Investors and Housing Affordability.* †

Carlos Garriga‡, Pedro Gete§, Athena Tsouderou¶

December 2019

Abstract

This paper studies the impact of housing investors on the dynamics of housing affordability, after the Global Financial Crisis. Using an instrumental variable approach, we find that the investors' purchases in the U.S. MSAs, increase the price-to-income ratio, especially in the bottom price-tier of the market, and in areas with large supply restrictions. However, these effects are short-lived. Investors cause a significant supply response, as they increase granting of new building permits. In the medium-term, investors' purchases lead to reductions in prices and improved affordability. These findings should be considered when designing policy regulations.

Keywords: Investors, Housing Prices, Affordability, Homeownership, Real Estate Investment.

*We thank Itzhak Ben-David, Henning Bohn, Daniel Cooper, Stefano Corradin, Morris Davis, Ignacio De la Torre, Anthony DeFusco, David Echeverry, Michael Ehrmann, Daniel Fernández Kranz, Andra Ghent, Juan-Pedro Gomez, Pere Gomis-Porqueras, Deeksha Gupta, Jonathan Halket, Lu Han, Marie Hoerova, Ivan Jaccard, Dirk Jenter, Nina Karnaukh, Daniel Kelly, Finn Kydland, Agnese Leonello, Jose Maria Liberti, David Ling, Haoyang Liu, David Marques, Clara Martinez-Toledano, Brian Melzer, Erwan Morellec, Charles Nathanson, Alex Popov, Michael Pries, Lev Ratnovski, Michael Reher, Rafael Repullo, Stephen L. Ross, Gerhard Runstler, Vahid Saadi, Glenn Schepens, Martin Schneider, Jean-David Sigaux, Jirka Slacalek, Alejandro Van der Ghote, Randal Verbrugge, Jeffrey Zabel, Fernando Zapatero, and seminar participants at Durham, ECB, NEOMA, Notre Dame, IE, Ohio State, and 2018 UEA, 2019 HULM, 2019 AREUEA International, 2019 AEFIN Finance Forum, and 2019 SED 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. athena.tsouderou@student.ie.edu.

Introduction

Housing affordability is one of the most critical policy challenges for most cities in the world (Favilukis, Mabille and Van Nieuwerburgh 2019). For example, in 2019, in more than 55% of the U.S. MSAs the median household needs at least three times their annual income to buy a median-priced home. Figure 1 shows that this level of affordability resembles the situation during the large housing boom of the 2000s.

Interestingly, it has become common to link affordability problems to the explosive growth of institutional investors in housing markets (see for example ACCE Institute 2018; the Wall Street Journal 2017; or the Guardian 2019). These concerns are supported by figures, as Figure 2, which shows that the MSAs that experienced the highest increase in the price-to-income ratio also had a corresponding contemporaneous increase in the market share of institutional investors. Officials in several cities have enacted or are discussing policies to block 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 wants to directly ban investors from buying and renting properties. Spain recently imposed a set of measures to penalize them. Berlin is considering to expropriate large, private, profit-seeking landlords (The Wall Street Journal 2019).

This paper contributes to the debate, whether housing investment is good or bad for affordability. Using a rich database covering the whole U.S. from 2009 to 2017, we quantify the effects of institutional investors in housing markets post-financial crisis. We exploit variation across MSAs in the presence of institutional investors, and perform various analyses, firstly in the cross-section and secondly dynamic using panel data.[1] Our results show causal evidence that the presence of institutional investors has been a significant driver of the housing price recovery since the financial crisis. However, our paper also shows an equilibrium response of supply: investors increase the supply of housing and eventually reverse the price growth. Investors eventually contribute to making houses more affordable to purchase.

Our identification strategy exploits heterogeneity across the U.S. MSAs in the market share of housing purchases by institutional investors. The challenge for our identification is to isolate shocks to investment in the housing market from shocks that drive both housing variables and the share of housing purchases by institutional investors. For example, an ordinary least squares (OLS) regression of price growth on investors' market share would be biased downwards

[1] We define *institutional* investors as legal entities who purchase homes, that is, they purchase under the name of an LLC, LP, Trust, REIT, etc.

if investors were attracted to areas where prices collapsed during the crisis.[2]

We use an instrumental variable approach to overcome the previous challenge. Beginning in December 2008, the Fed's Quantitative Easing (QE) 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. In such a scenario, investors who would traditionally invest in safe assets switch their money to riskier assets. Large amounts are invested in housing. In particular, knowledgeable investors searching for a stable yield, are more likely to consider buying properties and renting them out, which would give them a stable rental yield. Consistent with this "search-for-yield" theory, Daniel, Garlappi and Xiao (2018) find that a low-interest-rate monetary policy increases investors' demand for high-dividend stocks, which is more pronounced among investors who fund consumption using dividend income. As rents are mostly stable, housing becomes an investment asset and a close substitute to safe yield-earning investments.

The QE produced a national shock to the risk-free rates. Our instrument exploits heterogeneity across MSAs in exposure to this shock, by measuring the MSAs' exposure to "investment attitude", in other words, high-earning, sophisticated, risk-seeking residents. The idea is to identify whether a particular MSA has a large share business income earned by the MSA residents, as a proxy for a large propensity to invest and or understand investments. Facing more attractive terms for investment in housing after 2008, investment-prone high-earning individuals are likely to direct more capital to the housing market, through new or existing legal entities. Our analysis of the institutional investors' investment strategy is in line with the view that investors were searching for yield.

Our instrument is the average share of value of business income over total income of high earners in 2007 in each MSA, calculated using tax returns from the IRS. We show that MSAs with greater exposure to business income of high earners experienced larger share of housing purchases by institutional investors after 2008. To create the panel version of this instrument we multiply the 2007 business income share by the certificate of deposits rate growth, which we use to quantify the effects of QE. Since the business income share of each MSA is determined prior to the QE, there is minimal risk of reverse causality. That is, it is unlikely that housing price growth during 2009-2017 affected the share of business income in 2007. Our key identification assumption is that, once we control for an array of factors and fixed effects, exposure to business income in 2007 is uncorrelated with other drivers of housing markets over 2009-2017.

We rigorously assess the validity of the instrumental variable. First, we control thoroughly for an array of local activity shocks, trends during the previous housing boom and bust, and local

[2]The national housing price index shows that average prices in the U.S. were falling from 2007 to 2012.

factors, making it unlikely that the error term reflects common movers of both investors and housing market variables. Second, we provide extensive evidence that in the pre-QE period the instrument does not correlate with either higher price growth, rent growth, new construction, or with other factors that drive the housing markets. For example, before 2008 patterns between MSAs with the highest and lowest exposure to the share of business income are parallel. Third, placebo tests confirm that the instrument only captures post-crisis shocks to investment in the housing markets. Fourth, an alternative instrument from Gete and Reher (2018) that operates through credit constraints to households, gives us similar results.

In the cross-section, we find that institutional investors significantly increase housing prices and price-to-income ratios over the 2009-2017 period. One standard deviation larger share of investors that the mean, increases price growth by eighty to ninety percent of the cross-sectional standard deviation. It also increases the price-to-income ratio by up to two standard deviations. The largest effects happen on the bottom price-tier of the market relative to the top-tier. During this period, the investors increase the supply of housing through a rise in new construction. The investors also bring more liquidity to the housing market, by reducing the share of houses that stay vacant.

To study the dynamic effects of investors, we estimate the response of housing prices and quantities to the institutional investors' purchases over time, by employing the methodology of Jordà (2005). This allows for separation of the dynamic adjustment in markets where the short-run supply elasticity is relatively inelastic. The results indicate that on impact the investors' purchases have a positive effect on house price growth lasting two years. Between the second and third year the impact on the growth rate becomes zero, and negative thereafter. The hump-shaped response of construction, combined with the relatively inelastic short-run housing supply rationalizes the observed patterns. As investors trigger new construction, the supply of housing increases and houses become cheaper and more affordable in the areas where investors hold a larger share of the market. Consistent with our theory, the fluctuations in prices are much larger in areas with low price elasticity of supply.

This paper contributes to the literature on housing affordability. Housing affordability is one of the most critical policy issues for most large cities in the world, and more research is required to inform policy (Deng, Qin and Wu 2019). In their influential paper, Gyourko, Mayer and Todd (2013) suggest that the inelastic supply of land along with an increasing number of high-income households at the national level in the U.S., leads to persistent high house prices in large MSAs and crowds out lower-income households. The effectiveness of various policies aimed at improving housing affordability is at the core of the recent literature. For example, Favilukis, Mabille and Van Nieuwerburgh (2019) study the welfare gains of different key policy

levers. Relative to this literature, our paper explores and provides evidence for a pro-market approach. By letting markets come to an equilibrium, the response of supply may restore the affordability levels in the medium-term.

Our paper also contributes to a new literature that focuses on the emergence of institutional investors in housing markets after the global financial crisis. Garriga, Gete and Tsouderou (2019), analyzing residential deeds in all U.S. MSAs, show that post-crisis institutional investors replaced the individual speculators that flooded the U.S. housing market during the previous housing boom. The authors show that the new institutional investors are mainly small and local, and their investment strategy is driven by search for yields. In a thorough analysis of institutional investors, using deeds data from a different source, Lambie-Hanson, Li and Slonkosky (2019) find that single-family institutional buyers contributed to an increase in price and rent growth. Their identification is based on the First Look program, a program that gave households and non-profits an opportunity to bid on real estate owned properties before they became available to investors.[3] Verbrugge and Gallin (2019) find that multi-unit landlords, when renegotiating rent contracts, set rent increases that exceed the inflation rate, aided by the law of large numbers and exploiting tenant moving costs. Graham (2019) finds that during the latest housing bust investors substitute for falling homeowner demand, and thereby dampen declines in housing prices.[4]

The rest of the papers in this literature have focused on single-family housing in specific areas in the U.S. For example, Gay (2015) documents that purchases and sales of single-family houses by institutional investors were four times more in the years 2012 and 2013 compared to 2011. He finds that institutional investors buy more properties in lower income areas of the Chicago Metro Area and sell them at a premium after renovations, which is related to low affordability of buying a house in those areas. Brunson (2019) finds that institutional investors paid a discount of about 8% to 11% per transaction in the Charlotte region post-crisis. Allen, Rutherford, Rutherford and Yavas (2017) focusing on the Miami-Dade County in Florida, find that an increase in the share of houses purchased by investors in a census block is related to an increase in house prices. Raymond et al. (2018) find that large institutional owners of single-family rentals file more frequently eviction notices compared to small landlords in Fulton County in Atlanta. Mills, Molloy and Zarutskie (2019) show that prices increased in neighborhoods where large buy-to-rent investors are concentrated. The previous papers however do not use causal identification.

[3]The findings are also included in a nontechnical report (Lambie-Hanson, Li and Slonkosky (2018)).

[4]A more established literature has studied foreign and out-of-town investment in the housing markets. See for example recent papers by Cvijanovic and Spaenjers (2018), Davids and Georg (2019), and Favilukis and Van Nieuwerburgh (2018).

Relative to the previous literature, our paper provides novel insights into the dynamic effects of institutional investment in housing markets. To the best of our knowledge this paper is the first to study the housing affordability dynamics in response to institutional investors, and the equilibrium response of supply. Moreover, by introducing and thoroughly defending a novel instrumental variable based on local "investment attitude" in all U.S. MSAs, our paper quantifies the contribution of investors to real effects in the housing markets, and differential effects for the different market segments.

