411 CBSAs and divisions in the data.

Housing prices, rents and supply elasticity

Our price and rent data at MSA-level from 1999 through 2017 come from Zillow. To measure housing prices, we use the Metro Zillow Home Value Index (ZHVI). The ZHVI measures the median monthly price for each MSA and has units of nominal dollars per month. Zillow imputes this price based on a proprietary machine learning model taking into account the specific characteristics of each home and recent sale listings for homes with similar characteristics. The median price is computed across all homes in an MSA, not only those that are currently for sale. Thus, unlike pure repeat-listing indices, the ZHVI is not biased by the current composition of for-sale properties. To measure housing prices specifically for single-family homes, we use the ZHVI Single-Family Homes Time Series. To measure the price of top tier and bottom tier homes we use the Zillow's Top Tier Index and Bottom Tier Index, which measure the median house price among homes in the top third and bottom third of the price distribution within an MSA respectively. To measure rents, we use the Metro Zillow Rent Index (ZRI). The ZRI measures the median quarterly rent for each MSA and has units of nominal dollars per month. Zillow imputes this rent using an analogous methodology to ZHVI. Importantly, the ZRI does not impute a property's rent from its price. To convert the prices and rents to annual, we take the last value of each year. Housing price growth is the percentage growth of housing prices from year $t-1$ to year t. Housing rent growth is the percentage growth of housing rents from year $t-1$ to year t.

The housing supply elasticities are originally estimated by Saiz (2010). The elasticities are based on the amount of developable land in the U.S. MSAs, which is calculated based on satellite-generated geographical data. We use the dataset provided by Favara and Imbs (2015) as our source of elasticity data.[29] The original data are at the MSA level (CBSA 2003 codes), and cover 275 MSAs. We crosswalk these to our 2017 CBSA and CBSA division codes.

Construction and vacancy data

Data on construction permits come from the Census Bureau's annual Residential Building Permits Survey. Statistics on construction authorized by building permits are based upon reports submitted by local building permit officials in response to a mail survey. When a report is not received, missing residential data are either obtained from the Survey of Use of Permits (SUP) or imputed. The SUP is used to collect information on housing starts. All other missing data are imputed. The imputations are based on the assumption that the ratio of current year authorizations to those of a year ago should be the same for both respondents and nonrespondents.

Our construction data cover the years 2000 to 2017 and they are collected initially at the county level. We then use the crosswalk file from Census Bureau to match the County FIPS codes to the Census Bureau 2017 core based statistical area (CBSA) and CBSA division codes. Then we aggregate the number of construction permits at the CBSA level. The permits are split into 1-unit, 2-units, 3-4 units and 5+ units, and they count the number of new buildings authorized. For our main construction variable we add up all the permits together, since our analysis includes the total housing market. The MSA-level data cover all the 411 CBSA codes.

Vacancy data come from the American Community Survey One-Year Estimates. Data are available annually and they cover 311 MSAs over the 2005-2017 period. We start from the original data at the county level: number of vacant housing units for homeowners and number

[29] The AER site from which we obtained the data is: https://www.aeaweb.org/articles?id=10.1257/aer.20121416, and the specific dataset is "hp_dereg_controls".

of total units for homeowners. We then crosswalk to the 2017 CBSA codes and CBSA division codes and sum the number of households in the counties within the MSAs. Starting from county-level data results in more accurate MSA values for the most recent CBSA codes. Owner vacancy rate is the share of the number of vacant housing units for homeowners over the total housing units for homeowners.

Tax report data

The main data source to construct our instruments comes from the Internal Revenue Services (IRS), in particular, the Statistics of Income (SOI). This dataset provides zip code data on administrative records of individual tax returns. The data excludes zip codes with less than 100 returns. Detailed description of the instruments is included in Section B.

Control variables

We also rely on the following data sources to get data at the county-year level and then aggregate to MSA-year level using the 2017 CBSA and CBSA division codes:

- Population: U.S. Census Bureau, from 1990 to 2017.
- Median Income: Zillow Median Household Income dataset, from 1990 to 2017.
- Unemployment and labor force participation: Bureau of Labor Statistics, from 1990 to 2017.
- Median age: American Community Survey One-Year Estimates, Census Bureau. The data only cover the 2005-2017 period. The data come in discrete age intervals that are 5 years apart. Based on the number of people in each age interval we find the interval that contains the median age, and take as the median age the mid-point of this interval.
- Employment by industry: County Business Patterns (CBP) dataset, from 2007 to 2017.
- Gross Domestic Product and wages: U.S. Department of Commerce's Bureau of Economic Analysis (BEA), from 2008 to 2017.
- Natural Amenities Scale: U.S. Department of Agriculture. The scale is constructed by combining six measures of climate, topography, and water area that reflect environmental qualities. These measures are warm winter, winter sun, temperate summer, low summer humidity, topographic variation, and water area.

These additional controls come from the following data sources:

- Migration: American Community Survey 5-year estimates, Census Bureau, MSA-level in 2007.
- Income tax rate: Tax Foundation, the top marginal tax rate for an individual, State-level in 2007.
- Entrepreneurship rank: CNBC America's top states for business in 2007. This index provides a ranking of 50 States based on 40 different measures of competitiveness from publicly available data.

To summarize, there are 332 MSAs with the full set of average housing variables and investors' market share for the years 2009-2017, control variables beginning in 2000, and tax-returns for the year 2007.

B Detailed Description of the Instrumental Variable

Our instrument approximates the average individual's tax returns by the zip code returns of a specific adjusted gross income (AGI) group. Since the Statistics of Income (SOI) dataset from the IRS does not provide returns at the individual level, the zip code AGI group level is the closest approximation to the average individual of each group within the zip code. AGI is defined as the total income minus adjustments to the income, that might be subject to change each year. The dataset splits the returns into six income groups. We specifically focus on the returns of the top two high earnings groups, which include people with annual AGI above $100,000.

Our instrument is the share of business income which measures the local attitude towards investment. Next, we describe in detail how we construct this instrument.

Share of business income

The share of business income instrument is concerned with the component of earnings associated to net business income. With the implementation of the QE housing becomes an attractive investment. High earners with high business income in each MSA are likely to be more knowledgeable about investments. They are more likely to pursue investments in general, and investments in residential real estate in particular.

To construct the instrument we calculate the average share of net business income of top earners in 2007 at zip code level as:

$$b_{z,2007} = \sum_{g=5}^{6} \mu_g \frac{\text{Net business income } (\$)_g}{\text{Adjusted gross income } (\$)_g},$$

where z denotes the zip code and $g \in \{5, 6\}$, denotes the AGI group. Group 5 consists of returns with AGI between \$100,000 and \$200,000, and group 6 consists of returns with AGI above \$200,000. The weight μ_g weights by the number of returns of each group. $\mu_g = N_g/(N_5 + N_6)$, where N represents the number of returns. All values refer to the 2007 returns.

We calculate the average share of business income of top earners in 2007 at the MSA level as:

$$b_{m,2007} = \sum_{z \in m} \omega_z k_z b_z,$$

where m denotes the MSA. k_z is the share of the zip code population that belongs to the MSA. This share comes from the Department of Housing and Urban Development (HUD) zip-CBSA and zip-CBSA division crosswalk files. k_z is one for most of the zip codes. ω_z weights by the number of returns of each zip code within the MSA: $\mu_z = N_z / \sum_{z \in m} N_z$.Our instrument $b_{m,2007}$ is used in the cross-sectional regression (1) to instrument for the average share of institutional investors in MSA m, using a 2-stage least square estimation methodology.

