Institutional investors in the market for single-family housing: Where did they come from, where did they go?

Sebastian Hanson*

Current draft: December 10, 2024†

Latest version available here

Abstract

Since 2012 the U.S. market for single-family homes has experienced a large and unprecedented rise in rental purchases by institutional investors or "Wall Street Landlords". My paper causally attributes the entry of these financially unconstrained investors to an increase in expected excess returns that is driven by (i) declining long-term interest rates and (ii) tighter household funding constraints. After entry of these investors into a local market, house prices and rents grow 2pp and 1.2pp faster per year. The majority of faster house price growth can be accounted for by market timing and would likely also have occurred in the absence of institutional investors. Faster rental growth cannot be accounted for by market timing, suggesting that institutional investors use their size to extract markups in rental markets. I replicate this evidence in a stylized model of the residential housing market in which funding-constrained households bid against unconstrained investors.

*shanson2@stanford.edu, Graduate School of Business, Stanford University. I am deeply grateful to Hanno Lustig for his guidance and continued encouragement. This paper greatly benefitted from discussions and comments from Greg Buchak, Susan Cherry, Antonio Coppola, Rebecca Diamond, Daniel Greenwald, Brian Higgins, Arvind Krishnamurthy, Christian Kontz, Monika Piazzesi, Joshua Rauh, Amit Seru, Jann Spiess and Vikrant Vig, as well as many useful questions and suggestions by seminar and conference participants at Stanford University, the 2023 European and American Meetings of the Urban Economics Association and the 2023 Transatlantic Doctoral Conference. I am deeply grateful to Meghan Phillips and Amber Brabender of CBRE Econometric Advisors for granting me access to their data and to SIEPR for financial support through a Bradley graduate fellowship. All errors are my own.

†First draft available online: December 9, 2022.

The U.S. market for single-family rentals is worth approximately 6.5 trillion U.S. dollars (CoreLogic, 2023; US Census, 2024) and has historically been dominated by small-scale investors, or "mom-and-pop" landlords. In the aftermath of the Great Recession the U.S. single-family rental market experienced a large and unprecedented rise in purchases by institutional investors or "Wall Street Landlords". In less than 10 years a small number of private equity firms and real estate investment trusts collectively acquired over 350,000 single-family homes. In Atlanta alone, these firms own over 50,000 homes.

Commentators and policymakers from both sides of the political aisle have raised concerns about this development. Common concerns are that these firms raise house prices and rents for households.[1] Whether these concerns are warranted is an open question that has proven challenging to answer.

While house prices and rents have risen disproportionately in markets where institutional investors operate, it is unclear whether this can be attributed to the *impact* of these firms. An alternative explanation is that institutional investors are simply *timing* their entry into markets that would also experience rapid growth in their absence. To disentangle the effects of market timing and market impact, we need to understand why these firms enter into some neighborhoods and not into others. However, what explains the entry (and absence) of institutional investors across time and space is not well understood.

This paper makes two contributions: First, I causally attribute the entry of institutional investors to an increase in expected returns driven by tighter household funding constraints. When households in a particular market have limited access to leverage, house prices cannot fully adjust upward when interest rates decline. Instead, declines in the risk-free rate are passed through into higher expected excess returns in equilibrium. In the aftermath of the housing bust these forces reinforced each other: long-term interest rates declined to all-time lows while lending standards increased to all-time highs. Together, these forces raised expected excess returns in areas that were particularly exposed to changes in lending standards to a sufficiently high level for institutional investors to enter.

Second, I decompose differences in house price and rental growth observed upon entry of institutional investors into a market timing and market impact effect. Compared to markets that institutional investors did not enter, house prices and rents grow 2.03pp and 1.16pp faster per year. Market timing can account for the majority of faster house price growth. Market timing, however, *cannot* account for faster rental growth. I interpret this as evidence to suggest that institutional investors use their size to extract rental markups.

[1] see e.g. New York Times, 10/19/2022 or House Committee hearing 06/28/2022, retrieved 07/16/2023.

I develop a portfolio choice model with funding-constrained households and unconstrained investors in the style of Frazzini and Pedersen (2014) to explain the entry, expansion and exit decisions of institutional investors. Agents invest in housing if the returns to housing exceed their opportunity cost of capital. Non-tradable benefits of home-ownership give households an additional incentive to own homes: in the absence of funding constraints they own the entire housing stock. Investors enter the housing market if household funding constraints - the gap between the value of a home to households and the funds at their disposal - are sufficiently tight. All else equal, household funding constraints are tighter if households have less wealth or access to leverage. As long as household wealth is less sensitive to changes in the interest rate than home values, which have high duration, household funding constraints also tighten when interest rates decline. Investors who buy homes in anticipation of high future returns raise house prices. Once in the market, incumbent investors benefit from returns to scale that make expansion profitable even if expected returns decrease.

The model delivers three directly testable predictions: First, expected excess returns and returns to scale explain when and where institutional investors enter, expand, and exit local housing markets. Second, expected excess returns are higher if households are poorer, more restricted in access to leverage and risk-free rates are lower. Third, house prices and rents increase upon entry of institutional investors into a market due to both market timing and potential market impact. Market timing and market impact are positively correlated and increasing in the share of institutional investor-owned housing stock.

To take the predictions of the model to the data, I need to confront three challenges. The first challenge is measuring the activity of institutional investors. I measure the activity of institutional investors at the property-level in CoreLogic housing deeds data. Most activity of institutional investors is not directly observable in the deeds data because they buy and hold properties through corporate subsidiaries. I solve this problem by tracing property ownership from subsidiaries back to the parent company through SEC filings and PropStream, a property investment platform.

The second challenge is measuring expected excess returns. I infer expected excess returns from transformations of the rent-price ratio, following an established literature in asset pricing (see, e.g. Campbell and Shiller, 1988; Campbell et al., 2009). I verify that my findings are not dependent on the method used to construct expected excess returns by constructing two alternative measures of expected excess returns that yield qualitatively identical results. The biggest obstacle in this exercise is limited availability of high-quality rental data. I address this challenge by using a quarterly repeat-sales rent index provided by CBRE Econometric Advisors. This rent index covers 814 submarkets of 69 metro areas from 2000Q1 to 2021Q4 that form the geographic unit of observation of my analysis.

The third challenge is that I need plausible exogenous variation in expected excess returns. To address this I construct a novel shift-share instrument for household access to leverage that is motivated by the predictions of my model. The instrument interacts a time-varying U.S.-wide shifter for lending standards with pre-determined exposure weights that vary locally in the cross-section to create exogenous variation in access to leverage. The aggregate shifter is constructed from the national 10th percentile of approved mortgage credit scores, a proxy for the minimum credit score necessary to access the mortgage lending market. The exposure weights are constructed by fixing the local distribution of household credit scores in a pre-period. The constructed variable represents the local share of households that would have had access to leverage given the pre-determined *local* distribution of household credit scores and changes in *aggregate* lending standards. If the predictions of my model are correct, areas with larger population shares of households that lose access to leverage should see larger increase in expected excess returns when lending standards increase.

I use the shift-share instrument to estimate a Two Stages Least Squares specification that jointly tests the relationship between household funding constraints, expected excess returns, and the investment decisions of institutional investors. The intuition of the joint test is that the first and second stage each test a separate prediction of my model. The first stage exploits exogenous variation in access to leverage to test whether household funding constraints explain expected excess returns. The second stage tests how institutional investors respond to expected excess returns.

To ensure that my results are not contaminated by unobserved confounders I saturate the regression with metro area-by-quarter and submarket fixed effects. This removes time-invariant heterogeneity across submarkets, such as housing stock characteristics, and restricts the variation to comparisons of submarkets within the same metro area each quarter. I further include a rich set of time-varying controls that potentially affect the investment decisions of institutional investors, such as distance to existing operations, income, foreclosure sales, property crime rates, house price growth, rental growth and vacancy rates. The identifying assumption is that, conditional on controls and fixed effects, differences in the initial distribution of credit scores do not affect the investment decisions of institutional investors for reasons other than higher expected excess returns.

Access to leverage is a strong predictor of expected excess returns in the first stage. All else equal, a 10pp (0.83sd) decrease in the share of households with access to leverage raises expected excess returns by 108bp (0.35sd). Consistent with market timing and market impact, expected excess returns are slightly lower in areas that are close to existing institutional investor-owned properties.

Expected excess returns and proximity to institutional investor-owned properties can causally explain the investment decisions of institutional investors. A 100bp increase in expected excess returns increases the probability of institutional investor investment by 5.2pp. An increase in distance to existing institutional investor operations lowers the probability of institutional investor investment, consistent with localized efficiency gains from scale (see, e.g. Highfield et al., 2021).

I further test whether household funding constraints explain variation in expected excess returns across time and space by estimating first stage specifications that drop time-fixed effects. This allows me to include controls for the risk-free rate and test how much of the variation that is otherwise absorbed by fixed effects can be explained by household funding constraints. Household funding constraints explain between 46% to 67% of total variation in expected excess returns. As predicted, declines in the risk-free rate are passed through into expected excess returns approximately one-to-one. A variance decomposition shows that the risk-free rate and its covariance with access to leverage account for over 85% of the time-series variation in expected excess returns within local housing markets.

Motivated by my model and previous findings I decompose differences in house price and rental growth observed upon entry of institutional investors into market timing and market impact effects. I use a difference-in-differences design to compare house price and rental growth upon entry of institutional investors into a local market. I place upper and lower bounds on the extent of market timing and potential market impact by saturating the comparison with controls for predictors of future returns. This allows me to compare two local markets, say A and B, that would have been equally attractive to investors from a market timing perspective, given observable predictors of future returns. If institutional investors enter in A, but not in B, and prices in A grow faster, this is evidence of market impact specific to institutional investors. This exercise requires that *conditional on predictors of future returns* the entry of institutional investors is driven only by variables that are orthogonal to future rent and house price growth. In practice, these variables are factors such as proximity to existing operations and characteristics of the housing stock (see also

Raymond, 2023; Gorback et al., 2023).

On average, house prices grow 2.03pp per year faster after institutional investors enter. The majority of faster house price growth can be accounted for by market timing. Relative to markets that would have been equally attractive to invest in, house prices do not grow substantially faster. This does not imply that institutional investors have no effect on house prices. Rather, it implies that they do not impact house prices in ways that are substantially different from other, smaller investors who would invest in this market in their absence (Garriga et al., 2023a). Consistent with my model predictions market timing and market impact effects are increasing in the share of institutional investor-owned housing stock.

Rents grow 1.16pp faster per year on average after institutional investors enter. However, market timing *cannot* account for faster rental growth. Compared to markets that would have been equally attractive to invest in, rents grow 1.21pp per year faster. These effects are also increasing in the share of institutional-investor owned housing stock. I interpret these findings as evidence that institutional investors use their size to extract markups in the rental market (see also Gurun et al., 2022; Austin, 2023; Gorback et al., 2023).

I further test whether the entry of institutional investors has effects on home-ownership rates, income, and the construction of new housing supply. Relative to the rest of the U.S., the areas that institutional investors enter are on persistent downward income trends and experience significant declines in home-ownership rates. Relative to areas that would be similarly attractive to invest in, income and home-ownership rates follow an identical trajectory. This implies that in equilibrium institutional investors, and investors who invest in similarly attractive markets in their absence, convert owner-occupied housing into rental housing, rather than replacing incumbent landlords (see also Lambie-Hanson et al., 2022; Gorback et al., 2023; Coven, 2023). It also implies that the observed increase in rents constitutes a real decline in rental affordability that is not driven by changes in local economic conditions, even though rental supply expands. I find little evidence that institutional investors impact new construction. The regions they enter have more elastic housing supply where the housing stock grows faster even before they enter. Contrary to a prevailing narrative, I find that while institutional investors enter markets that are more diverse compared to the U.S. average, they have significantly larger white population shares compared to submarkets that would have been similarly attractive to invest in.

Lastly, I conduct a calibration exercise that allows me to test model predictions out-of-sample. Between 2021Q4, when my sample coverage ends, and 2023Q4 the U.S. 10-year treasury yield increased by around 350 basis points. In response institutional investors have begun selling homes back to the household sector. I calibrate model parameters to match observed dynamics of house prices, rents and the share of institutional investor-owned housing stock in the data. Consistent with the out-of-sample evidence, the calibration exercise predicts that institutional investors sell homes back to the household sector when household funding constraints relax or interest rates rise.

Literature The U.S. housing market in the aftermath of the housing bust of '08 has seen the emergence of a range of new actors, such as iBuyers (Buchak et al., 2020), algorithmic investors (Raymond, 2023) and various new investors in the single-family rental sector, from small and medium sized real estate investors (Garriga et al., 2023a,b) to large institutional investors, or "Wall Street Landlords" (Mills et al., 2019; Gurun et al., 2022; Austin, 2023; Coven, 2023; Gorback et al., 2023).

The literature has proposed two explanations for the emergence of institutional investors. The first attributes the entry of institutional investors to technological innovations that made single-family rentals profitable in areas where characteristics of the housing stock are particularly suited to large scale rental operations (see, e.g. Mills et al., 2019; Gorback et al., 2023; Raymond, 2023). The second explanation argues that the entry of these firms is driven by credit rationing and an excess supply of housing due to foreclosure sales in the aftermath of the housing bust (see, e.g. Allen et al., 2018; Mills et al., 2019; Smith and Liu, 2020; Lambie-Hanson et al., 2022).

In contrast, I attribute the local entry, expansion and exit decisions of institutional investors to variation in expected excess returns that is informed by the tightness of household funding constraints. This positions the credit rationing narrative within a broader literature on funding constraints in asset pricing (e.g. Black, 1972; Frazzini and Pedersen, 2014) and highlights the importance of interest rates. Viewed through the lens of this mechanism the entry of institutional investors is an equilibrium response to a new environment of low long-term interest rates and increased lending standards in the aftermath of the housing bust (see also Gete and Reher, 2018; D'Acunto and Rossi, 2021).[2] Importantly, this mechanism, unlike others, not only explains the entry of institutional investors, but also predicts their exit decisions. Conditional on expected excess returns, I also find that characteristics, such as the age of the housing stock, are

[2]This mechanism is distinct from reaching for yield (e.g. Korevaar, 2023; Garriga et al., 2023b). Investors reaching for yield move into riskier assets when risk-free rates decline to earn higher returns, holding risk compensation constant (e.g. Campbell and Sigalov, 2022). Here, tighter funding constraints increased the return on the asset class, holding risk constant.

important determinants of institutional investor investment decisions (see also Coven, 2023; Gorback et al., 2023; Raymond, 2023).

Compared to prior work studying the contribution of investors to the recovery of house prices in the aftermath of the housing bust (see, e.g. Allen et al., 2018; Smith and Liu, 2020; Lambie-Hanson et al., 2022; Ganduri et al., 2023) this paper asks whether institutional investors impact house price dynamics in ways that other investors do not. I find little evidence to support this view. I also find no evidence to suggest that their entry impacts local economic conditions or that their effect on rental supply differs from other investors who invest in similarly attractive markets in their absence. However, I find evidence to suggest that they raise rents (see also Gurun et al., 2022; Austin, 2023; Gorback et al., 2023).

Lastly, this paper connects to two broader strands of literature on housing in macro-finance and asset pricing (see, e.g. Davis and Van Nieuwerburgh, 2015; Piazzesi and Schneider, 2016, for an overview). The first studies how credit supply affects the real economy (see, e.g. Mian and Sufi, 2011; Gete and Reher, 2018; Kaplan et al., 2020; Greenwald and Guren, 2021; Lilley and Rinaldi, 2021; Adelino et al., 2022). The second studies the returns to residential real estate (see, e.g. Eichholtz et al., 2021; Demers and Eisfeldt, 2022). Household funding constraints not only explain expected and realized housing returns, they also explain the negative correlation between housing premia and risk-free rates documented by Campbell et al. (2009). The entry of institutional investors, driven by credit rationing and low interest rates, has real effects on housing affordability.

The remainder of this paper is structured as follows: Section 1 presents my model that derives empirically testable predictions. Section 2 describes my data and variable construction. Section 3 tests the relationships between household funding constraints, expected excess returns and the activity of institutional investors. In section 4 I decompose differences in house price and rental growth observed upon entry of institutional investors into market timing and market impact effects. Section 5 presents the calibration exercise I use to test my predictions out of sample. I conclude with final remarks in section 6.

1 Model

Introducing funding constraints into canonical Asset Pricing models distorts the relationship between risk and expected excess returns (Black, 1972). Unconstrained investors who trade on funding constraints can earn excess returns above and beyond compensation for bearing risk (Frazzini and Pedersen, 2014). Similar intuition can explain the entry of institutional investors into the single-family rental market. I formalize this intuition in a stylized model that delivers testable predictions for expected excess returns and the activity of institutional investors across time and space that I later take to the data. Appendix A provides further details, proofs and extensions of the model.

Setup There are two types of agents, a mass of atomistic households h occupying an equal mass of housing units in total supply Q and a finite number $I \in \mathbb{N}$ of investors i. Agents $j \in \{h, i\}$ have uncertain future income $\tilde{y}^j_{t+1}$ and decide whether to invest in risky housing or a risk-free asset in infinitely elastic supply with interest rate R^f. Their investment horizon is approximately three to five years, consistent with a typical horizon for home-ownership and rental investments. The quantity of housing each agent demands is denoted by q^j_t and restricted to be non-negative, consistent with a lack of short-sale markets for housing. There are no divisibility constraints on housing.[3] The preferences of each agent type j are standard mean-variance preferences over future net worth $\tilde{W}^j_{t+1}$ with absolute risk aversion A^j.

$$\max_{q^j_t \geq 0} \mathbb{E}_t \left[\tilde{W}^j_{t+1}\right] - \frac{A^j}{2} Var_t(\tilde{W}^j_{t+1})$$

The future price of a unit of housing supply $\tilde{P}_{t+1}$ is uncertain and jointly distributed with income, following an exogenous distribution with mean $\mathbb{E}_t\left[P_{t+1}, \mathbf{y}_{t+1}\right]$ and variance-covariance matrix Σ. For the purpose of exposition I assume that income and house prices are uncorrelated. In the appendix I relax this assumption. Households inelastically pay rent d to the owner(s) of their home at an exogenously determined rental rate. In the appendix I endogenize rents as bargaining outcome between owners and residents.

Household problem Households earn non-tradable benefits of home-ownership worth ν per quantity of housing supply owned.[4] Household future wealth inclusive of private benefits from home-ownership $\tilde{W}^h_{t+1}$ follows

$$\tilde{W}^h_{t+1} = q^h_t \left(\tilde{P}_{t+1} + d + \nu - R^f P_t\right) + R^f \left[W^h_t - d\right] + \tilde{y}^h_{t+1}$$

[3] This is to say, a household can own a part q^h_t of their home and rent the other part from investors.

[4] These can include both pecuniary and non-pecuniary benefits, such as tax deductibility of mortgage payment, access to collateral (Schmalz et al., 2017) or e.g. rules on pet ownership.

Households can borrow at the risk-free rate but face a constraint on their access to leverage $L \in [0, 1)$ that limits the amount of housing a household can buy to $\frac{1}{1-L}$ times their wealth.

$$q_t^h P_t \leq \frac{W_t - d}{1 - L}$$

In this setup with representative households the leverage constraint captures both intensive and extensive margins of credit constraints, depending on interpretation. The intensive margin interpretation is that of a Loan-to-Value constraint that requires households to keep an equity stake equal to $1 - L$ of the house price per unit of housing owned. The extensive margin interpretation is that lenders shut out a fraction $1 - L$ of households from the mortgage market, for example due to unmodeled characteristics such as credit scores.

Investor problem Investors are unaffected by borrowing constraints but incur operating costs $\mathcal{O}(q_t^i) = q_t^i(\alpha - \frac{\beta}{2} q_t^i)$ that feature economies of scale β and search costs $\mathcal{C}(\Delta q_t^i) = \frac{\gamma}{2}(\Delta q_t^i)^2 P_t^h$. An investor's future wealth $\tilde{W}_{t+1}^i$ follows

$$\tilde{W}_{t+1}^i = q_t^i \left(\tilde{P}_{t+1} + d - R^f P_t \right) - \mathcal{O}(q_t^i) - \mathcal{C}(\Delta q_t^i) + R^f W_t^i + \tilde{y}_{t+1}^i$$

The operating costs reflect both real financial costs of capital and labor for management and maintenance, as well as monitoring costs that arise from agency conflicts between landlord and tenant. The parameters α and β are positive and satisfy $\beta > A^i \sigma_p^2$ such that the marginal efficiency gains of an additional property in an investor's portfolio outweigh the costs of additionally incurred risk. The search cost reflects the costs of screening homes for suitability and finding households willing to sell or buy a home.

Equilibrium Equilibrium is a collection of house prices P_t and non-negative positions in housing (q_t^h, q_t^i) such that $i)$ households and investors choose optimal quantities of housing subject to their budget and borrowing constraints; and $ii)$ the housing market clears:

$$\frac{I q_t^i(P_t)}{Q} + q_t^h(P_t) = 1$$

Household and investor demand Household and investor demand $q_t^j(P_t)$ follow directly from their first order conditions:

$$q_t^h(P_t|W_t^h) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - (R^f + \lambda)P_t}{A^h\sigma_p^2}\right\}$$
$$q_t^i(P_t|q_{t-1}^i) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha - (R^f - \gamma q_{t-1}^i)P_t}{A^i\sigma_p^2 + \gamma P_t - \beta}\right\}$$

where λ denotes the Lagrange multiplier on the household borrowing constraint measuring the gap between the value of a home to a household and the funds at their disposal. Demand is increasing in expected cash flows and decreasing in price P_t and opportunity costs of capital R_t^f. Investor demand is increasing in scale q_{t-1}^i and scale efficiencies β.

