Speculative Dynamics of Prices and Volume[*]

Anthony A. DeFusco[†]   **Charles G. Nathanson**[‡]   **Eric Zwick**[§]

April 17, 2018

Abstract

We present a dynamic theory of prices and volume in housing cycles. In our framework, predictable price increases endogenously attract short-term buyers more strongly than long-term buyers. Short-term buyers amplify volume by selling more frequently, and they destabilize prices through positive feedback. Our model predicts a lead–lag relationship between volume and prices, which we confirm in the 2000–2011 U.S. housing bubble. Using data on 50 million home sales from this episode, we document that much of the variation in volume arose from the rise and fall in short-term investment.

The United States underwent an enormous housing market cycle between 2000 and 2011, shown in Figure 1. The rise and fall in house prices caused several problems for the U.S. economy. During the boom, a surge in housing investment drew resources into construction from other sectors (Charles, Hurst and Notowidigdo, 2017) and contributed to a capital overhang that slowed the economic recovery from the subsequent recession (Rognlie, Shleifer and Simsek, 2017). During the bust, millions of households lost their homes in foreclosure, and falling house prices led many others to cut consumption (Mayer, Pence and Sherlund, 2009; Mian, Rao and Sufi, 2013; Mian, Sufi and Trebbi, 2015; Mian and Sufi, 2014; Guren and McQuade, 2015). Large real estate cycles are not unique to the U.S. (Mayer, 2011) or to this time period (Case, 2008; Glaeser, 2013). Given the economic costs of these recurring episodes, understanding their cause is critical for economists and policymakers.

This paper presents theory and evidence that speculation is a key driver of real estate cycles.[1] Building on Cutler, Poterba and Summers (1990) and De Long, Shleifer, Summers and Waldmann (1990), we present a model in which extrapolation—the belief that asset prices continue to rise after recent gains—causes house prices to go through a boom and bust cycle in response to a positive demand shock. In a departure from the existing extrapolation literature, we relax the assumption of Walrasian price discovery to allow the possibility that homes listed for sale may not sell immediately. This innovation allows the model to describe the other salient features of Figure 1: the pronounced lead–lag relationship between transaction volume and prices, and the sharp rise in unsold listings as volume falls and price growth slows. We then document new facts on the composition of buyers and sellers during the recent U.S. housing bubble using transaction-level data between 1995 and 2014 for 115 metropolitan statistical areas (MSAs) that represent 48% of the U.S. housing stock. The facts highlight the aggregate role of speculation in driving the bubble, rule out some alternative explanations, and confirm several predictions of the model.

In each period of our model, potential buyers interact with movers who are trying to sell their house immediately. Movers post list prices, and then as many transactions as possible clear at the posted prices. The average transacted price, or *market price*, diverges from the Walrasian price due to a subset of movers who engage in extrapolative price-posting: rather than set list prices optimally, they simply list at the market price they

---

[1]Harrison and Kreps (1978, p. 323) define speculation as follows: "Investors exhibit speculative behavior if the right to resell a stock makes them willing to pay more for it than they would pay if obliged to hold it forever."

expect under extrapolative beliefs. The remaining movers set list prices optimally because they are fully informed about market fundamentals. In evaluating whether to buy, potential buyers extrapolate market prices to estimate the future price at which they can sell. Potential buyers differ in the expected time until moving, meaning that some potential buyers have short horizons while others have long horizons. This heterogeneity captures the difference in previous models between long-term investors and short-term arbitrageurs.[2]

We solve for the equilibrium dynamics of prices and volume following a one-time, permanent demand shock to a steady state. After the shock, informed movers sharply raise their list prices while extrapolative movers do so more gradually. During this time, a disproportionate share of buyers are short-horizon investors hoping to exploit rising prices for a quick capital gain. As they turn to sell, this influx of short-horizon investors leads to a sharp increase in subsequent listings and volume. Eventually, so many houses are listed that not all of them can sell at the extrapolative movers' price. This glut leads informed movers to cut their list prices to the extrapolative list price, which continues to rise as it traces out the path expected under extrapolative beliefs. The slowdown in price growth causes demand to fall, leading volume to decline as unsold listings accumulate. Once demand falls enough, informed movers cut their list prices below the extrapolative list price, causing the market price to decline for the first time. The resulting bust in market prices occurs on low volume and continues until the entire stock of unsold listings sells.