For our dynamic analysis that uses a panel specification, we use the panel version of the instrument. The time-varying instrument captures the exposure of an MSA to the QE over time. We construct the time-varying instrument as follows:

$$b^p_{m,t} = b_{m,2007} \times CD_{t-1},$$

where CD_t is the growth in the one-year certificate of deposits rate from year $t-1$ to t. In our panel data t ranges from 2009 to 2017. The investors' share is used with one year lag in the panel specification (2) .

Having the business income share fixed in 2007, ensures that the exposure to the QE is predetermined, and not affected by the housing market variables post 2008. Figure A8 plots the time series of an average one-year CD rate. CD_t is a national shock that is also unrelated to each of the local housing markets. This methodology constructs instruments that are likely to satisfy the exclusion restriction. Our multiple tests in Section 6 provide strong evidence in this direction.

C Theory: Housing Purchases, Prices and Affordability

Traditionally the literature on affordability has emphasized how supply restrictions increase house prices above the fundamental value. This appendix presents a stylized macro-housing model, based on Garriga, Manuelli, and Peralta-Alva (2019), that allows for a sharp characterization on how housing purchases by different types of buyers influence house prices and affordability in the presence of frictions in the credit market.

C.1 Economic Environment

The economy has a population of heterogeneous home buyers/investors denoted by $i = 1, 2, ..., I$. The size of each group of potential buyers is represented by N_i where $\sum_{i=1}^{I} N_i = \overline{N}$. Each individual type i has preferences defined over non-housing consumption (numeraire) and housing services, $\sum_{t=0}^{\infty} \beta^t [\ln c_{it} + \gamma_i \ln h_{it}]$, where the discount rate is the same for all agents, $\beta \in (0, 1)$, The intensity of the housing demand, captured by the parameter $\gamma_i \geq 0$, can vary in the cross-section of buyers. For tractability, it is convenient to consider the case where income levels by type are fixed over time $y_{it} = y_i$ for all t.

To illustrate the effects of credit frictions on affordability, it is sufficient to assume that all buyer types have access to the same credit conditions. House purchases are financed via a mortgage subject to loan-to-value requirements denoted by $\phi_{it} = \phi_t \in [0, 1]$. Since mortgage loans are long-term contracts, it is necessary to differentiate the stock and the flow of credit. Let B_{it} denote the stock of collateralized debt at the beginning of period t, b_{it+1} is the new loan originated this period, and by r_t^m the interest rate on the mortgage loan. The model focus is on equilibria that satisfies $r_t^d - r_t^m \geq 0$.[30] Collateralized credit is provided by lenders outside the model (i.e., international capital markets), whereas non-collateralized loans are provided domestically with households making deposits in a financial intermediary, D_{it}, earning r_t^d. The law of motion for the mortgage balance B_{it} for individual i is given by

$$B_{it+1} = b_{it+1} + (1 - \triangle) B_{it}, \qquad \forall i, t \tag{A1}$$

where $0 \leq \triangle \leq 1$ is the fraction of debt that needs to be amortized each period. The collateral

[30] That is, the domestic interest rate exceeds the rate at which the rest of the world is willing to hold some fraction of mortgage-backed assets. The collateral constraint prevents arbitrage opportunities and restricts the amount of foreign borrowing. Garriga, Manuelli, and Peralta-Alva (2019) show that this condition was satisfied in the United States during the early 2000s housing boom.

constraint on borrowing is given by

$$b_{it+1} \leq \phi_t p_t^h h_{it} - (1-\triangle)B_{it}, \qquad \forall i, t. \tag{A2}$$

The optimization problem for a consumer/investor of type i is given by

$$U_i = \max \textstyle\sum_{t=0}^{\infty} \beta^t [\ln c_{it} + \gamma_i \ln h_{it}], \tag{A3}$$

$$s.t. \quad c_{it} = y_i - (r_t^m + \triangle)\, B_{it} + b_{it+1} + (1+r_t^d)D_{it} - D_{it+1} + p_t^h(h_{it} - h_{it+1}), \qquad \forall i, t$$

subject to Equations $(A1)$-$(A2)$ and the standard non-negativity constraints. The left-hand side of the budget constraint captures spending on consumption, whereas the right-hand side captures earnings, servicing mortgage debt, and the flow of resources associated to borrow and adjust the portfolio of deposits and housing.

C.2 Equilibrium Affordability with a Fixed Housing Supply

For the case that the housing supply is fixed $\{\overline{H}\}$, given a path for credit conditions $\{r_t^m, \phi_t\}_{t=0}^{\infty}$ and income endowments $\{y_i\}_{i=1}^{I}$, an equilibrium is constituted by prices paths $\{p_t^h, r_t^d\}_{t=0}^{\infty}$ and sequences of individual decisions $\{\{c_{it}, D_{it+1}, b_{it+1}, h_{it+1}\}_{i=1}^{I}\}_{t=0}^{\infty}$ that (i) solves each household's optimization problem, and (ii) clear markets.

The relevant condition that determines house prices is

$$\textstyle\sum_{i=1}^{I} N_i h_{it} = \overline{H}, \qquad \forall t. \tag{A4}$$

Solving for the optimal housing demand for individual i yields an asset pricing equation that depends on the traditional fundamental component, often referred to as the user cost, as well as credit component.

$$p_t^h = \underbrace{\frac{U_{h_{it}}}{U_{c_{it}}} + \frac{p_{t+1}^h}{1+r_{t+1}}}_{\text{Fundamental}} + \underbrace{\phi_t p_t^h \left(\frac{r_t^d - r_t^m}{1+r_t^d}\right)}_{\text{Credit}}, \qquad \forall t. \tag{A5}$$

The ability to borrow captured by the loan-to-value, ϕ_t, as well as the relative cost of borrowing captured by the spread of interest rates, $r^d_t - r^m_t > 0$, can make the equilibrium house price trade above the fundamentals component. For the preference specification, the housing demand for each type i is given by

$$h_{it} = \gamma_i \frac{c_{it}}{p^h_t \left[1 - \triangle^h_{t+1} - \triangle^\phi_{t+1}\right]}, \qquad \forall i, t. \tag{A6}$$

where the terms $\triangle^h_{t+1} = (p^h_{t+1}/p^h_t)/(1 + r^d_{t+1})$ and $\triangle^\phi_{t+1} = \phi_t(r^d_t - r^m_t)/(1 + r^d_t)$ represent the option value of reselling tomorrow and the gains associated to the spread between rates tied to the ability to leverage the purchase. Replacing Equation ($A6$) in the housing equilibrium condition Equation ($A4$) yields

$$H^d(p^h) = \textstyle\sum_{i=1}^{I} N_i \left[\gamma_i \dfrac{c_{it}}{p^h_t \left[1 - \triangle^h_{t+1} - \triangle^\phi_{t+1}\right]}\right] = \overline{H}, \qquad \forall t. \tag{A7}$$

The lack of differences in credit conditions, makes the illustration of how credit affects affordability more straightforward.[31] In this case, it is possible to rewrite Equation ($A7$) as follows

$$p^h_t = \frac{1}{\left[1 - \triangle^h_{t+1} - \triangle^\phi_{t+1}\right]} \textstyle\sum_{i=1}^{I} N_i \dfrac{\gamma_i c_{it}}{\overline{H}}, \qquad \forall t. \tag{A8}$$

Notice that the equilibrium house price depends on the housing purchases by each type. That depends on the relative intensity of purchases, $\gamma_i c_{it}$, and the size of each group, N_i, relative to the stock of housing, $\overline{H}$. At this point, it is important to highlight that equilibrium house prices depend on the housing spending of each type, proxy by their consumption.[32] This is a useful point to separate the case where financial frictions do not affect prices from the case that they do.