Proposition 1. *An investor owning $q_{t-1}^i < Q$ units of housing supply expands (shrinks) their property portfolio if and only if house prices P_t are below (above) a threshold $\overline{P}_t^i(q_{t-1}^i)$ that is increasing in q_{t-1}^i and equal to*

$$\overline{P}_t^i(q_{t-1}^i) = \frac{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha}{R^f} + \frac{\beta - A^i\sigma_p^2}{R^f}q_{t-1}^i \tag{1}$$

Proposition 1 follows directly from the investor demand function. All else equal, the marginal value of an additional property in the portfolio of an incumbent exceeds the value of the first for a prospective new entrant.

An equivalent interpretation of this setup views the housing market as a market for housing services instead of as market for assets that generate cash flows. In this interpretation the market is cleared by the expected return to investing in capital that provides housing services at which the joint supply of capital by households and investors equals demand. Defining the returns to housing as $R_{t+1} \equiv \frac{\mathbb{E}_t[P_{t+1}]+d}{P_t}$ allows to restate Proposition 1 in terms of expected excess returns to investing in housing capital. This places a lower bound on expected excess returns that is decreasing in q_{t-1}^i.

Corollary 1. *An investor owning $q_{t-1}^i < Q$ units of housing supply expands (shrinks) their property portfolio if and only if the log expected excess returns to investing in housing capital $rx_{t+1} \equiv \log\frac{\mathbb{E}_t[P_{t+1}]+d}{P_t} - \log R^f$ are above (below) a threshold $\underline{rx}_{t+1}(q_{t-1}^i)$ equal to*

$$\underline{rx}_{t+1}(q_{t-1}^i) = \log\left(1 + \frac{\alpha - (\beta - A^i\sigma_p^2)q_{t-1}^i}{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha + (\beta - A^i\sigma_p^2)q_{t-1}^i}\right) \tag{2}$$

Proposition 2. *An equilibrium in which investors own no homes can only exist if*

$$\nu + \alpha > A^h \sigma_p^2 \tag{3}$$

Proposition 2 summarizes the necessary condition for existence of an equilibrium in which the entire housing stock is owned by households and investors prefers to invest all their wealth in the risk-free asset. Households have a comparative advantage at bearing housing risk, as they receive non-tradable benefits and do not incur operating costs for the same amount of risk. This advantage has to be sufficiently large that unconstrained household demand clears the market above the price where investors contemplating entry would enter $\overline{P}_t^i(0)$.

Proposition 3. *An investor owning $q_{t-1}^i < Q$ units of housing supply expands their property portfolio if households are sufficiently funding constrained, that is if*

$$\frac{W_t^h - d}{1 - L} < \left(1 - \frac{I q_{t-1}^i}{Q}\right) \overline{P}_t^i(q_{t-1}^i) \tag{4}$$

Equation (4) is a sufficient condition for investors to expand or enter a market. If households are sufficiently funding constrained that their demand would clear the market for $Q - I q_{t-1}^i$ units of housing supply - the total quantity of housing they initially own - below $\overline{P}_t^i(q_{t-1}^i)$, investors buy (additional) homes. Holding all else constant, this condition is decreasing in household wealth W_t, access to leverage L, and interest rates R^f. In the appendix I formally show that this condition is equivalent to a bound on the value of a home to households relative to funds at their disposal. The larger the difference between the two, the more likely investors are to enter or expand.

In practice, household wealth and expectations of future house prices are functions of the risk-free rate. In this case condition (4) is only decreasing in the risk-free rate if household wealth is less sensitive to changes in the interest rate than home values. Given that housing is a high duration asset compared to the wealth of the marginal home-buyer in the areas I study, this is generally true. Appendix A.2.2 expands on this point further.

Equation (4) is not a necessary condition because returns to scale could be sufficiently large that for a large enough investor the value of an additional property in their portfolio exceeds the value of a home to a household. In this case investors expand their portfolio irrespective of household funding constraints. Else, investors sell homes back to households if household funding constraints are sufficiently slack.[5] I expand on this in appendix A.

Proposition 4. *Investors raise house prices. Their impact is increasing in the total share of investor-owned housing stock* $\frac{Iq_t^i}{Q}$.

Whether many small investors or few large investors have different impacts on house prices depends on the parameters of the model and is left as an empirical question.

Remark 1. *The impact and quantity of housing owned or bought by investors is correlated with expectations of future house prices.*

This remark cautions that observing higher house price growth in areas with higher investor presence is due to both their impact on house prices and their timing of purchases to coincide with higher expected house price growth. This suggests that comparing house price and rental growth across markets that institutional investors entered and those they did not will capture both market timing and market impact effects.

Testable Predictions The model maps into three testable predictions. First, expected excess returns and returns to scale explain when and where institutional investors enter, expand and exit from local housing markets. Second, household funding constraints explain expected excess returns *within* and *across* time and space. Expected excess returns are higher if households are poorer, more restricted in access to leverage and risk-free rates are lower. Third, house prices and rents increase upon entry of institutional investors due to both market timing and potential market impact. Market timing and market impact are correlated and increasing in the share of institutional investor-owned housing stock.

2 Data

In this section I explain how I measure the key objects of my empirical analysis: the activity and ownership of institutional investors, house prices, rents and expected excess returns. My geographical unit of observation are 794 submarkets of 67 U.S. metro areas. The main results of my empirical findings cover the time period from 2006Q1 to 2021Q4. Appendix B reports summary statistics and additional details on the construction of my data.

[5] In fact, one firm in my sample explicitly operates on this premise as a business model. Home Partners of America offers households to buy a home on their behalf and lease it back to them as part of a 5-year rental agreement with an embedded call option that allows households to purchase the home.

Rental Data My rental data is provided by CBRE Econometric Advisors who offer a high-quality repeat-sales rent index for multi-unit apartment buildings that covers 814 submarkets of 69 U.S. metro areas at quarterly frequency from 2000Q1 onwards.[6] This rental index has three advantages over other commonly used rental data. First, it offers longer high-frequency time-series coverage with a high level of granularity than any other rental index. Second, the repeat-sales methodology captures current market conditions better than otherwise used rental surveys. This also eliminates bias that might arise from changes in the composition of units by looking at rent changes *within* units. Third, CBRE also measure vacancy rates. This allows me to construct rental returns that account for expected vacancies and test whether the entry of institutional investors impacts vacancy rates as well as rents. Appendix B compares the CBRE index to Zillow rent estimates and shows that they are tightly correlated in the time-period where coverage overlaps.

Institutional Investors My sample of institutional investors covers 10 of the largest private equity firms and real estate investment trusts with publicly verifiable operations in this sector. I measure the ownership and purchase activity of these firms at the property-level in CoreLogic deeds records. This requires tracing firm activity through subsidiaries that are listed as the immediate owners on the housing deeds. For publicly traded firms required to file with the SEC I obtain subsidiary names from SEC filings. For privately held firms I obtain subsidiary names through court filings, news coverage and PropStream, a commercial real estate platform. I match the subsidiaries to buyer and seller names in the deeds records to measure purchases and ownership of each firm across time and space.

Table 1 gives an overview over these firms, their largest shareholders and my estimates of their holdings in 2001Q4, 2011Q4 and 2021Q4. Because there has been considerable consolidation between firms in this space (Gurun et al., 2022; Austin, 2023) and firms do not disclose their subsidiaries frequently it is hard to accurately track firm boundaries over time. For this reason I restrict my analysis to aggregate institutional investor activity.

Appendix figure B.3 plots the estimated distribution of institutional investors across counties in 2021Q4. As of 2021Q4 institutional investors own properties in 367 submarkets across 48 of the 67 metro areas in my sample. Their activity is concentrated in the Sun Belt, with particularly large property holdings in Atlanta (>52,000 homes), Phoenix (>26,000) and Charlotte (>23,000). The distribution of the share of housing stock owned by

[6] A submarket is a geographic delineation custom to CBRE, comparable to a neighborhood or voting district in size. E.g. Harlem and the Upper West Side are two submarkets of the New York City metro area. CBRE provide a crosswalk that maps 14,822 zipcodes into the 814 submarkets.

Table 1: Estimate of institutional investor property portfolios

Investor	# Properties Owned			Shareholders (2021)
	2001Q4	2011Q4	2021Q4	
Invitation Homes	31	132	75602	Vanguard, Cohen & Steers, Blackstone
Progress Residential	57	54	59393	Pretium Partners (PE)
American Homes 4 Rent	34	185	46690	Vanguard, Blackrock, Norges Bank
FirstKey Homes	6	13	36417	Cerberus Capital (PE)
Tricon Residential	39	36	26381	CI Investments, T. Rowe, Cohen & Steers
Home Partners of America	0	0	17853	Blackstone (PE)
Vinebrook Homes	0	2	17816	Privately held by founder
Amherst Residential	1	5	15253	Amherst Group
My Community Homes	0	0	2634	KKR (PE)
Camillo Homes	0	1	1366	Camillo Properties (PE)

Note: *Table lists estimated number of single-family housing units owned by 10 of the largest institutional investors as of 2001Q4, 2011Q4 and 2021Q4 in the universe of CoreLogic housing deeds and their largest shareholders in 2021. Data are from CoreLogic, SEC filings, PropStream and industry reports.*

institutional investors is heavily skewed. In the 50th percentile institutional investors own 0.89% of the total housing stock, in the 75th percentile they own 1.67%, and in the 95th percentile they own 4.1%. These shares might appear small, but constitute a significant fraction of the single-family rental stock, given that on average 84% of single-family homes are owner-occupied (US Census, 2024).

Other data I complement the rental and ownership data with zipcode-level quarterly house price indices from Zillow beginning in 2000Q1, quarterly credit bureau aggregates with coverage beginning in 2006Q1, annual average adjusted gross income estimates from the IRS Statistics of Income, annual FBI county crime statistics (if available) and further data from the ACS and 2010 U.S. Census. I linearly interpolate annually measured variables to quarterly frequency. I aggregate all data to the submarket$\times$quarter level using a crosswalk provided by CBRE.

Final dataset After data cleaning and aggregating the final dataset is an unbalanced panel covering 717 submarkets of 64 metro areas from 2006Q1 to 2021Q4. The cross-sectional coverage grows over time, from 613 submarkets in 2006Q1 to 717 submarkets in 2021Q4.

2.1 Measuring expected excess returns

An obvious challenge to testing how agents respond to changes in expected excess returns is that expected excess returns are by their very nature unobservable:

$$\mathbb{E}_t^*\left[rx_{nt\to t+h}\right] = \mathbb{E}_t^*\left[\sum_{h=1}^{T}\frac{d_{nt+h}-\tau_{nt+h}}{P_{nt}} + \frac{P_{nt+T}}{P_{nt}}\right] - R_{t\to T}^f$$

where $\mathbb{E}_t^*\left[rx_{nt\to T}\right]$ denotes total (subjective) expected excess returns to rental investments in submarket n from t to T, $\mathbb{E}_t^*\left[\sum_{h=1}^{T}\frac{d_{nt+h}-\tau_{nt+h}}{P_{nt}}\right]$ denotes expected rental yields less operating costs, $\mathbb{E}_t^*\left[\frac{P_{nt+T}}{P_{nt}}\right]$ denotes expected appreciation and $R_{t\to T}^f$ denotes the risk-free rate.

To address this concern I construct three complementary measures of expected excess returns that differ in their assumptions about agents' expectations. For each measure I construct operating costs before taxes following Demers and Eisfeldt (2022) and use the 10-year treasury yield as risk-free rate benchmark.

Campbell and Shiller (1988) The first approach uses standard transformations of the net rent-price ratio dp_{nt}^{τ} and forecasts of (net) rental growth Δd_{nt}^{τ} from a Vector Autoregression to construct a measure of total expected excess returns $\widehat{\mathbb{E}}_t^*[rx_{nt}]$ (as in Campbell et al., 2009)

$$\widehat{\mathbb{E}}_t^*[rx_{nt}] = (1-\rho_n)\left\{k_n + dp_{nt}^{\tau} + \mathbb{E}_t^*\left[\sum_{j=0}^{\infty}\rho_n^j\Delta d_{nt+j+1}^{\tau}\right] - \mathbb{E}_t^*\left[\sum_{j=0}^{\infty}\rho_n^j r_{t+j+1}^f\right]\right\} + u_{nt}$$

where ρ_n and k_n denote constants that depend on the sample average rent-price ratio and u_{nt} denotes some measurement error introduced by construction.[7]

Figure 1 plots the time-series of constructed expected excess returns against the 10-year treasury yield and institutional investor purchases. The time-series of expected excess returns and 10-year treasury yields are tightly negatively correlated. The entry of institutional investors coincides with the peak of expected excess returns.

Excess rental yields If agents are agnostic about future house price growth or differences in expected appreciation are absorbed by fixed effects the (net) rent-price ratio in excess of the risk-free rate, denoted dpx_{nt}, is a sufficient measure of expected excess returns

$$dpx_{nt} = dp_{nt}^{\tau} - r_t^f + u_{nt}$$

[7]For this exercise I restrict my sample to submarkets for which I observe at least 50 consecutive quarters of house price and rental data. This leads to a slightly smaller sample of 671 submarkets in 65 metro areas.

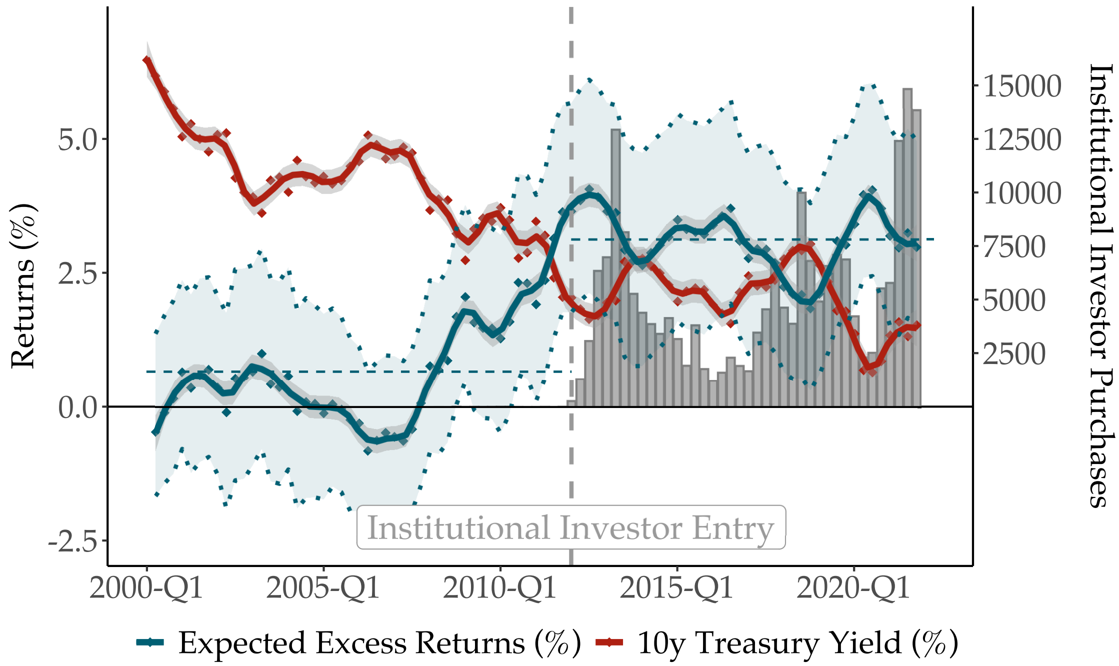

Figure 1: Expected excess returns, 10y treasury and institutional investor purchases.

Figure plots time-series of the 25th, 50th and 75th percentile of expected excess returns $\widehat{\mathbb{E}}_t^*[rx_{nt}]$ estimated from a Campbell-Shiller decomposition (blue) and 10-year treasury yields (red) on left axis, quarterly institutional investor purchases (grey) on right axis. Solid blue line marks median(smoothed) expected excess returns. Shaded area between dotted lines in blue covers interquartile range of expected excess returns. Dashed blue lines denote average expected excess returns before and after entry of institutional investors.

Realized returns If agents know the true data generating process of excess returns, realized excess returns over the next h periods are equal to expected excess returns plus some noise $\eta_{nt\to t+h}$ and measurement error $u_{nt\to t+h}$. This means that realized excess returns themselves are a noisy but unbiased measure of expected excess returns.

$$rx_{nt\to t+h} = \mathbb{E}_t^*[rx_{nt\to t+h}] + \eta_{nt\to t+h} + u_{nt\to t+h}$$

3 Excess returns and household funding constraints

Motivated by the predictions of my model in section 1 I assume a data-generating process

$$\mathbb{P}\left(inv_{nt}^{buy} > 0\right) = \beta_1 \mathbb{E}_t^*[rx_{nt}] + \beta_2 dist(inv_{m(n)t-1}^{own}) + \Gamma' X_{nt} + \epsilon_{nt} \tag{5}$$

where inv_{nt}^{buy} is an indicator for institutional investor purchases, $\mathbb{E}_t^*[rx_{nt+1}]$ denotes expected excess returns to rental investments in submarket n at time t, $dist(inv_{m(n)t-1}^{own})$ is a measure of distance to existing institutional investor operations in the metro area, X_{nt} is a vector of controls and fixed effects, and ϵ_{nt} is an error term that encompasses the combined effect of unmodeled factors that might also affect the probability of institutional investor investment. The coefficients of interest are β_1 and β_2, the change in probability of institutional investor purchases given a change in expected excess returns and distance to the nearest submarket where institutional investors already operate, which proxies for economies of scale.

3.1 Identification strategy

Identifying the coefficients of interest requires an identification strategy that addresses the endogeneity of expected excess returns. Expected excess returns are equilibrium outcomes that depend on the outcome variable, potentially introducing reverse causality. Additionally, the constructed measures of expected excess returns are (by design) contaminated by measurement error.[8]

I construct a novel shift-share instrument for household access to leverage L^{IV}_{nt} that is motivated by the model predictions: If expected excess returns depend on household access to leverage, exogenous variation in access to leverage allows to identify β_1 by estimating a Two Stages Least Squares specification:

$$I(inv^{buy}_{nt} > 0) = \beta_1 \widehat{\mathbb{E}}^*_t \left[rx_{nt}\right] + \beta_2 dist(inv^{own}_{m(n)t-1}) + \Gamma' X_{nt} + \varepsilon_{nt} \tag{6}$$

$$\widehat{\mathbb{E}}^*_t \left[rx_{nt}\right] = \theta_1 \, L^{IV}_{nt} \qquad + \theta_2 dist(inv^{own}_{m(n)t-1}) + \Lambda' X_{nt} + u_{nt} \tag{7}$$

This specification amounts to a joint test of the model. The second stage tests whether institutional investors respond to changes in expected excess returns. The first stage tests whether household funding constraints explain expected excess returns.

Instrument design My shift-share instrument L^{IV}_{nt} combines a time-varying aggregate shifter for household access to leverage with exposure weights (shares) fixed in a pre-period that vary in the cross-section of submarkets. The aggregate shifter is constructed from the nationwide 10th percentile of approved mortgage credit scores in the New York Federal Reserve Consumer Credit Panel, denoted by $score^{p10}_t$. The exposure weights are constructed by fixing the distribution of credit scores in 2006Q1, the first quarter I observe credit scores. The intuition of the instrument is that the 10th percentile of approved

[8]As long as the measurement error is uncorrelated with the constructed measure of expected excess returns ("classical") an instrumental variable approach identifies the coefficients of interest.