The most novel feature of this cycle is the middle period in which prices continue to rise despite falling volume and accumulating unsold listings. We refer to this period as the *quiet*. The quiet characterizes the joint behavior of prices, volume, and listings in the U.S. housing market between 2005 and 2007. This pattern contrasts with the focus in existing literature on the contemporaneous correlation between prices and volume in the housing market (Stein, 1995; Genesove and Mayer, 2001; Ortalo-Magné and Rady, 2006) and in bubbles (Scheinkman and Xiong, 2003; Barberis, Greenwood, Jin and Shleifer, 2017). The average correlation is positive, but masks the changing relationship between prices and volume during different parts of the cycle, driven in our model by the absence of Walrasian price discovery.[3]

---

[2]The closest model to ours is Hong and Stein (1999), which features momentum traders who follow simple price-based trading rules and newswatchers who receive new information slowly over time but do not infer additional information from prices. We discuss connections between our model and theirs below.

[3]Barberis, Greenwood, Jin and Shleifer (2017) present an extrapolative, Walrasian model of bubbles which, like our model, generates endogenous trading and a joint dynamic of prices and volume. Though their model delivers a positive relationship between past returns and volume, it does not naturally generate

The model features endogenous entry of short-term, speculative buyers into the market, who drive volume during the boom by selling shortly after they enter. Using transaction-level data collected by CoreLogic on 50 million home sales between 1995 and 2014, we exploit the panel structure of this data to link shifts in the distribution of realized holding periods to dynamic patterns in volume and prices during the U.S. cycle. Because realized holding times do not perfectly capture holding times expected upon purchase, we supplement this analysis with survey data from the National Association of Realtors (NAR) on heterogeneity in expected holding horizons in the cross-section and over time.

We present four facts concerning the composition of aggregate volume. First, much of the 2000–2005 rise in volume comes from short-term investment, with 42% of the national volume increase arising from the growth in sales of homes held for less than 3 years. The rise in short-term sales also explains much of the variation in volume across MSAs and across ZIP codes within each MSA. Second, short-term sales decline during the 2005–2006 quiet and subsequent bust, matching when short-term buyers begin to exit in the model. Third, a sharp rise in non-occupant purchases explains much of the variation in volume across and within MSAs between 2000 and 2005. This fact matches the model's prediction of a rising share of "speculative buyers"—those who would not buy absent expected capital gains—during the boom. Fourth, non-occupant buyers disproportionately contribute to the growth in short-term volume, further suggesting that speculative motives drive their trading behavior. Consistent with these facts, the NAR survey data reveals wide variation in expected holding times, shorter expected holding times among investors, and increases in the short-term buyer share following recent price gains.[4]

The most likely alternative explanation for the rise of short-term volume is that this activity reflects move-up purchases enabled by higher home equity during the boom (Stein, 1995; Ortalo-Magné and Rady, 2006). While surely part of the story, this channel is unlikely to explain the full rise in short-term volume for two reasons. First, following Bayer, Geissler, Mangum and Roberts (2011) and Anenberg and Bayer (2013), we use the names of buyers

---

a lead–lag relationship between prices and volume, which they note was a puzzling phenomenon during the Dotcom bubble. Section 5 has more detail on connections between our models.

[4]These empirical results build on a number of prior studies of the housing bubble, surveyed in Section 5. Our contribution is documenting the role of speculation in driving aggregate market dynamics, in particular, with complete market-level information across many MSAs. These findings complement those in Haughwout, Lee, Tracy and van der Klaauw (2011), who use data from mortgage accounts to show investors were important for aggregate mortgage growth during the boom and aggregate delinquencies during the bust.

and sellers to evaluate the frequency of within-MSA moves. Only 24% of all transactions and 31% of short-term transactions appear to be associated with move-up purchases within an MSA. Second, short-term volume sharply increases even among transactions where the seller used low leverage initially when buying. Another possible explanation for the rise of short-term sales is the construction of new homes and entry of sophisticated real estate arbitrageurs. However, only 14% of short-term purchases were made by companies such as developers or incorporated real estate investors. Of the remaining short-term purchases, 73% were from people buying only one or two homes, who seem unlikely to be professional developers given the small number of houses they buy.

To further test the predictions of our model, we then present four facts about the joint price–volume relationship. First, the lead–lag relationship between prices and volume in our model holds both at the national level and across MSAs, and is more pronounced in MSAs that experience larger price booms. In the cross-section of MSAs, prices correlate most strongly with a 24-month lag of volume. Second, price booms are correlated with volume booms, both across MSAs and across ZIP codes within MSAs. Price booms are similarly correlated with short-term volume booms and non-occupant purchase booms.[5] Third, rising listings enable the 2000–2005 increase in volume, but falling listings do not drive the subsequent decline in volume. Instead, volume declines due to a slowdown in the rate at which listings sell. This dichotomy matches the model's predictions about how speculative demand determines volume during the boom and quiet. Finally, the house price cycle is larger in MSAs where the 2000 rate of existing sales is greater. As implied in our model, a higher frequency of short-term buyers increases both steady-state volume and the amplitude of the price response to the demand shock. Other theories, including housing search and Walrasian behavioral finance models, struggle to explain many of these facts.