The rest of the paper is organized as follows: Section 1 describes our data. Section 2 presents our cross-sectional analysis. Section 3 presents our dynamic analysis. Section 4 rigorously assesses the validity of our instrument. Section 5 concludes. The Appendix has extra information about the variables and more robustness tests.

1 Data

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

We use a rigorous methodology to identify investors. First we distinguish between individual and non-individual buyers based on the buyer name. Second, we filter out buyers that are relocation companies, non profit organizations, construction companies and national and regional authorities, as well as banks, Ginnie Mae, Fannie Mae, Freddie Mac and other mortgage loan companies and credit unions, and the state taking ownership of foreclosed properties.

We use the dollar value of purchases by investors instead of the number of purchases to calculate the investors' market share, since 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. The dollar value of purchases reflects more accurately the presence of individuals and institutional investors in the single-family and multi-family markets. 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. The total local market value accounts for economic shocks in each location that affect residential purchases.

In our robustness checks we use alternative measures of the presence' of investors based on the number of properties or the number of units purchased. For example a purchase of a 10-unit apartment building counts as 10 units. The number of units is coded by Zillow. The online appendix A describes our coding of this variable when there are missing or incomplete data in ZTRAX.

The main data source to construct our instrument is 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. Our instrument approximates the average individual's tax returns by the zip code returns of a specific income group. Since the dataset does not provide returns at the individual level, the zip code income group level is the closest approximation to the average individual of each group within the zip code. We specifically focus on the returns of the top earnings groups, which include people with annual adjusted gross income above \$100,000, and we perform robustness for this cut-off. Our instrument is the average share of business income over total income of high earners, in each MSA in 2007. We weight by the total income of high-earners to aggregate to MSA level. To construct the panel version of the instrument, we retrieve the average one-year certificate of deposits (CD) rate from Bankrate, a consumer financial services company.

The housing variables come from Zillow. Housing prices come from the Zillow Home Value Index, a dollar-denominated, smoothed, seasonally adjusted measure of the median estimated home value across a given region and housing type. This index is constructed using estimated monthly sale prices not just for the homes that sold, but for all homes even if they didn't sell in that time period, which addresses the bias created by the changing group of properties that sell in different periods of time.

Specifically, we use the Zillow Home Value Index for all homes, single-family 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. 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. Housing rents come from the Zillow Rent Index for all homes, which is constructed using a similar methodology to the Zillow Home Value Index.

To calculate the price-to-income (rent-to-income) ratio, we divide the median housing price (annual rent) by the median household income from Zillow in each MSA-year. To calculate the price-to-income ratio for different market segments, we use instead the individual adjusted gross

income from the SOI. The bottom-tier price-to-income ratio is the ratio of the 17th percentile of housing prices from Zillow over the 17th percentile of individual adjusted gross income from the SOI. The median price-to-income ratio is the ratio of the median housing price over the median individual adjusted gross income. The top-tier price-to-income ratio is the ratio of the 83rd percentile of housing prices over the 83rd percentile of individual adjusted gross income.

We collect the number of new construction permits from the Census Bureau's annual Building Permits Survey. The permits distinguish among single-family, two-family, three- or four-family, and five-or-more family buildings. Finally, we collect MSA-year level controls, which are population from the U.S. Census Bureau, unemployment rate from the U.S. Bureau of Labor Statistics, and median household income from Zillow.

A more detailed description of the data sources is included in the online appendix A. 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. There are 317 MSAs in our MSA-year balanced panel, in terms of housing prices and investors' market share. Table 1 contains summary statistics of the key variables in our study.

2 Investors and Affordability in the Cross-Section

The evidence presented in previous section is suggestive that the growth in the market share of purchases of institutional investors could have been an important driver in the growth of housing prices after the financial crisis. To measure the contribution of institutional investors to housing outcomes, we exploit the regional variation across MSAs in the share of investors, using a specification that exploits the cross-sectional differences. The idea is to identify similar areas where the differential effect in housing outcomes comes from the impact of institutional investors. Formally, the initial specification in the cross-section takes the form of:

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

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 average log number of building permits, and the average log vacancy rate. $Inst_{m,09-17}$ is the average share of institutional investors' dollar value of purchases over the total purchases in MSA m over the same period. The specification has two different types of controls. The term C_m summarizes the MSA-specific controls, which include the more

recent experience in terms of the housing boom and bust and the standard effects in terms of population, income and unemployment. More specifically, these controls are the 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 in the specification includes state dummies to account for the time-invariant state-specific influences, i.e. sunbelt states versus Midwest.

A direct OLS estimation of the specification (1) using investors as an exogenous variable is likely to generate biased estimates. This is because the estimates might capture "reverse causality": investors might target MSAs where prices fell more after the crisis and were slow to pick up, which would potentially give them higher capital gains. As a result, the OLS estimate for housing price growth would be biased downward. Estimates might also suffer from "omitted variable bias": local economic shocks can drive both housing market dynamics and the share of investors. A positive shock to an MSA's economic activity would increase amenities and thus price growth, while raising the attractiveness of the houses as investment opportunities. In this case the OLS estimate for housing price growth would be biased upward. Regardless of the direction of the bias, we aim to overcome it by using an instrument for the investors' market share of purchases.

2.1 The instrumental variable: Attitude towards investing

Our instrument for the investors' market share measures the exposure of an MSA to variation of nationwide risk-free investments. Towards the end of 2008 the Fed's QE reduced the supply of safe assets in the market, reduced risk-free rates and consequently the amounts allocated in safe investments. In theory, we expect that investors facing lower risk-free rates increase their demand for alternative higher-yield investments, such as investment in housing. Investors are looking for a stable yield, measured in terms of housing rents (Jordà et al. 2019). The instrument measures investment attitudes in the population living in a particular MSA. The idea is to establish whether a particular MSA has a large share business income earned by the MSA residents, as a proxy for a large propensity to invest and or understand investments. Facing more attractive terms for investment in housing after 2008, investment-prone high-earning individuals directed more capital to the housing market. Formally, 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 $100,000 in their tax returns.[5] The instrument captures an MSA's exposure to high earners with attitude

[5] As robustness tests, we have also constructed instruments using the average share of busness income in the MSA and different moments of the distribution. The results, not reported here, hold

towards investments, where the exposure is measured with predetermined variables unrelated to the factors the literature has identified as drivers of housing variables.

Table $A3$ assesses the relevance of the instrument, showing the results of the first stage of the 2-stage least squares (2SLS) regression based on the specification (1). After controlling for the relevant MSA-level controls, and state dummies, our instrument is significantly correlated with the investors' purchases. We reject that the instrument is weak using the rule of thumb that the Kleibergen and Paap (2006) Wald F statistic should be larger than ten.

In Section 4 we discuss multiple tests that all suggest that the instrument satisfies the exclusion restriction for our housing dependent variables. Our key identification assumption is that, once we control for an array of factors and fixed effects, exposure to business income in 2007 is uncorrelated with other drivers of housing markets over 2009-2017.

2.2 Cross-sectional results

We use the cross-sectional specification(1) as the baseline, essentially to perform the tests of exclusion restriction for our instrumental variable, since the variation comes from the cross-section of the MSAs. We also build on the cross-sectional results to then calculate the dynamic effects of investors in the next section.

Table 2 summarizes the contribution of institutional investors to the housing price growth and the price-to-income ratio, by price tier, over the period 2009-2017.[6] The first column, showing the IV estimation of the effects on the bottom-tier real price growth, shows a significant coefficient of the share of investors of 0.29. This means, one percentage point increase in the average market value share of institutional investors' purchases lead to an increase in the average annual bottom-tier real housing price growth of 0.29 percentage points. To put this into perspective, one standard deviation higher share of housing purchases by institutional investors than the national mean, lead to 2.3% higher annual bottom-tier real housing price growth over the period 2009-2017. This increase corresponds to 91% of the cross-sectional standard deviation of the bottom-tier real price growth over this period. The lower part of Table 2 shows the results for the standardized share of investors and standardized dependent variables, for easier comparison and derivation of the economic significance of the results.

for different versions of the instrumental variable.

[6]For comparison purposes, we restrict the sample of all regressions in Table 2to the MSAs for which we have Zillow housing prices for all price-tiers.

The second and third columns of Table 2 show that investors had significant and positive effects on the price growth of the median and top price-tier segments of the market. However, the largest effects were on the bottom price tier of the market. One standard deviation higher share of housing purchases by institutional investors than the national mean, lead to an increase of 83% of the cross-sectional standard deviation of the median price growth and 77% of the top-tier price growth over 2009 to 2017.

The fourth column of Table 2 shows the effects of the investors' share of purchases on the bottom-tier price-to-income ratio. One percentage point increase in the average market value share of institutional investors' purchases lead to an increase in the average bottom-tier price-to-income ratio of 1.5. To put this into perspective, one standard deviation higher share of housing purchases by institutional investors than the national mean, lead to an increase of 2.1 cross-sectional standard deviations of the bottom-tier price-to-income ratio over 2009 to 2017. The two last columns show that the corresponding increases for the median and top-tier price-to-income ratios are 1.7 and 1.9 standard deviations.

Overall, investors increase price growth and price-to-income ratio for all price segments. The effects are larger for the bottom price-tier of the market.

Table $A1$ shows the results of the estimation of (1) for the median housing price growth using OLS. The OLS estimation in the first column gives a coefficient of 0.03, significant at the 5% level. In the second column we estimate (1) using our instrumental variable for the share of investors and for the full sample of MSAs where we have median housing prices from Zillow. After accounting for biases, the coefficient of the share of investors becomes 0.24, significant at the 1% level. The smaller coefficient of the OLS estimation is consistent with downward bias of the OLS, since the prices were falling significantly up to 2012, and investors were likely to select areas were prices collapsed.

Table $A2$ shows that the institutional investors have a positive effect on quantities by increasing the number of new building permits. Investors also reduce the vacancy rates for both owner-occupied houses and rentals, which indicates that they bring more liquidity to the housing market.

3 Dynamic Real Effects of Investors

The cross-sectional analysis in the previous section shows evidence that institutional investors increase housing demand and prices, however, they also stimulate supply. The results

have shown that MSAs with higher share of investors' purchases experience higher housing price growth and authorize a higher number of new building permits over the period 2009-2017. In theory, however, the increase in supply would be expected to lead to a drop in prices in longer horizons.

This section studies how the response of housing prices and quantities to the institutional investors' purchases change over time. Jordà (2005) introduces a projection method that estimates impulse response functions directly, without specifying or estimating the unknown true multivariate process. This method estimates local projections based on sequential regressions of the dependent variable shifted forward.[7]

We estimate the dynamic real effects of investors with the following specification:

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

where t indexes years and m MSAs. $y_{m,t}$ denotes our housing variables: real housing price growth rate from year $t-1$ to year t, real housing price growth rate specifically for top tier and bottom tier houses, single-family and single-unit properties, the price-to-rent ratio growth rate from year $t-1$ to year t, and the log number of new construction permits. $Inst_{m,t}$ is the institutional investors' share of dollar value of purchases over the total market value for the year t in MSA m. $C_{m,t}$ summarizes the time-varying MSA-specific controls that include the population growth rate, the median income growth rate, and the unemployment rate change, from year $t-1$ to year t. These controls capture local variations that can affect the housing market and also influence the decisions of investors to select a specific location. We include these controls lagged for one year, since they are likely to be themselves outcomes of the investors' presence. The location fixed effects α_m hold constant the time-invariant MSA-specific influences, and the time fixed effects b_t account for the time-varying factors common to all MSAs. Since the specification includes a lagged dependent variable $y_{m,t-1}$, the growth response is allowed to be temporary.