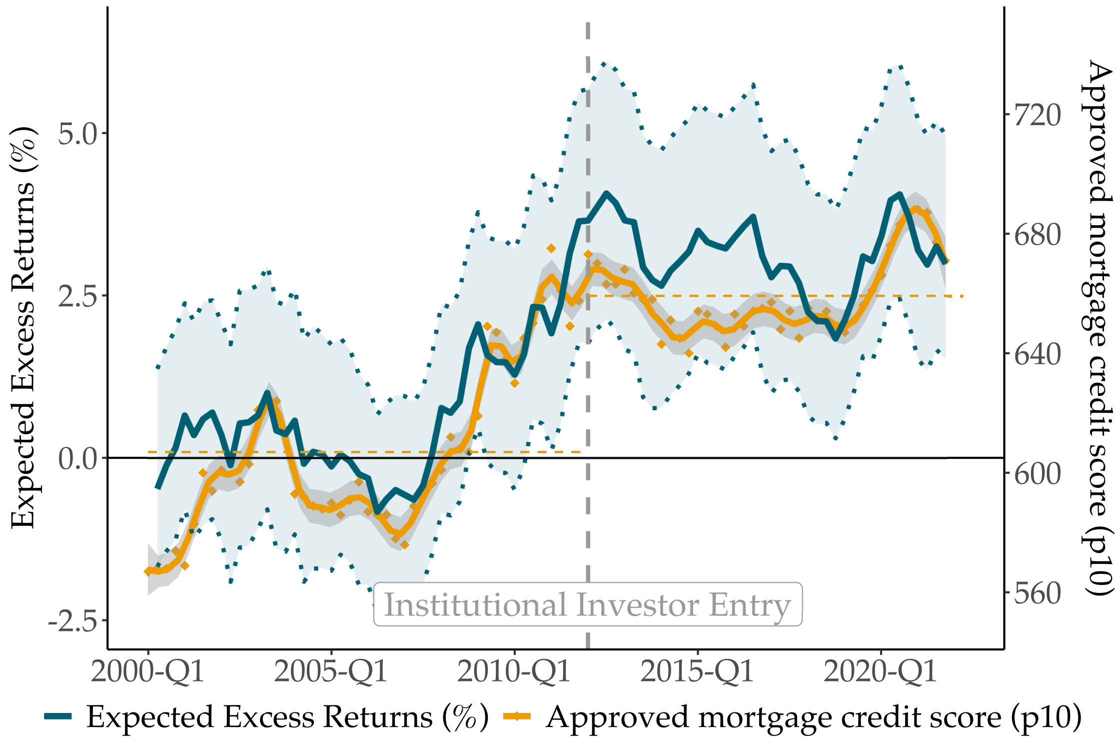

Figure 2: Expected excess returns and 10th percentile of approved mortgage credit scores.
Note: *Figure plots time-series of the 25th, 50th and 75th percentile of expected excess returns* $\widehat{\mathbb{E}}_t^* [rx_{nt}]$ *estimated from a Campbell-Shiller decomposition (blue) and (smoothed) 10th percentile of approved mortgage credit scores (yellow). Solid blue line marks median (smoothed) expected excess returns.Shaded area between dotted lines in blue covers interquartile range of expected excess returns. Dashed yellow lines denote average 10th percentile of approved mortgage credit scores before and after entry of institutional investors.*

mortgage credit scores is a plausible approximation of the minimum credit score necessary for households to access the mortgage lending market.

Figure 2 plots the time-series of $score_t^{p10}$ against the time-series of expected excess returns $\widehat{\mathbb{E}}_t^* [rx_{nt}]$. Two points are noteworthy. First, the two time series are highly correlated. Second, between 2000Q1 and 2021Q4 this bound varied from a low of 567 in 2000Q1 to a high of 688 in 2020Q4. Households with credit scores between these bounds should have gained or lost access to the mortgage lending market, depending on the prevailing regime of lending standards.[9]

To construct the instrument I divide credit scores into 10-digit bins $[\underline{k}, \overline{k})$ where $\underline{k}$ and

[9] Appendix figure C.1 provides further evidence that these patterns are driven by tightening lending standards by comparing the 10th percentile of approved mortgage credit scores to responses from the Federal Reserve Senior Loan Officer Opinion Survey. In further anecdotal evidence, The New York Times, 10/03/2014 (retrieved 07/20/2023) reported that increasing lending standards in the aftermath of the housing bust precluded the former chair of the Federal Reserve from refinancing his mortgage.

$\overline{k}$ denote the lower and upper bound of each bin k. The aggregate shifter is an indicator variable L_t^k that takes the value 1 if the lower bound of a bin lies above $score_t^{p10}$ and 0 otherwise. The exposure weights w_{n0}^k are constructed as the local population share of households inside each bin in 2006Q1.[10] The shift-share instrument L_{nt}^{IV} is constructed as the inner product of initial credit score population shares and the aggregate shifter L_t^k:

$$\text{Credit access } L_{nt}^{IV} = \sum_k \underbrace{w_{n0}^k}_{\text{Exposure weights}} \times \underbrace{I(\underline{k} > score_t^{p10})}_{\text{Aggregate shifter } L_t^k}$$

The constructed variable represents the share of households in submarket n that would have had access to credit in quarter t given the *local* distribution of household credit scores in 2006Q1 and changes in *aggregate* lending standards that restrict access to the mortgage lending market to households with credit scores above $score_t^{p10}$.

This instrument has several advantages over other instruments for credit supply used in the literature. The first advantage is that it provides repeated high-frequency variation in credit supply across time and space within metro areas. Secondly, by varying the extensive margin of credit supply it creates substantially larger variation in credit supply than intensive margin changes, such as changes in the conforming loan limit (see, e.g. Greenwald and Guren, 2021; Lilley and Rinaldi, 2021; Adelino et al., 2022).

Identifying assumption Recent econometric work on shift-share instruments establishes that the validity of shift-share instruments can be justified through either the exogeneity of shares w_{n0}^k (as in Goldsmith-Pinkham et al., 2020) or aggregate shocks L_t^k (as in Borusyak et al., 2021) conditional on included controls. Given the large cross-section this setting naturally lends itself to the framework of Goldsmith-Pinkham et al. (2020) that applies for a fixed number of time periods t and "industries" (credit score bins k), and an increasingly larger sample of locations (submarkets n).

In this setup the identifying assumptions require that i) the instrument is relevant, which is a directly testable prediction of the model; and ii) that the credit score bins w_{n0}^k are uncorrelated with the structural error term ε_{nt} in equation (6), *after* controlling for covariates X_{nt}, for credit score bins that would lose or gain access to leverage over time:

$$\mathbb{E}\left[\varepsilon_{nt} w_{n0}^k | X_{nt}\right] \text{ for all } k \text{ where } L_t^k \neq \text{const.} \tag{8}$$

[10]Given that households with credit scores below 560 (above 690) are assumed to never (always) have access to the mortgage lending market I aggregate the bins below (above) this bound into one bin.

This assumption does *not* require that the shares are unconditionally exogenous or randomly assigned. Rather, it requires that the assumption holds conditional on the vector of included controls that can include time-varying controls as well as granular fixed effects.

While this is a substantially weaker assumption than unconditional exogeneity of the shares, it still requires careful consideration of what could constitute a potential violation of the exclusion restriction. Equation (8) will not hold if the distribution of 2006Q1 credit scores affects the investment decisions of institutional investors for reasons other than higher expected excess returns that are not accounted for by the vector of controls X_{nt}. This motivates the included covariates in my baseline specification that I expand on further when I discuss potential threats to identification.

Baseline specification The dependent variable in my baseline specification is an indicator that takes the value 1 if institutional investors buy at least 0.5% of all properties transacted in a submarket in a quarter and 0 otherwise. I construct expected excess returns $\widehat{\mathbb{E}}_t^*[rx_{nt}]$ through a Campbell-Shiller decomposition. To control for distance to existing operations I interact an indicator for existing institutional investor operations in the metro area $I(inv_{m(n),t-1}^{own} > 0)$ with the log distance to the nearest submarket $\tilde{n}$ with institutional investor operations in the metro area. This variable is equal to $\log(1+dist_{n,\tilde{n},t})$ if institutional investors already operate in the metro area but not in submarket n and 0 otherwise.

$$dist(inv_{m(n)t-1}^{own}) = I(inv_{m(n),t-1}^{own} > 0) \times \log(1 + dist_{n,\tilde{n},t-1})$$

The vector of controls X_{nt} includes submarket and metro area$\times$quarter fixed effects, as well as time-varying controls for the level of log average adjusted gross income, vacancy rates, transaction share of foreclosure sales, county property crime rates, and year-on-year changes (plus up to three lags) for all variables included in levels as well as log house price and log rental growth. To account for potential auto-correlation in investment decisions I also include up to eight lags of the dependent variable. For variables that could plausibly be affected by reverse causality, such as house price and rental growth I control for their previous quarter values and subsequent lags of year-on-year growth.[11]

The baseline regression weights all submarkets equally.[12] I cluster standard errors two-ways by submarket and quarter. Clustering by submarket accounts for residual serial

[11] I control for the previous quarter value of vacancy rates, the share of foreclosure sales, log house price growth and log rental growth and their subsequent year-on-year growth rates.

[12] Weighting by estimated housing stock or population yields qualitatively identical results.

correlation in expected excess returns. Clustering by time allows for areas with different distributions of '06Q1 credit scores to experience other common shocks.

Threats to identification The identifying assumption requires that following a change in $score_t^{p10}$ the probability of institutional investors investing in submarkets with a different distribution of 2006Q1 credit scores - relative to their own mean and other submarkets in the metro area in quarter t, conditional on time-varying controls included in X_{nt} - does not change for reasons other than changes in expected excess returns.

Conceptually there are two potential sources of confounding variation that also affect the investment decisions of institutional investors and threaten identification: Firstly, operative considerations, such as costs of management, maintenance and delinquent tenants. For example, households with lower credit scores are more likely to live in areas with older housing stock that requires more costly maintenance. These variables are mostly time-invariant, such as the composition of the housing stock, or determined at the metro area, such as wages, economic conditions and tenant protection laws. As such, these variables are largely absorbed by submarket and metro$\times$quarter fixed effects. The inclusion of time-varying controls for income, county property crime rates and vacancy rates (and their growth rates) directly accounts for any residual variation that is not controlled for by submarket and metro-quarter fixed effects.

The second threat to identification are unobserved beliefs or private information about future returns that are not reflected in the constructed measures of expected excess returns and correlated with the 2006Q1 distribution of credit scores. For example, if areas with more subprime borrowers in 2006Q1 experienced more foreclosure sales during the housing bust, this could induce agents to expect higher future returns *beyond* the extent to which this is captured by the constructed proxies for expected excess returns. The inclusion of time-varying controls for foreclosure sales, income, house price and rental growth (and their lags) that predict future returns directly addresses this concern. Note that informational advantages, for example due to local political connections, do not constitute a violation of the exclusion restriction as long as they are uncorrelated with the 2006Q1 distribution of credit scores.

3.2 Instrumental Variable estimates

Table 2 reports the estimated baseline specification in (6) and (7). Columns 1 and 2 report the first stage (FS) estimates. Columns 3 and 4 report the second stage (IV) estimates.

Table 2: Instrumental Variable estimates for equations (6) and (7)

	$\widehat{\mathbb{E}}_t^*\left[rx_{nt}\right]$ (%, annualized) (FS)		$I(inv_{nt}^{buy}>0)$ (IV)	
	(1)	(2)	(3)	(4)
$\widehat{\mathbb{E}}_t^*\left[rx_{nt}\right]$ (%, annualized)			0.045***	0.052**
			(0.0111)	(0.0160)
$dist(inv_{m(n),t-1}^{own})$	-0.023	-0.025*	-0.010**	-0.008*
	(0.0165)	(0.0106)	(0.0032)	(0.0031)
Credit Access $L_{nt}^{IV}(\%)$	-0.164***	-0.108***		
	(0.0187)	(0.0151)		
Observations	39,624	39,624	39,624	39,624
First stage F-Stat.			77.01	51.57
Adjusted R^2	0.983	0.990		
Within R^2	0.136	0.502		
Fixed effects	Metro × Quarter	Metro × Quarter	Metro × Quarter	Metro × Quarter
	Submarket	Submarket	Submarket	Submarket
Add. Controls		Yes		Yes
Dep. var. lags			8	8

Note: *Table reports $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}_1, \hat{\theta}_2$ for a Two Stages Least Squares specification of the form:*

$$I(inv_{nt}^{buy}>0) = \beta_1 \widehat{\mathbb{E}}_t^*\left[rx_{nt}\right] + \beta_2 dist(inv_{m(n)t-1}^{own}) + \Gamma' X_{nt} + \varepsilon_{nt}$$
$$\widehat{\mathbb{E}}_t^*\left[rx_{nt}\right] = \theta_1\, L_{nt}^{IV} \qquad + \theta_2 dist(inv_{m(n)t-1}^{own}) + \Lambda' X_{nt} + u_{nt}$$

First stage F-Stat. denotes Kleibergen-Paap F-statistic. Standard errors in parentheses clustered two-ways by submarket and quarter. ($^{+}$,*,**,***) *denote significance at 0.10, 0.05, 0.01, 0.001 level.*

Columns 1 and 3 only control for fixed effects and lags of the second stage dependent variable. Columns 2 and 4 include the full vector of controls.

The IV estimates in columns 3 and 4 confirm that expected excess returns and distance to existing operations are strong predictors of institutional investor activity. All else equal, a 1pp increase in expected excess returns increases the probability of institutional investors investing in a submarket by 5.2pp. Institutional investors are 1.9pp less likely to invest in a submarket that is 10 miles away from existing operations, consistent with localized scale economies or gravity effects.

The estimated coefficients are stable across all specifications. The first stage F-Statistic is significantly above commonly recommended thresholds for relevance. Including the full vector of time-varying controls has no qualitative or quantitative effect on the IV estimates.

The first stage in columns 1 and 2 confirms that the instrument generates substantial variation in expected excess returns, even after absorbing metro×quarter and submarket fixed effects. All else equal, a 10pp decrease in the share of households with access to the mortgage lending market raises expected excess returns by 108bp. Consistent with both market timing and market impact effects on house prices, submarkets closer to existing operations have lower expected excess returns. Including the full vector of controls explains a substantially larger part of the residual variation in expected excess returns in the first stage, but does not qualitatively alter the estimated first stage coefficients.

Appendix table C.1 estimates analogous specifications using excess rental yields and realized 3-year ahead returns as alternative measure of expected excess returns. The findings are qualitatively identical. Realized returns are approximately four times as dispersed as excess rental yields, which leads to larger (smaller) coefficient estimates in the first (second) stage.

3.3 Expected excess returns across time and space

In this section I test whether household funding constraints also explain expected excess returns *across* time and space and whether declines in the risk-free rate pass-through into expected excess returns. I drop fixed effects from the first stage specification in (7) and include additional controls for the 10-year treasury yield and the 10th percentile of approved mortgage credit scores $score_t^{p10}$ that are else absorbed by time-fixed effects.

$$\widehat{\mathbb{E}}_t^*[rx_{nt}] = \theta_1 \text{ Credit Access } L_{nt}^{IV} + \theta_2 \text{ Income } y_{nt} + \theta_3 \text{ 10y-Treasury } r_t^f + \Lambda' X_{nt} + u_{nt} \tag{9}$$

Table 3 reports the estimated coefficients. Columns 1, 2 and 3 only include the shift-share instrument, local income and lags of the second stage dependent variable as controls that vary across time and space. Columns 4, 5 and 6 include the full vector of time-varying controls. Columns 1 and 4 include no fixed effects, columns 2 and 4 include quarter fixed effects, columns 3 and 6 include submarket fixed effects.

Household funding constraints explain expected excess returns within and across time and space: Expected excess returns are high if households are poor, restricted in access to leverage and risk-free rates are low. Consistent with the model predictions declines in the risk-free rate pass-through into expected excess returns approximately one-to-one.[13]

The baseline regression in column 1 explains over 46% of the total variation in expected excess returns. Including quarter-fixed effects in column 2 has nearly no effect on the estimated coefficients or the amount of total variation explained, suggesting that the included variables already capture most variation in expected excess returns due to aggregate shocks. Including submarket fixed effects column 3 explains over 95% of

[13] Campbell et al. (2009) document a similar negative correlation between real rates and housing premia in a variance decomposition of the rent-price ratio for the period from 1975 to 1997.

Table 3: Expected excess returns across time and space

	Dependent variable: $\widehat{\mathbb{E}}_t^*\left[rx_{nt}\right]$ (%, annualized) $(\mu = 2.57, \sigma = 3.07)$					
	(1)	(2)	(3)	(4)	(5)	(6)
Credit Access $L_{nt}^{IV}(\%)$	-0.068***	-0.074***	-0.046*	-0.076***	-0.078***	-0.034*
	(0.013)	(0.013)	(0.020)	(0.014)	(0.015)	(0.015)
$\log$(Av. adj. gross income$_{nt}$)	-2.332***	-2.205***	-2.191***	-1.944***	-2.007***	-0.475^{+}
	(0.290)	(0.314)	(0.294)	(0.310)	(0.319)	(0.244)
10-year Treasury yield$_t$ (%)	-1.117***		-1.107***	-1.194***		-1.035***
	(0.078)		(0.071)	(0.076)		(0.037)
Observations	39697	39697	39697	39697	39697	39697
Adjusted R^2	0.467	0.476	0.958	0.495	0.505	0.976
Within R^2	0.468	0.356	0.799	0.495	0.392	0.887
Fixed effects		Quarter	Submarket		Quarter	Submarket
Distance control	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls				Yes	Yes	Yes

Note: *Table reports $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ for specifications of the form:*

$$\widehat{\mathbb{E}}_t^*[rx_{nt}] = \theta_1 \text{ Credit Access } L_{nt}^{IV} + \theta_2 \text{ Income } y_{nt} + \theta_3 \text{ 10y-Treasury } r_t^f + \Lambda' X_{nt} + u_{nt}$$

Standard errors in parentheses clustered two-ways by submarket and quarter. μ and σ denote overall mean and standard deviation of the dependent variable. ($^{+}, ^{*}, ^{**}, ^{***}$) denote significance at 0.10, 0.05, 0.01, 0.001 level.

the total variation. Household funding constraints can explain about 80% of variation in expected excess returns across time within submarkets. Including the full vector of time-varying controls in columns 4, 5 and 6 only slightly increases the amount of variation explained within and across time and space. The only coefficient that changes is income, which mostly explains cross-sectional variation across submarkets.

Appendix table C.2 estimates analogous specifications using excess rental yields and realized 3-year ahead returns as alternative measure of expected excess returns. The results are qualitatively very similar. As in table C.1 the coefficients differ in magnitude because realized future returns are substantially more volatile than rental yields. Excess rental yields and realized future returns are decreasing in household access to leverage and the risk-free rate. The baseline regressions in column 1 explain 59% to 66% in total variation across time and space, and over 70% to 80% of time-series variation within submarkets.

Variance decomposition Combining the estimates of table 3 with the (residualized) variance-covariance matrix of the explanatory variables after absorbing fixed effects yields

Table 4: Variance decomposition of expected excess returns across time and space

Variance	Within Quarter	Within Submarket	Total	Within Quarter (%)	Within Submarket (%)	Total (%)
	(1)	(2)	(3)	(4)	(5)	(6)
$\theta_1^2 Var(L_{nt}^{IV})$	0.537	0.076	0.819	0.072	0.038	0.089
$\theta_2^2 Var(y_{nt}^{IV})$	0.800	0.000	0.785	0.107	0.000	0.085
$\theta_1\theta_2 Cov(L_{nt}^{IV}, y_{nt})$	0.962	-0.001	0.743	0.129	-0.000	0.081
$\theta_3^2 Var(r_t^f)$		1.207	1.534		0.593	0.167
$\theta_2\theta_3 Var(y_{nt}, r_t^f)$		-0.007	-0.447		-0.003	-0.049
$\theta_1\theta_3 Var(L_{nt}^{IV}, r_t^f)$		0.515	1.133		0.253	0.123
Σ	2.299	1.790	4.566	0.308	0.880	0.497
$Var\left(\widehat{\mathbb{E}}_t^* \left[rx_{nt}\right]\right)$	7.452	2.034	9.192			

Note: *Table decomposes the variance of expected excess returns $\widehat{\mathbb{E}}_t^* \left[rx_{nt}\right]$ within-quarter, within-submarket and overall into the variance explained explained by Credit Access L^{IV}, income y and risk-free rates r^f.*

a variance decomposition of expected excess returns across and within time and space:

$$Var\left(\widehat{\mathbb{E}}_t^* \left[rx_{nt}\right]\right) = \sum_{x_{nt}^i, x_{nt}^j \in \{L_{nt}^{IV}, y_{nt}, r_t^f\}} \theta_i\theta_j Cov(x_{nt}^i, x_{nt}^j) + \sum_{x_{nt}^i, x_{nt}^j \notin \{L_{nt}^{IV}, y_{nt}, r_t^f\}} \theta_i\theta_j Cov(x_{nt}^i, x_{nt}^j) + Var(u_{nt})$$

where $Var\left(\widehat{\mathbb{E}}_t^* \left[rx_{nt}\right]\right)$ denotes the (residual) variance of expected excess returns within quarter or submarket after absorbing fixed effects, θ_i denotes the estimated coefficients in columns 4, 5 and 6 of table 3, $Cov(x_{nt}^i, x_{nt}^j)$ denotes the (residual) covariance of explanatory variables x_{nt}^i, x_{nt}^j and $Var(u_{nt})$ denotes the (residual) variance not explained by the explanatory variables.

Table 4 quantifies the relative contributions of credit access, income and the risk-free rate to explaining expected excess returns. Columns 1 and 4 decompose the variance of expected excess returns across submarkets (within quarter). Columns 2 and 5 decompose the variance of expected excess returns across time (within submarket). Columns 3 and 6 decompose the total variance of expected excess returns across submarkets and time. Columns 1, 2 and 3 report the total variance decomposition. Columns 4, 5 and 6 report the share of variance explained by each variable.

Credit access, income and the risk-free explain around 50% of the total variance of expected excess returns across time and space. Credit access and the risk-free rate contribute approximately equal parts. Credit access and income explain around 30% of the cross-sectional variation (within quarter). Within submarket the risk-free rate and its covariance with access to leverage explain around of 85% of the time-series variation (within geography) in expected excess returns.

These findings are further corroborating evidence in support of my mechanism. The historical absence of institutional investors in the time-series can be mostly explained by historical long-term interest rates and access to leverage. At least within this sample, over 60% of the time-series variation is driven by the risk-free rate alone. This suggests that the risk-free rate is a key determinant of expected excess returns (see also Campbell et al., 2009).

4 Market timing and market impact

This section decomposes differences in price and rental growth observed upon entry of institutional investors into a submarket into market timing and market impact. I begin with an unconditional, "naive" comparison across submarkets that institutional investors entered that recovers both market timing and market impact effects. I then saturate the comparison with controls for predictors of future returns in order to purge market timing effects. This allows to place bounds on the extent of market timing and market impact.