The paper proceeds as follows. Section 1 presents the theoretical results. Section 2 details the data we use. Section 3 describes the composition of buyers and sellers over the 2000–2011 U.S. housing cycle, while Section 4 relates these changes in volume to changes in house prices. We compare our paper to the existing literature in Section 5 and conclude in Section 6.

---

[5]Gao, Sockin and Xiong (2017) argue that this correlation between house price booms and non-occupant volume booms is causal using state level capital gains taxes as an instrument for investment activity.

1 Model

1.1 Setup

We present a discrete-time model of a city with a fixed amount of perfectly durable housing, normalized to have measure one. Agents go through a life cycle with three possible phases: potential buyer, stayer, and mover.

Each period, a measure A_t of potential buyers arrive. Some are matched to take-it-or-leave-it offers from movers and decide whether to buy a house. The potential buyer value function is:

$$\mathcal{V}_{i,t}^{pb} = \max\left(0, \mathcal{V}_{i,t}^s - P_{i,t}\right) \tag{1}$$

where 0 is the value of not buying, $P_{i,t}$ is the price of housing the potential buyer faces, and $\mathcal{V}_{i,t}^s$ is the value of being a stayer. Non-buying potential buyers exit permanently.

Upon buying, the potential buyer becomes a stayer and earns idiosyncratic flow utility δ_i (received at the beginning of next period) from the house, but cannot sell it. Stayers face an idiosyncratic probability $\lambda_i > 0$ of becoming movers. The stayer value function is:

$$\mathcal{V}_{i,t}^s = \rho\delta_i + \rho(1-\lambda_i)\mathrm{E}\left(\mathcal{V}_{i,t+1}^s \mid \mathcal{F}_{i,t}\right) + \rho\lambda_i\mathrm{E}\left(\mathcal{V}_{i,t+1}^m \mid \mathcal{F}_{i,t}\right), \tag{2}$$

where $\mathcal{V}_{i,t}^m$ is the value of being a mover, $\mathcal{F}_{i,t}$ is the information set available to agent i at time t, and ρ is the discount factor corresponding to a risk-free discount rate r. The characteristics δ_i and λ_i are distributed independently and identically across potential buyers according to the time-invariant pareto density $f_\delta(\delta_i) = \epsilon\delta_0^\epsilon/\delta_i^{\epsilon+1}$ where $\delta_0, \epsilon > 0$, and a general density $f_\lambda(\lambda_i)$.[6]

Movers choose list prices and then match to potential buyers, with potential buyers in descending order of willingness-to-pay matched to movers in ascending order of list prices. A transaction occurs when the list price is below the matched willingness-to-pay, and the transaction price equals the list price. Our model follows Hong and Stein (1999), Mankiw and Reis (2002), and Guren (2018) in that some movers are inattentive and do not adjust their list prices in response to abnormal levels of demand, while other movers observe current

---

[6]Heterogeneity in δ is necessary only to produce the demand curve that is a decreasing function of the price. A similar demand curve would hold in a model with risk-averse agents and homogeneous δ, which may better describe the stock market because dividends are the same for all owners.

demand and choose their list prices in response to it. *Extrapolative movers* post $E(P_t \mid \mathcal{F}_{i,t})$, their expectation of the average transacted price at t, which we denote P_t. Their value is $\mathcal{V}_{i,t}^{xm}$. *Informed movers* have complete information and choose the list price according to:

$$\mathcal{V}_{i,t}^{nm} = \max_{P_{i,t}} \pi_t(P_{i,t})P_{i,t} + (1 - \pi_t(P_{i,t}))\rho\left(E\left(\mathcal{V}_{i,t+1}^{nm}|\mathcal{F}_{i,t}\right) - k\right), \tag{3}$$

where $\pi_t(P_{i,t})$ is the probability of selling at price $P_{i,t}$ and k is the flow cost from owning a house if the mover does not sell. Among movers posting the same price, informed movers match to potential buyers before extrapolative movers. Movers face a constant probability β of becoming informed. Combining these types gives the mover value function:

$$\mathcal{V}_{i,t}^{m} = \beta\mathcal{V}_{i,t}^{nm} + (1 - \beta)\mathcal{V}_{i,t}^{xm}. \tag{4}$$

The model contains three features common in the housing search literature: a lockup period in which homeowners do not sell, a Poisson hazard of the expiration of this lockup period, and a loss of housing flow benefits upon lockup expiration (Wheaton, 1990; Caplin and Leahy, 2011; Burnside, Eichenbaum and Rebelo, 2016). In contrast to this literature, the Poisson hazard may differ across agents in our model. This hazard influences whether potential buyers choose to buy, leading the composition of buyers to vary over the cycle.