- **No financial frictions:** The lack of financial frictions imply that $\triangle^\phi_{t+1} = 0$. This can occur when $\phi_t = 0$ or $r^d_t = r^m_t$, and in this case house prices trade at their fundamental

[31] Notice that the segmentation between the market for deposits and mortgages does not depend on individual specific characteristics. In other words, the cost of borrowing or the ability to borrow are not type specific. When the loan-to-value, ϕ_{it}, or the cost of borrowing, r^m_{it}, vary by individual type i, then, the equilibrium house price is determined by solving $p^h_t = \overline{H}_t^{-1} \sum_{i=1}^{I} [1 - \triangle^h_{t+1} - \triangle^\phi_{it+1}]^{-1} N_i \gamma_i c_{it}$, which in this case depends on the entire distribution of credit restrictions by type, $\triangle^\phi_{it+1}$. In the empirical specification used in the next section, the only relevant information is the quantity of housing purchases by type, as the financing aspects are implicitly captured by their purchase decision.

[32] Guren et al. (2020) explore how the dispersion of consumption across MSAs affects the dynamics of house prices. Since house prices and consumption are endogenous variables, they rely on an IV approach to resolve the issue of endogeneity. Our empirical analysis follows the same strategy but uses a different type of instrument.

value. In general, consumption is determined by *current income*, y_{it}, *permanent income*, $\overline{y}_i$, and potentially through *housing wealth*, $p_t^h(h_{it} - h_{it+1})$.[33]

- **Financial frictions:** The presence of frictions in the mortgage market, changes in loan-to-value requirements, ϕ_t, or borrowing costs, r_t^m, affect house prices via a *direct effect* via a credit channel, $\triangle_{t+1}^{\phi}$, as expressed in Equation ($A6$), and the *indirect effect* via housing spending. This is the extent that credit conditions can affect house prices and worsen affordability.

Steady state equilibrium: Because current prices depend on the expectation of the future housing resale, it is convenient to explore the steady state with constant house prices, $p_t^h = p_{t+1}^h = p^h$. Without aggregate movements in income, the equilibrium interest rate on deposits is determined by the rate of time preference, $r^d = (1-\beta)/\beta$. In a steady state equilibrium, each type maintains a stable mortgage balance, $B_{it+1} = B_{it}$, that implies a constant level of borrowing, $b_i = \triangle B_i$, to maintain the home equity level or principal unchanged. The level of borrowing is determined by the equilibrium house price, $B_i = \phi p^h h_i$, and consumption is given by $c_i = y_i - r^m B_i = y_i - \phi r^m (p^h h_i)$. Replacing this expression in the pricing equation gives

$$p^h = \frac{1+r^d}{(1-\phi)r^d + \phi r^m} \textstyle\sum_{i=1}^{I} N_i \frac{\gamma_i}{\overline{H}}(y_i - r^m \phi p^h h_i). \tag{A9}$$

In general the challenge is the dependence of house prices on both sides of the equation. However for this class of preferences one can solve for the equilibrium steady state house price:

$$p^h = \frac{r^d}{\underbrace{(1-\phi)r^d + \phi r^m}_{\text{Direct Effect}} + \underbrace{\phi r^m (1+r^d)}_{\text{Indirect Effect}}} \underbrace{\left(\frac{1+r^d}{r^d}\right) \textstyle\sum_{i=1}^{I} N_i \frac{\gamma_i y_i}{\overline{H}}}_{\text{Fundamentals}}. \tag{A10}$$

In the model, a relaxation on credit conditions increase house prices increase via the *direct effect* in the pricing expression, but can reduce disposable consumption and reduce prices via the *indirect effect*.[34] When all individuals have identical preferences with respect to housing, $\gamma_i = \gamma$, it is direct to show how credit makes house values to exceed income $p^h\overline{H}/Y > (1+r^d)/r^d$ where $Y = \sum_{i=1}^{I} N_y y_i$, making housing less affordable. Similarly, if during the recovery of the

[33] This is not the case as with this particular functional form, the wealth effects are zero as emphasized by Berger et al. (2018). But it is not difficult to extend this model to allow income fluctuations that generate trading in the housing market to smooth income volatility.

[34] The negative effect of debt service cost is reduced as $r^m \to 0$. Hence, in economies in which the underlying real cost of borrowing is low, the equilibrium prices deviate more from the fundamental component. This additional term can often be referred to as the housing premium, with borrowing being one of the several components that can affect the additional return of housing.

financial crises the institutional investors are less affected by credit constraints and their income recovers faster than other individuals, house prices are more likely to trade at their fundamental value improving affordability.

C.3 Equilibrium Affordability with Housing Supply Frictions

This extension shows how credit conditions can interact with traditional supply restrictions. A stylized way to capture their interaction is to formalize the supply of housing with a production function with diminishing returns to scale, $H_t = z_t^h S_t^{\alpha_f} \overline{L}^{1-\alpha_f}$ where $\alpha_f \in (0, 1]$ due to the presence of land being a fixed factor, $\overline{L} = 1$.[35] Assume that the relative price of structures is measured in terms of consumption goods with a price normalized to 1, but the supply restrictions can increase the cost of the input above the market price by a factor $\tau \geq 1$. The regulation can also affect the number of housing units produced by each unit of structure, so the term $\alpha_f = \alpha - \xi$ can be viewed as the span of production subject to production frictions, ξ.

Consider a representative real estate developers that takes the regulation (τ, ξ) as given and solves

$$\max_{S_t} \{p_t^h z_t^h S_t^{\alpha_f} - \tau S_t\}, \tag{A11}$$

where the optimal demand of input, in terms of structures, equals $S_t = \alpha_f p_t^h H_t / \tau$. Replacing this expression in the production function yields a housing supply schedule for any house price.

$$H^s(p_t^h) = \left(\frac{\alpha_f^{\alpha_f}}{\tau^{\alpha_f}} z_t^h \right)^{\frac{1}{1-\alpha_f}} \left(p_t^h\right)^{\frac{\alpha_f}{1-\alpha_f}} . \tag{A12}$$

Notice that increases of house prices increase the supply of new housing units, but the regulation limits the quantity produced. Holding the demand schedule constant, $H^d(p_t^h)$, a restricted supply makes house prices to trade above the frictionless level. In this particular case, since the overhead cost per unit of construction, τ, is constant, the housing supply elasticity, $\Sigma_{H,p}$, is only affected by the friction that reduces the span of control, $\alpha_f/(1 - \alpha_f)$. Replacing the housing supply function, and solving for the steady state yields an expression that connects credit constraints with housing supply frictions.

$$p^h = \left(\frac{\tau^{\alpha_f}}{\alpha_f^{\alpha_f} z^h} \right) \left[\frac{r^d}{(1-\phi) r^d + \phi r^m + \phi r^m (1 + r^d)} \left(\frac{1 + r^d}{r^d} \right) \textstyle\sum_{i=1}^{I} N_i \gamma_i y_i \right]^{1-\alpha_f} . \tag{A13}$$

[35] It is convenient to abstract from zoning restrictions that might affect the optimal quantity of land used in new residential units.