4.1 Methodology

I use an event-study approach to compare the dynamics of outcome variables, such as house prices and rents, upon entry of institutional investors relative to submarkets they did not enter.

$$Y_{nt} = \alpha_n + \alpha_t + \sum_{\ell=-L}^{-2} \delta_\ell^{pre} \times entry_{n,t-\ell} + \sum_{\ell=0}^{K} \delta_\ell \times entry_{n,t-\ell} + e_{nt} \quad (10)$$

where Y_{nt} denotes the outcome variable, $entry_{n,t-\ell}$ denotes an indicator for institutional investor entry ℓ quarters before time t and δ_ℓ, the coefficient of interest, summarizes the average difference in the outcome variable ℓ quarters after entry into a submarket.

Recent work in the econometric literature establishes that estimating specifications of the form in (10) by OLS can yield biased estimates (see, e.g. de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Baker et al., 2022). In particular, Baker et al. (2022) shows that if entry is staggered over time and the impact of institutional investors is dynamically increasing over time, estimating specifications as in (10) by OLS yields estimates with opposite signs to the true coefficient because submarkets entered earlier serve as effective comparison units. To address this issue I estimate a difference-in-differences event study specification using the `did`-estimator of Callaway and Sant'Anna (2021).[14]

Given that institutional investors enter into neighborhoods with higher expected future returns the estimated coefficients will capture market timing, both market timing and potential market impact of institutional investors.

$$\delta_\ell = \delta_\ell^{timing} + \delta_\ell^{impact}$$

Market timing δ_ℓ^{timing} denotes changes in the outcome variable due to forces unrelated to the activity of institutional investors that any investor would be able to act upon. Market impact δ_ℓ^{impact} denotes differences in the outcome variable due to the impact of institutional investors that other investors would not have.

I decompose observed differences in the outcome variable δ_ℓ into market timing and market impact by saturating the comparison in (10) with controls for predictors of future returns.[15] This alters the comparison in (10) from an unconditional comparison to a conditional comparison. Rather than comparing submarkets that institutional investors enter to all submarkets, I compare them to submarkets that are similarly attractive to invest in, given observable predictors of future returns at the time of their entry decision. The intuition of this approach is that these predictors of future returns are observable to any investor and therefore purge market timing effects that other investors would act upon in the absence of institutional investors. If the controls fully account for expected future returns any remaining differences in the outcome variable - as far as they exist - are due to the market impact δ_ℓ^{impact} of institutional investors.

Identifying assumption The identifying assumption for this exercise to recover the "true" market impact effect δ_ℓ^{impact} is a conditional parallel trends assumption: *Conditional on predictors of future returns observed at time of entry* the outcome variables Y_{nt} would follow parallel trends in the absence of institutional investors.

This assumption does *not* require that entry decisions are made at random, nor does it

[14]For further details on the `did`-estimator I refer the interested reader to Baker et al. (2022) and Callaway and Sant'Anna (2021).

[15]The `did` estimator of Callaway and Sant'Anna (2021) controls for covariates by estimating a specification analogous to (10) but weights differences in the outcome variable by the likelihood of entry given the value of controls observed at the time of entry.

require that the outcome variable would follow parallel trends unconditionally. Rather, it requires that conditional on observed predictors of future returns the entry decisions of institutional investors are driven by variables that are independent of the outcome variable. In practice, two important predictors of entry decisions besides expected excess returns are proximity to existing institutional investor operations and the suitability of the housing stock (see also e.g. Gorback et al., 2023; Raymond, 2023).

This assumption is, by design, untestable. However, several robustness tests can provide evidence in support of the identifying assumption. First, inspecting pre-trends before entry allows to test whether the conditional parallel trends assumption holds prior to entry. Second, placebo tests that move the time of entry to some time before actual entry occurred should not identify any market impact effects before actual entry occurs. Similarly, market timing and market impact should be increasing in the share of housing stock bought by institutional investors and close to zero in submarkets where they only own a small share of the housing stock. Third, alternative estimators, such as the synthetic differences-in-difference estimator of Arkhangelsky et al. (2021), should recover similar estimates.

4.2 Baseline specification

My baseline specification controls for the levels of expected excess returns implied by a Campbell-Shiller decomposition, log income, credit access, vacancy rates, previous year transactions and foreclosure sales scaled by total housing stock, year-on-year changes in log house prices and log rents, 3-year changes in log house prices and log rents, log income, log housing stock, credit access, and vacancy rates, and up to 3 lags of annual foreclosure sales scaled by total housing stock.[16] For all variables that are plausibly affected by reverse causality, such as house price and rental growth I control for their value in the quarter prior to entry and previous changes. This not only directly controls for observable expected excess returns, but also for predictors of future rents and house prices that would allow agents to extrapolate from further information, such as past foreclosure sales or momentum. The sample for this exercise is limited to a balanced panel of 671 submarkets for which I observe all control variables over the sample period.

I define a submarket as entered by institutional investors if institutional investors own

[16] For this exercise there is no need to fix the distribution of local credit scores in a pre-period. Instead, I construct credit access as the product of prior quarter population shares w_t^k and the aggregate shifter L_t^k observed at the time of entry.

more than 100 single-family housing units or more than 1% of the total single-family housing stock at any time. I define the time of entry as the first quarter in which they own at least 5 units and 0.1% of the housing stock in a submarket.[17] I cluster standard errors at the submarket level. This follows standard practice to cluster at the level of "assignment" and allows for shocks that are correlated within submarkets over time.

Balance of covariates Table 5 compares observable characteristics and outcome variables between submarkets that institutional investors entered and those they did not in the quarter prior to entry. Because entry is staggered over time there is no way to directly construct a balance table that compares covariates in the period before entry. Instead, I estimate a regression of the form

$$C_{nt} = \alpha_t + \Delta \times entry_{nt-1} + \gamma X_{nt} + \epsilon_{nt}$$

where C_{nt} denotes the variable of interest, X_{nt} is the vector of controls for predictors of future returns observed at time of entry and Δ is the coefficient of interest. To ensure that Δ is identified exclusively through within-quarter comparisons of submarkets not yet entered by institutional investors I estimate the regression on a restricted sample that only keeps observations of not-yet-entered submarkets.

Columns 1 and 2 report the unconditional averages of each variable for submarkets that institutional investors enter and those they did not. Column 3 reports the total standard deviation of the variable of interest. Columns 4 and 5 report the unconditional and conditional difference in the variable of interest Δ.

Several points are noteworthy. An unconditional comparison finds that institutional investors enter into submarkets with substantially lower rents, house prices and price-rent ratios that experienced larger house price drops during the housing bust. Conditional on predictors of future returns there is no significant difference in house prices, rents or house price drops during the housing bust. This confirms that the included controls enable a comparison of similar neighborhoods.

However, there are notable differences in non-financial characteristics that shed light on the source of variation driving entry decisions. Even conditional on future return predictors, institutional investors enter submarkets with significantly younger housing stock and more elastic housing supply. This is consistent with institutional investors

[17] My results are broadly robust to varying these thresholds. I use different thresholds for entry time and classification as "entered" to avoid false positives and mis-measurement of entry time.

Table 5: Covariate balance with staggered entry

	Summary Statistics			Δ at time of entry	
	No Entry	Entry	SD	Δ	Δ adj. for controls
Time-Invariant Characteristics					
Median year of construction	1977.41	1985.47	12.13	8.06 (0.82)	4.18 (0.62)
Median age	36.60	35.72	3.72	-0.87 (0.41)	-0.16 (0.30)
White pop. share	57.07	54.25	20.80	-2.82 (2.10)	3.25 (1.13)
Black pop. share	13.20	16.65	15.40	3.46 (1.45)	-2.86 (1.19)
Asian pop. share	7.69	4.95	8.09	-2.74 (0.48)	0.12 (0.44)
Hispanic pop. share	19.06	21.29	17.09	2.23 (1.71)	-0.44 (1.76)
Saiz elasticity	1.70	1.94	0.84	0.25 (0.08)	0.11 (0.06)
Boom growth (%)	56.65	53.81	36.80	-2.83 (5.09)	-4.79 (3.84)
Bust drop (%)	21.30	27.83	12.98	6.53 (1.62)	0.36 (1.01)
Comparison variable ($entry - 1$)					
House Price ('000)	361.28	227.88	296.98	-133.39 (12.31)	20.35 (13.34)
Rent (monthly)	1324.55	1021.00	577.15	-303.54 (27.56)	14.26 (25.13)
Home ownership rate (%)	58.60	64.31	14.96	5.71 (0.59)	4.01 (0.74)
Price-rent ratio	20.93	16.98	7.91	-3.95 (0.30)	-0.09 (0.32)
Number of Submarkets	337	334			

Note: *Table compares submarkets in the quarter prior to entry of institutional investors by estimating a specification of the form:* $C_{nt} = \alpha_t + \Delta \times entry_{nt-1} + \gamma X_{nt} + \epsilon_{nt}$. *Standard errors in parentheses clustered two-ways by submarket and quarter.*

preferentially investing in areas with more elastic housing supply and lower operating costs (see also, e.g. Gorback et al., 2023; Raymond, 2023), which my model also predicts (see appendix A.2.3).

There is an interesting finding in the balance of demographic characteristics: Compared to all other submarkets in the U.S., institutional investors enter submarkets that are younger and more diverse. Compared to submarkets that would be similarly attractive to invest in the submarkets they operate in have a significantly larger white population share. This runs contrary to a prevailing narrative that their entry disproportionately affects minority neighborhoods.

Threats to identification Even if there is no evidence to reject the identifying assumption, the results of this exercise are limited by the extent to which expectations of future returns are accounted for by the included predictors of future returns. Unobservable predictors of future returns, to the extent that they exist, remain unaccounted for. For example, private information about a future metro line expansion that is not reflected in expected excess returns at the time of entry could violate the conditional parallel trends assumption.

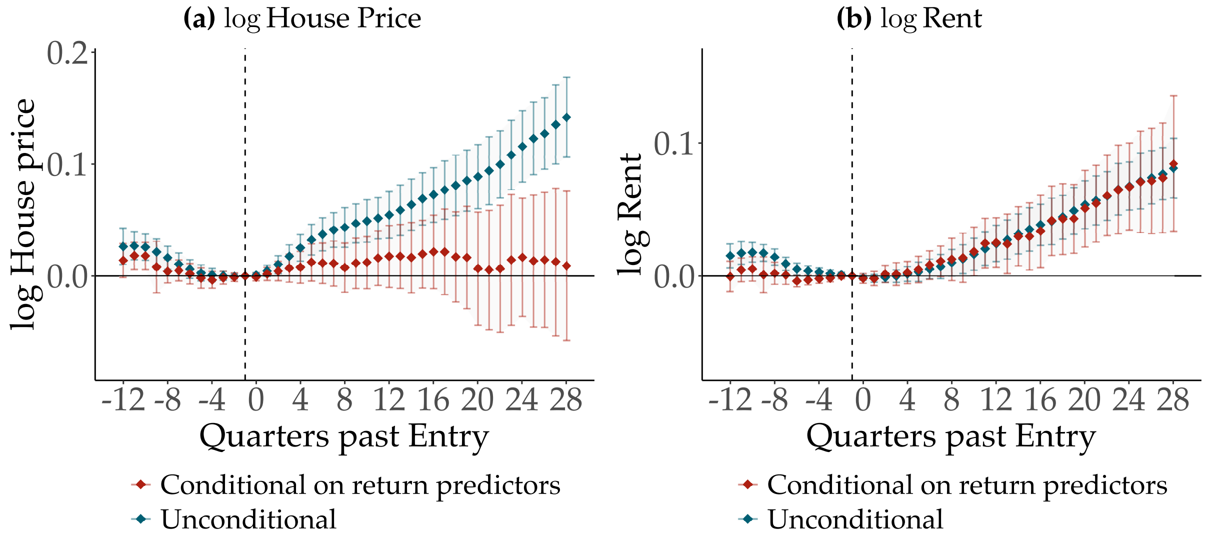

Figure 3: Market timing and market impact of institutional investors.
Figure shows difference-in-differences event study coefficients for equation (10) for quarterly log house prices and log rents. Coefficients in blue report estimates from an unconditional comparison of submarkets. Coefficients in red report estimates from a comparison conditional on predictors of future returns observed in the quarter prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

In practice, given the large number of submarkets in which institutional investors operate, it is unlikely that examples such as the above are important drivers of entry decisions. It is very difficult to be highly selective and operate at a large scale across many geographies at the same time. Additionally, it is not clear that smaller local investors might not have informational advantages over institutional investors, which would bias my estimates downward. Nevertheless, a conservative interpretation of my estimates is that these are lower bounds on the extent of market timing and upper bounds on potential market impact effects.

4.3 Differences-in-difference estimates

Figure 3 plots my estimates of (10) for log rents and house prices in the full sample of submarkets. The total difference in the outcome variable δ_ℓ that captures the sum of market timing and market impact is plotted in blue. Estimates of market impact δ_ℓ^{impact} are plotted in red.

On average, house prices grow 2.03pp per year faster upon entry of institutional investors. Market timing can account for the entire difference in observed house price growth. This does not imply that they do not impact prices, but rather that they do not impact prices differently from other investors or home-buyers who would invest in this market in their absence.

Rents grow 1.16pp per year faster, which cannot be accounted for by market timing. This suggests that higher rental growth observed upon entry is driven by market impact unique to institutional investors (see also Gurun et al., 2022; Austin, 2023). Intuitively, in the absence of institutional investors a large enough number of smaller investors (Garriga et al., 2023a) would have similar effects on house prices. However, they would likely not be able to affect the rental market.[18] In section 4.4 I confirm that in similarly attractive neighborhoods owner-occupied housing stock is also converted into rental units in the absence of institutional investors by looking at the trajectory of home-ownership rates.

Given that I use a repeat sales index for multi-unit apartment buildings to measure rents these findings cannot reflect changes in the quality of the rental housing stock. Indeed, the fact that the rent of a close (but likely inferior) substitute to single-family rental housing increases in response to a new source of rental supply is strong evidence to suggest that the entry of institutional investors has a causal impact on the rental market. In section 4.4 I confirm that the impact on rents is not driven by changes in local economic conditions by comparing the trajectory of income.

There are two further potential explanations why there is strong evidence of market timing with respect to future house price growth and very little evidence of market timing with respect to rental growth: firstly, the returns to predicting house price growth are substantially larger than the returns to predicting rental growth. Secondly, rental growth is also harder to predict than house price growth (see, e.g. Campbell et al., 2009).

Heterogeneity If institutional investors have an impact on housing markets and act on expected future returns market timing and market impact should be increasing in the share of institutional investor-owned housing stock. To test this hypothesis I sort submarkets that institutional investors entered into three bins, based on the ex-post share of institutional investor-owned housing stock: A bottom bin for submarkets below the 25th percentile (0.50% of total housing stock owned), a middle bin for submarkets between the 25th and 75th percentile (1.67% of total housing stock owned) and a top bin for submarkets above the 75th percentile. I then repeat the exercise above for each bin separately in appendix figure D.1. Panel (a) plots log house prices, panel (b) plots log rents.

[18] Figuratively, consider 1000 mom-and-pop investors investing in one rental property each versus one or two large investors investing in 1000 rental properties.

Market timing and market impact effects are increasing in the share of institutional investor-owned housing stock. In the bottom bin house prices (rents) grow 0.52pp (0.63pp) per year faster, which can be entirely accounted for by market timing. I find no evidence to suggest that institutional investors impact rents or house prices in submarkets where they only own a small share of the housing stock.

In the middle bin house prices and rents grow 1.90pp and 1.05pp pear year faster, respectively. Market timing can account for the entire difference in house price growth. However, market timing cannot account for the observed difference in rental growth. In fact, the market impact estimates δ_ℓ^{impact} are insignificantly larger than the combined effect of market timing and market impact δ_ℓ.

In the top bin house prices and rents grow 2.58pp and 1.43pp per year faster. Market timing can account for at least 1.22pp per year of the observed difference in house price growth, leaving up to 1.36pp per year higher house price growth that can potentially be attributed to market impact. As before, market timing cannot account for the observed difference in rental growth, suggesting that differences in rental growth are driven by market impact.

These findings further corroborate the predictions of my model and the narrative described above. The absence of market impact on rents or house prices in the bottom bin, as well as a lack of pre-trends in each subsample further supports the identifying assumption.

Robustness checks Appendix figure D.2 conducts additional robustness tests for findings with regard to house prices in panel (a) and with respect to rents in panel (b).

The first column conducts a placebo test that moves the time of entry forward by three years for every submarket. The placebo test finds no evidence of market impact effects before actual entry occurs. The second column uses the synthetic control and synthetic difference-in-differences framework of (Arkhangelsky et al., 2021) to estimate the market impact effect δ_ℓ^{impact}. Unlike the differences-in-difference approach above that conditions on observable predictors of future returns this approach selects comparison submarkets based on the pre-entry trajectories of the outcome and control variables for each submarket. This exercise yields qualitatively indistinguishable findings to figure 3.

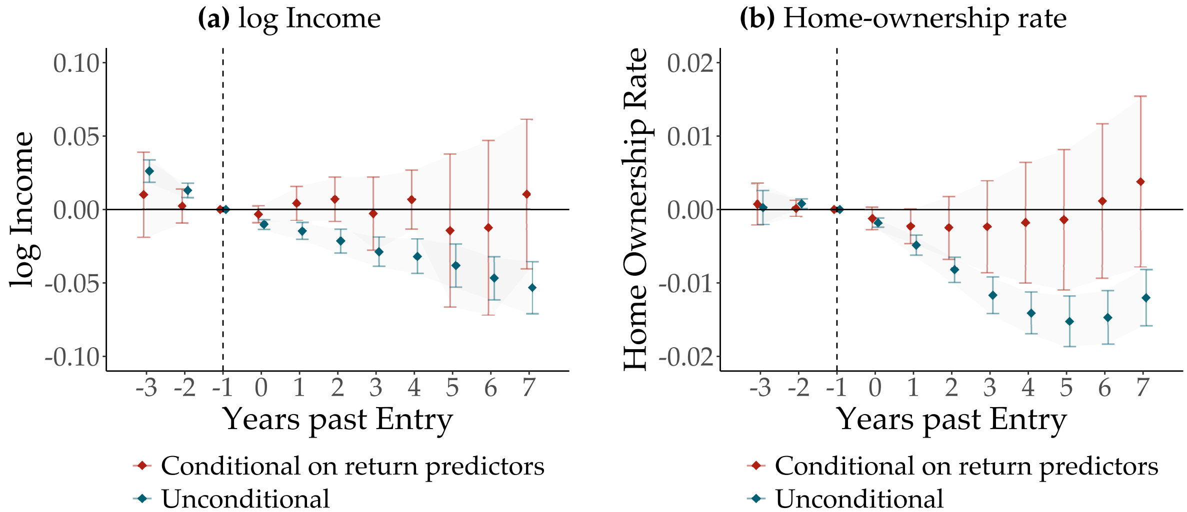

Figure 4: Home-ownership rates, income and institutional investor entry.
Figure shows difference-in-differences event study coefficients for equation (10) for annual home-ownership rates and log income. Coefficients in blue report estimates from an unconditional comparison of submarkets. Coefficients in red report estimates from a comparison conditional on predictors of future returns observed in Q4 of year prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

4.4 Other outcome variables

Repeating the exercise above for other outcome variables, such as income or home-ownership rates, provides additional context on the potential impact of institutional investors and the areas where they operate.

Home ownership rates and income Figure 4 repeats the previous exercise for annual home-ownership rates and average adjusted gross income estimates from the American Community Survey and IRS Statistics of Income.[19] Appendix figure D.3 repeats the exercise for bins sorted by the ex-post share of institutional investor-owned housing stock.

Institutional investors enter into submarkets that are on a persistent downward income trend relative to the rest of the U.S.. There is no indication that the entry of institutional investors has an effect on income, or by extension, local economic conditions. Compared to submarkets that would have been similarly attractive to invest in incomes are on identical trajectories. This implies that the increase in rents documented in figure 3 signifies a real decline in housing affordability that is not driven by local economic conditions.

[19]The ACS only releases 5-year estimates of zipcode level home-ownership rates. I use the last surveyed year as year of measurement, i.e. I use the reported values from the 2011-2015 ACS for the year 2015.

On average, home-ownership rates decline by 1.52pp within 5 years of institutional investor entry. However, there is very little difference compared to submarkets that would have been similarly attractive to invest in. This has two implications for the interpretation of previous findings.

First, in equilibrium institutional investors displace home-owners and replace them with tenants, rather than replacing incumbent landlords (see also Lambie-Hanson et al., 2022; Coven, 2023). Appendix Table D.1 corroborates this further by comparing zipcode and submarket-level changes in the home-ownership rate from 2011 to 2021 to the change in the share of institutional investor-owned housing stock. On average, a 1pp increase in the share of housing owned by institutional investors is associated with a 0.76pp decrease in home-ownership rates.

Second, this finding provides important context on how non-institutional rental supply would respond in the absence of institutional investors. The fact that home-ownership rates in areas that are similarly attractive to invest in follow identical trajectories implies that non-institutional investors have similarly elastic rental supply and would also convert owner-occupied homes into rental units in the absence of institutional investors. Appendix figure D.3 corroborates this further by repeating the exercise above for bins sorted by the ex-post share of institutional investor-owned housing stock.