All agents use extrapolative expectations to forecast future prices. Such extrapolative expectations match a growing body of survey evidence on how investors predict prices in asset markets (Case, Shiller and Thompson, 2012; Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014). Let P_t denote the average of all transacted prices in period t. Define the moving average of past price changes, $\omega_t \equiv \mu^t\omega_0 + (1 - \mu)\sum_{j=0}^{t-1}\mu^j(P_{t-j} - P_{t-j-1})$, where $0 < \mu < 1$. Agents expect that prices evolve linearly as a function of past price growth:

$$P_t^x \equiv E(P_t) = P_{t-1} + \gamma\omega_{t-1}, \tag{5}$$

where $0 < \gamma < 1$ reflects the strength of extrapolation.[7]

At t, information available to potential buyers about past prices is limited to ω_{t-1}. Extrapolative movers observe the history of average market prices but not other market fun-

---

[7]This functional form of extrapolation is the discrete-time analog of the belief specification in Barberis, Greenwood, Jin and Shleifer (2015), for which Glaeser and Nathanson (2017) provide a microfoundation.

damentals. Extrapolative movers choose P_t^x as their listing price, while informed movers adjust their listing prices to maximize (3). The additional information available to informed movers fully reveals the current state of the housing market. This information includes the current number of listings L_t, the number of potential buyers A_t, and the number of stayers of each type $S_t(\lambda)$.

Lemma 1 in Appendix A applies the law of iterated expectations to derive a term structure of expectations given by:

$$E(P_{t+\tau}) = P_t + g(\mu, \gamma, \tau)\omega_t, \tag{6}$$

where $g(\cdot)$ is positive, increasing in γ, and concavely increasing in τ. This term structure leads past price growth to disproportionately attract short-horizon potential buyers, who drive the dynamics of market prices and volume.[8]

To decide whether to buy, potential buyers at t must forecast $\mathcal{V}_{i,t+\tau}^m$, which depends in general on both $P_{t+\tau}$ and the probability of selling upon becoming a mover. However, if potential buyers expect to sell immediately at the market price upon becoming a mover, then $E(\mathcal{V}_{i,t+\tau}^m \mid \mathcal{F}_{i,t}) = E(P_{t+\tau} \mid \mathcal{F}_{i,t})$ and we can solve the model in closed form for each potential buyer i using equation (6). To achieve this goal and simplify our analysis, we impose the following restrictions on β and k:

(a) $\beta < A_t$,

(b) $k > E(\mathcal{V}_{i,t+1}^{nm} \mid \mathcal{F}_{i,t}) - \rho^{-1} \max \pi_t^{-1}(1)$ for each informed mover i, and

(c) $\Pr(\beta = 1 \mid \mathcal{F}_{i,t}) = 1$ for each potential buyer i.

The bound on β in part (a) guarantees that there is sufficient demand among potential buyers for all informed movers to sell each period. The bound on k in part (b) is an equilibrium condition stating that holding costs are sufficiently large so as to lead informed movers to choose a list price that will guarantee immediate sale. In Appendix B, we solve for this bound in equilibrium and numerically calculate it in the calibration. Part (c) is a

---

[8]Consistent with the concavity of g with respect to τ, past asset returns do influence annualized expected capital gains more strongly over short versus long future horizons. Armona, Fuster and Zafar (2016) find that 1-year-ahead expectations of house price growth are nearly five times more sensitive to perceived past price changes than annualized expectations of price growth 2-to-5 years in the future. Graham and Harvey (2003) and Vissing-Jorgensen (2004) show that, in the US between 1998 and 2003, expected (excess) stock returns over the next year responded more strongly to recent (excess) returns than did expected returns over the next ten years. Frankel and Froot (1990) find that a 1% increase in the exchange rate over the past week increases expectations of next week's appreciation by 0.13% and decreases expectations of weekly appreciation over the next 12 months by 0.01%.

restriction on beliefs; it states that all potential buyers think that all movers are informed. This mistake is a form of overconfidence in which potential buyers fail to forecast their own possible inattention and thus naively believe that they will be informed when they go to sell. Together, these restrictions imply that potential buyers believe that all movers immediately sell, which is sufficient to guarantee $\mathrm{E}(\mathcal{V}^m_{i,t+\tau} \mid \mathcal{F}_{i,t}) = \mathrm{E}(P_{t+\tau} \mid \mathcal{F}_{i,t})$ and allows us to solve the model.