There are two important cases to highlight. When $\alpha_f = 1$, the price of housing is entirely determined by the cost of construction and supply frictions, implying that the purchases of investors can only affect rental markets. When $\alpha_f \in (0, 1)$, there are important interactions between both types of frictions.

The empirical strategy in this paper uses the cross-sectional variation and exploits the idea that the effects of investors' purchases on house prices need to be large in areas with a low housing supply relative to areas with high elasticity. Similarly, housing spending of investors can depend on expectations of future capital gains, cash-flow generated from supplying housing in the rental market, wealth effects and portfolio decisions not formalized in the model, or cost-saving technologies that make rental supply less costly. The empirical strategy uses an instrument variable approach to avoid endogeneity issues.

Extra Figures (NOT FOR PUBLICATION)

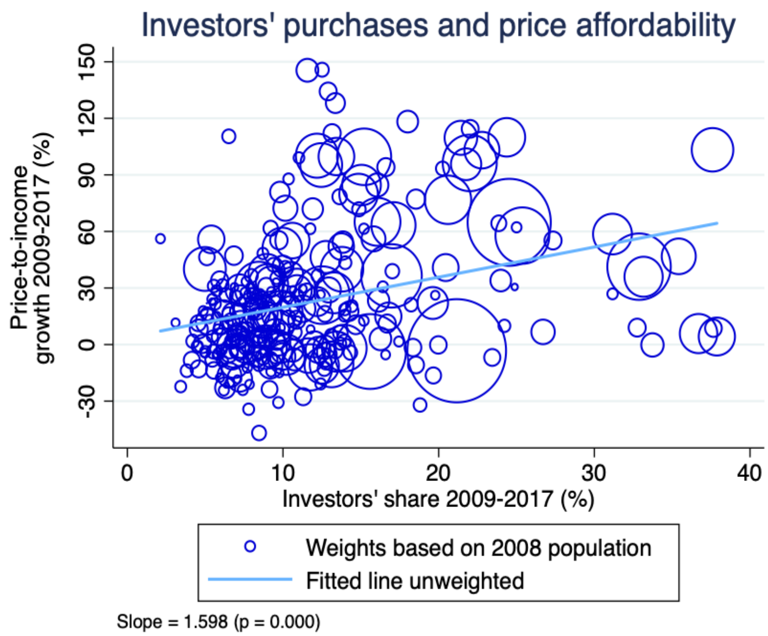


Figure A1. Affordability and institutional investors in the U.S. The figure plots the average share of institutional investors' purchases in the years 2009 to 2017 against the growth of the bottom-tier price-to-income ratio from 2009 to 2017 in the U.S. MSAs. Each circle represents an MSA, and the size of the circle is analogous to the MSA population in 2008.

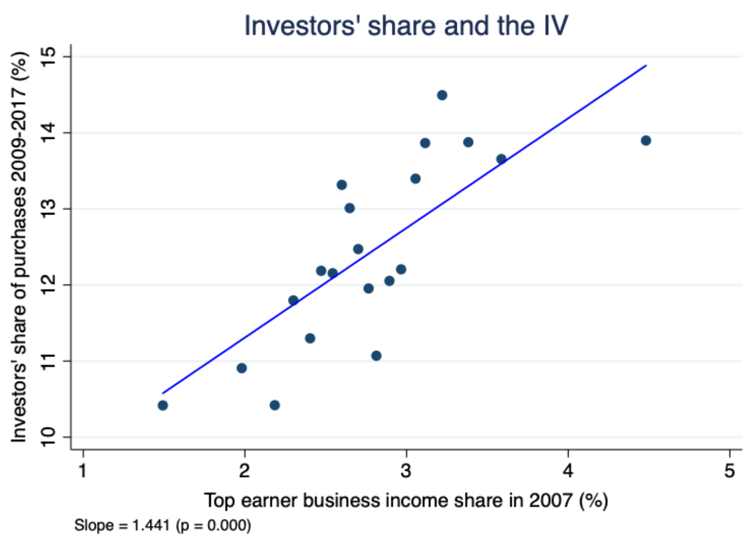


Figure A2. Institutional investors' market share and the instrumental variable. This figure plots the average share of value of business income over total income of top earners in an MSA in 2007, against the 2009-2017 average market share of institutional investors' purchases in each MSA. The top earners are the ones who reported adjusted gross income of 100,000 U.S. dollars or higher in their tax returns. The MSAs are binned by percentiles so that each point represents around 15 MSAs. The figure controls for the controls in the baseline specification in Table 3.

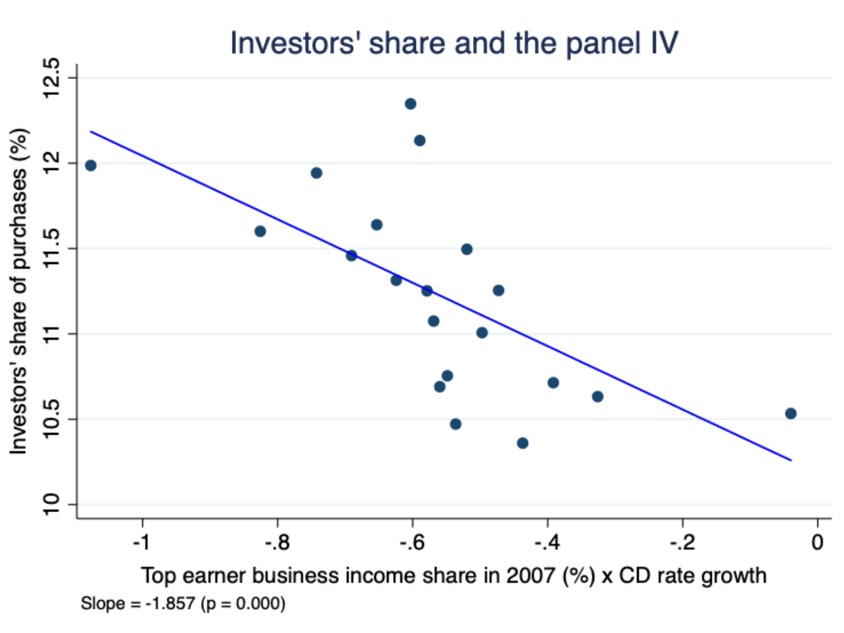


Figure A3. Institutional investors' market share and the instrumental variable. This figure plots the share of value of business income over total income of top earners in an MSA in 2007 multiplied by the CD rate growth, against the market share of institutional investors' purchases each year in each MSA. The top earners are the ones who reported adjusted gross income of 100,000 U.S. dollars or higher in their tax returns. The MSAs are binned by percentiles so that each point represents around 15 MSAs. The figure controls for the controls in the panel specification in Table A5.