The local projections of the response of the housing variables to the investors' share over time are given by the vector of estimates $\{\beta_1^{(i)}\}$, 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 our estimation

[7] Mian, Sufi and Verner (2017) and Favara and Imbs (2015) apply this method.

we cluster standard errors by MSA to allow for within-MSA correlation throughout the sample period.[8]

To overcome potential bias in an OLS estimation, in addition to using the usual controls related to the housing market and location and time fixed effects, we use the panel version of our instrumental variable to instrument investors' share. The instrument is based on the 2007 share of business income in each MSA. MSAs with higher exposure to business income would experience higher investment in housing over time as the CD rate drops. For the reasons we explained earlier, as the CD rate drops, investors have incentives to switch their fixed income investments to the housing markets. To construct the instrument we multiply the 2007 local exposure to business income by the CD rate growth. The 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 also unrelated to local factors affecting the housing markets, as we assess in Section 4. The exposure is also predetermined, fixed in 2007, which minimizes the possibility of reverse causality.

Table *A4* shows the first stage of the 2-stage least square estimation, confirming the relevance of our instruments in the panel specification.

Table 3 shows the instrumental variable estimation of specification (2) for housing price growth. Figure 3 plots the estimated impulse response of prices. Our results show that immediately after the investors' purchases prices have positive growth, that continues after two years. Three years after the investors' purchases the price growth becomes negative and continues to be negative 6 years later.

Figure 4 plots the estimated impulse response of construction, measured by new building permits. The response of construction is positive at impact, and has a hump shape which picks after two years. The response stays positive after three years before it becomes zero. Figure 5 shows that, consistent with our theory, investors cause movements in prices only in MSAs with low price elasticity of supply. In MSAs where there is more developable land and fewer geographical restrictions to build new properties, we don't find evidence that investors have an effect on prices.

Consistent with the cross-sectional evidence, we find that investors have larger effects on the bottom-tier of the market. Figure 6 plots the estimated impulse responses for the top and bottom price tiers.

[8]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.

Table 4 shows that while investors have a positive and significant effect on price-to-income and rent-to-income ratios, this effect gets smaller in the medium term, until it disappears after six years. Figure 7 plots the response of the affordability variables.

Table $A5$ shows that our results are robust to using alternative instrumental variables for the share of investors. Table $A6$ shows that our results are robust to using alternative measures of investors' share based on number of purchases and number of units.

To quantify more accurately the effect on single-family homes, we restrict our variables of investors' share of purchases and housing prices to the single-family segment of the market. Table $A7$ shows that the response of prices to investors is exactly as statistically significant in the single-family segment as in the total market. Moreover, we restrict the investors' share of purchases to single-family homes and one-unit properties, for example a purchase of a condominium in a large apartment building. It is very likely that small investors purchase single-unit properties, instead of entire buildings. For this analysis we use the usual price index for all homes.[9] The lower panel of table $A7$ shows the results of the analysis for single-unit properties, which are again as statistically significant as the results for the total market.

Finally, 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. [10] Likely, our baseline controls, population, income, unemployment, MSA and year fixed effects, already capture a large part of the variation in housing demand.

To summarize, the positive effects of the investors's share of purchases on housing prices are offset in the medium-term by the response of construction. Consistent with the supply response, the impact on prices is magnified in areas with low housing supply elasticity.

4 Validity of the Instrument

In this section we defend the validity of our instrumental variable, by assessing in particular the exclusion restriction. We have already shown, in section 2.1, that the instrumental variable is relevant, as it is strongly correlated with the investors' share of purchases. Figure 8 provides visual support of the strong correlation between the instruments and the share of investors'

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

[10] We don't report the tables of these results, as they are similar to the previous results.

purchases over 2009-2017.

We address the exclusion restriction of the instrumental variable with the following exercises: (1) parallel trends analysis; (2) extensive local economy controls; (3) placebo tests; (4) inspection of correlation with standard drivers of housing markets; and (5) sensitivity of results to alternative instrumental variables.

4.1 Parallel trends

Figure 9 plots the annual log number of construction permits, annual rent growth, annual price growth and institutional investors' share for MSAs ranking in the top and bottom 25% of exposure to business income. The year 2008 is the critical year when the Fed implemented the first wave of unconventional monetary policy, which led to a large drop in interest rates. In the figures we notice a substantial divergence in the post-2008 housing variables between MSAs with high versus low exposure. However, prior to the shock, there are parallel dynamics between the high and low exposure groups. That is, the instrument appears to only be driving the investors, prices, rents and construction in the post-crisis period.

4.2 Local economy controls

To rule out the possibility that local economic conditions drive the results, we reestimate our baseline instrumental variables specification from Table 2 in Table *A*9 after controlling for a range of local business-cycle variables. In particular, Table *A*9 controls for four measures of contemporaneous economic activity in an MSA: average annual unemployment rate change, labor force participation growth, real gross domestic product (GDP) per capita growth, and median hourly wage per capita growth from 2009 to 2017.

Regardless of which measure we use, Table *A*9 shows that the point estimate for the effect of investors' share on price growth is consistently between 0.22 and 0.26 and statistically significant. Moreover, the various business-cycle measures all enter with the correct sign. This suggests that the regional business cycles and the institutional investors' market share are both important for price growth, but they operate independently.

4.3 Placebo tests

For the next exercise we perform placebo tests. In Figure 10 we visually inspect the impact of the instrument on the annual real housing price growth over 2009–2017 and the placebo pre-crisis period 2001-2006. 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 shown in the top panel of the figure. This evidence suggests that the instrument is not contaminated by pre-crisis price growth.

To assess the intuition from Figure 10, we conduct various placebo tests over the 2000–2006, 2001–2006, and 2000–2005 periods.[11] We ask if, when using a specification analogous to (1), the exposure to deposits and business income can explain housing price growth over any of these periods. We should expect no effect of our instruments on pre-crisis price growth because the instrument corresponds to shocks to investment in the U.S. housing market linked to the Fed's QE, unrelated to other drivers of housing prices. The placebo point estimates in Table $A10$ are insignificant across periods. This result suggests that the instrument is truly capturing post-crisis positive shocks in housing market investment.

Table $A12$ contains the results of placebo tests for our panel analysis, that show evidence that our instrumented investors do not contribute to price growth in other time periods pre-crisis. The corresponding visual evidence is provided in Figure $A2$.

4.4 Correlation with standard drivers of the housing market

We inspect the correlation of the instrumental variable with standard drivers of the housing market. Table $A11$ regresses the local share of business income on a variety of pre-crisis trends and MSA controls. 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.

While it is impossible to directly test the exclusion restriction, the results in Table $A11$ suggest that the instrument satisfies it as there is no relevant correlation between common

[11] 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.

drivers of housing price growth and other housing market variables and exposure to business income. Importantly, in our specifications we include an expansive set of controls.

4.5 Sensitivity to alternative instruments

Finally, we assess the sensitivity of the results to our instrumental variable. Our alternative instrument is based on credit constraints to households, and was used as an instrument for loan denials post-2009 in Gete and Reher (2018). This instrument exploits heterogeneity across MSAs in exposure to banking institutions that suffered regulatory shocks following the Dodd-Frank Act, approved in 2010. MSAs with greater exposure to these credit supply shocks experienced larger contraction in credit towards borrowers. In the market for credit, contraction of lending leads institutions to lower their deposit rates. This increases demand for alternative investments, such as investments in housing. The contraction of credit supply to households affects the share of institutional investors, not only through the market for credit, but also through demand for homeownership. Constrained households lower the demand for homeownership, which is likely to (mechanically) increase the share of institutional investors in the housing market, by reducing the number of total housing purchases.

This instrument exploits MSA exposure to lenders subject to a Comprehensive Capital Analysis and Review (CCAR) stress test from 2011 onwards. Formally, the instrument is the value of deposits' share in 2008 for lenders that underwent the CCAR stress test between 2011 and 2017 in MSA m multiplied by the difference in denial propensity between stress-tested and non stress-tested lenders in year $t-1$. We use the 2008 bank distribution, determined prior to Dodd-Frank, to minimize the risk of reverse causality. The CCAR test, like other stress tests, is meant to ensure that the largest bank holding companies have enough capital to weather a financial crisis, but as a side-effect it encourages those institutions to tighten their standards in mortgage markets (Calem, Correa, and Lee 2016).

Table $A5$ shows the estimates of (2), using the credit denial instrument. While the local projections are larger in magnitude, the results confirm that the response of prices is initially positive and becomes significantly negative in the medium term. Table $A5$ also shows results using both our baseline instrumental variable, based on business income, and the alternative one, based on credit constraints. Our results are robust to the use of both instrumental variables.

5 Conclusions

The explosive growth of investors in residential housing markets after the latest Global Financial Crisis, has been central to many affordability debates. Cities around the world have acknowledged the importance of investors and have designed policies to block their participation and improve affordability. This paper provides evidence that a pro-market solution to the affordability problems may be effective.

By analyzing 85 million housing transactions in the U.S. MSAs, this paper showed that the immediate response of price-to-income ratios to the investors' purchases is positive and economically significant. Coming out of the financial crisis, housing investors gave rise to new housing demand. The investors surged in the U.S. MSAs and worsened housing affordability in the short-term. 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.

However, the presence of investors triggered a strong equilibrium response of supply. One to three years after the investors' purchases, there was a substantial positive effect on new building permits. The presence of investors reversed the growth of price-to-income ratios. After five to six years the price-to-income ratio response became zero. In the medium term the investors helped to restore affordability at the initial levels. Consistent with the previous finding, this paper shows evidence that investors did not cause price increases in MSAs where there are loose supply restrictions.

The dynamic results can inform policies regarding regulating investment in housing markets. By letting markets come to an equilibrium, additional supply of housing may follow the demand from investors, which, in the medium-term, may restore the affordability levels.

References

ACCE Institute: 2018, Wall street landlords turn american dream into a nightmare.

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.

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

Brunson, S.: 2019, 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. Retrieved from https://www.businessinsider.com/california-pass-rent-control-bill-to-address-affordable-housing-crisis-2019-9?IR=T.

Calem, P., Correa, R. and Lee, S. J.: 2016, Prudential policies and their impact on credit in the United States.

Cvijanovic, D. and Spaenjers, C.: 2018, 'We'll always have Paris': Out-of-country buyers in the housing market, *Kenan Institute of Private Enterprise Research Paper* (18-25).

Daniel, K., Garlappi, L. and Xiao, K.: 2018, Monetary policy and reaching for income.

Davids, A. and Georg, C.-P.: 2019, The cape of good homes: Exchange rate depreciations, foreign demand and house prices.

Deng, Y., Qin, Y. and Wu, J.: 2019, Superstar cities and the globalization pressures on affordability, **45**.

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

Favilukis, J., Mabille, P. and Van Nieuwerburgh, S.: 2019, Affordable housing and city welfare.

Favilukis, J. and Van Nieuwerburgh, S.: 2018, Out-of-town home buyers and city welfare.