1.2   Equilibrium

Combining the potential buyer and stayer value functions, we iterate forward to derive a cutoff flow utility above which a potential buyer matched to a price $P_{i,t}$ is willing to buy.

Proposition 1 (Potential buyer demand and horizon). *Potential buyer i matched to price* $P_{i,t}$ *buys when*

$$\delta_i \geq rP_{i,t} - \phi(\lambda_i; \mu, \gamma, r)\omega_{t-1}, \tag{7}$$

where ϕ is positive and increasing in λ_i.

Proof. See Appendix A. □

The first term on the right side of this cutoff captures the user cost of holding a house for one period and the second term captures the benefit from expected capital gains. The function $\phi(\cdot)$ is increasing in the probability of becoming a mover, λ_i, both because of discounting and because extrapolation from past price changes has a concave term structure. Thus, speculative motives magnify total demand when buyers expect capital gains and attenuate total demand when buyers expect capital losses. Furthermore, the force of this motive increases as the buyer's horizon shortens, and expected capital gains stimulate demand more for buyers with low flow utility.

Total demand among all potential buyers as a function of P and λ is given by:

$$D_t(P, \lambda) = A_t f_\lambda(\lambda) \left(1 - F_\delta \left[rP - \phi(\lambda; \cdot)\omega_{t-1}\right]\right). \tag{8}$$

Integrating over the distribution of λ yields a demand function:

$$D_t(P) = \int_0^1 D_t(P, \lambda)d\lambda. \tag{9}$$

8

Total listings include past unsold listings and the flow of stayers who become movers:

$$L_t = I_{t-1} + \int_0^1 \lambda S_{t-1}(\lambda) d\lambda, \tag{10}$$

where I_{t-1} is the stock of unsold inventory from $t-1$ and S_{t-1} is the stock of stayers from $t-1$. The expression for L_t has two key implications for the dynamics of volume. First, volume today depends on volume before, as past buyers become current sellers. Second, the number of listings today depends both on the level of past volume and on the expected holding periods among past buyers. The larger the number of past buyers with short horizons, the larger the flow of current listings.

Proposition 2 derives equilibrium volume and prices:[9]

Proposition 2 (Equilibrium prices and volume).

$$V_t = \begin{cases} L_t, & D_t(P_t^x) > L_t \\ D_t(P_t^x), & D_t(P_t^x) \in [\beta L_t, L_t] \\ \beta L_t, & D_t(P_t^x) < \beta L_t \end{cases} \tag{11}$$

$$P_t = \begin{cases} (1-\beta)P_t^x + \beta D_t^{-1}(L_t), & D_t(P_t^x) > L_t \\ P_t^x, & D_t(P_t^x) \in [\beta L_t, L_t] \\ D_t^{-1}(\beta L_t), & D_t(P_t^x) < \beta L_t. \end{cases} \tag{12}$$

Proof. See Appendix A. □

We partition equilibrium into three possible phases in the relationship between supply L_t and demand at the extrapolative price $D_t(P_t^x)$. When $D_t(P_t^x) > L_t$, there is excess demand relative to supply at the extrapolative price. We refer to this period as the *boom*. During the boom, volume includes all extrapolative movers, who post P_t^x, and all informed movers, who post the highest price that would guarantee sale given the total number of listings, $D_t^{-1}(L_t) > P_t^x$. Prices during the boom equal a weighted average of the informed and extrapolative movers' posted prices, and all listings sell.

---

[9]When the number of potential buyers exactly equals the number of movers, $A_t = L_t$, the formula for the equilibrium price P_t when $D_t(P_t^x) = L_t$ differs slightly from (12). The proof of Proposition 2 in Appendix A gives the expression for this special case.

When $\beta L_t \leq D_t(P_t^x) \leq L_t$, demand at the extrapolative price is between supply from informed movers and supply from all movers. We refer to this period as the *quiet*. During the quiet, volume equals all informed movers plus a share of extrapolative movers. Everyone posts P_t^x. Informed movers do not post lower prices than extrapolative movers because they are guaranteed to sell and therefore maximize value at this price.

When $D_t(P_t^x) < \beta L_t$, supply from informed movers exceeds demand at the extrapolative price. We refer to this period as the *bust*. During the bust, volume equals all informed movers, who post the highest price that would guarantee sale given the total number of listings, $D^{-1}(\beta L_t) < P_t^x$. Extrapolative movers post P_t^x but do not sell.