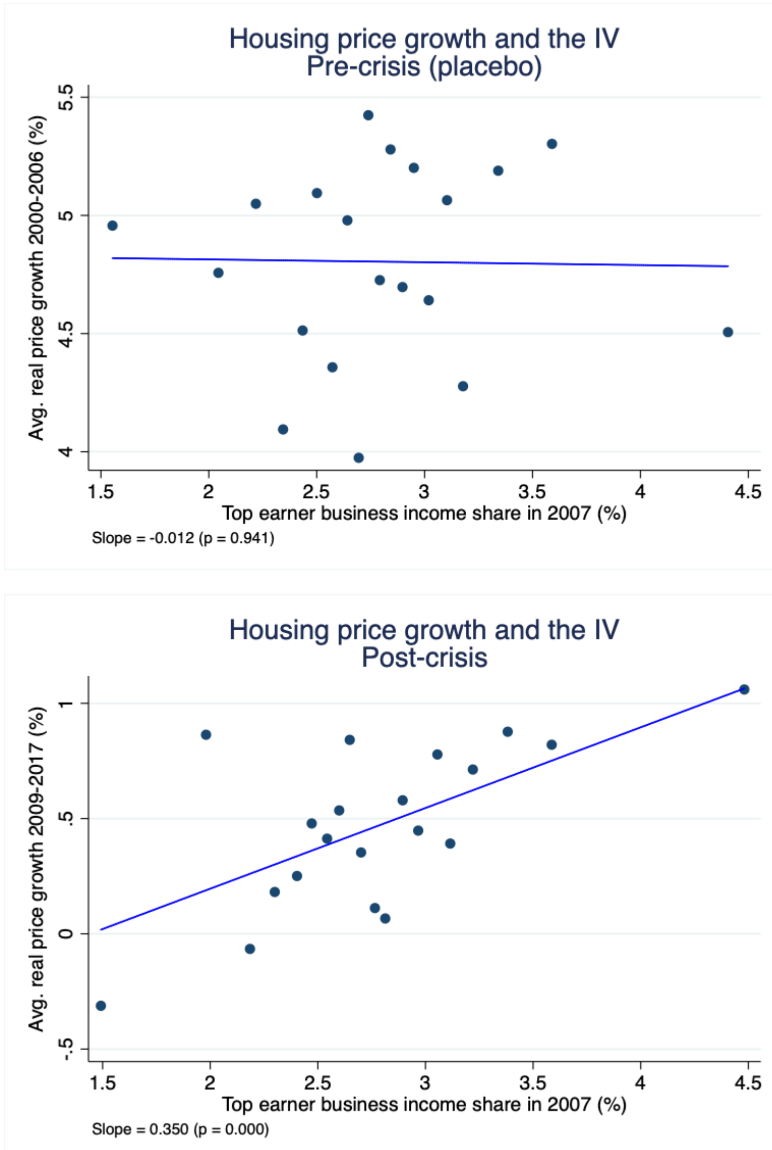
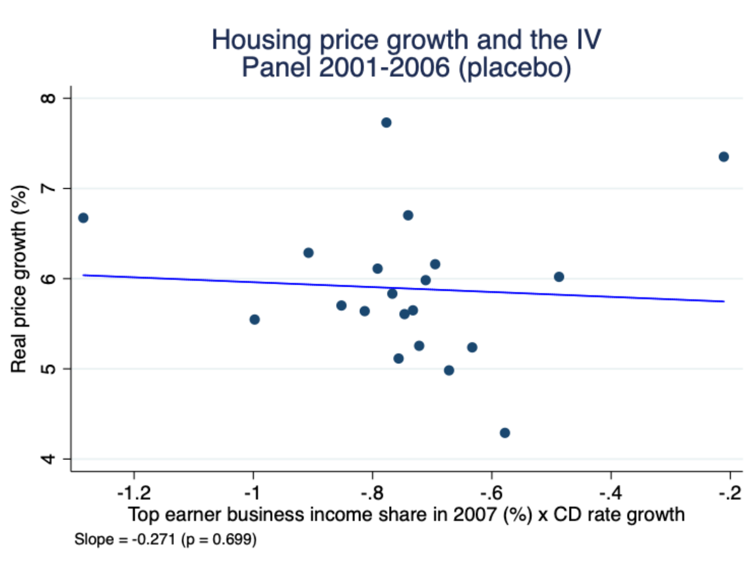


Figure A4. Housing price growth against the instrument for investors pre- and post-2008. The top panel plots the 2000–2006 average annual real housing price growth against the average share of business income over total income of top earners in each MSA in 2007. The bottom panel plots the 2009–2017 average annual real housing price growth against the same instrument. The top panel controls are the ones used in the placebo specification in Table 5. The bottom panel controls are the ones used in the baseline specification in Table 3. Figure A5 in the online appendix performs the same visual exercise for the panel version of the instrument.



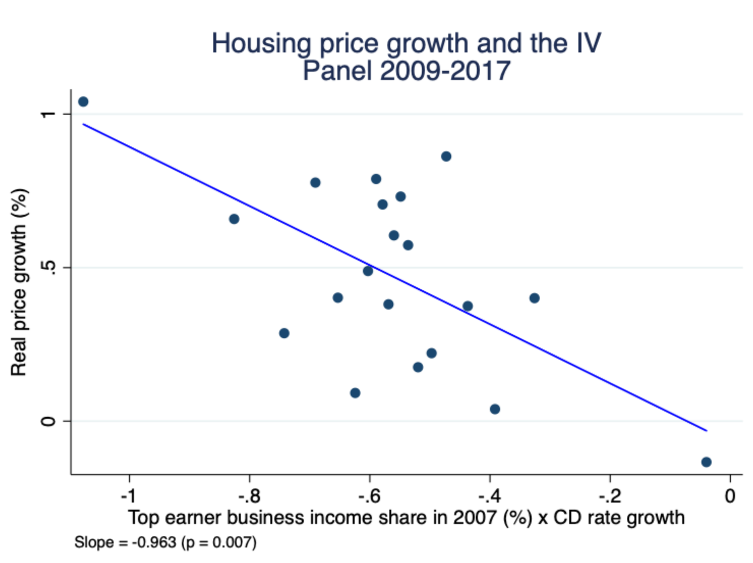


Figure A5. Pre- and post-2008 housing price growth against the panel instrument for investors. The top panel plots the annual real price growth over the 2001-2006 period against the panel instrument: the average share of business income over total income of top earners in each MSA in 2007 multiplied by the lagged CD rate growth. The bottom panel plots the annual real housing price growth over the 2009-2017 period against the same instrument. The top panel controls are the ones used in the placebo panel specification in Table A7. The bottom panel controls are the ones used in the panel specification in Table A5.