Garriga, C., Gete, P. and Tsouderou, A.: 2019, The arrival of buy-and-hold investors to housing markets.

Gay, S.: 2015, Investors effect on household real estate affordability.

Gete, P. and Reher, M.: 2018, Mortgage supply and housing rents, *The Review of Financial Studies* **31**(12), 4884–4911.

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

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

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

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M. and Taylor, A. M.: 2019, The rate of return on everything, 1870–2015, *The Quarterly Journal of Economics* **134**(3), 1225–1298.

Khwaja, A. I. and Mian, A.: 2008, Tracing the impact of bank liquidity shocks: Evidence from an emerging market, *The American Economic Review* **98**(4), 1413–1442.

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

Lambie-Hanson, L., Li, W. and Slonkosky, M.: 2018, Investing in Elm Street: What happens when firms buy up houses?, *Federal Reserve Bank of Philadelphia Economic Insights* pp. 9–14. Retrieved from https://www.philadelphiafed.org/-/media/research-and-data/publications/economic-insights/2018/q3/eiq318-elmstreet.pdf?la=en.

Lambie-Hanson, Lauren, W. L. and Slonkosky, M.: 2019, Leaving households behind: Institutional investors and the U.S. housing recovery.

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.

Raymond, E. L., Duckworth, R., Miller, B., Lucas, M. and Pokharel, S.: 2016, Corporate landlords, institutional investors, and displacement: Eviction rates in single-family rentals.

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 Guardian: 2019, Buyer beware: Amsterdam seeks to ban buy-to-let on newbuild homes. Retrieved from https://www.theguardian.com/world/2019/mar/18/buyer-beware-amsterdam-seeks-to-ban-buy-to-let-on-newbuild-homes.

The Wall Street Journal: 2017, Meet your new landlord: Wall street. Retrieved from 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. Retrieved from https://www.wsj.com/articles/in-berlin-a-radical-proposal-to-combat-rising-rents-expropriate-big-landlords-11554202800.

Verbrugge, R. and Gallin, J.: 2019, A theory of sticky rents: Search and bargaining with incomplete information.

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

Figures

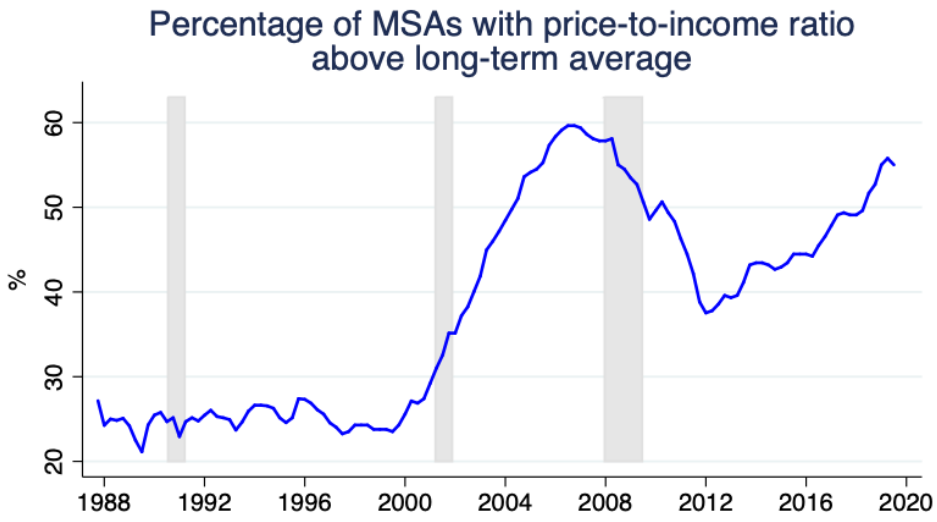

Figure 1. Housing affordability in the U.S. relative to long-term average. The figure plots the share of MSAs where price-to-income ratio is above three, which has been the national average in the U.S. over the period 1987-2019. That is, the share of MSAs where the median housing price is higher than three times the median annual household income, from the last quarter of 1987 to the second quarter of 2019. More specifically, the national average has been between 2.77 and 2.80 during all quarters up to 2001. It increased sharply after that and reached 3.00 in 2006, and peaked at 3.12 in 2010. In 2019 the long-term mean price-to-income ratio reached a new peak at 3.14. We plot quarters for which data are available for at least 300 MSAs. The gray areas illustrate the U.S. Recessions. The price-to-income ratios come from Zillow.

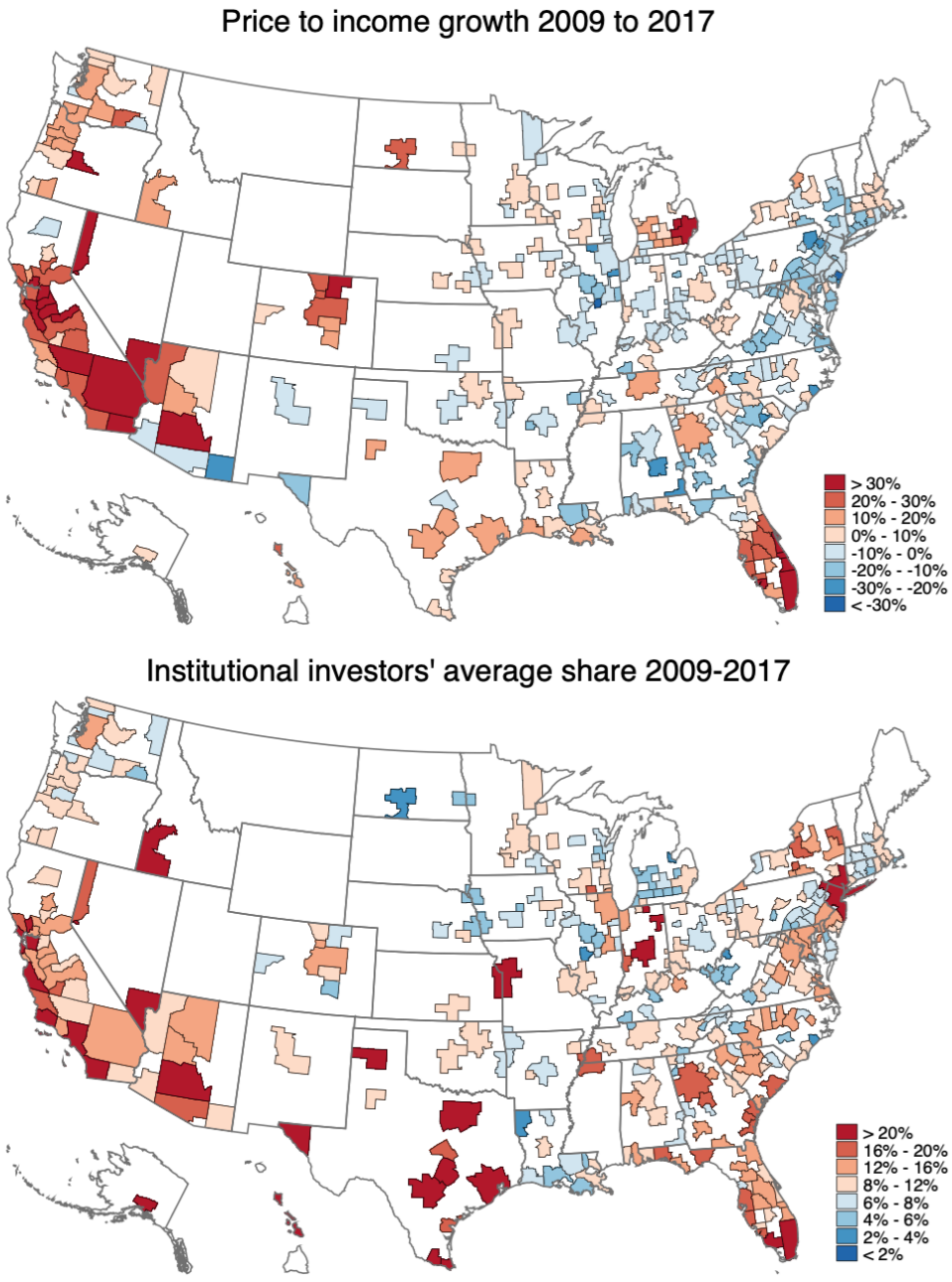

Figure 2. Affordability and institutional investors in the U.S. The top map shows the percentage growth of price-to-income ratio from 2009 to 2017 in each MSA. The bottom map shows the average market share of dollar purchases by institutional investors over 2009 to 2017 in each MSA. Source: ZTRAX, Zillow.

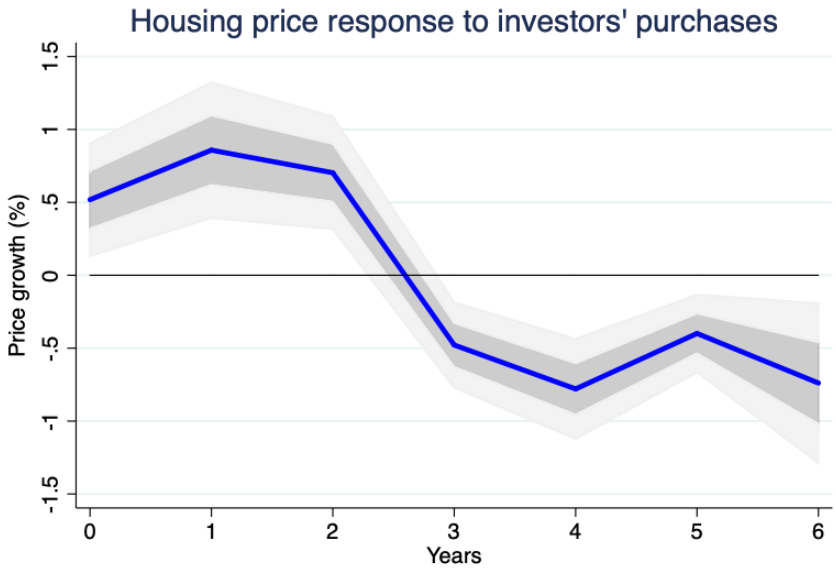

Figure 3. Dynamics of housing prices after investors' purchases. The figure plots the estimated local projections from sequential regressions of the real housing price growth on the instrumented past investors' share of purchases. Section 3 contains the methodology that follows Jordà (2005). The dark shaded area shows one standard deviation above and below the mean, and the light shaded area the 95% confidence interval. Table 3 contains the results.