Vacancy rates and housing supply Appendix figure D.4 repeats the previous exercise for vacancy rates and the log housing stock for each bin sorted on ex-post share of institutional investor-owned housing stock.

Panel (a) compares the log housing stock measured in the CoreLogic deeds data upon entry of institutional investors. Compared to all other submarkets the housing stock grows 0.15pp to 0.23pp faster per year. Relative to submarkets that would have been similarly attractive to invest in the housing supply grows at most 0.08pp per year faster. There is a strong pre-trend of housing supply growing at a faster rate even before entry, consistent with institutional investors preferentially investing in areas with more elastic housing supply. This cautions against a causal interpretation that institutional investors could have meaningful impact on the construction of new housing supply.

Panel (b) compares quarterly vacancy rates in CBRE rental data upon entry of institutional investors. These vacancy rates reflect the vacancy rates of a close substitute to single-family rentals that would be competing for the same tenants. Compared to all submarkets that institutional investors did not enter vacancy rates are about 0.5pp higher before entry and then decline by 1pp to 2pp after entry. Compared to submarkets that would have been similarly attractive to invest in there is little change in vacancy rates and the effects go in opposite directions for the middle bin compared to the top bin. In the middle bin vacancy rates slightly increase. In the top bin vacancy rates slightly increase. This makes it difficult to argue that institutional investors impact rental markets through quantity rationing, as in textbook models of oligopolistic competition.

4.5 Realized excess returns

This section inspects the realized excess returns to institutional investor investment decisions. I regress realized excess returns after operating costs $rx_{nt \to t+12}$ over 12 quarters from time t onto an indicator for institutional investor purchases $I(inv_{nt}^{buy} > 0)$.[20] In order to quantify the relative contributions of timing and locational choice I saturate the regression with increasingly granular fixed effects.

$$rx_{nt \to t+12} = \alpha_{(n,t)} + \beta I(inv_{nt}^{buy} > 0) + \varepsilon_{nt}$$

Table 6 reports the results. On average, institutional investors realized excess returns of 11.38% p.a. over the three years following their purchases, significantly exceeding the unconditional average excess return in the sample period. The inclusion of quarter and metro×quarter fixed effects reveals that the majority of realized excess returns is not idiosyncratic to individual submarkets but driven by wider aggregate and metro area trends. Notwithstanding, even *within* metro area-quarter institutional investors realize 0.91% p.a. higher excess returns on average.

Appendix figure D.5 decomposes realized excess returns into rental yields, appreciation and risk-free rates and repeats the exercise above for each component individually. The majority, around 70% of total realized excess returns are driven by appreciation. Holding the time of entry into the housing market or a metro area fixed, the contribution of rental yields to total excess returns increases. Conditional on the time of entry into a metro area, rental yields explain up to 50% of higher excess returns.

[20] Alternative investment horizons yield qualitatively identical results.

Table 6: Realized excess returns to investment decisions of institutional investors

	Dependent variable: $rx_{nt\rightarrow t+12}$ (in %, annualized)				
	(1)	(2)	(3)	(4)	(5)
$I(inv_{nt}^{buy} > 0)$	7.098***	2.806***	0.910**	2.585***	0.995***
	(0.8590)	(0.3017)	(0.3100)	(0.3681)	(0.2265)
Const.	4.282***				
	(0.7764)				
Observations	52,473	52,473	52,431	52,473	52,431
Adjusted R^2	0.05	0.59	0.89	0.65	0.93
Fixed effects		Quarter	Metro × Quarter	Quarter Submarket	Metro × Quarter Submarket

Note: *Table reports $\hat{\beta}$ for specifications of the form: $rx_{nt\rightarrow t+12} = \alpha_{(n,t)} + \beta I(inv_{nt}^{buy} > 0) + \varepsilon_{nt}$.*
Standard errors in parentheses clustered two-ways by submarket and quarter. $(^{+}, ^{*}, ^{**}, ^{***})$ *denote significance at 0.10, 0.05, 0.01, 0.001 level.*

5 Calibration exercise

This section performs a stylized calibration exercise to verify and illustrate the intuition of the model in section 1. The calibrated model generates rental market and house price dynamics, as well as institutional investor ownership patterns, that are consistent with the model predictions and empirical evidence. Further detail on the calibration procedure is provided in appendix E.

Given paths of rents, income, access to leverage, wealth and interest rates and the share of institutional investor-owned housing stock $\left\{d_t, y_t, L_t, W_t^h, R_t, \frac{q_t^i}{Q}\right\}_{t=1}^{T}$ I repeatedly solve an extension of the static model in section 1 for equilibrium quantities $\{q_t^i, q_t^h\}$ and prices $\{P_t, d_t\}$. I calibrate parameter values to match observed paths of house prices, rental market impact and institutional investor ownership in submarkets that institutional investors enter.

Income, house prices, access to leverage, rents, interest rates and the share of institutional investor-owned housing stock are directly measurable in the data. Rental market impact is estimated analogously to section 4. I estimate the path of wealth W_t to feed into the calibration in the Survey of Consumer Finances by computing the average wealth of households within a one standard-deviation range of income and house prices or rents of a representative submarket. Because the Survey of Consumer Finances is released in three-year waves I calibrate the model every three years from 2002 to 2021 and assume that one period lasts three years.

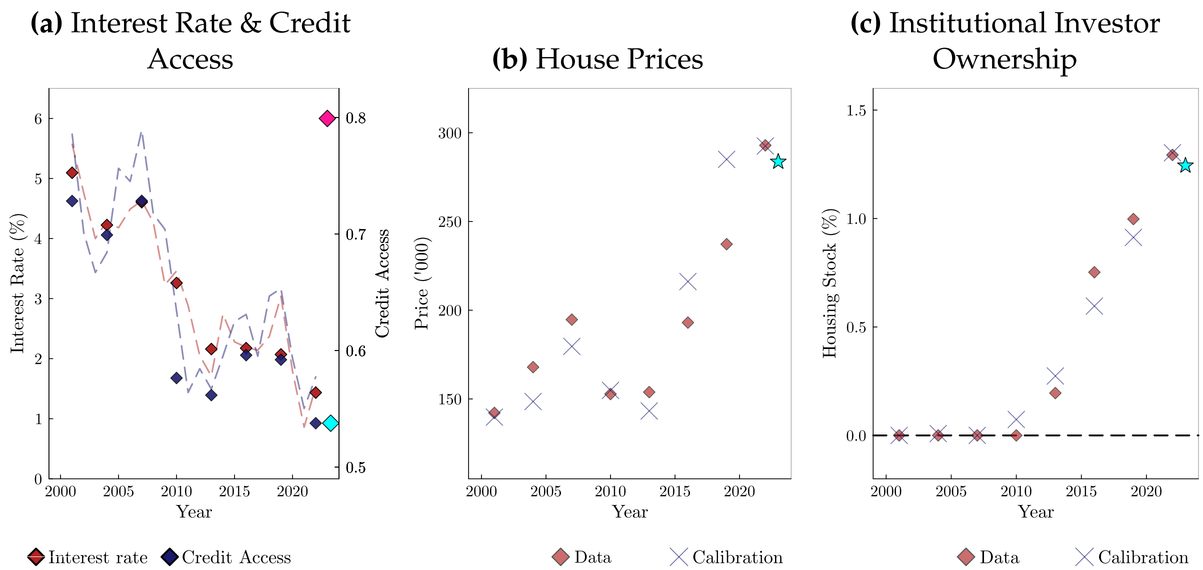

Figure 5: Calibration exercise and out-of-sample moments.
Figure plots data points and calibrated moments from the calibration exercise detailed in section 5 and appendix E. Panel (a) plots the path of interest rates and credit access fed into the calibration exercise. In-sample data points shaded in dark. Out-of-sample data points shaded brighlty. Panels (b) and (c) plot calibrated house prices and share of institutional investor-owned housing stock. Diamonds denote in-sample data points, crosses denote calibrated moments. Stars shaded brightly denote calibrated out-of-sample moments.

Out-of-sample counterfactuals I use the calibrated model to test whether the model can match out-of-sample findings. Between 2021Q4, when my sample coverage ends, and 2023Q4 the U.S. 10-year treasury yield increased by around 350 basis points. In response institutional investors have begun selling homes back to the household sector.[21]

In figure 5 I feed an increase in interest rates into the calibrated model after the sample period ends, holding all else equal. Panel (a) plots the path of interest rates and access to leverage that I feed into the calibration to estimate the model, as well as the out-of-sample data points I use to simulate an increase in interest rates. Panels (b) and (c) plot calibrated house prices and share of institutional investor-owned housing stock.

Consistent with the out-of-sample evidence, an increase in interest rates leads to a small decline in house prices and share of institutional investor-owned housing stock. Incumbent investors only make small adjustments to their portfolio because returns to scale and search costs incentivize them to remain in the market, even though expected excess returns decrease substantially.

[21] See e.g. Fortune, July 27 2023: Invitation Homes - the largest owner of U.S. homes - was a net seller for the third straight quarter.

6 Concluding remarks

In conclusion, this paper documents that household funding constraints and declining long-term interest rates can explain both expected and realized excess housing returns and the absence, entry and exit decisions of institutional investors across time and space. Institutional investors trade on household funding constraints: They buy homes when and where household funding constraints are tight (expected excess returns are high) and sell them back to the household sector when and where household funding constraints are slack (expected excess returns are low).

Viewed through the lens of this mechanism the emergence of institutional investors in the single-family rental market is an equilibrium response to a new environment of low long-term interest rates and increased lending standards in the aftermath of the housing bust. I formalize this intuition in a stylized model that yields predictions consistent with empirical in-sample and out-of-sample evidence.

In many ways institutional investors do not appear to be much different from other real estate investors. There is little evidence to support the view that institutional investors impact house prices, home-ownership rates, rental vacancy rates, local economic conditions, or construction of new housing supply differently from other investors who invest in similarly attractive markets in their absence. However, the evidence suggests that they impact rental markets in ways that other investors do not and make rental housing less affordable.

These findings have significant implications for policymakers and regulators concerned about housing affordability. In particular, my findings suggest that the first order concern for regulators should be the impact of institutional investors on rents, not on house prices. If there are scale efficiencies to large scale ownership of rental housing, these are not passed on to tenants in the form of lower rents. How regulators can incentivize institutional investors to pass these scale efficiencies on to tenants instead of raising rents is an open question that suggests there is much to be gained from further research.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2022, Credit Supply and House Prices: Evidence from Mortgage Market Segmentation.

Allen, Marcus T., Jessica Rutherford, Ronald Rutherford, and Abdullah Yavas, 2018, Impact of Investors in Distressed Housing Markets, *The Journal of Real Estate Finance and Economics* 56, 622–652.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, 2021, Synthetic difference-in-differences, *American Economic Review* 111, 4088–4118.

Austin, Neroli, 2023, Keeping up with the blackstones: Institutional investors and gentrification, SSRN Working Paper 4269561.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang, 2022, How much should we trust staggered difference-in-differences estimates?, *Journal of Financial Economics* 144, 370–395.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *The Journal of Business* 45, 444–55.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, 2021, Quasi-Experimental Shift-Share Research Designs, *The Review of Economic Studies* 89, 181–213.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2020, Why is intermediating houses so difficult? Evidence from iBuyers.

Callaway, Brantly, and Pedro H.C. Sant'Anna, 2021, Difference-in-differences with multiple time periods, *Journal of Econometrics* 225, 200–230, Themed Issue: Treatment Effect 1.

Campbell, John Y, and Robert J Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* 1, 195–228.

Campbell, John Y., and Roman Sigalov, 2022, Portfolio choice with sustainable spending: A model of reaching for yield, *Journal of Financial Economics* 143, 188–206.

Campbell, Sean D., Morris Davis, Joshua Gallin, and Robert F. Martin, 2009, What moves housing markets: A variance decomposition of the rent-price ratio, *Journal of Urban Economics* 66, 90–102.

CoreLogic, 2023, Residential real estate: Largest us asset class but not biggest economic driver.

Coven, Joshua, 2023, The impact of institutional investors on homeownership and neighborhood access, SSRN Working Paper 4554831.

Davis, Morris A., and Stijn Van Nieuwerburgh, 2015, Chapter 12 - Housing, Finance, and the Macroeconomy, in Gilles Duranton, J. Vernon Henderson, and William C. Strange, Editor, *Handbook of Regional and Urban Economics*, Volume 5 of *Handbook of Regional and Urban Economics*, 753–811 (Elsevier).

de Chaisemartin, Clément, and Xavier D'Haultfœuille, 2020, Two-way fixed effects estimators with heterogeneous treatment effects, *American Economic Review* 110, 2964–96.

Demers, Andrew, and Andrea L. Eisfeldt, 2022, Total returns to single-family rentals, *Real Estate Economics* 50, 7–32.

D'Acunto, Francesco, and Alberto G Rossi, 2021, Regressive Mortgage Credit Redistribution in the Post-Crisis Era, *The Review of Financial Studies* .

Eichholtz, Piet, Matthijs Korevaar, Thies Lindenthal, and Ronan Tallec, 2021, The Total Return and Risk to Residential Real Estate, *The Review of Financial Studies* 34, 3608–3646.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Ganduri, Rohan, Steven Chong Xiao, and Serena Wenjing Xiao, 2023, Tracing the source of liquidity for distressed housing markets, *Real Estate Economics* 51, 408–440.

Garriga, Carlos, Pedro Gete, and Athena Tsouderou, 2023a, The economic effects of real estate investors, *Real Estate Economics* 51, 655–685.

Garriga, Carlos, Pedro Gete, and Athena Tsouderou, 2023b, Investors in housing markets: A story of two booms, SSRN Working Paper 3962420.

Gete, Pedro, and Michael Reher, 2018, Mortgage supply and housing rents, *Review of Financial Studies* 31, 4884–4911.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, 2020, Bartik instruments: What, when, why, and how, *American Economic Review* 110, 2586–2624.

Goodman-Bacon, Andrew, 2021, Difference-in-differences with variation in treatment timing, *Journal of Econometrics* 225, 254–277, Themed Issue: Treatment Effect 1.

Gorback, Caitlin, Franklin Qian, and Zhu Zipei, 2023, Impact of institutional owners on housing markets.

Greenwald, Daniel L, and Adam Guren, 2021, Do credit conditions move house prices?, Working Paper 29391, National Bureau of Economic Research.

Gurun, Umit G, Jiabin Wu, Steven Chong Xiao, and Serena Wenjing Xiao, 2022, Do Wall Street Landlords Undermine Renters' Welfare?, *The Review of Financial Studies* 36, 70–121.

Highfield, Michael J., Lily Shen, and Thomas M. Springer, 2021, Economies of Scale and the Operating Efficiency of REITs: A Revisit, *The Journal of Real Estate Finance and Economics* 62, 108–138.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante, 2020, The housing boom and bust: Model meets evidence, *Journal of Political Economy* 128, 3285–3345.

Korevaar, Matthijs, 2023, Reaching for yield and the housing market: Evidence from 18th-century amsterdam, *Journal of Financial Economics* 148, 273–296.

Lambie-Hanson, Lauren, Wenli Li, and Michael Slonkosky, 2022, Real estate investors and the U.S. housing recovery, *Real Estate Economics* 50, 1425–1461.

Lilley, Matthew, and Gianluca Rinaldi, 2021, Credit supply and house prices: Evidence from conforming loan limits.

Mian, Atif, and Amir Sufi, 2011, House prices, home equity-based borrowing, and the us household leverage crisis, *American Economic Review* 101, 2132–56.

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

Piazzesi, Monika, and Martin Schneider, 2016, Housing and macroeconomics, Working Paper 22354, National Bureau of Economic Research.

Raymond, Lindsey, 2023, The market effects of algorithms.

Schmalz, Martin C, David A Sraer, and David Thesmar, 2017, Housing collateral and entrepreneurship, *The Journal of Finance* 72, 99–132.

Sinai, Todd, and Nicholas S. Souleles, 2005, Owner-Occupied Housing as a Hedge Against Rent Risk*, *The Quarterly Journal of Economics* 120, 763–789.

Smith, Patrick S., and Crocker H. Liu, 2020, Institutional Investment, Asset Illiquidity and Post-Crash Housing Market Dynamics, *Real Estate Economics* 48, 673–709.

US Census, Bureau, 2024, American housing survey (ahs).

A Appendix: Model

A.1 Proofs

In section 1 I assume that income and house prices are uncorrelated. Relaxing this assumption leads to household and investor demand

$$q_t^h(P_t|W_t^h) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - (R^f + \lambda)P_t - A^h\sigma_{py^h}}{A^h\sigma_p^2}\right\}$$

$$q_t^i(P_t|q_{t-1}^i) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha - (R^f - \gamma q_{t-1}^i)P_t - A^i\sigma_{py^i}}{A^i\sigma_p^2 + \gamma P_t - \beta}\right\}$$

If income is positively correlated with house prices this makes housing less attractive as it creates additional risk exposure.

Proposition 1. Proposition 1 follows directly from the investor demand function and is only mentioned for the sake of completeness. Investors add additional units of housing supply to their portfolio if and only if $q_t^i(P_t|q_{t-1}^i) > q_{t-1}^i$. Rearranging investor demand yields the price at which an investor adds an additional unit of housing supply to their portfolio in equation (1).

$$\overline{P}_t^i(q_{t-1}^i) = \frac{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha - A^i\sigma_{py^i}}{R^f} + \frac{\beta - A^i\sigma_p^2}{R^f}q_{t-1}^i \tag{1}$$

Corollary 1. Corollary 1 follows from adopting the perspective of the housing market as a market in which households and the investor compete to invest in housing capital that provides housing services. Let $R_{t+1} \equiv \frac{\mathbb{E}_t[P_{t+1}]+d}{P_t}$ denote the expected returns to investing in housing capital. Substituting this expression into household and investor demand yields

$$q_t^i(R_{t+1}|q_{t-1}^i) = \max\left\{0, \frac{\frac{\mathbb{E}_t[P_{t+1}]+d}{R_{t+1}}\left(R_{t+1} - R^f + \gamma q_{t-1}^i\right) - \alpha - A^i\sigma_{py^i}}{\gamma\frac{\mathbb{E}_t[P_{t+1}]+d}{R_{t+1}} + A^i\sigma_p^2 - \beta}\right\}$$

$$q_t^h\left(R_{t+1}|W_t^h\right) = \max\left\{0, \frac{\nu}{A^h\sigma_p^2} + \frac{\mathbb{E}_t\left[P_{t+1}\right] + d}{R_{t+1}}\frac{R_{t+1} - R^f - \lambda}{A^h\sigma_p^2} - \frac{\sigma_{py^h}}{\sigma_p^2}\right\}$$

Corollary 1 follows directly from following analogous steps to the above.

$$\underline{rx}_{t+1}(q_{t-1}^i) = \log\left(1 + \frac{\alpha + A^i\sigma_{py^i} - (\beta - A^i\sigma_p^2)q_{t-1}^i}{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha - A^i\sigma_{py^i} + (\beta - A^i\sigma_p^2)q_{t-1}^i}\right) \tag{2}$$

Proposition 2. For an equilibrium in which investors optimally invest all their wealth in the risk-free asset to exist, household demand has to be able to clear the market at a price above $\overline{P}^i_t(0)$. Relaxing the household constraint to $\lambda = 0$ yields equation 3.

$$\frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - R^f \overline{P}^i_t(0)}{A^h \sigma_p^2} \geq 1 \text{ iff } \nu - A^h\left[\sigma_p^2 + \sigma_{py^h}\right] > -\alpha - A^i \sigma_{py^i} \tag{3}$$

Proposition 3. Combining proposition 1 that investors expand if house prices are below $\overline{P}^i_t(q^i_{t-1})$ with market clearing $q^h_t(P_t) + \frac{I q^i_t(P_t)}{Q} = 1$ requires that

$$1 - \frac{I q^i_{t-1}}{Q} > \frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - A^h \sigma_{py^h} - \left[R^f + \lambda\left(\overline{P}^i_t(q^i_{t-1})\right)\right] \overline{P}^i_t(q^i_{t-1})}{A^h \sigma_p^2} = q^h_t(\overline{P}^i_t(q^i_{t-1}))$$

This requires to distinguish between two cases: If

$$\frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - A^h \sigma_{py^h} - R^f \overline{P}^i_t\left(\frac{Q}{I}\right)}{A^h \sigma_p^2} > 0 \tag{11}$$

equation (4) is a necessary and sufficient condition for investors to expand: even if investors own the entire housing stock, households value a unit of housing supply higher than investors. The only case in which investors buy more homes is if households are sufficiently funding constrained, i.e. if $\lambda\left(\overline{P}^i_t(q^i_{t-1})\right)$ is sufficiently large. Evaluating household demand at the constraint yields equation (4) of proposition 3.

If equation (11) does not hold, equation (4) is a sufficient but not necessary condition for investors to expand. In this case there is a $\underline{q}^i_{t-1} < \frac{Q}{I}$ for which investors expand even if households are not funding constrained. In this case equation (4) is a necessary and sufficient condition for investors to expand only if $q^i_{t-1} < \underline{q}^i_{t-1}$.

In cases where (4) is a necessary and sufficient condition for the investor to expand, an investor owning $q^i_{t-1} > 0$ units of housing supply will sell homes back to the household sector whenever household funding constraints are sufficiently slack.