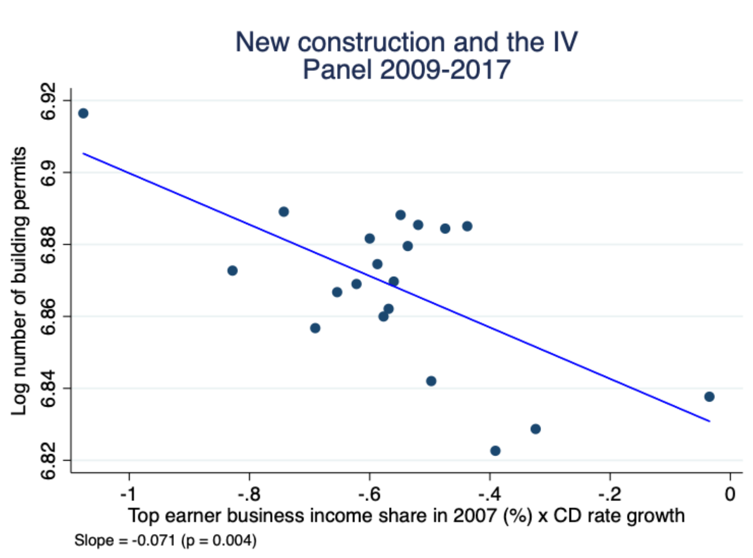


Figure A6. Pre- and post-2008 building permits against the panel instrument for investors. The top panel plots the log number of building permits over the 2001-2006 period against the panel instrument: the average share of business income over total income of top earners in each MSA in 2007 multiplied by the lagged CD rate growth. The bottom panel plots the log number of building permits over the 2009-2017 period against the same instrument. The top panel controls are the ones used in the placebo panel specification in Table A7. The bottom panel controls are the ones used in the panel specification in Table A5.

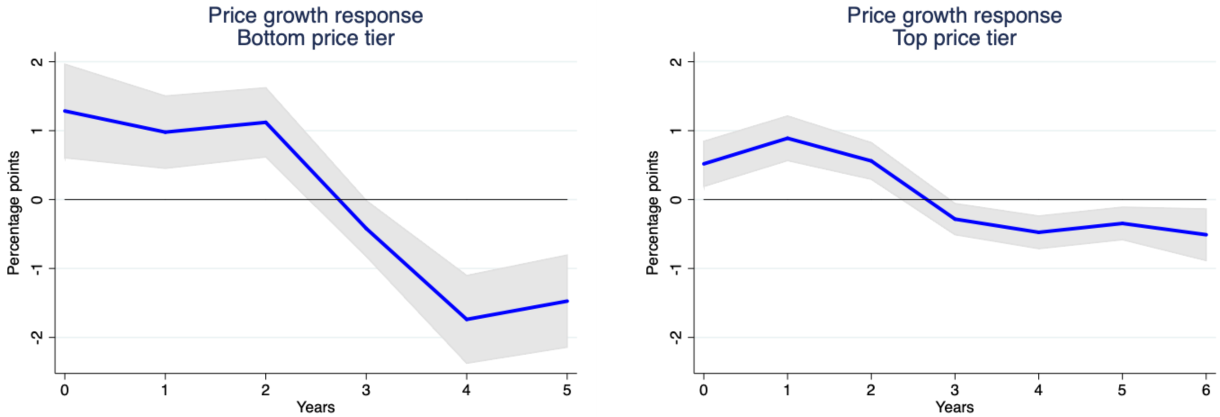


Figure A7. Dynamics of housing prices after investors' purchases by tier. The figure plots the estimates from sequential regressions of the real housing price growth on the instrumented past investors' share of purchases for top and bottom price-tier houses. Top tier houses are houses in the top third, and bottom tier houses are houses in the bottom third of the house value distribution within an MSA. We estimate the impulse responses for the full panel data from 2009 to 2017. The shaded areas show the 90% confidence interval.

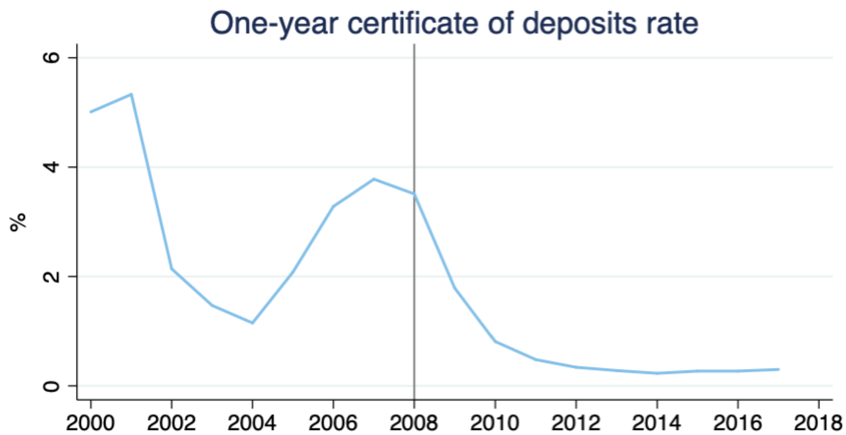


Figure A8. Interest rate used in the construction of the panel instrument. The figure plots the one-year CD rate, annually from 2000 to 2018. The vertical line in 2008 indicates the beginning of the Fed's QE. Source: Bankrate.

Extra Tables (NOT FOR PUBLICATION)

Table A1. Affordability results excluding top MSAs

	Price growth$_{m,09-17}$			Price-to-income ratio$_{m,09-17}$		
	Bottom Tier	Mid Tier	Top Tier	Bottom Tier	Mid Tier	Top Tier
Sample without top 10 MSAs						
Investors' share$_{m,09-17}$	0.306***	0.246***	0.184***	1.540***	0.540***	0.307***
	(0.102)	(0.083)	(0.070)	(0.378)	(0.161)	(0.096)
F-test of excluded instruments	18.970	19.141	19.627	18.970	19.141	19.627
Underidentification p-value	0.000	0.000	0.000	0.000	0.000	0.000
Observations	286	322	318	286	322	318
Sample without top 20 MSAs						
Investors' share$_{m,09-17}$	0.313***	0.252***	0.188**	1.728***	0.579***	0.342***
	(0.112)	(0.089)	(0.076)	(0.429)	(0.175)	(0.101)
F-test of excluding instruments	16.782	17.205	17.734	16.782	17.205	17.734
Underidentification test p-value	0.000	0.000	0.000	0.000	0.000	0.000
Observations	276	312	308	276	312	308

Heteroskedasticity robust standard errors are in parentheses. Top MSAs are the ones with the largest dollar purchases by top 1% institutional investors. These include the 20 largest investors in single-family rentals, and the apartment REITs in the S&P 500 Real Estate Sector. Prices are inflation adjusted. All models include state dummies, MSA-level controls and the instrumental variable as in Table 3. The weak identification F statistic is the Kleibergen and Paap Wald F statistic. The underidentification test is from Kleibergen and Paap (2006). Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level.

Table A2. County level results

	Bottom Tier	Mid Tier	Top Tier
	Price growth$_{m,09-17}$		
Investors' share$_{m,09-17}$	0.361***	0.264**	0.144**
	(0.136)	(0.125)	(0.069)
Observations	601	691	676
	Price-to-income ratio$_{m,09-17}$		
Investors' share$_{m,09-17}$	1.675***	0.659***	0.478***
	(0.566)	(0.222)	(0.096)
Observations	601	691	676

Heteroskedasticity robust standard errors are in parentheses. Bottom tier houses are houses in the bottom third, and top tier in the top third of the house value distribution within a County. Investors' share is the average annual market share of purchases by institutional investors in County c over 2009-2017. All models include state dummies and county-level controls: population growth, income growth, unemployment rate change and real housing price growth over the periods 2000-2006 and 2006-2007, and the log number of construction unit permits in 2007. Price growth is inflation adjusted. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in County c in the year 2007. Each observation is a County. ***significant at the 1% level; **significant at the 5% level.