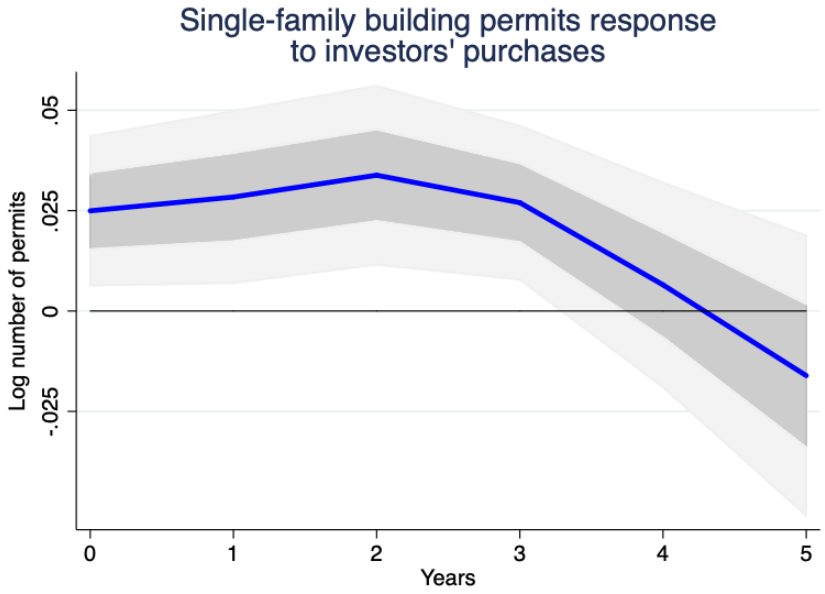

Figure 4. Dynamics of housing construction after investors' purchases. The figure plots the estimated local projections from sequential regressions of the log number of single-family building permits on the instrumented past investors' share of purchases. The dark shaded area shows one standard deviation above and below the mean, and the light shaded area the 95% confidence interval. Table A8 contains the results.

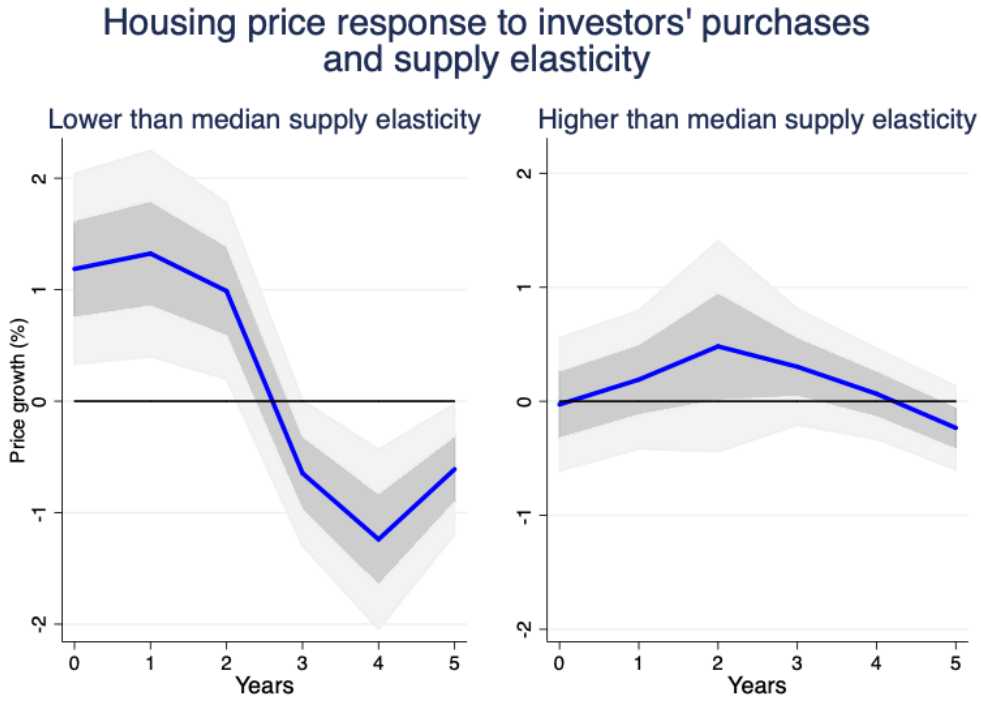

Figure 5. Dynamics of housing prices after investors' purchases and supply elasticity. The figure plots the estimated local projections from sequential regressions of the real housing price growth on the instrumented past investors' share of purchases. The left figure shows the response for MSAs that have supply elasticity below the median, and the right figure for MSAs that have supply elasticity above the median. The housing supply elasticity comes from Saiz(2010), and the median value for the MSAs in our sample is 2.2. The dark shaded areas show one standard deviation above and below the mean, and the light shaded area the 95% confidence interval.

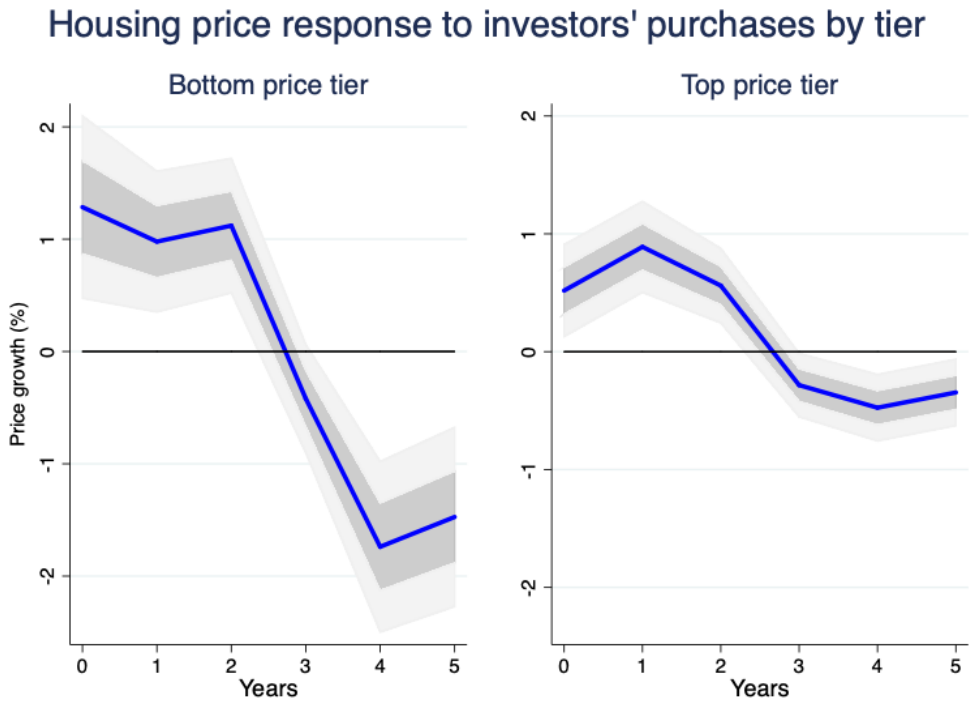

Figure 6. Dynamics of housing prices after investors' purchases by tier. The figure plots the estimated local projections from sequential regressions of the real housing price growth on the instrumented past investors' share of purchases for top and bottom price-tier houses. Top tier houses are houses in the top third, and bottom tier houses are houses in the bottom third of the house value distribution within an MSA. The dark shaded areas show one standard deviation above and below the mean, and the light shaded areas the 95% confidence interval.

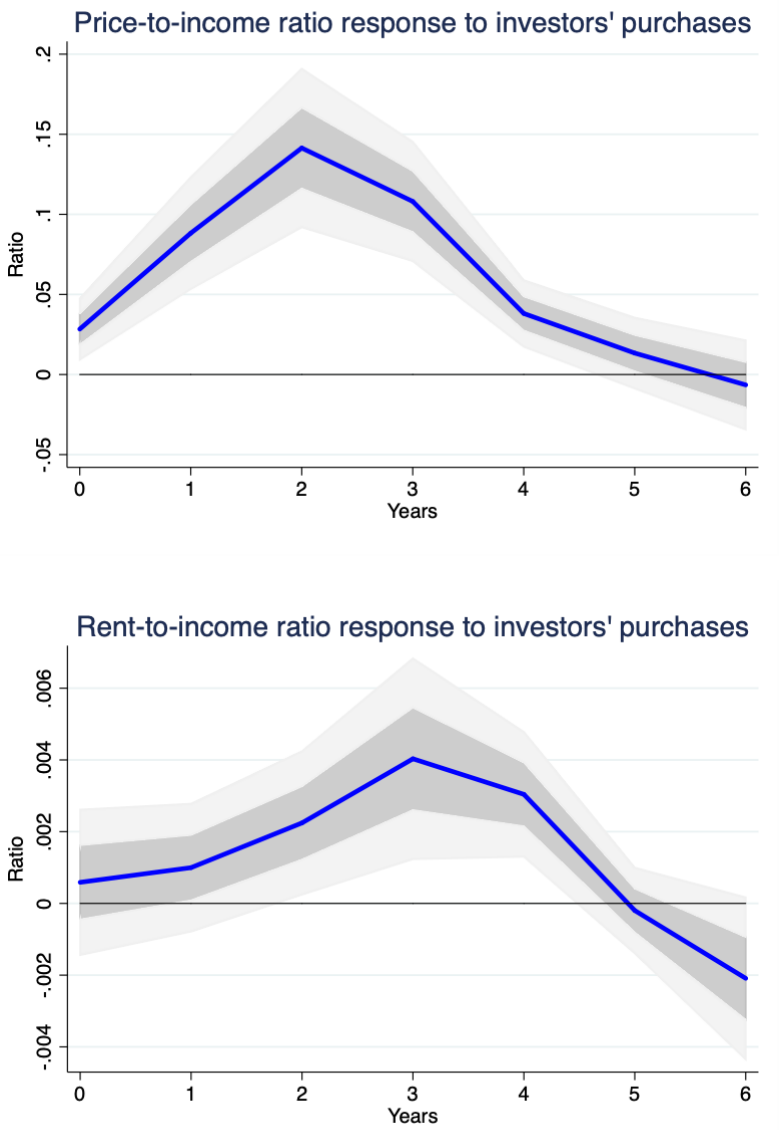

Figure 7. Dynamics of housing affordability after investors' purchases. The top figure plots the estimated local projections from sequential regressions of the housing price-to-income ratio, and the bottom figure the estimated local projections of the housing rent-to-income ratio, on the instrumented 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. The dark shaded areas show one standard deviation above and below the mean, and the light shaded areas the 95% confidence interval. Table 4 contains the results.

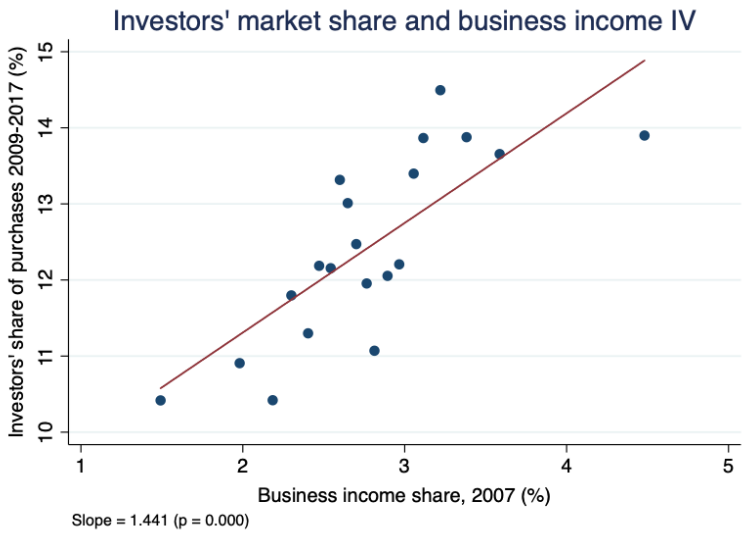

Figure 8. Institutional investors' market share and the instrumental variable. This figure plots the average share of value of business income over total income of top earners in an MSA in 2007, against the 2009-2017 average market share of institutional investors' purchases in each MSA. The top earners 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 2.