Given V_t, we can solve for two additional state variables. Inventories of unsold listings can accumulate during both the quiet and the bust when extrapolative movers post prices that are too high to guarantee sale:

$$
I_t = L_t - V_t = \begin{cases} 0, & D_t(P_t^x) > L_t \\ L_t - D_t(P_t^x), & D_t(P_t^x) \in [\beta L_t, L_t] \\ (1 - \beta)L_t, & D_t(P_t^x) < \beta L_t. \end{cases} \tag{13}
$$

The number of stayers of type λ equals past stayers who do not receive moving shocks plus new buyers:

$$
S_t(\lambda) = (1 - \lambda)S_{t-1}(\lambda) + V_t(\lambda), \tag{14}
$$

where $V_t(\lambda)$ equals volume to potential buyers of type λ. Because cheaper houses are matched to potential buyers first, potential buyers are rationed according to the highest equilibrium list price, which is the price posted by informed movers: $P_t^n = D_t^{-1}(L_t)$ during the boom and $P_t^n = P_t$ during the quiet and bust. As a result, $V_t(\lambda) = V_t \frac{D_t(P_t^n, \lambda)}{D_t(P_t^n)} = D_t(P_t^n, \lambda)$.

1.3 Calibration

To investigate the model's predictions over a boom-bust cycle, we calibrate the model and study how equilibrium prices and quantities adjust to a one-time, unanticipated, permanent shock to the number of potential buyers arriving each period, A_t.

Initial and Terminal Conditions

We follow Barberis, Greenwood, Jin and Shleifer (2017) and simulate a finite-horizon version of the model that features a fixed liquidation price P^T received by all stayers and movers in the final period T. We simulate the model at a quarterly frequency and set $T = 48$, which corresponds to a 12-year window. Potential buyers know P^T when deciding whether to buy. Imposing the finite horizon allows us to focus on a single boom-bust episode.

We define an efficient steady state of this finite-horizon version of the model to be one in which state variables are constant over time and there are no unsold listings, $I_t = 0$. Appendix B gives the conditions under which a unique efficient steady state exists and characterizes this equilibrium along with its implied liquidating price. To study how prices and volume respond to a demand shock, we initialize the model in this efficient steady state at $t = 0$ with an initial arrival rate of A^i. At $t = 1$, the arrival rate jumps to $A^f > A^i$ and we study the joint dynamics of prices and volume as prices converge to the level that would be consistent with the new efficient steady state associated with this higher level of demand.

Parameter Values

We must specify the distribution of potential holding periods $f_\lambda(\cdot)$ and the eight other parameters of the model: μ, γ, β, r, ϵ, δ_0, A^i, and A^f. This section briefly discusses our method for calibrating these parameters. Appendix B contains further details.

Since the distribution of potential holding periods is the key force in our model, we measure $f_\lambda(\cdot)$ non-parametrically in the data, rather than specifying an explicit functional form. We use data from the National Association of Realtors' (NAR) Investment and Vacation Home Buyer Survey, which annually asks a representative sample of recent home buyers the "length of time [the] buyer plans to own [the] property." The responses to this question map directly into values of λ_i, so we measure $f_\lambda(\cdot)$ as the distribution of these values across survey respondents.

Another important ingredient in the model is the term structure of extrapolation, governed by the parameters μ and γ. We calibrate these parameters with survey evidence on expectation formation over different horizons in the housing market. Armona, Fuster and Zafar (2016) estimate that a 1% increase in perceived house price growth over the last year causally increases expected annualized house price growth by 0.2% over the next year and

11

by 0.05% over the next two-to-five years. We map these estimates into our model using (6), yielding $\mu = 0.71$ and $\gamma = 0.40$. We set $\beta = 0.17$, which implies an upper bound of 1.5 years on the expected time for a mover to sell during the bust.[10]

The model's information structure means that market participants observe prices with a quarterly lag. We use a quarterly discount rate of $r = 0.015$, which implies an annual discount rate of 6 percent. The elasticity of demand ϵ is 0.6, a value in the range of estimates suggested by Hanushek and Quigley (1980). The remaining three parameters, δ_0, A^i, and A^f, are chosen to normalize the initial price level to $P_0 = 1$ and generate a final price of $P^T = 1.6$, which means that we study a demand shock that raises steady-state prices by 60%.

1.4   Results

Figure 2 displays results from simulating the model at the chosen parameter values. Panel (a) plots total transaction volume and average market prices. The demand shock leads to a large boom and bust. Average market prices overshoot the new steady-state price by 60 percent, rising to 2.6, then falling to 1.4 before recovering to the new steady-state price of 1.6. Volume doubles, reaching its peak 8 quarters prior to the peak in market prices. Prices fall on relatively low volume.