In order to make the connection between household funding constraints and investor investment decisions even more explicit, I re-express equation 4 in terms of the value of a home to households relative to the funds at their disposal. Let V^h denote the value of a home to a household that follows from unconstrained household demand $q^h(1)$

$$V^h = \frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - A^h[\sigma_{py^h} + \sigma_p^2]}{R^f}$$

Dividing equation (4) by V^h yields an equivalent expression in terms of $\frac{W_t^h - d}{1-L}/V^h$:

$$\frac{\frac{W_t^h - d}{1-L}}{V^h} < \frac{1 - \frac{Iq_{t-1}^i}{Q}}{1 + \Lambda(q_{t-1}^i)}$$

The left hand side of this expression measures the tightness of household funding constraints in terms of the ratio of funds at the disposal of households relative to the value of a home to a household. The right hand side is a constant that is decreasing in q_{t-1}^i.[22] Investors buy (sell) homes when the value of a home increases (decreases) relative to the funds at the disposal of households.

Proposition 4. Proposition 4 follows directly from market clearing and the fact that demand is downward sloping in price. Let $J(P_t) = q_t^h(P_t) + \frac{Iq_t^i(P_t)}{Q}$ denote the joint demand of households and investors for a unit of housing supply. Since $J(P_t) \geq q^h(P_t)$ it follows that the joint demand of households and investors clears the market at a price above the price where household demand alone would clear the market.

A.2 Extensions

A.2.1 Rent bargaining

I give the investor ability to raise rents by introducing a Nash bargaining game between owners and residents of a home. Let u denote the value of residency, equal to the value of local wages and amenities and opportunity costs of finding an outside option. Let $\underline{d}$ denote the user cost of housing. Owner and resident bargain over the surplus created by residency $u - \underline{d}$ by choosing rents d.

$$(u - d)^{(1-\chi)}(d - \underline{d})^{\chi}$$

I assume that the owner's bargaining weight $\chi = \delta \frac{q_t^i}{Q}$ is increasing in the share of the housing stock owned by the investor for some $\delta \in [0, 1]$. This is a simple, reduced form way to model potential impact on rental markets motivated by the fact that owning a larger share of the housing stock gives an investor more bargaining power due to i) the fact that they have more repeated interactions with individual residents that allow them to approach individual negotiations more aggressively and ii) the fact that residents have

[22] equal to $\Lambda(q_{t-1}^i) = \frac{\nu + \alpha - A^h[\sigma_{py^h} + \sigma_p^2] + A^i \sigma_{py^i} - (\beta - A^i \sigma_p^2) q_{t-1}^i}{\mathbb{E}_t[P_{t+1}] + d - \alpha - A^i \sigma_{py^i} + (\beta - A^i \sigma_p^2) q_{t-1}^i}$

fewer local outside options. This leads to rents

$$d = \underline{d} + \chi(u-c) = \underline{d}\left(1 + \frac{\mu}{2}\frac{q_t^i}{Q}\right)$$

where $\mu = \frac{2\delta(u-c)}{c}$ is a constant introduced for convenience. Intuitively, if households own the entire housing stock they charge themselves rent equal to the user cost of housing. This places a lower bound $\underline{d}$ on the rental rate. Conversely, if a single investor owns the entire housing stock they will set rents as high as the parameter δ permits, up to point where residents are indifferent between renting and finding an outside option if $\delta = 1$.

The ability of investors to raise rents enters into both the investor and the household demand function. It gives households and investors an additional incentive to invest in homes - investors have the possibility to extract higher rents, while households insure themselves against the investor extracting higher rents (see Sinai and Souleles, 2005).

$$q_t^i(P_t|q_{t-1}^i) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + \underline{d} - \alpha - (R - \gamma q_t^i)P_t - A^i\sigma_{py^i}}{A^i\sigma_p^2 + \gamma P_t - \beta - \frac{\mu}{Q}\underline{d}}\right\}$$

$$q_t^h(P_t|W_t^h) = \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + \underline{d}\left(1 + \frac{\mu}{2}\frac{q_t^i(P_t)}{Q}\right) + \nu - (R^f + \lambda)P_t - A^h\sigma_{py^h}}{A^h\sigma_p^2}\right\}$$

The results and propositions follow analogously to the previous section, except that demand and equilibrium rents d are functions of the user cost of housing $\underline{d}$.

Unlike for house prices, where the impact is increasing in the total share of investor owned housing stock, the impact is increasing in the share of housing stock owned by an individual investor. This means that while it is unclear whether a larger or small number of investors has different impacts on house prices, a smaller number of investors will have larger impact on rents.

A.2.2 Duration

Holding all else constant a decline in the risk-free rate R^f will always induce investors to buy more homes. However, household wealth and expectations of future house prices commonly also depend on the level of the risk-free rate. In this case whether a decline in the risk-free rate induces additional investor purchases depends on how sensitive to changes in the interest rate household wealth and expectations of future house prices are.

Let $W_t^h(R^f)$ and $\mathbb{E}_t\left[P_{t+1}(R^f)\right]$ denote household wealth and expectations of future house prices as functions of the risk-free rate and δ^W, δ^P their respective derivatives. In order for a decline in the risk-free rate to induce the investor to buy additional homes it has to be the case that the left hand side of equation (4) increases less than the right hand side:

$$\frac{\partial}{\partial R^f}\frac{W_t^h(R^f)-d}{1-L} > \frac{\partial}{\partial R^f}\left(1-\frac{Iq_{t-1}^i}{Q}\right)\left[\frac{\mathbb{E}_t\left[P_{t+1}(R^f)\right]+d-\alpha-A^i\sigma_{py^i}+(\beta-A^i\sigma_p^2)q_{t-1}^i}{R^f}\right]$$

Trivially, if $\frac{\partial W_t^h(R^f)}{\partial R^f} = \frac{\partial \mathbb{E}_t\left[P_{t+1}(R^f)\right]}{\partial R^f} = 0$ as in section 1 this is always the case. If not, the equation above can be expressed in terms of the respective durations of expected future house prices and household wealth $D^p = -\delta^P \frac{1}{\mathbb{E}_t\left[P_{t+1}(R^f)\right]}$ and $D^W = -\delta^W \frac{1}{W_t^h(R^f)-d}$.

$$D^W < \left(D^P + \frac{\overline{P}_t^i(q_{t-1}^i)}{R^f\mathbb{E}_t\left[P_{t+1}(R^f)\right]}\right)\frac{\frac{\mathbb{E}_t\left[P_{t+1}(R^f)\right]}{R^f}}{\frac{W_t^h-d}{1-L}} \times \left(1-\frac{Iq_{t-1}^i}{Q}\right)$$

This places an upper bound on the duration of household wealth for which a decline in the risk-free rate means that the sufficient condition in (4) is more likely to be satisfied. The bound is increasing in the duration of expected future house prices and decreasing in households funds and the share of housing stock owned by investors.

At the boundary where equation (4) holds with equality, that is where all else equal investors want to neither buy or sell homes, a decline in the interest rate induces investors to buy homes if

$$D^W < \frac{\mathbb{E}_t\left[P_{t+1}(R^f)\right]}{\mathbb{E}_t\left[P_{t+1}(R^f)\right]+d-\alpha-A^i\sigma_{py^i}+(\beta-A^i\sigma_p^2)q_{t-1}^i}D^P + \left(\frac{1}{R^f}\right)^2$$

which, for most specifications, requires that household wealth has lower duration than expected future house prices.

A.2.3 Endogenous housing supply

The model can easily be extended to include endogenous construction of new housing supply. Instead of Q households occupying an equal number of housing units let H denote the mass of households and the total housing supply be equal to

$$Q + \Delta Q(P_t) = Q + \max\{0, -B + \eta P_t\}$$

where Q denotes existing housing supply and ΔQ denotes new construction of housing supply that depends on B, a minimum house price above which new construction is profitable for builders and η, a parameter that depends on the elasticity of housing supply. Equilibrium market clearing now requires that

$$Hq^h(P_t) + Iq^i(P_t) = Q + \Delta Q(P_t)$$

which means that equation (4) is replaced by

$$\frac{W_t^h - d}{1 - L} < \frac{Q - Iq_{t-1}^i + \Delta Q(\overline{P}_t^i(q_{t-1}^i))}{H} \overline{P}_t^i(q_{t-1}^i)$$

implying that the sufficient condition for (additional) investor purchases is more likely to be met if housing supply is more elastic (η is large, B is small).

A.2.4 Idiosyncratic risk

Another commonly cited advantage to owning a large number of properties is the ability to pool risk across properties (Piazzesi and Schneider, 2016). I assume that housing risk consists of a systemic risk component with variance σ_p^2 and an additional idiosyncratic component with variance σ_ε^2 that is diversifiable for investors but not diversifiable for households. Abstracting away from divisibility concerns, this implies that the variance of investor and household wealth is equal to

$$\begin{aligned} Var(\tilde{W}_{t+1}^h) &= (q_{t+1}^h)^2 \left[\sigma_p^2 + \sigma_\varepsilon^2\right] + q_t^h \sigma_{py^h} + \sigma_{y^h}^2 \\ Var(\tilde{W}_{t+1}^i) &= (q_{t+1}^i)^2 \sigma_p^2 + q_t^i \left[\sigma_{py^i} + \sigma_\varepsilon^2\right] + \sigma_{y^i}^2 \end{aligned}$$

which leads to demand functions

$$\begin{aligned} q_t^h(P_t|W_t^h) &= \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d + \nu - (R+\lambda)P_t - A^h \sigma_{py^h}}{A^h\left[\sigma_p^2 + \sigma_\varepsilon^2\right]}\right\} \\ q_t^i(P_t|q_{t-1}^i) &= \max\left\{0, \frac{\mathbb{E}_t\left[P_{t+1}\right] + d - \alpha - (R - \gamma q_{t-1}^i)P_t - A^i\left[\sigma_{py^i} + \frac{1}{2}\sigma_\varepsilon^2\right]}{A^i \sigma_p^2 + \gamma P_t - \beta}\right\} \end{aligned}$$

This does not affect any of the core results of the model, but places a stricter bound on nontradable benefits and operating costs in equation (3) to support an equilibrium where households own the entire housing stock in proposition 2.

$$\nu - A^h[\sigma_p^2 + \sigma_\varepsilon^2 + \sigma_{py^h}] > -\alpha - A^i\left[\sigma_{py^i} + \frac{1}{2}\sigma_\varepsilon^2\right]$$

All other propositions follows analogously to above.

A.2.5 Heterogeneous households

The model can be extended to allow for heterogeneous households in order to accommodate a household landlord sector. Non-tradable benefits of home-ownership only accrue to the residence of a household. This means that next period household wealth follows

$$W_{t+1}^h = q_t^h\left(P_{t+1} + d - R^f P_t\right) + R^f(W_t^h - d) + y_{t+1}^h + \nu\left[q_t^h I(q_t^h \leq 1) + I(q_t^h > 1)\right]$$

which leads to individual household demand

$$q^h(P_t|W_t^h, L^h) = \begin{cases} \frac{\mathbb{E}_t[P_{t+1}]+d-\left[R^f+\lambda(P_t)\right]P_t}{A^h\sigma^2} & \text{if } P_t < \frac{\mathbb{E}_t[P_{t+1}]+d-A^h\sigma_p^2}{R^f} \\ \min\left\{\frac{\mathbb{E}_t[P_{t+1}]+d+\nu-\left[R^f+\lambda(P_t)\right]P_t}{A^h\sigma^2}, 1\right\}, & \text{if } \frac{\mathbb{E}_t[P_{t+1}]+d-A^h\sigma_p^2}{R^f} < P_t < \frac{\mathbb{E}_t[P_{t+1}]+d+\nu-A^h\sigma_p^2}{R^f} \\ \frac{\mathbb{E}_t[P_{t+1}]+d+\nu-\left[R^f+\lambda(P_t)\right]P_t}{A^h\sigma^2} & \text{if } P_t > \frac{\mathbb{E}_t[P_{t+1}]+d+\nu-A^h\sigma_p^2}{R^f} \end{cases}$$

where λ denotes the Lagrange multiplier on the leverage constraint. In this case aggregate household demand depends on the distribution of wealth and access to leverage across households $f(W_t^h, L^h)$.

$$\mathbf{q}^h(P_t) = \int_0^1 \int_0^\infty q^h(P_t|W_t^h, L^h) f(W_t^h, L^h) dW_t^h dL^h$$

The intuition follows analogously, except that household demand now is a function of the distribution of wealth and access to leverage across households $f(W_t^h, L^h)$. Market clearing requires that

$$\mathbf{q}^h(P_t) + I q_t^i(P_t) = Q$$

If households are heterogeneous and the constraint binds for some households, these households end up renting part of their residence from less constrained households or from investors. Specifically, the sufficient condition for the household sector to own the entire housing stock now is equal to $\mathbf{q}^h(\overline{P}_t^i(0)) \geq Q$.

B Appendix: Data

B.1 House price and rental data

Rental Data As noted in section 2, the availability of high frequency geographically disaggregated rental data reaching back in time is a challenge. To the best of my knowledge the CBRE Torto-Wheaton Same Store Rent index for multi-family housing units is the only time series of geographically disaggregated rental data that reaches back further than the early 2010s. Figure B.1 plots log rents from the CBRE EA multifamily rental index against log rents from Zillow and shows that both measures are highly correlated in the time window where coverage overlaps (2014-2022).

House price data I use the 2022 release of monthly zipcode-level home price time series from Zillow as baseline for house price data and aggregate these prices using the zipcode-to-submarket crosswalk. This vintage is the last release that does not use Zestimates to construct house price indices. I drop four submarkets (two in Detroit, one in Memphis, one in Jacksonville) where rent-price ratios in both Zillow and CBRE rental data are implausibly high due to very low house prices from my sample.

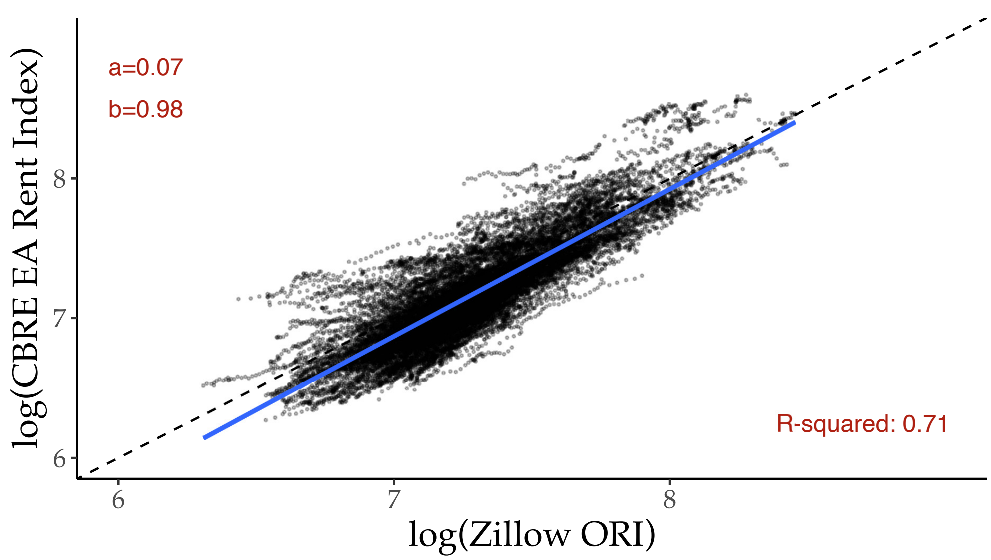

Figure B.1: Comparison of rental prices. *Figure plots log CBRE rental index against log Zillow Overall Rent index for the time period of 2014Q1-2022Q2 where coverage overlaps.*

B.2 Institutional investors

Measuring institutional investor activity Measuring institutional investor activity and ownership requires identifying individual institutional investors at the transaction level. Since many institutional investors act through subsidiaries this is not entirely straightforward.

I identify institutional investor subsidiaries through two complementary approaches. Firstly, for Real Estate Investment Trusts that are required to file with the SEC I match owner names to lists of subsidiaries obtained from exhibit 21.1 of SEC 10-K filings. Secondly, PropStream is a real estate investment software that identifies common owners across properties based on physical and online mailing addresses and phone numbers. I use PropStream to identify subsidiaries of institutional investors by matching the provided public identifiers of large property owners in PropStream to publicly available identifiers of institutional investors. This gives me a list of subsidiaries and their parent companies which I use to match names of buyers and sellers recorded in the deeds records.

Exhibit 21.1

Subsidiaries of the Registrant
Effective 12/31/2019

Name	State of Formation
2013-1 IH Borrower G.P. LLC	Delaware
2013-1 IH Borrower L.P.	Delaware
2013-1 IH Equity Owner G.P. LLC	Delaware
2013-1 IH Equity Owner L.P.	Delaware
2013-1 IH Property Holdco L.P.	Delaware
2015-3 IH2 Borrower TRS LLC	Delaware
2015-3 IH2 Property Holdco L.P.	Delaware
2015-3 IH2 Equity Owner G.P. LLC	Delaware
2015-3 IH2 Equity Owner L.P.	Delaware
2015-3 IH2 Borrower G.P. LLC	Delaware
2017-1 IH Property Holdco L.P.	Delaware
2017-1 IH Equity Owner G.P. LLC	Delaware
2017-1 IH Equity Owner L.P.	Delaware
2017-1 IH Borrower GP LLC	Delaware

Figure B.2: Excerpt of 13F listing of subsidiaries, Invitation Homes (2019)

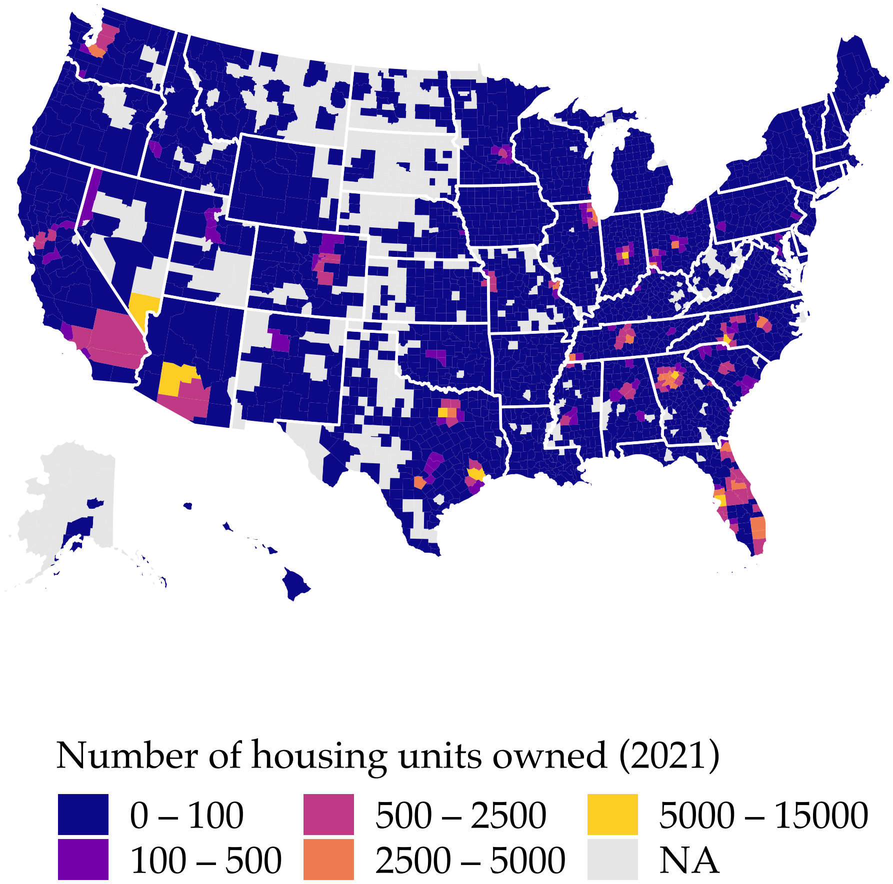

Figure B.3: Geographic distribution of institutional investors, 2021Q4. *Figure shows geographic distribution of institutional investors. Counties with an estimated housing stock of less than 1000 single family housing units shaded in gray. Source: CoreLogic housing deeds.*

I proceed to construct a panel of institutional investor purchases and institutional investor-owned housing stock by aggregating transaction level data. The first step in my analysis are preliminary data cleaning steps in the raw deeds data. Starting from 456,993,121 observation records, I restrict my analysis to deeds records of properties listed as single-family homes, townhomes or duplexes.[23] I discard observations that are marked as pending, missing transaction dates or zipcodes, or include mobile homes on the property. I also discard zipcodes with fewer than 50 total deeds records.

In the next step I assemble zipcode-by-quarter panels of institutional investor purchases and ownership, proceeding zipcode by zipcode. Constructing the panel of transactions is straightforward. For each zipcode I restrict the deeds records to transactions marked as

[23]Property Indicator Codes $\in \{10, 21\}$ and/or Land Use Code $\in \{102, 115, 163\}$.

arms-length purchases and drop transactions with missing, zero or negative sale price. I then proceed to count the number of transactions that involve subsidiaries of institutional investors listed in table 1 as buyer and aggregate this number at the zipcode$\times$quarter level. This allows me to back out both the number of properties bought by institutional investors and their market share of purchases in a neighborhood in each quarter.