Table A3. Construction results excluding top MSAs

	All	Single-unit	2-4 units	5+ units
Sample without top 10 MSAs				
Investors' share$_{m,09-17}$	0.052***	0.049**	0.102*	0.150***
	(0.020)	(0.019)	(0.055)	(0.047)
Observations	322	322	320	317
Sample without top 20 MSAs				
Investors' share$_{m,09-17}$	0.057***	0.053**	0.121**	0.164***
	(0.021)	(0.020)	(0.060)	(0.052)
Observations	312	312	310	307

Heteroskedasticity robust standard errors are in parentheses. Top MSAs are the ones with the largest dollar purchases by top 1% institutional investors. These include the 20 largest investors in single-family rentals, and the apartment REITs in the S&P 500 Real Estate Sector. Single-unit refers to permits for the construction of single-unit properties, 2-4 units refers to permits for the construction of properties that have between 2 and 4 units, and 5+ units refers to permits for the construction of properties of 5 units or more. All models include state dummies, MSA-level controls and the instrumental variable as in Table 3. The underidentification test is from Kleibergen and Paap (2006). Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table A4. First stage panel: Investors' share and the instrumental variable

	Investors' share$_{m,t-1}$
Top earner business income share$_{m,07}$×CD rate growth$_{t-2}$	-1.857***
	(0.368)
MSA-year controls	Yes
MSA fixed effects	Yes
Year fixed effects	Yes
R-squared	0.691
Observations	2,842

Standard errors clustered by MSA are in parentheses. The controls are the housing price growth, population growth, median income growth and unemployment rate change, all lagged by one year. The sample period is 2009-2017. Each observation is an MSA-year. ***significant at the 1% level.

Table A5. Housing price growth in response to investors' purchases

	Price growth$_{m,t+i}$						
	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
Top price-tier							
Investors' share$_{m,t-1}$	0.52***	0.89***	0.56***	-0.28**	-0.48***	-0.35**	-0.51**
	(0.20)	(0.20)	(0.16)	(0.14)	(0.14)	(0.14)	(0.23)
Observations	2,804	2,492	2,180	1,868	1,556	1,243	932
Mid price-tier							
Investors' share$_{m,t-1}$	0.52***	0.86***	0.70***	-0.48***	-0.78***	-0.40***	-0.74***
	(0.20)	(0.24)	(0.20)	(0.15)	(0.18)	(0.14)	(0.28)
Observations	2,842	2,525	2,207	1,891	1,575	1,258	942
Bottom price-tier							
Investors' share$_{m,t-1}$	1.29***	0.98***	1.12***	-0.42*	-1.74***	-1.47***	-2.63**
	(0.41)	(0.32)	(0.31)	(0.25)	(0.39)	(0.41)	(1.02)
Observations	2,547	2,260	1,974	1,690	1,406	1,118	837

Standard errors clustered by MSA are in parentheses. All models include location and time fixed effects and controls: the lagged dependent variable, and population growth, median household income growth and unemployment rate change, all lagged by one year. Prices are inflation adjusted. The IV is the average share of business income over total income of the top earners in MSA m in 2007 multiplied by the lagged CD rate growth. The sample period is 2009-2017. Each observation is an MSA-year. The Kleibergen and Paap (2006) underidentification test has p-value of 0.001, and the Kleibergen and Paap Wald F statistic is 25.475 for the mid-tier market panel regression ($i = 0$). Table A4 contains the first stage of the IV regression. Table A13 contains the dynamic results using alternative measures of the investors' presence. ***significant at 1%; **significant at 5%; *significant at 10%.

Table A6. Affordability measures in response to investors' purchases

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
			Price-to-income ratio$_{m,t+i}$				
Investors' share$_{m,t-1}$	0.03***	0.09***	0.14***	0.11***	0.04***	0.01	-0.01
	(0.01)	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)
Observations	2,844	2,527	2,210	1,892	1,576	1,259	944
			Rent-to-income ratio$_{m,t+i}$				
Investors' share$_{m,t-1}$	0.06	0.10	0.22**	0.40***	0.30***	-0.02	-0.21*
	(0.10)	(0.09)	(0.10)	(0.14)	(0.09)	(0.06)	(0.12)
Observations	2,580	2,293	2,006	1,719	1,432	1,144	858

Standard errors clustered by MSA are in parentheses. All models include location and time fixed effects and controls: the lagged dependent variable, and population growth, median income growth and unemployment rate change, all lagged by one year. The IV is the average share of business income over total income of the top earners in MSA m in 2007 multiplied by the lagged CD rate growth. The sample period is 2009-2017. Each observation is an MSA-year. The investors' share is divided by 100 in the regressions of rent-to-income to adjust the coefficients. ***significant at 1%; **significant at 5%; *significant at 10%.

Table A7. Placebo panel: Housing price growth and investors' share pre-crisis

	Price growth$_{m,t}$			
Panel period	2001-2005	2001-2006	2001-2005	2001-2006
Investors' share$_{m,t-1}$	-1.465	-0.007	-0.034	-0.791
	(0.945)	(0.406)	(0.567)	(0.665)
Estimation	IV	IV	IV	IV
Instrumental variable period	2001-2005	2001-2006	2009-2013	2009-2014
Observations	1,585	1,906	1,584	1,905

Standard errors clustered by MSA are in parentheses. The specifications include location and time fixed effects and MSA-year level controls: the real housing price growth, population growth, median income growth and unemployment rate change from time $t-2$ to $t-1$. Prices are for the median house and are inflation adjusted. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in MSA m in the year 2007 multiplied by the lagged CD rate growth. In the first two columns the instruments are constructed using CD rate growth$_{m,t-1}$, so the CD rate is contemporaneous to the panel variables. In the last two columns the instruments are constructed using CD rate growth$_{m,t+7}$, so the instrument is identical to the baseline panel specification, which begins in the year 2009. Each observation is an MSA-year.

Table A8. Estimation controlling for labor demand shifts by industry

	Price growth$_{m,09-17}$					
	coef.	s.e.	coef.	s.e.	coef.	s.e.
Investors' share of purchases$_{m,09-17}$	0.236***	(0.087)	0.221***	(0.082)	0.224***	(0.085)
Employment growth by industry$_{m,08-17}$						
Health Care & Social Assistance	-0.001	(0.003)	-0.004	(0.004)	-0.004	(0.004)
Retail Trade	0.296**	(0.138)	0.274**	(0.132)	0.275**	(0.137)
Accommodation & Food Services	0.029	(0.094)	0.018	(0.091)	0.009	(0.094)
Manufacturing	-0.001	(0.006)	-0.008*	(0.005)	-0.009*	(0.005)
Professional, Scientific, Tech. Services	0.003	(0.003)	0.003	(0.002)	0.003	(0.002)
Administrative, Support, Waste Mgmt.	-0.001*	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)
Finance & Insurance	0.002	(0.002)	0.002	(0.002)	0.002	(0.002)
Wholesale Trade			0.032	(0.034)	0.031	(0.036)
Other Services			0.096**	(0.042)	0.094**	(0.043)
Transportation & Warehousing			0.025**	(0.012)	0.023**	(0.012)
Information			0.004	(0.003)	0.005	(0.003)
Educational Services					-0.000	(0.001)
Management of Companies					-0.001	(0.001)
Real Estate, Rental & Leasing					0.003	(0.003)
Arts, Entertainment & Recreation					-0.000	(0.001)
1st stage F-test of excluded instruments	16.993		18.165		17.050	
Underidentification test p-value	0.000		0.000		0.000	
Observations	332		332		330	