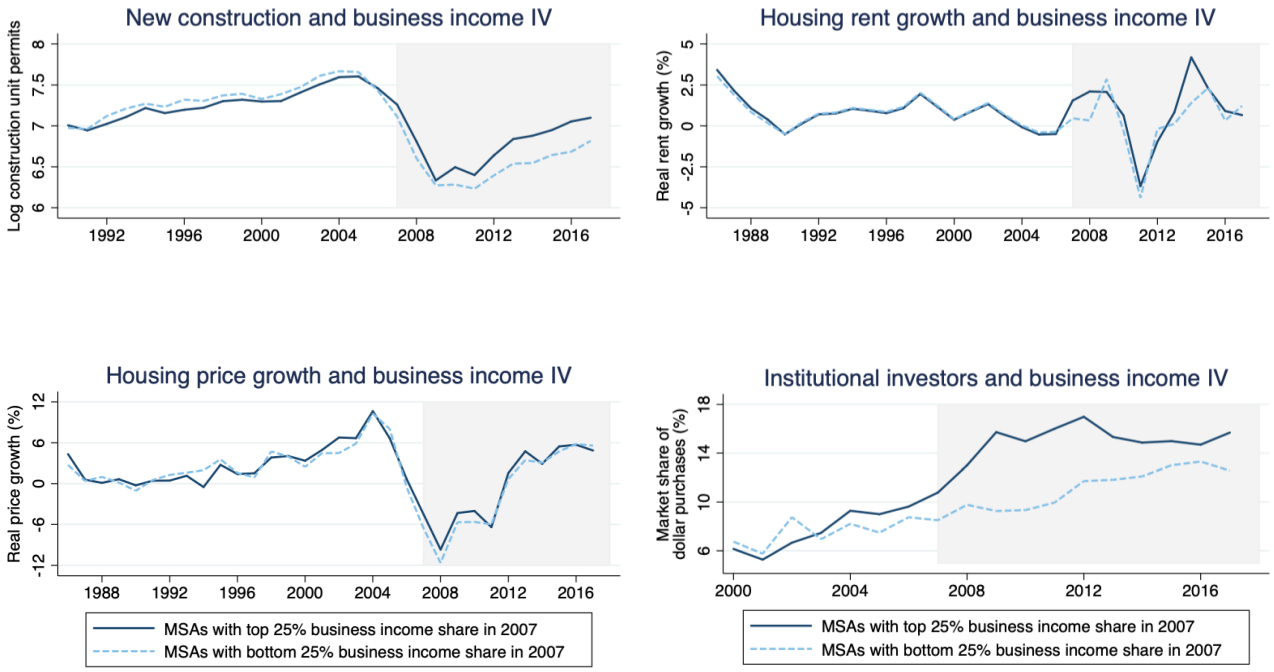

Figure 9. Parallel trends. The figures plot the time series of (1) the log number of new construction unit permits, (2) real rent growth, (3) real price growth and (4) market share of institutional investors' dollar purchases for MSAs ranking in the top and bottom 25% of exposure to our instrumental variable: the 2007 average share of business income over total income of the top earners in an MSA. The gray shaded area shows the period after the end of 2007, to denote the introduction of the Fed's QE in 2008.

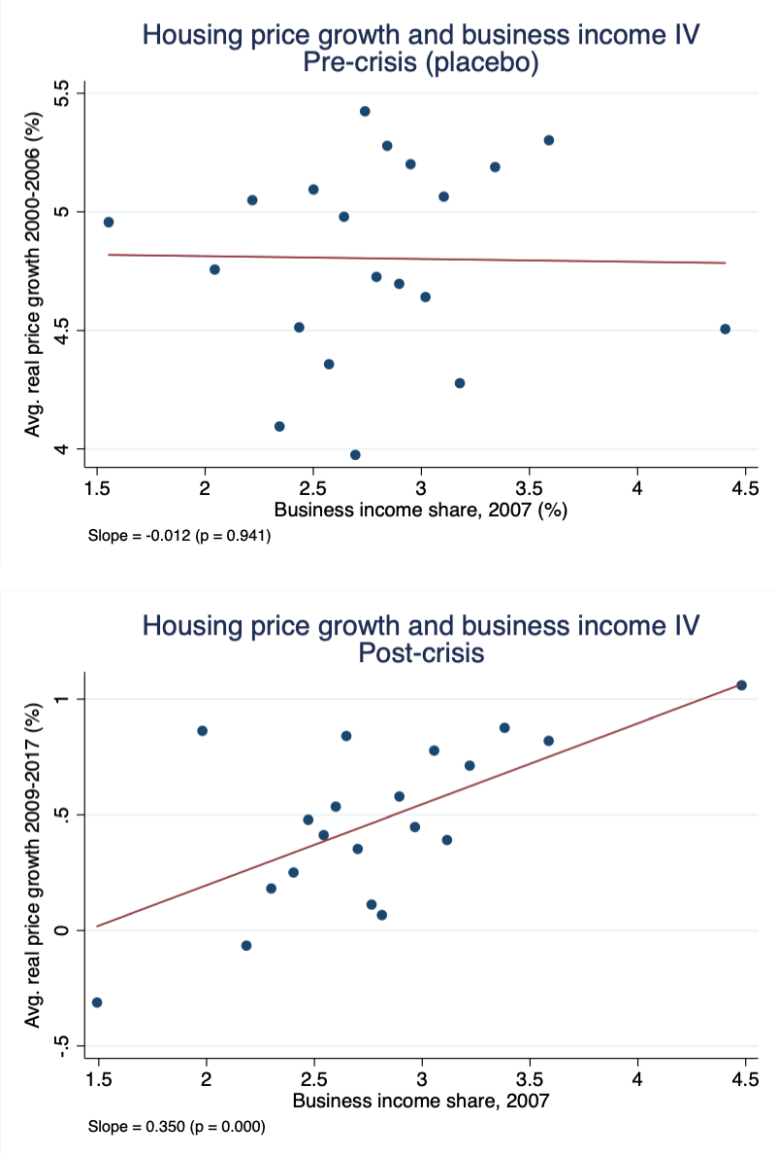

Figure 10. Housing price growth against the instrument for investors pre- and post-2008. The top panel plots the 2000–2006 average annual real housing price growth against the average share of business income over total income of top earners in each MSA in 2007. The bottom panel plots the 2009–2017 average annual real housing price growth against the same instrument. The top panel controls are the ones used in the placebo specification in Table A10. The bottom panel controls are the ones used in the baseline specification in Table 2. Figure A2 in the online appendix performs the same visual exercise for the panel version of the instrument.

Tables

Table 1. Summary statistics (panel data)

	Obs	Mean	SD	Min	Max
Investors' share of dollar purchases (%)	2,997	11.50	8.40	0.65	75.95
Real housing price growth (%)	2,901	0.47	6.67	-25.51	36.47
Top tier real price growth (%)	2,853	0.46	5.61	-24.92	28.41
Bottom tier real price growth (%)	2,610	0.13	9.87	-53.03	34.09
Single-family real price growth (%)	2,881	0.51	6.62	-25.44	36.52
Multi-family real price growth (%)	2,217	-0.04	8.61	-45.13	33.83
Real housing 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 building permits	2,997	6.46	1.36	1.10	10.58
Log single-unit building permits	2,997	6.42	1.37	0.69	10.55
Log vacancy rate owned	2,554	0.67	0.55	-2.60	2.49
Log vacancy rate rentals	2,558	1.91	0.53	-0.67	3.90
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
Business income share $(\%)_{07} \times$CD rate growth$_{t-1}$	2,997	-0.57	0.76	-4.98	1.58

This table presents summary statistics of the key variables in our estimations from 2009 to 2017. Detailed description of the variables and data sources is included in the Appendix A.

Table 2. Housing price affordability and investors' share

	Real price growth$_{m,09-17}$			Price-to-income ratio$_{m,09-17}$		
	Bottom Tier	Median	Top Tier	Bottom Tier	Median	Top Tier
Investors' share$_{m,09-17}$	0.293***	0.184***	0.153**	1.525***	0.547***	0.311***
	(0.089)	(0.063)	(0.062)	(0.336)	(0.146)	(0.088)
			Standardized			
Investors' share$_{m,09-17}$	0.909***	0.827***	0.768**	2.108***	1.679***	1.884***
	(0.276)	(0.285)	(0.308)	(0.464)	(0.447)	(0.534)
Observations	303	303	303	303	303	303

Heteroscedasticity robust standard errors are in parentheses. 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 bottom part of the table shows the estimated effects of the standardized independent variable on the standardized dependent variables. All models include state dummies and MSA-level controls: 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. 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 underidentification test of Kleibergen and Paap (2006) has p-value smaller than 0.0001, and the weak identification test Kleibergen and Paap Wald F statistic is 19.235. Table $A3$ contains the first stage of the IV regression. Each observation is an MSA. ***significant at the 1% level; **significant at the 5% level.

Table 3. Housing price growth in response to investors' purchases

	Real housing price growth$_{m,t+i}$						
	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
Total market							
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
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
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. 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 certificate of deposits 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 total market panel regression ($i=0$). Table $A4$ contains the first stage of the IV regression. Tables $A5$ and $A6$ contain the dynamic results using alternative IVs and alternative measures of the investors' presence. ***significant at 1%; **significant at 5%; *significant at 10%.

Table 4. 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 certificate of deposits rate growth. The sample period is 2009-2017. Each observation is an MSA-year. The coefficients of the regressions of rent-to-income on the share of investors are multiplied by 100 in this table. ***significant at 1%; **significant at 5%; *significant at 10%.

ONLINE APPENDIX (Not for Publication)

A Detailed Description of Database

In this section 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 new 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.[12] 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 A13 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(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-

[12] 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.[13] 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[14] 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

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

[14] 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.

To classify non-institutional buyers into *individual investors* we start from the ZTRAX classification of buyer names as names of individuals. We calculate the number of purchases of each individual name within the MSA within the given year and the year before. We define individuals that are investors as individuals who purchase more than one property within the MSA in the given year and the year before.

We calculate the market share of investors as the dollar value of investors' purchases (either individual or institutional, or local institutional investors) divided by the dollar value of all purchases, by institutional and individual investors and homeowners. 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 A13 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.

The holding duration is the duration between the purchase and sale of the property. We define long-term investments the purchases that are held for more than one year, and short-term investments the purchases that are sold within a year.

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 in the data. Tables A9 to A12 contain descriptives of the ZTRAX data (before merging with other variables).

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.[15] 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

[15] 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 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 occupied households and number of owner-occupied households. 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. Rental vacancy rate is the share of the number of vacant rental units over the total rental units.

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.

Other 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: Data come from the U.S. Census Bureau, from 1990 to 2017.
- Median age: Data on median age come from the American Community Survey One-Year

Estimates. The data only cover the 2005-2017 period. The momentum analysis use as control the MSA log median age in 1997. We approximate the median age using the U.S. Census 1997 total county population data. 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.

- Median Income: Data on come from the Zillow Median Household Income dataset, from 1990 to 2017.
- Unemployment rate: Data on unemployment rate come from the Bureau of Labor Statistics, from 1990 to 2017.

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 Variables

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. 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 4 provide strong evidence in this direction.