Measuring ownership requires a measure of the housing stock at any point in time. To construct my measure of the housing stock I look at the first recorded transaction of each property. If this transaction is marked as resale of an existing property, I assume the property has been in the housing stock before sample coverage begins. If it is marked as new construction sale it enters the housing stock at the time of transaction, which leads to a slowly increasing housing stock over time. I impute the owner of each property at the end of each quarter by extrapolating buyer and seller names forward and backward. For any given property and point in time in my sample, I impute the current owner as the buyer listed on the most recent deeds record of that property. If the buyer name is missing, I impute the current owner from the seller name of the next transaction involving the property. I also use the seller name to impute the owner of the property prior to the first recorded transaction of the property for deeds that are marked as resale transaction in the first recorded deed. I then measure institutional investor ownership by matching the imputed owner names to the list of subsidiaries.

Lastly, I aggregate my zipcode$\times$quarter panels up to the submarket$\times$quarter level. Figure B.3 shows the number of institutional investor-owned properties across counties in 2021Q4.

B.3 Summary statistics

Table B.1: Summary Statistics

(a) Full sample, 2000Q1-2021Q4

	Coverage	N	Mean	SD	p5	p95	Wtd. Mean	Wtd. SD
Single Family Housing (Units)	1996-2021	88660	57219	45245	8322	132797		
Arms Length Transactions	1996-2021	88660	859	786	41	2373		
Average adjusted gross income (IRS)	2002-2021	65120	73442	44904	34345	151478	68403	33327
Cred. Score bw. 580-680 (% of Population)	2006-2021	48840	23	4.8	15	31	23	4.4
Minority Population Share (2010 Census %)	2010	814	43	22	13	84	40	22
Black Population Share (2010 Census %)	2010	814	15	17	1.3	51	14	15
Median Age (2010 Census)	2010	814	36	3.8	30	42	37	3.6
Average monthly rent (USD)	2000-2021	64539	1178	544	641	2208	1161	477
Average house price ('000 USD)	2000-2021	69804	307	251	103	758	292	215
Rent-to-Income ratio	2002-2021	58961	0.21	0.081	0.12	0.36	0.22	0.078
Price-to-Income ratio	2002-2021	63679	4.3	2.2	2.2	8.8	4.4	2.2
Home Ownership Rate (%, ACS)	2011-2021	8954	59	15	31	78	65	12

(b) Split sample, 2012Q1-2021Q4

	Coverage	N	Mean	SD	p5	p95	Wtd. Mean	Wtd. SD
Institutional Investor Submarket: No								
Single Family Housing (Units)	2012-2021	16529	51310	48279	7298	135153		
of which Institutional Investor-owned (Units)	2012-2021	16529	9.3	18	0	50	11	20
in 2021Q4	2021$Q4$	419	16	23	0	68	17	24
of which Institutional Investor-owned (%)	2012-2021	16529	0.025	0.05	0	0.13	0.018	0.038
in 2021Q4	2021$Q4$	419	0.046	0.077	0	0.23	0.03	0.055
Arms Length Transactions	2012-2021	16529	688	666	73	1958		
of which Institutional Investor Purchases (%)	2012-2021	16496	0.049	0.26	0	0.23	0.036	0.2
Average adjusted gross income (IRS)	2012-2021	16529	100673	64476	41242	234136	90803	47911
Cred. Score bw. 580-680 (% of Population)	2012-2021	14853	21	5.1	13	30	21	4.5
in 2006Q1	2006Q1	419	22	4.3	16	30	22	4
Minority Population Share (2010 Census %)	2010	419	43	21	14	83	39	22
Black Population Share (2010 Census %)	2010	419	14	16	1.2	50	12	15
Median Age (2010 Census)	2010	419	37	3.7	31	42	38	3.2
Average monthly rent (USD)	2012-2021	16529	1588	722	780	2920	1564	624
Average house price ('000 USD)	2012-2021	16529	453	370	122	1192	426	327
Rent-to-Income ratio	2012-2021	16529	0.22	0.095	0.099	0.4	0.23	0.098
Price-to-Income ratio	2012-2021	16529	4.6	2.4	2	9.4	4.6	2.4
Home Ownership Rate (%, ACS)	2012-2021	4141	55	16	25	76	62	14
Institutional Investor Submarket: Yes								
Single Family Housing (Units)	2012-2021	15078	67258	41573	24047	145974		
of which Institutional Investor-owned (Units)	2012-2021	15078	413	561	1	1442	526	670
in 2021Q4	2021Q4	379	813	938	102	2599	1007	1102
of which Institutional Investor-owned (%)	2012-2021	15078	0.65	0.79	0.0026	2.1	0.61	0.74
in 2021Q4	2021Q4	379	1.3	1.3	0.17	4.1	1.2	1.2
Arms Length Transactions	2012-2021	15078	1164	802	326	2806		
of which Institutional Investor Purchases (%)	2012-2021	15078	1.1	2.1	0	5	1	1.9
Average adjusted gross income (IRS)	2012-2021	15078	68251	28341	36499	117754	67865	24889
Cred. Score bw. 580-680 (% of Population)	2012-2021	13562	25	4.3	18	32	25	4.1
in 2006Q1	2006Q1	379	26	3.7	19	32	26	3.7
Minority Population Share (2010 Census %)	2010	379	44	22	13	85	42	22
Black Population Share (2010 Census %)	2010	379	16	17	1.8	49	15	15
Median Age (2010 Census)	2010	379	36	3.8	30	41	36	3.7
Average monthly rent (USD)	2012-2021	15078	1116	319	711	1737	1130	316
Average house price ('000 USD)	2012-2021	15078	243	115	105	468	248	117
Rent-to-Income ratio	2012-2021	15078	0.21	0.074	0.12	0.35	0.21	0.073
Price-to-Income ratio	2012-2021	15078	3.7	1.5	2.2	6.6	3.8	1.7
Home Ownership Rate (%, ACS)	2012-2021	3773	64	11	43	78	66	9.8

Summary statistics of complete data coverage, including all submarkets and quarters of which some are not covered in the main analysis. Panel (a) covers the full sample from 2000Q1 to 2021Q4. Panel (b) covers the split sample from 2012Q1 to 2021Q4.

C Appendix: Excess returns and household funding constraints

C.1 Instrument Design

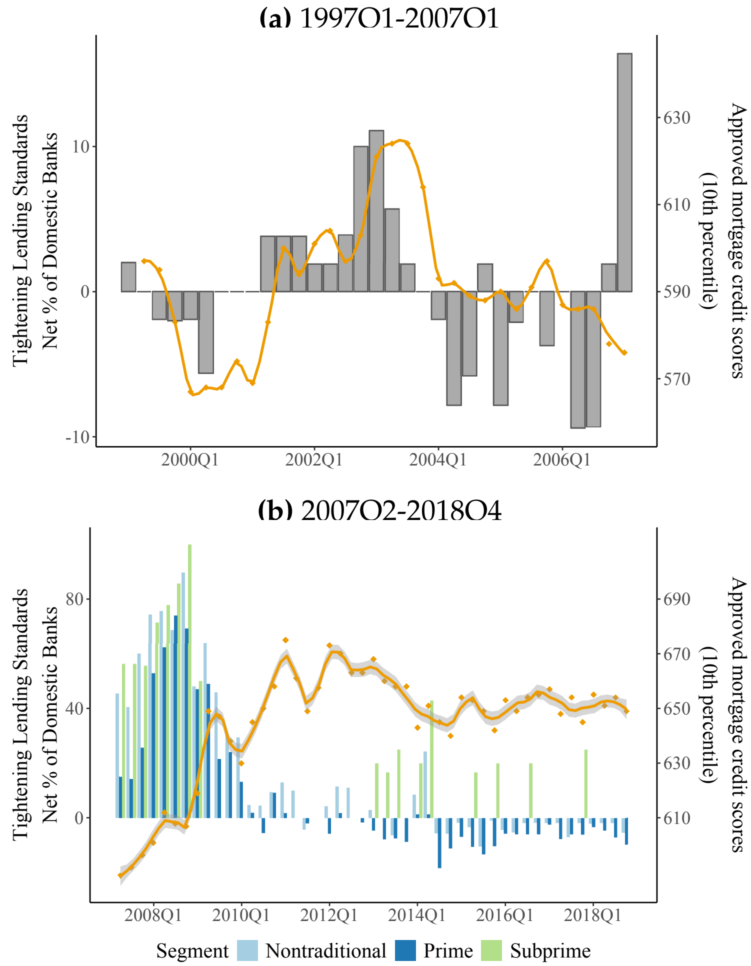

Figure C.1: 10th percentile of approved mortgage credit scores and Federal Reserve Board Senior Loan Officer Opinion Survey.
Figure plots the net percentage of domestic banks reporting a tightening of lending standards on the left axis in the Federal Reserve Senior Loan Officer Opinion Survey on the left axis. Right axis plots the quarterly 10th percentile of approved mortgage applicants' credit scores as measured by New York Federal Reserve Consumer Credit Panel. Panel (a) covers time period from 1997Q1 to 2007Q1. Panel (b) covers time period from 2007Q2 to 2018Q1. Split sample is due to a change in the methodology of the FRB Survey of Senior Loan Officers.

C.2 Instrumental Variable Estimates

Table C.1: Instrumental Variable estimates: Alternative measure of expected excess returns

(a) Excess rental yields dpx_{nt}

	dpx_{nt} (%, annualized)		$I(inv^{buy}_{nt} > 0)$ (IV)	
	(1)	(2)	(3)	(4)
dpx_{nt} (%, annualized)			0.060***	0.093**
			(0.0149)	(0.0304)
$dist(inv^{own}_{m(n),t-1})$	-0.018	-0.018+	-0.010**	-0.009**
	(0.0151)	(0.0104)	(0.0033)	(0.0032)
Credit Access $L^{IV}_{nt}(\%)$	-0.109***	-0.060***		
	(0.0140)	(0.0113)		
Observations	41,091	41,091	41,091	41,091
First stage F-Stat.			60.43	28.49
Adjusted R^2	0.968	0.982		
Within R^2	0.076	0.480		
Fixed effects	Metro × Quarter	Metro × Quarter	Metro × Quarter	Metro × Quarter
	Submarket	Submarket	Submarket	Submarket
Add. Controls		Yes		Yes
Dep. var. lags			8	8

(b) Realized excess returns $rx_{nt\to t+12}$

	$rx_{nt\to 12}$ (%, annualized)		$I(inv^{buy}_{nt} > 0)$ (IV)	
	(1)	(2)	(3)	(4)
$rx_{nt\to 12}$ (%, annualized)			0.007***	0.006**
			(0.0017)	(0.0019)
$dist(inv^{own}_{m(n),t-1})$	-0.350***	-0.283***	-0.007+	-0.008*
	(0.0858)	(0.0782)	(0.0037)	(0.0035)
Credit Access $L^{IV}_{nt}(\%)$	-0.909***	-0.714***		
	(0.0809)	(0.0845)		
Observations	32,589	32,589	32,589	32,589
First stage F-Stat.			126.46	71.33
Adjusted R^2	0.948	0.956		
Within R^2	0.240	0.353		
Fixed effects	Metro × Quarter	Metro × Quarter	Metro × Quarter	Metro × Quarter
	Submarket	Submarket	Submarket	Submarket
Add. Controls		Yes		Yes
Dep. var. lags			8	8

Note: *Table reports* $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}_1, \hat{\theta}_2$ *for a Two Stages Least Squares specification of the form*

$$I(inv^{buy}_{nt} > 0) = \beta_1 \widehat{\mathbb{E}}^*_t[rx_{nt}] + \beta_2 dist(inv^{own}_{m(n)t-1}) + \Gamma' X_{nt} + \varepsilon_{nt}$$
$$\widehat{\mathbb{E}}^*_t[rx_{nt}] = \theta_1 L^{IV}_{nt} + \theta_2 dist(inv^{own}_{m(n)t-1}) + \Lambda' X_{nt} + u_{nt}$$

using excess rental yields dpx *and realized 3-year ahead excess returns* $rx_{nt\to t+12}$ *as alternative measure of expected excess returns* $\widehat{\mathbb{E}}^*_t[rx_{nt}]$. *First stage F-Stat. denotes Kleibergen-Paap F-statistic. Standard errors in parentheses clustered two-ways by submarket and quarter.* ($^+$, *, **, ***) *denote significance at 0.10, 0.05, 0.01, 0.001 level.*

C.3 Expected excess returns across time and space

Table C.2: Expected excess returns across time and space: Alternative measures of expected excess returns

(a) Excess rental yields dpx_{nt}

	Dependent variable: $dpx_{nt\to t+12}$ (%, annualized) ($\mu = 1.22, \sigma = 2.14$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Credit Access L^{IV}_{nt} (%)	-0.076***	-0.084***	-0.028	-0.084***	-0.088***	-0.005
	(0.008)	(0.008)	(0.020)	(0.009)	(0.010)	(0.011)
$\log$(Av. adj. gross income$_{nt}$)	-0.887***	-0.702***	-2.429***	-0.652***	-0.696***	-0.654**
	(0.158)	(0.164)	(0.302)	(0.165)	(0.172)	(0.223)
10-year Treasury yield$_t$ (%)	-1.181***		-1.192***	-1.250***		-1.111***
	(0.087)		(0.081)	(0.077)		(0.041)
Observations	41091	41091	41091	41091	41091	41091
Adjusted R^2	0.627	0.657	0.922	0.672	0.694	0.962
Within R^2	0.628	0.419	0.819	0.672	0.482	0.911
Fixed effects		Quarter	Submarket		Quarter	Submarket
Distance control	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls				Yes	Yes	Yes

(b) Realized excess returns $rx_{nt\to t+12}$

	Dependent variable: $rx_{nt\to t+12}$ (%, annualized) ($\mu = 4.42, \sigma = 8.36$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Credit Access L^{IV}_{nt} (%)	-0.166***	-0.141***	-0.738***	-0.166***	-0.152***	-0.311**
	(0.017)	(0.017)	(0.154)	(0.018)	(0.018)	(0.098)
$\log$(Av. adj. gross income$_{nt}$)	0.547	-0.311	1.825	0.473	0.127	8.679**
	(0.507)	(0.477)	(3.037)	(0.424)	(0.428)	(3.000)
10-year Treasury yield$_t$ (%)	-2.603***		-2.450***	-2.157***		-1.765***
	(0.677)		(0.681)	(0.472)		(0.485)
Observations	32589	32589	32589	32589	32589	32589
Adjusted R^2	0.670	0.744	0.734	0.784	0.821	0.836
Within R^2	0.670	0.159	0.705	0.784	0.414	0.819
Fixed effects		Quarter	Submarket		Quarter	Submarket
Distance control	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls				Yes	Yes	Yes

Note: *Table reports $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ for specifications of the form*

$$\widehat{\mathbb{E}}^*_t[rx_{nt}] = \theta_1 \text{ Credit Access } L^{IV}_{nt} + \theta_2 \text{ Income } y_{nt} + \theta_3 \text{ 10y-Treasury } r^f_t + \Lambda' X_{nt} + u_{nt}$$

using excess rental yields dpx and realized 3-year ahead excess returns $rx_{nt\to t+12}$ as alternative measure of expected excess returns $\widehat{\mathbb{E}}^*_t[rx_{nt}]$. μ and σ denote overall mean and standard deviation of the dependent variable. Standard errors in parentheses clustered two-ways by submarket and quarter. ($^+$,*,**,***) denote significance at 0.10, 0.05, 0.01, 0.001 level.

D Appendix: Market timing and Market impact

D.1 Heterogeneity

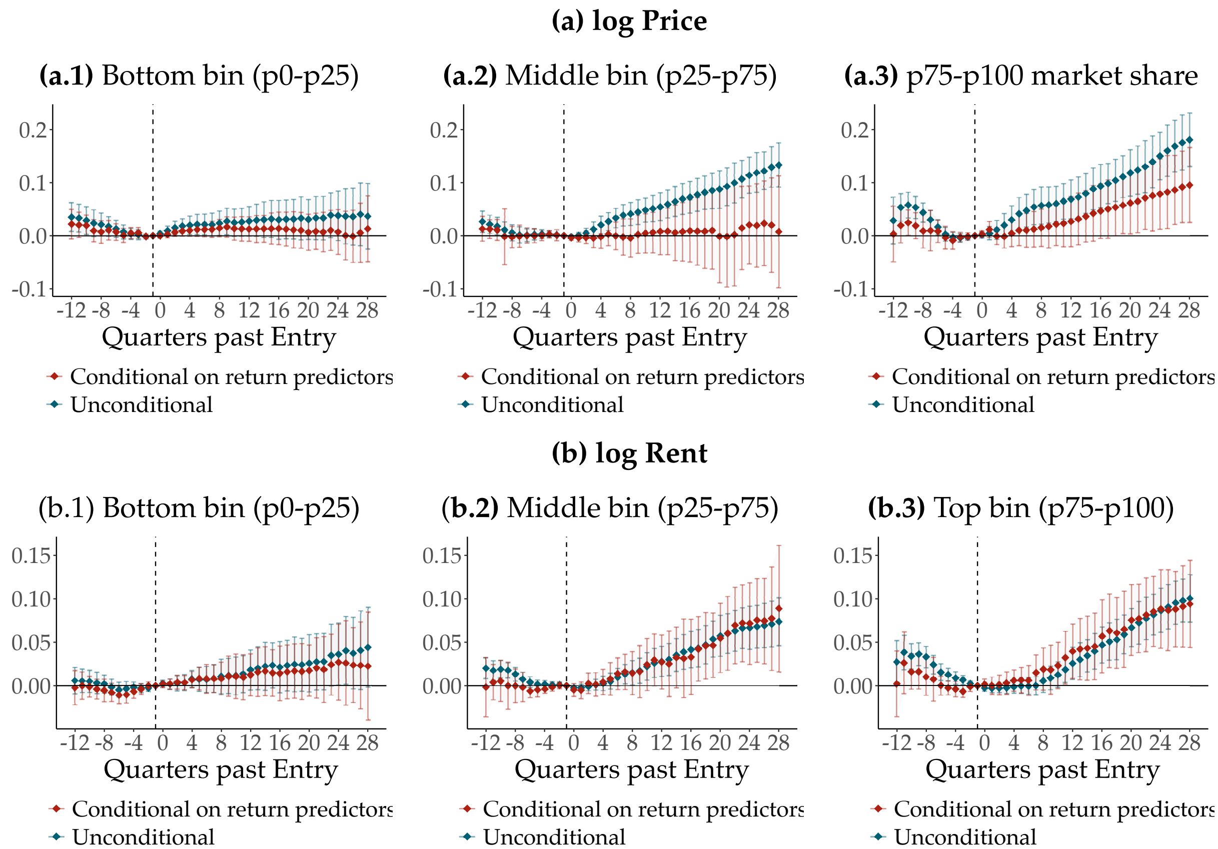

Figure D.1: Market timing and market impact by share of housing stock owned.
Figure shows difference-in-differences event study coefficients for equation (10) for quarterly log house prices and log rents by bins sorted on ex-post share of institutional investor-owned housing stock. Coefficients in blue report estimates from an unconditional comparison of submarkets. Coefficients in red report estimates from a comparison conditional on predictors of future housing returns observed in the year prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

D.2 Robustness exercises

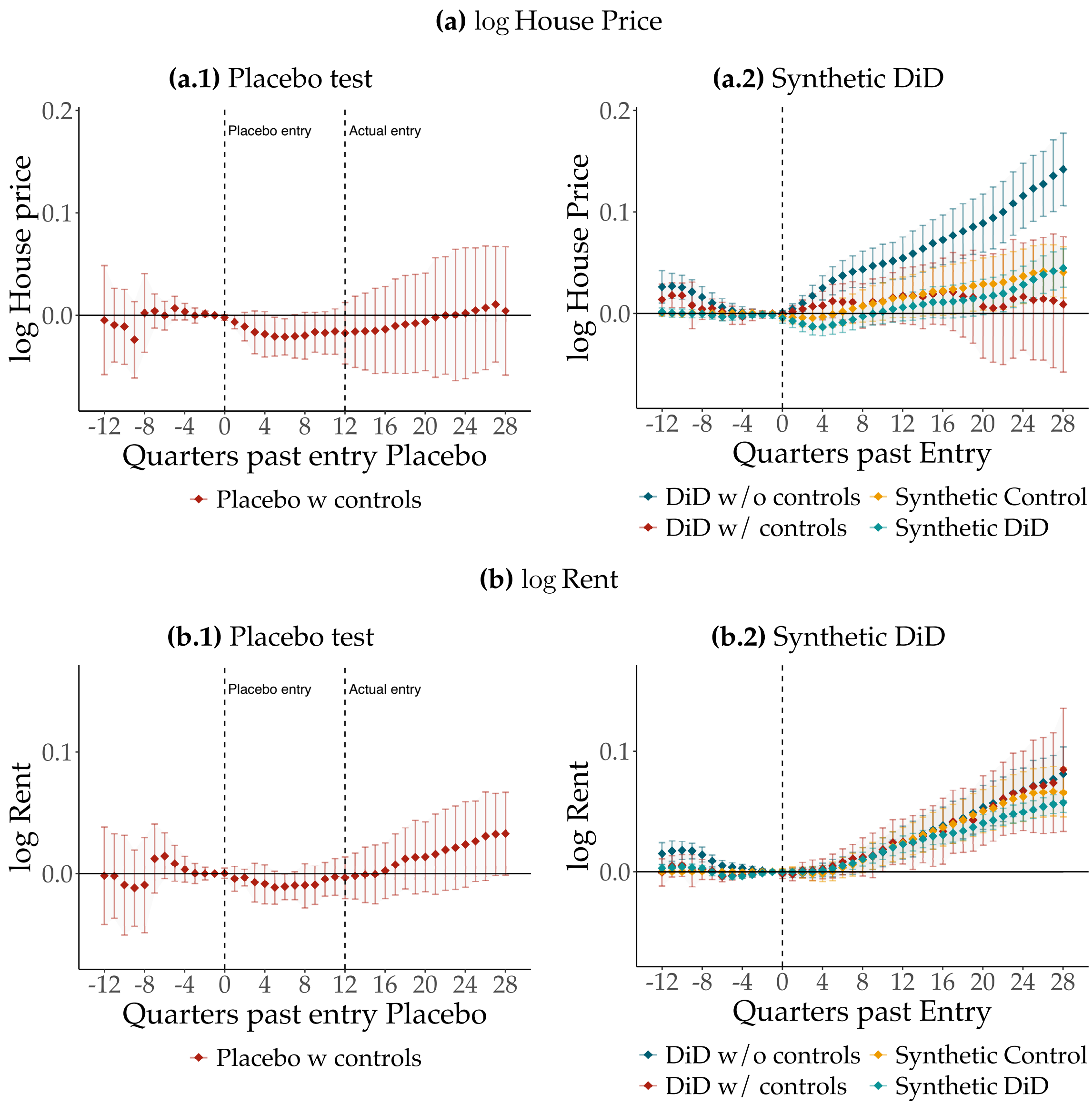

Figure D.2: Market impact of institutional investors, robustness exercises.