Heteroskedasticity robust standard errors are in parentheses. The specifications control for the average annual growth in the number of employees in various industries - based on the North American Industry Classification System (NAICS) 2 digit sector codes - that predominate the labor market of MSAs, over 2008-2017. Prices are inflation adjusted. The specifications include MSA-level controls, state dummies and the instrumental variable as in Table 3. The underidentification test is that of Kleibergen and Paap (2006) and the F statistic is the Kleibergen and Paap Wald F statistic. Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table A9. Dynamic results controlling for labor demand shifts by industry

	Price growth$_{m,t+i}$						
	$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$
Investors' share$_{m,t-1}$	0.50***	0.83***	0.69***	-0.40***	-0.74***	-0.35***	-0.71***
	(0.19)	(0.23)	(0.19)	(0.14)	(0.18)	(0.13)	(0.27)
Observations	2,756	2,440	2,135	1,834	1,529	1,224	914

Standard errors clustered by MSA are in parentheses. The specifications include location and time fixed effects and controls as in Table A5, and are estimated using our IV. Additional controls are the lagged growth rate of employment in the main industries - based on the NAICS 2 digit sector codes - within the MSAs: Health Care & Social Assistance, Retail Trade, Accommodation & Food Services, Manufacturing, Professional, Scientific and Technical Services, Administrative and Support and Waste Management, Finance and Insurance, Wholesale Trade, Other Services, and Transportation and Warehousing. Prices are for mid-tier houses and are inflation adjusted. The sample period is 2009-2017. Each observation is an MSA-year. The Kleibergen and Paap (2006) underidentification test has p-value of 0.001, and the Kleibergen and Paap Wald F statistic is 27.562 for the panel regression ($i = 0$). ***significant at 1%.

Table A10. The instrumental variable and its predictors

	Top earner business income share$_{m,07}$	
Median age$_{m,07}$	0.030*	0.011
	(0.016)	(0.017)
Immigrants as % of population$_{m,07}$	0.032***	-0.001
	(0.008)	(0.010)
Income tax rate for top earners$_{m,07}$	0.055***	0.053***
	(0.018)	(0.016)
Entrepreneurship rank$_{m,07}$	-0.0001	0.001
	(0.003)	(0.003)
Natural amenity index$_{m,07}$		0.121***
		(0.022)
R-squared	0.113	0.223
Observations	280	277

Heteroskedasticity robust standard errors are in parentheses. The outcome variable is our instrument for the investors' share of purchases: the average share of business income over total income of the top earners in MSA m in the year 2007. Each observation is an MSA.

Table A11. The instrumental variable and drivers of housing markets

	Top earner business income share$_{m,07}$
Avg. median age change$_{m,00-06}$	0.019
	(0.048)
Avg. homeownership rate change$_{m,00-06}$	-0.002
	(0.055)
Median age change$_{m,07}$	-0.014
	(0.051)
Homeownership rate change$_{m,07}$	-0.024
	(0.066)
MSA-level controls	Yes
State dummies	Yes
R-squared	0.528
Observations	288

Heteroskedasticity robust standard errors are in parentheses. All variables are normalized to have zero mean and standard deviation of one. The outcome variable is our instrument for the investors' share of purchases: the average share of business income over total income of the top earners in MSA m in the year 2007. The controls are as in Table 3. Each observation is an MSA.

Table A12. Single-family properties

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
	Single-family price growth$_{m,t+i}$						
Investors' single-family share$_{m,t-1}$	0.61***	1.07***	0.86***	-0.64***	-0.98***	-0.47***	-1.10**
	(0.22)	(0.28)	(0.24)	(0.19)	(0.22)	(0.17)	(0.46)
Observations	2,830	2,514	2,197	1,882	1,567	1,250	936
	Price growth$_{m,t+i}$						
Investors' single-unit share$_{m,t-1}$	0.59***	1.05***	0.88***	-0.61***	-1.01***	-0.51***	-1.07**
	(0.21)	(0.27)	(0.24)	(0.18)	(0.22)	(0.18)	(0.43)
Observations	2,842	2,525	2,207	1,891	1,575	1,258	942

Standard errors clustered by MSA are in parentheses. The fixed effects and controls are as in Table A5. The top panel uses single-family prices and the bottom panel prices for all homes, from Zillow. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in MSA m in 2007 multiplied by the lagged CD rate growth. The sample period is 2009-2017. Each observation is an MSA-year. ***significant at the 1% level; **significant at the 5% level.

Table A13. Alternative measures of investors

	Price growth$_{m,t+i}$						
	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
Share of number of purchases$_{m,t-1}$	0.85***	1.72***	1.53***	-1.13***	-1.91***	-0.86***	-2.10**
	(0.30)	(0.48)	(0.44)	(0.39)	(0.52)	(0.32)	(1.05)
Observations	2,842	2,525	2,207	1,891	1,575	1,258	942
Share of number of units$_{m,t-1}$	0.74***	1.30***	1.15***	-0.83***	-1.35***	-0.64**	-1.12**
	(0.28)	(0.39)	(0.38)	(0.29)	(0.39)	(0.26)	(0.54)
Observations	2,842	2,525	2,207	1,891	1,575	1,258	942

Standard errors clustered by MSA are in parentheses. The investors' share of number of purchases denotes the market share of the count of purchases by institutional investors. Each purchase counts as one purchase, independent of the type of property, that is, one single-family detached home, one apartment building, etc. The investors' share of number of units denotes the market share of the count of units purchased by institutional investors. For example a purchase of a 10-unit apartment building counts as 10 units. The number of units is coded by ZTRAX. The online appendix describes our coding of this variable when there are missing or incomplete data from ZTRAX. The fixed effects and controls are as in Table A5. The instrument for the investors' share of purchases is the average share of business income over total income of the top earners in MSA m in 2007 multiplied by the lagged CD rate growth. The sample period is 2009-2017. Each observation is an MSA-year. ***significant at the 1% level; **significant at the 5% level.

Table A14. Land use and buildings' classification

Single-family: single family residential; townhouse; row house; mobile home; cluster home; seasonal, cabin, vacation residence; bungalow; zero lot line; patio home; manufactured, modular, prefabricated homes; garden home; planned unit development; rural residence; residential general; inferred single family residential.
Multi-family: condominium; cooperative; landominium; duplex (2 units, any combination); triplex (3 units, any combination); quadruplex (4 units, any combination); apartment building (5+ units); apartment building (100+ units); high-rise apartment; garden apartment, court apartment (5+ units); mobile home park, trailer park; dormitory, group quarters (residential); fraternity house, sorority house; apartment (generic); multifamily dwelling (generic any combination 2+); boarding house rooming house apt hotel transient lodging; residential condominium development (association assessment); residential income general (multi family).

This table shows the classification of homes into single-family and multi-family based on the ZTRAX land use standard codes.[36]

[36] We excluded from the data the following land use standard codes that do not refer to homes: "residential common area", "timeshare", "residential parking garage" and "miscellaneous improvement".