Panels (b)-(d) shed further light on the mechanism driving these results. Panel (b) disaggregates the average market price into its two components: the extrapolative price posted by extrapolative movers, and the market-clearing price posted by informed movers. Panel (c) plots the inventory of unsold listings each period. Panel (d) plots the share of buyers with expected holding periods less than three years.

The rise in prices occurs in two stages. In the first stage, the boom, both prices and volume increase. During the boom, informed movers post higher prices than extrapolative movers and all listings sell. Extrapolative movers underreact to the demand shock as the price they post is a function only of prior prices, even though demand at the extrapolative price exceeds total listings. Informed movers react to this sluggish behavior and increase their prices to the level that clears the market, equating volume and listings. Combing these

---

[10]During the bust, the probability that a mover sells in a given period is β, as only informed movers sell. Once the bust ends, the probability of selling rises above β, as some extrapolative movers sell. Thus, an upper bound on the expected time until sale is $1/\beta = 6$ quarters.

effects leads to an initial underreaction in market prices, similar in spirit to the effect of slow information diffusion in Hong and Stein (1999). However, in our model the underreaction derives from extrapolation among uninformed movers. As prices continue to rise, extrapolative expected capital gains increase and demand among potential buyers grows. Demand growth leads informed movers to raise list prices above 1.6, the new steady-state price, and eventually the market price exceeds this level. The growth in demand is disproportionately high among short-term buyers, whose share of market activity expands. Entry of these buyers also leads the supply of listings to rise in subsequent periods, enabling rising volume during the boom.

The boom does not continue indefinitely. Eventually, the shifting composition of buyers toward short horizons leads listings to rise enough that, for informed movers to guarantee sale, they must reduce their prices. At this point, informed movers match the price being posted by extrapolative movers, so market prices continue to rise but at slower rates. The slowdown in average price growth leads to a fall in expected capital gains, causing demand growth among short-term buyers to slow. Transaction volume falls and some extrapolative movers fail to sell. The fall in volume and reduction in informed price-setting mark the beginning of the quiet. During this period, prices are rising, volume is falling, an inventory of unsold listings accumulates, and the composition of buyers skews away from short-term.

As unsold listings accumulate in the quiet, it becomes harder for informed movers to guarantee sale by posting the extrapolative price. The bust begins as they choose to undercut extrapolative movers, causing average market prices to fall. During the bust, prices fall on very low volume relative to the boom. Volume is low both because falling average prices lead to further declines in expected capital gains—which discourage short-term buyers—and market prices are high relative to the steady-state flow of housing services—which discourage long-term buyers. Once prices stabilize, expected capital gains rise and volume recovers. The large stock of unsold listings, which gradually sell over several periods, prevents informed movers from raising prices too quickly and thus slows the recovery of market prices.

The simulation produces many features of empirical price and volume dynamics in the housing market: an overshooting of prices, a rise in volume along with prices, a lead–lag relation between prices and volume, a rise in listings during the boom and an even sharper rise during the quiet, and a price crash on low volume. It also makes new predictions: a rise in the short-term buyer share during the boom, a fall in this share during the quiet, and

a rise in volume coming primarily from short-term sellers. We now turn to confirm these patterns in the data.

2   Data

In the remainder of the paper, we provide empirical evidence linking shifts in the distribution of realized holding periods over the course of the 2000–2011 U.S. housing cycle to dynamic patterns in volume and prices that directly mirror the patterns implied by our model. We focus on the housing market both because of its macroeconomic relevance and because the availability of comprehensive, asset-level microdata permits a uniquely rich analysis of holding periods and the details of buyers and sellers.

To conduct our analysis, we use data on individual housing transactions provided by CoreLogic, a private vendor that collects and standardizes publicly available tax assessments and deeds records from municipalities across the U.S. Our main analysis sample spans the years 1995–2014 and includes data from 115 metropolitan statistical areas (MSAs), which together represent 48% of the U.S. housing stock.

We include all transactions of single-family homes, condos, or duplexes that satisfy the following filters: (a) the transaction is categorized by CoreLogic as occurring at arm's length, (b) there is a nonzero transaction price, and (c) the transaction is not coded by CoreLogic as being a nominal transfer of title between lenders following a foreclosure. We then drop a small number of duplicate transactions where the same property is observed to sell multiple times at the same price on the same day or where multiple transactions occur between the same buyer and seller at the same price on the same day. Appendix C specifies the exact steps followed to arrive at a final sample of 51,080,640 transactions. Given the geographic coverage of these data and their source in administrative records, our analysis sample serves as a proxy for the population of transactions in the U.S. during the sample period.