Households' credit constraints

In this subsection we describe the alternative instrumental variable we use in our robustness tests, which is based on households's credit constraints. This instrument exploits heterogeneity across MSAs in exposure to banking institutions that suffered regulatory shocks following the Dodd-Frank Act, approved in 2010, following Gete and Reher (2018). Formally, the instrument is the value of deposits' share in 2008 for lenders that underwent the CCAR stress test between 2011 and 2017 in MSA m multiplied by the difference in denial propensity between stress-tested and non stress-tested lenders in year $t-1$.

Lenders' propensity to deny

Following the methodology of Khwaja and Mian (2008), we estimate a fixed effect for a given lender or group of lenders. Specifically, let L denote the set of lenders we observe in HMDA, and consider a partition of L into disjoint subsets $l_1, l_2, ..., l_n$. In this case we partition lenders according to whether or not they underwent the CCAR test between 2011 and 2017. Then $l_1 = \{\text{Tested}\}, l_2 = \{\text{Not Tested}\}$.

To extract a credit supply shock experienced by lenders of set l_j, we estimate the probability of loan denial at the application level, $\Pr(\text{Denied}_{i,m,t,l_j} = 1)$, as a linear probability model:

$$\Pr(\text{Denied}_{i,m,t,l_j} = 1) = \sum_j \Lambda_{t,l_j} + \gamma X_{i,m,t,l_j} + \alpha_{m,t} + \alpha_{m,l_j}, \tag{A1}$$

where our focus is on the Λ_{t,l_j}, which is a vector of fixed effects for lenders of set l_j in year t.[16] The controls in X_{i,m,t,l_j} account for the characteristics of borrowers: income, requested loan-to-income, and race of borrower i applying for a loan from lender type l_j in MSA m in year t.[17] The terms $\alpha_{m,t}$ and α_{m,l_j} control for lender, time, and regional shocks. The value $\alpha_{m,t}$ is the coefficient on an indicator variable which equals 1 if the borrower applies from MSA m in year t and equals 0 otherwise. Likewise the indicator variable α_{m,l_j} equals 1 if the borrower applies from MSA m to a lender of type l_j and equals 0 otherwise.

The vector Λ_{t,l_j} captures the lender specific component of denial rates. For example, it may reflect a higher cost of funds or greater regulatory risk borne by lenders of set l_j in a given year. Importantly, Λ_{t,l_j} does not confound either borrower or regional effects, since these are already captured by X_{i,m,t,l_j} and the pair $(\alpha_{m,t}, \alpha_{m,l_j})$, respectively. To emphasize this interpretation,

[16] We estimate the Λ_{t,l_j} using a series of indicator functions for whether the application was received by lenders of set l_j in year t. The reference category will be applications to lenders of some set l_r in some year t_r.

[17] We use 21,709,935 observations to estimate (A1) over 2007-2014.

we refer to Λ_{t,l_j} as the propensity to deny.

Stress-test shock

We estimate first (A1) using the partition $L = \{\text{Tested, Not Tested}\}$ to get the time-varying part of our instrumental variable, which is unrelated to factors affecting the local housing markets. To construct the panel version of the instrumental variable, we multiply the time-varying propensity to deny a loan, by the pre-determined exposure of each MSA to stress-tested lenders:

$$S_{m,t} = (\Lambda_{t,\text{Tested}} - \Lambda_{t,\text{Not Tested}}) \times \text{Stress-Test Share}_{m,08}. \tag{A2}$$

We define stress-tested lenders as those which underwent a CCAR test between 2011 and 2017, and Stress-Test $\text{Share}_{m,2008}$ as the 2008 value of deposits' share of these lenders. In words, $S_{m,t}$ captures the relative stringency of stress-tested lenders in a given year $(\Lambda_{t,\text{Tested}} - \Lambda_{t,\text{NonTested}})$ and the degree to which this tightening is felt in a given MSA, as measured by the deposit share of the tested lenders, Stress Test $\text{Share}_{m,08}$. Much of the temporal variation in credit tightness occurred after 2010.

We use the stress test shock instruments to estimate (2), using our full panel data from 2009 to 2017. Finally, we lag the stress-test shock by one period, following Gete and Reher (2018) and Favara and Imbs (2015).

Extra Figures

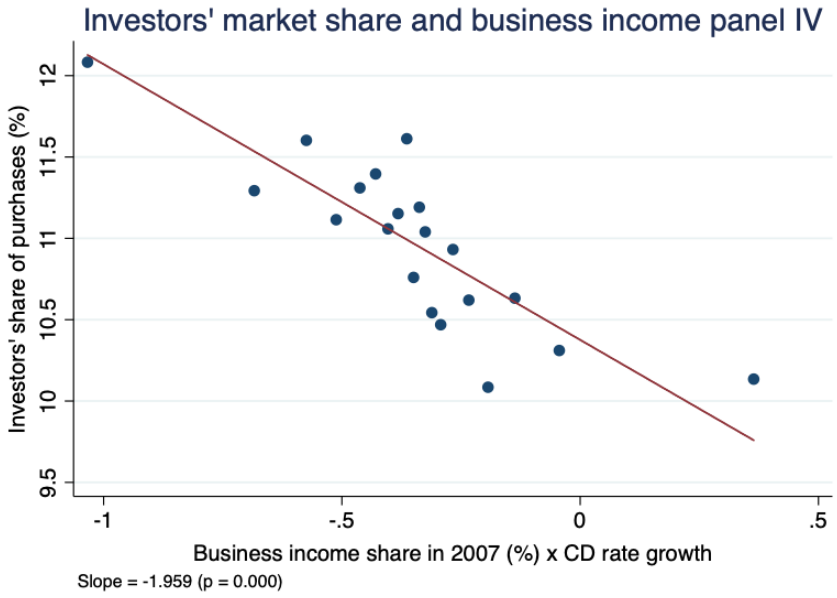

Figure A1. Institutional investors' market share and the instrumental variable. This figure plots the share of value of business income over total income of top earners in an MSA in 2007 multiplied by the CD rate growth, against the market share of institutional investors' purchases each year in each 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 3.

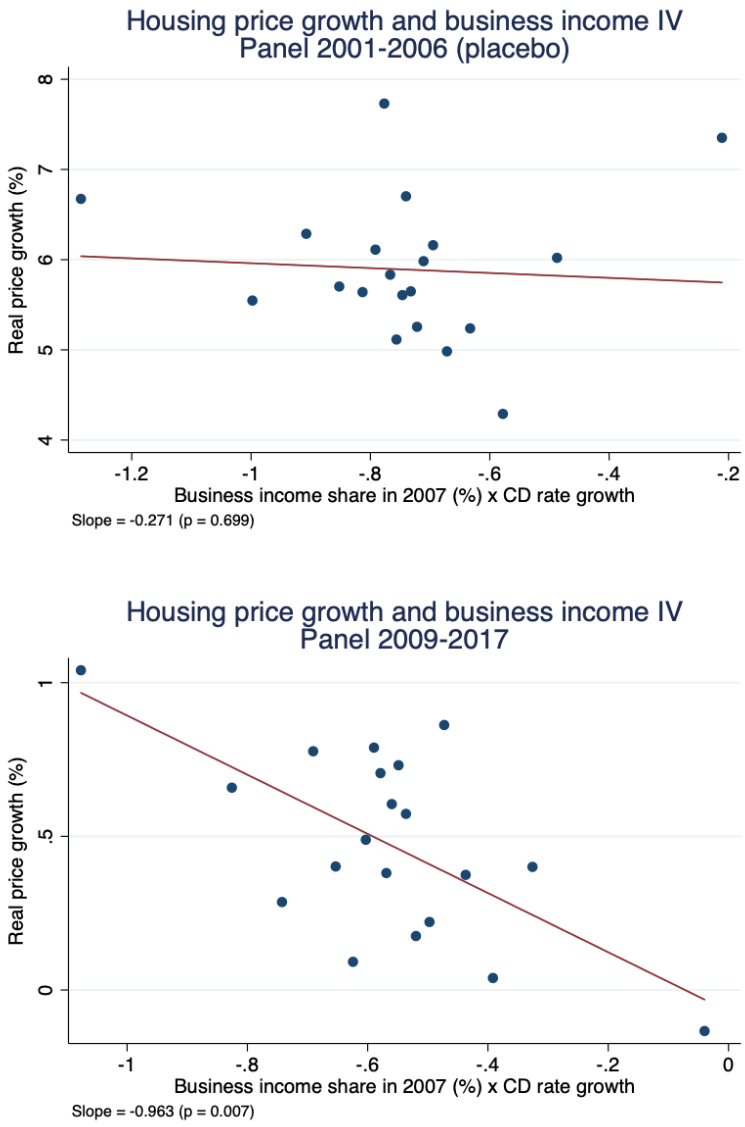

Figure A2. Pre- and post-2008 housing price growth against the panel instrument for investors. The top panel plots the annual real price growth over the 2001-2006 period against the panel instrument: the average share of business income over total income of top earners in each MSA in 2007 multiplied by the lagged certificate of deposits 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 A12. The bottom panel controls are the ones used in the panel specification in Table 3.

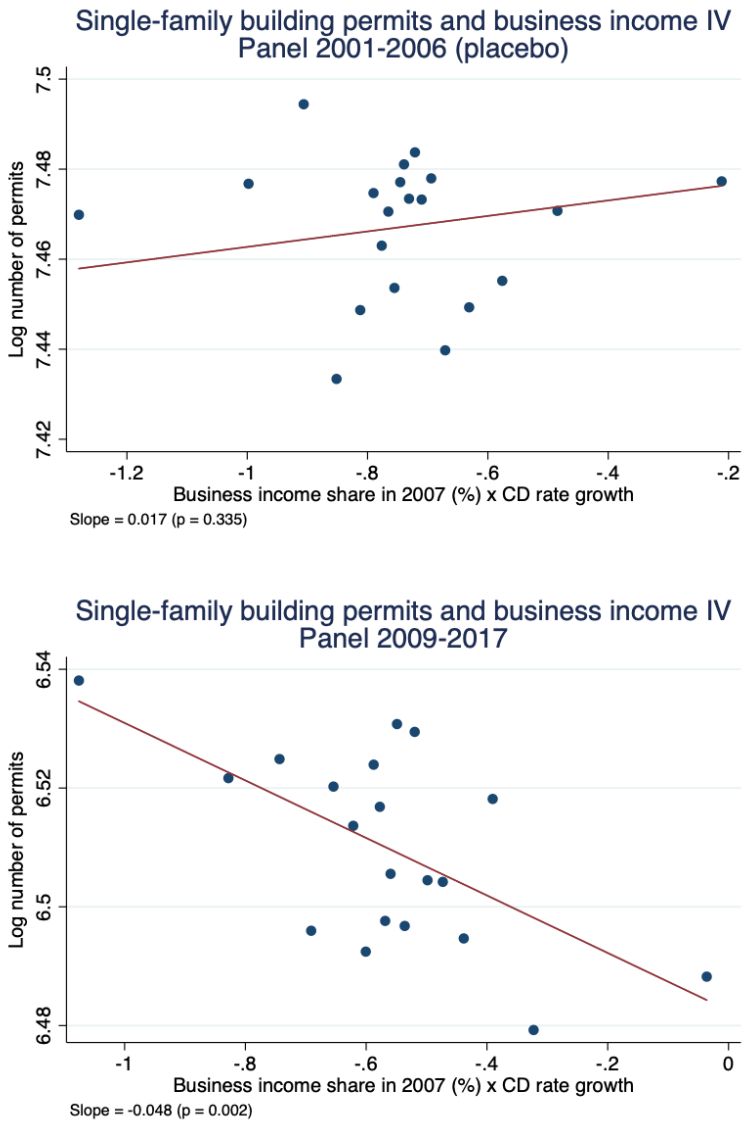

Figure A3. Pre- and post-2008 building permits against the panel instrument for investors. The top panel plots the log number of building permits for single-family homes over the 2001-2006 period against the panel instrument: the average share of business income over total income of top earners in each MSA in 2007 multiplied by the lagged certificate of deposits rate growth. The bottom panel plots the log number of building permits for single-family homes over the 2009-2017 period against the same instrument. The top panel controls are the ones used in the placebo panel specification in Table A12. The bottom panel controls are the ones used in the panel specification in Table 3.