Figure shows estimates of market impact for two robustness exercises. First column shows difference-in-differences event study coefficients for equation (10) for quarterly log house prices and log rents in a placebo exercise that pretends the time of entry occurred 12 quarters (3 years) earlier. Second column shows difference-in-differences event study coefficients of (10) and estimates of market impact using the synthetic control and synthetic difference-in-differences framework of (Arkhangelsky et al., 2021) for quarterly log house prices and log rents. Coefficients in blue report estimates from an unconditional comparison. Coefficients in red report estimates from a comparison conditional on predictors of future housing returns observed in the year prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

D.3 Other outcome variables

D.3.1 Income and home-ownership rates

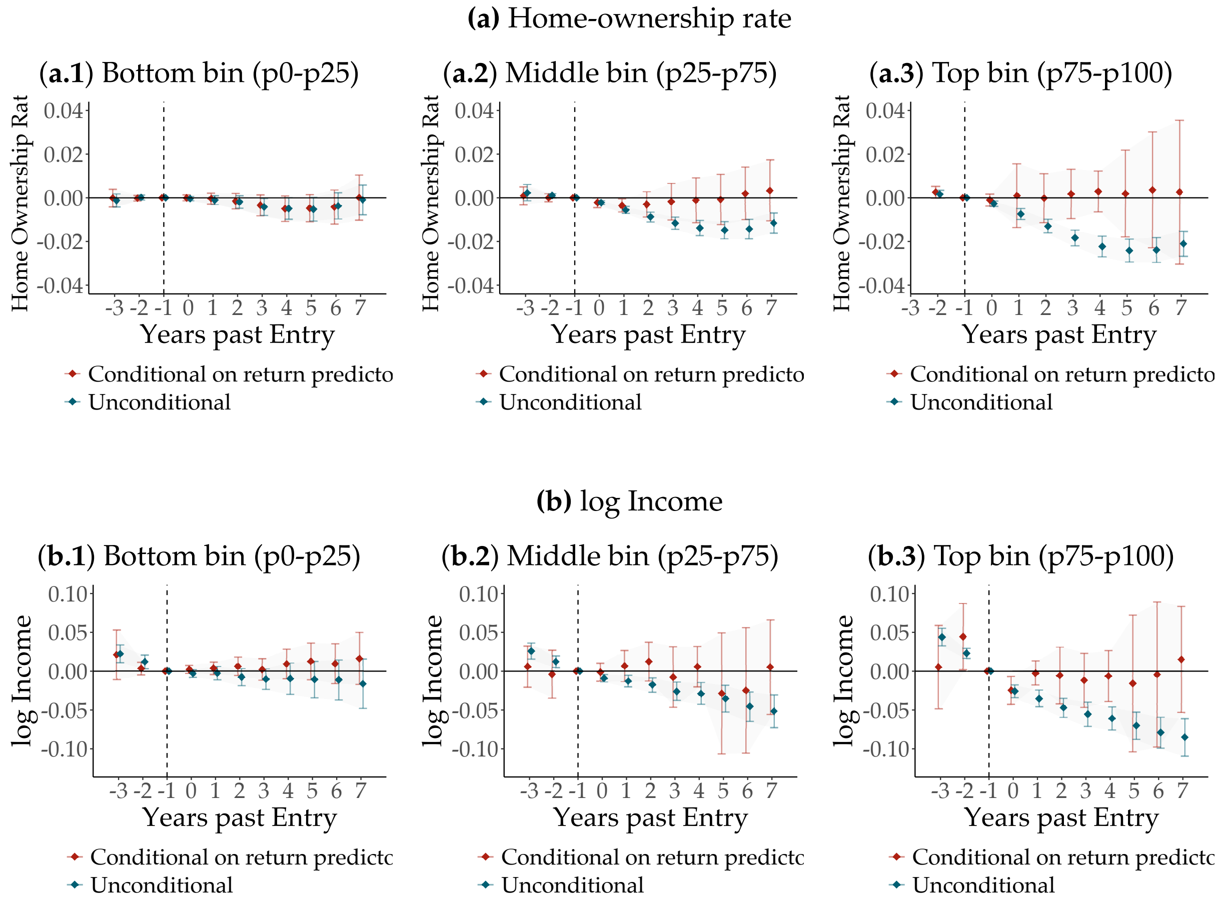

Figure D.3: Income, home-ownership rates and institutional investor entry.
Figure shows difference-in-differences event study coefficients for equation (10) for annual home-ownership rate estimates from the ACS and average log income estimates from the IRS Statistics of Income by bins sorted on ex-post share of institutional investor-owned housing stock. Coefficients in blue report estimates from an unconditional comparison of submarkets. Coefficients in red report estimates from a comparison conditional on predictors of future housing returns observed in Q4 of the year prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

Table D.1: Change in institutional investor-owned housing stock and home ownership rates

	Dependent variable: Δ_{11}^{21}Home Ownership rate (%)			
	Submarket (69 Metro areas)		Zipcode (US)	
	(1)	(2)	(3)	(4)
Δ_{11}^{21}WSL-owned housing stock (%)	-0.468^{***}	-0.516^{***}	-0.891^{***}	-0.756^{***}
	(0.088)	(0.121)	(0.099)	(0.112)
Const.	-1.560^{***}		-0.207^{*}	
	(0.094)		(0.112)	
Fixed effects		Metro		CBSA
Observations	814	814	40,294	40,294
Adjusted R^2	0.050	0.188	0.003	0.010

Note: *Table reports $\hat{\beta}$ for specifications of the form*

$$\Delta_{11}^{21}\textit{Home Ownership rate} = \alpha_{m(n)} + \beta\Delta_{11}^{21}\textit{WSL-owned housing stock} + \varepsilon$$

where Δ_{11}^{21}WSL-owned housing stock denotes the change in institutional investor or Wall Street Landlord (WSL)-owned housing stock. Standard errors in parentheses report heteroscedasticity robust standard errors.

Housing stock and vacancy rates

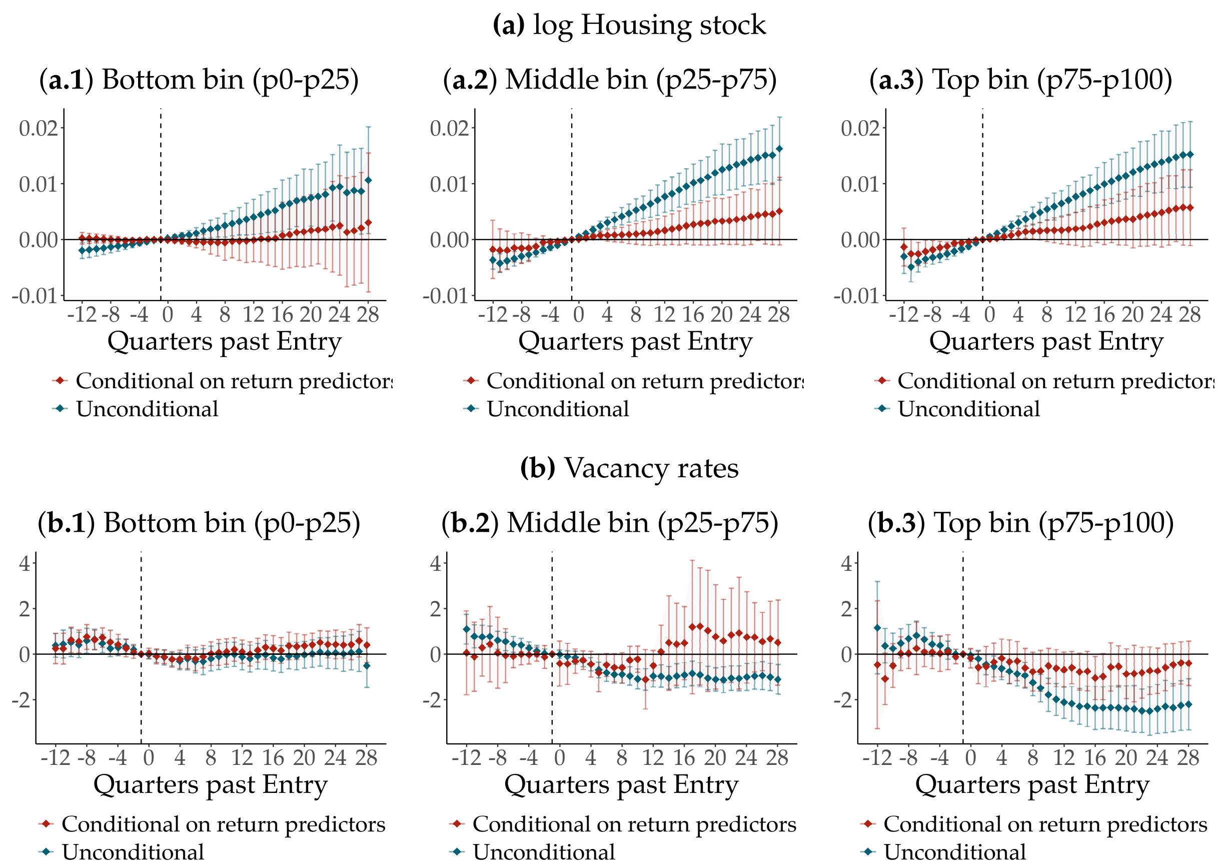

Figure D.4: Housing supply, vacancy rates and institutional investor entry.
Figure shows difference-in-differences event study coefficients for equation (10) for quarterly vacancy rate estimates from CBRE and estimated log housing stock in CoreLogic deeds by bins sorted on ex-post share of institutional investor-owned housing stock. Coefficients in blue report estimates from an unconditional comparison of submarkets. Coefficients in red report estimates from a comparison conditional on predictors of future housing returns observed in the quarter prior to entry. Pointwise confidence intervals (90%) based on standard errors clustered at the level of assignment (submarket).

D.4 Excess return decomposition

I decompose excess returns into (net) rental yields $y^d_{nt\to t+12}$, appreciation $y^p_{nt\to t+12}$ and the risk-free rate, where I calculate net rental yields as rental income less operating costs before taxes following Demers and Eisfeldt (2022) over the next 3 years divided by the purchase price.

$$\text{Excess return } rx_{nt\to t+12} = \underbrace{\frac{\sum_{k=0}^{11} d^\tau_{nt}}{P_{nt}}}_{y^d_{nt\to t+12}} + \underbrace{\frac{P_{nt+12} - P_{nt}}{P_{nt}}}_{y^p_{nt\to t+12}} - r^f_t$$

I regress excess returns rx_{nt} and their components y^d, y^p, r^f_t onto an indicator for institutional investor-purchases and fixed effects.

$$rx^k_{nt} = \alpha_{(n,t)} + \beta^k I(inv^{buy}_{nt}) + \varepsilon_{nt} \tag{12}$$

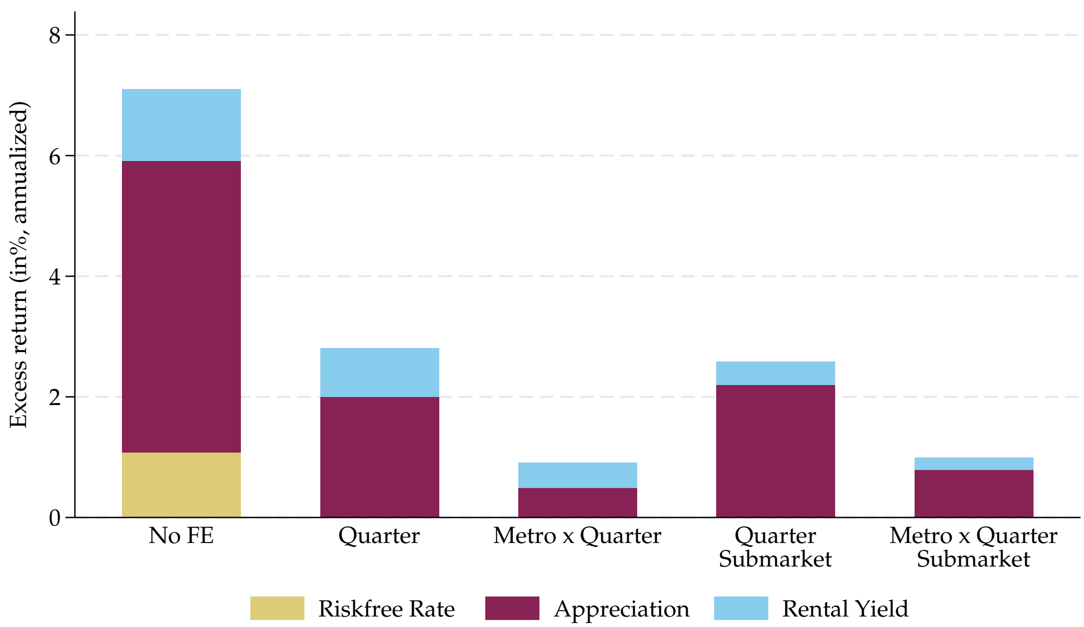

Figure D.5: Decomposition of realized excess returns.

Figure plots $\hat{\beta}^k$ for specifications of the form: $rx^k_{nt} = \alpha_{(n,t)} + \beta^k I(inv^{buy}_{nt}) + \varepsilon_{nt}$ where rx^k denotes realized excess returns components (net rental yields, appreciation and risk-free rate).

E Calibration exercise

The calibration exercise repeatedly solves the one-period static model by feeding in sample paths of rents, household net worth and income and solving for equilibrium prices and quantities. I calibrate model parameters to match observed levels of prices and share of institutional investor-owned housing stock for a representative neighborhood affected by the entry of institutional investors.

Moment construction To construct the moments that I feed into the calibration I combine the dataset constructed in section 2 with household wealth data from the Survey of Consumer Finances. Because the SCF comes in 3-year waves I calibrate the model such that one period lasts 3 years for each year from 2001 to 2022 in which an SCF wave is released (8 periods in total).

The moments I feed into the model are constructed as follows: I restrict my sample to submarkets where institutional investors own at least 1% of the total housing stock at some point. In this sample I estimate the median house price, rent, average adjusted gross income and share of households with credit scores above the 10th percentile of approved mortgage credit scores in the Senior Loan Officer Opinion Survey each period. I use these moments to identify representative households in the Survey of Consumer Finances. In each 3-year SCF wave from 2001 to 2022 I keep households aged between 30 and 40 with reported income and home values or rents within a one standard deviation range of the sample median for that year. I estimate the average net worth of these households in the SCF to use as the wealth of a representative household in the model.

Model calibration Compared to the model in section 1 I make the following changes: As in the extension in appendix A.2.1 I allow for investors to raise rents by a factor $\frac{\mu}{2}\frac{q_t^i}{Q}$ over the user cost of housing $\underline{d}_t$. This is a reduced form way of modeling impact of institutional investors on rents, e.g. due to bargaining power that is increasing in the share of investor-owned housing stock. I assume that non-tradable benefits ν and constant operating costs α are linear functions of $\underline{d}$, that is each period $\nu = v\underline{d}_t$ and $\alpha = a\underline{d}_t$. I specify the law of motion for household wealth to include current period income after taxes and consumption y_t that households use to pay rent.

$$W_{t+1}^h = q_t^h \left(P_{t+1} + d + \nu - R^f P_t\right) + R^f(W_t^h + y_t^h - d) + \tilde{y}_{t+1}^h \tag{13}$$

where y_t^h denotes the current period income after taxes and consumption, $\tilde{y}_{t+1}^h$ denotes uncertain future income and d denotes current period rent. I set household income after taxes and consumption equal to $\frac{1}{3} \times 0.85 \times$ representive household AGI for each sample year, consistent with a 15% average tax rate and a housing expenditure share of one third. I also allow for idiosyncratic house price risk as in the extension in appendix A.2.4.

I assume that only a fraction $\omega = 15\%$ of households participate in the housing market each period, such that equilibrium market clearing requires that

$$\underbrace{\omega\left(1 - \frac{q_{t-1}^i}{Q}\right)}_{\text{Init. share owned by particip. households}} + \underbrace{\frac{q_{t-1}^i}{Q}}_{\text{Init. investor share}} = \underbrace{\frac{q_t^i(P_t)}{Q}}_{\text{Investor demand}} + \underbrace{\omega q_t^h(P_t)}_{\text{particip. household demand}} \tag{14}$$

This is consistent with an average annual turnover rate of 5% and helps to match the gradual adjustment of investor housing portfolios.

Solving the model requires an assumption about expected future house prices. I assume that agents expect future house prices to converge to the present value of 40 years worth of rental cash flows discounted at a discount rate $r_t = r_t^f + hrp$:

$$\mathbb{E}_t\left[P_{t+1}^h\right] = \frac{d_t(1+g_t)\left[1 - \left(\frac{1}{1+r_t^f+hrp-g_t}\right)^{40}\right]}{1 - \left(\frac{1}{1+r_t^f+hrp-g_t}\right)} \tag{15}$$

where hrp denotes a housing risk premium parameter that I calibrate as part of the exercise and g_t denotes the current period average rental growth rate.

Exogenous parameters I hold the housing supply and mass of households constant at 50000 housing units and households. For simplicity I assume that there is only one investor. The household coefficient of absolute risk aversion A^h is set to $2/W_t^h$ each period in order to maintain a constant coefficient of relative risk aversion of $\phi = 2$. The investor's coefficient of absolute risk aversion A^i is held constant at $\phi/10^7$. I set the systemic volatility of house prices σ_p equal to 2×10^5 and the idiosyncratic volatility of house prices σ_ε^2 equal to 10^5. The volatility of household income after taxes and consumption is set to 5×10^4 and the volatility of investor income from other investments equal to 5×10^6. The correlations of household and investor income with systemic house price risk are set to 0.25 and 0.33.

Calibration I solve for equilibrium prices, quantities and user costs of housing $\underline{d}_t$ that clear the market and lead to equilibrium rents d_t. Because $\underline{d}_t$ is not observable in the data this requires solving for a fixed point to the investor problem where

$$\left(1 + \frac{\mu}{2}\frac{q_t^i(\underline{d}_t)}{Q}\right)\underline{d}_t = d_t \tag{16}$$

I solve for the fixed point by starting from a guess of $\underline{d}_t^{old} = d_t$ and then iteratively update $\underline{d}_t^{new} = \left(1 + \frac{\mu}{2}\frac{q_t^i(\underline{d}_t^{old})}{Q}\right)\underline{d}_t^{old}$ until $\underline{d}_t$ approximately satisfies 16. Trivially, if $q_t^i(d_t) = 0$ the search for the fixed point is not necessary and $\underline{d}_t = d_t$.

I calibrate parameter values to minimize the distance between calibrated moments and data points fed into the calibration exercise. Table E.1 summarizes the calibrated parameters.

Table E.1: Calibrated parameters

Moments from the data	Source
Risk-free rate R_t	10-year Treasury yield
Access to Leverage L_t	Households w. cred. score above $score_t^{p10}$
Rent d_t	Median log rent
Income (after taxes and consumption) y_t	$\frac{1}{3} \times 0.85\times$ median AGI
Wealth W_t	Survey of Consumer Finances
Targeted Moments	
House prices P_t	Median log house price
Investor Ownership q_t^i/Q	% of investor-owned housing stock
Rental markups μ_{t+h}	DiD estimates for sample of rep. submarkets
Fixed Parameters	
Housing supply Q	50000
Relative Risk Aversion ϕ	2
Turnover ω	0.15
Systemic house price risk σ_p	2×10^5
Idiosyncratic house price risk σ_ε	10^5
Household income risk σ_{y^h}	5×10^4
Investor income risk σ_{y^i}	5×10^6
Correlation between $\tilde{y}_{t+1}^h$ and $\tilde{P}_{t+1}$	0.25
Correlation between $\tilde{y}_{t+1}^i$ and $\tilde{P}_{t+1}$	0.33
Calibrated Parameters	
Non-tradable benefit ν	$0.2 \times \underline{d}_t$
Constant variable cost α	$0.5 \times \underline{d}_t$
Scale economies β	8.02
Rental Markup μ	8
Search cost γ	0.00415
Housing risk premium hrp	0.015