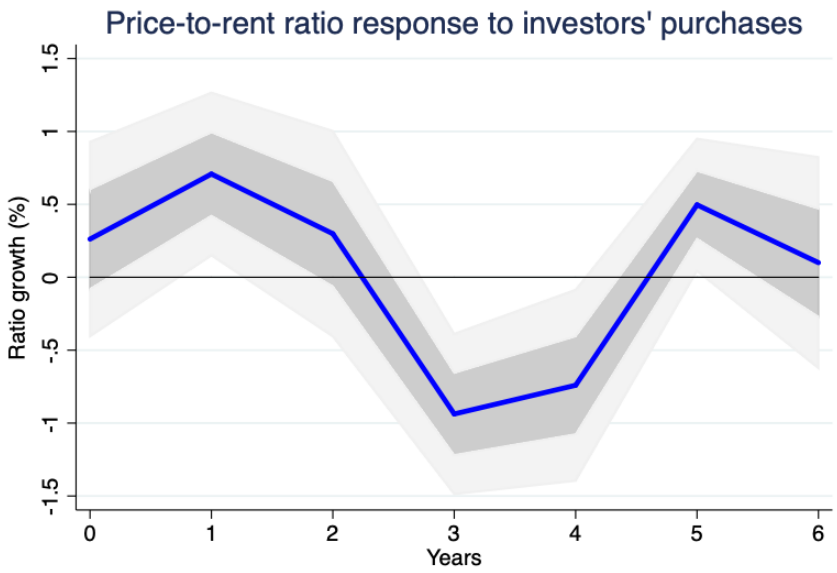

Figure A4. Dynamics of housing price-to-rent ratio growth after investors' purchases. The figure plots the estimated local projections from sequential regressions of the housing price-to-rent ratio growth on the instrumented past investors' share of purchases. The price-to-rent ratio is calculated as the median housing price over the median housing rent in an MSA, and the growth is the annual percentage growth. The dark shaded area shows one standard deviation above and below the mean, and the light shaded area the 95% confidence interval.

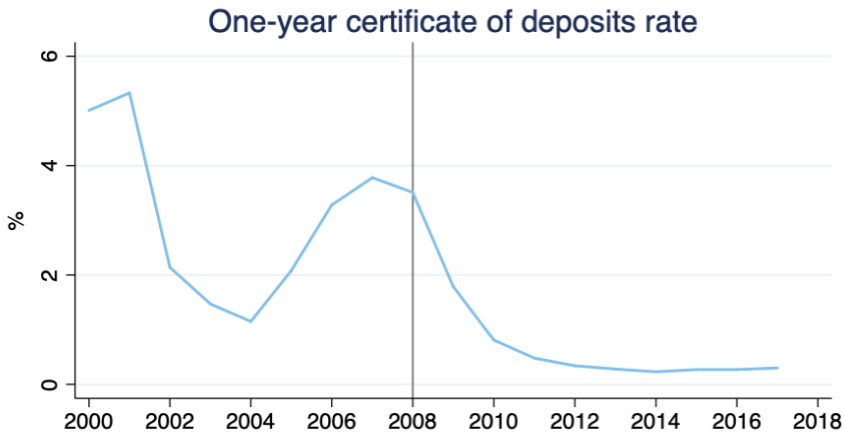

Figure A5. Interest rate used in the construction of the panel instrument. The figure plots the one-year CD rate, annually from 2000 to 2018. Source: Bankrate.

Extra Tables

Table A1. Housing price growth and investors' share

	Avg. real housing price growth$_{m,09-17}$	
Avg. investors' share$_{m,09-17}$	0.034**	0.243***
	(0.015)	(0.083)
Underidentification test (p-value)		0.000
Weak identif. test F statistic		19.430
Estimation	OLS	IV
Observations	332	332

Heteroscedasticity robust standard errors are in parentheses. All models include state dummies and MSA-level controls: 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. 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 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. Table $A3$ contains the first stage of the IV regression. ***significant at the 1% level; **significant at the 5% level.

Table A2. Housing supply and liquidity

	Avg. log building permits$_{m,09-17}$	Avg. log owner-occupied vacancy rate$_{m,09-17}$	Avg. log rental vacancy rate$_{m,09-17}$
Avg. investors' share$_{m,09-17}$	0.057**	-0.058**	-0.057***
	(0.024)	(0.023)	(0.020)
MSA-level controls	Yes	Yes	Yes
State dummies	Yes	Yes	Yes
Estimation	IV	IV	IV
Observations	332	327	327

Heteroscedasticity robust standard errors are in parentheses. The controls are as in Table A1. 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. ***significant at the 1% level; **significant at the 5% level.

Table A3. First stage: Investors' share and the instrumental variable

	Avg. investors' share of purchases$_{m,09-17}$
Business income share$_{m,07}$	1.441***
	(0.327)
MSA-level controls	Yes
State dummies	Yes
R-squared	0.689
Observations	332

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

	Investors' share$_{m,t-1}$
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. Alternative panel IVs

	Real housing price growth$_{m,t+i}$						
	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
IV2							
Investors' share$_{m,t-1}$	1.44***	1.41***	1.92***	0.48**	-1.29***	-0.38*	-0.97**
	(0.35)	(0.32)	(0.44)	(0.22)	(0.35)	(0.22)	(0.44)
Observations	2,536	2,253	1,969	1,687	1,405	1,122	840
IV1 & IV2							
Investors' share$_{m,t-1}$	1.45***	1.33***	1.62***	0.15	-1.20***	-0.42**	-0.89***
	(0.34)	(0.26)	(0.33)	(0.17)	(0.27)	(0.17)	(0.31)
Observations	2,536	2,253	1,969	1,687	1,405	1,122	840

Standard errors clustered by MSA are in parentheses. The fixed effects and MSA-year level controls are as in Table 3. The alternative instrument (IV2) for the investors' share of purchases is the credit denial instrument of Gete and Reher (2018): the mortgage application share of stress-tested lenders in MSA m in 2008 multiplied by the difference in denial propensity between stress-tested and non stress-tested lenders in year $t-1$.Stress-tested lenders are those subject to the CCAR test between 2011-2017. The bottom part of the table uses two instrumental variables: our baseline instrument (IV1) of the average share of business income over total income of the top earners in MSA m in 2007 multiplied by the lagged one-year CD rate growth; and the alternative IV2. In the case of the two IVs, the Hansen J statistic overidentification test has p-value of 0.901. The sample period is 2009-2017. Each observation is an MSA-year. ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table A6. Alternative measures of investors

	Real housing price growth$_{m,t+i}$						
	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
Inv. share of no. 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
Inv. share of no. 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 3. 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 certificate of deposits 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 A7. Single-family properties

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
	Single-family real housing price growth$_{m,t+i}$						
Inv. 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
	Real housing price growth$_{m,t+i}$						
Inv. 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 3. 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 certificate of deposits 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 A8. New building permits and rents

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
	Log new single-family building permits$_{m,t+i}$						
Investors' share$_{m,t-1}$	0.25***	0.28***	0.34***	0.27***	0.07	-0.16	-0.39
	(0.09)	(0.11)	(0.11)	(0.10)	(0.13)	(0.18)	(0.25)
Observations	2,850	2,533	2,216	1,899	1,582	1,264	948
	Log new building permits$_{m,t+i}$						
Investors' share$_{m,t-1}$	0.26***	0.28***	0.33***	0.24**	0.05	-0.15	-0.37
	(0.09)	(0.10)	(0.11)	(0.10)	(0.14)	(0.19)	(0.25)
Observations	2,850	2,533	2,216	1,899	1,582	1,264	948
	Real housing rent growth$_{m,t+i}$						
Investors' share$_{m,t-1}$	0.67**	0.39**	0.51	0.42*	-0.11	-0.92***	-0.72*
	(0.33)	(0.18)	(0.36)	(0.23)	(0.22)	(0.28)	(0.38)
Observations	2,580	2,293	2,006	1,719	1,432	1,144	858

Standard errors clustered by MSA are in parentheses. The coefficients and standard errors for the log building permits are multiplied by 10 in this table. The fixed effects and controls are as in Table 3. 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 certificate of deposits rate growth. The dependent variable in the first panel is the log number of building permits for single-family homes, whereas in the second panel for all residential building types, single- and multi-family. The sample period is 2009-2017. Each observation is an MSA-year. ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table A9. Robustness to including other local drivers

	Avg. real housing price growth$_{m,09-17}$			
Avg. investors' share of purchases$_{m,09-17}$	0.259***	0.252***	0.223***	0.247***
	(0.084)	(0.085)	(0.083)	(0.084)
Avg. unemployment rate change$_{m,09-17}$	-3.072**			
	(1.268)			
Avg. labor force participation growth$_{m,09-17}$		-0.006		
		(0.208)		
Avg. real per cap. GDP growth$_{m,09-17}$			0.204	
			(0.126)	
Avg. per cap. wage growth$_{m,09-17}$				0.001
				(0.189)
Estimation	IV	IV	IV	IV
MSA-level controls	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
Underidentification test (p-value)	0.000	0.000	0.000	0.000
Weak identification test F statistic	20.747	21.058	19.276	21.231
Observations	332	331	332	332

Heteroscedasticity robust standard errors are in parentheses. Avg. unemployment rate change$_{m,09-17}$ denotes the average unemployment rate change in MSA m over 2009-2017. Avg. labor force participation growth$_{m,09-17}$, Avg. real per cap. GDP growth$_{m,09-17}$ and Avg. per cap. wage growth$_{m,09-17}$ denote the average annual growth rate of those variables in MSA m over 2009-2017. The controls and instrumental variable are 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.

Table A10. Placebo: Housing price growth and investors' share pre-crisis

	Avg. real housing price growth$_{m,[t_1,t_2]}$		
$[t_1, t_2]$	2000-2006	2001-2006	2000-2005
Avg. 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

Heteroscedasticity 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 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 A11. The instrumental variable and drivers of housing markets

	Business income share$_{m,07}$
Avg. unempl. rate change$_{m,09-17}$	0.090
	(0.105)
Avg. median age change$_{m,00-06}$	0.026
	(0.046)
Avg. homeownership rate change$_{m,00-06}$	0.009
	(0.045)
Median age change$_{m,08}$	-0.056
	(0.049)
Homeownership rate change$_{m,08}$	0.073
	(0.056)
Real housing rent growth$_{m,08}$	-0.022
	(0.057)
MSA-level controls	Yes
State dummies	Yes
R-squared	0.533
Observations	280

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

	Real housing 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 $t-1$. 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 certificate of deposits 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 A13. 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.[18]

[18] 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".