We supplement these data with national and MSA-level housing stock counts from the U.S. Census, national counts of sales and listings of existing homes from the NAR, and national and MSA-level nominal house-price indices from CoreLogic. We also use the Investment and Vacation Home Buyer Survey from the NAR mentioned in Section 1.3; further details on these data appear in Appendix B.

3 The Composition of Buyers and Sellers

3.1 Variation in Expected Holding Times

Our model implies that recent price changes will differentially draw in short-term investors who amplify volume by selling more frequently and destabilize prices through positive feedback. The magnitude of these effects depends on the degree of heterogeneity in the distribution of expected holding times among prospective investors. While scarce data are available concerning the expected holding times of investors, the best data we are aware of, which come from the housing market, suggest that investment horizons vary considerably across individuals and commove strongly with recent price changes.

Figure 3, Panel (a) documents the substantial cross-sectional heterogeneity in expected holding times among respondents to the Investment and Vacation Home Buyers Survey. Each bar reports an equal-weighted average of the share of recent buyers who report a given expected holding time across survey years. Averages are reported separately by property type. Two facts stand out. First, the vast majority of recent homebuyers (roughly 80%) report knowing what their expected holding time will be. Second, there is wide variation in expected holding times among those who report. About half of the expected holding times are between 0 and 11 years and are distributed somewhat uniformly over that range. The survey question groups the remaining half of the responses into a single expected holding time of greater than or equal to 11 years; however, there may be substantial variation within that group as well. Expected holding times also vary in an intuitive way across property types. Recent buyers of investment properties report substantially shorter expected holding periods than recent buyers of primary residences or vacation homes.

There is also significant variation in the time series. To demonstrate this variation, we construct a "short-term buyer share," the fraction of respondents (other than those reporting "don't know") who report an expected holding time of less than 3 years or had already sold their property by the time of the survey. Across survey years, the short-term buyer share varies from 26% to 41% for investment properties, from 10% to 22% for primary residences, and from 13% to 34% for vacation properties. The weighted average of the short-term buyer share across property types varies from 13% to 26%.

This variation over time is not random. Figure 3, Panel (b) shows the short-term buyer share moves closely with recent price appreciation in the housing market. A simple regression

of the pooled short-term buyer share on the equal-weighted average year-over-year change in the nominal quarterly FHFA U.S. house price index during the survey year yields a statistically significant coefficient estimate of 0.82.[11] This coefficient implies that a recent nominal gain of 10% in house prices is associated with an increase in the short-term buyer share of 8.2 percentage points. Nominal house price appreciation was 11% in the US in 2005 and much larger in some metropolitan areas. Thus, changes in house prices over the last cycle may have induced significant shifts in the distribution of expected holding times among homebuyers at different points in the cycle. Because the NAR data on expected holding periods is only available from 2008, we turn to transaction-level data on realized holding periods over the last cycle to investigate this possibility.

3.2 The Composition of Buyers and Sellers over Time

The key mechanism that generates time variation in transaction volume in our model is that changes in expected capital gains over the course of the housing cycle differentially attract buyers with shorter versus longer expected holding periods. This phenomenon, which follows from the cutoff rule for market entry in Proposition 1, implies that large swings in volume should be accompanied by parallel changes in the distribution of realized holding periods among those who sell their homes at various points in the cycle.

Figure 4 presents a simple yet compelling illustration of the time variation in realized holding periods during the 2000–2011 U.S. housing cycle. For each transaction in our sample of CoreLogic deeds transfers, we define the holding period as the number of days since the last transaction involving the same property. We then group all transactions with holding periods of less than 5 years into bins of 1, 2, 3, 4, or 5 years and count the number of transactions falling into each bin. Figure 4 plots these bin counts by year for each year between 2000 and 2011.

During the boom years of 2000–2005, there is a clear compression in the distribution of realized holding periods toward shorter holding periods. This pattern then reverses as national house prices peak in 2006 and begin to fall in the subsequent years. The increase in transaction volume at short holding periods during the boom years represents a nontrivial portion of the overall increase in volume during this period. For example, total volume

---

[11]Unlike the CoreLogic indices available to us that we use elsewhere in the paper, the FHFA house price index covers the period 2015–2016. For this reason we use the FHFA index in Figure 3.

across all holding periods (including those greater than 5 years) increased from 2,766,902 transactions in 2000 to 3,835,049 transactions in 2005. During the same period, total volume in the 1-, 2-, and 3-year bins increased from 484,666 transactions to 928,611, which implies that these three groups alone account for 42 percent of the total increase in volume between 2000 and 2005.

Although these patterns are consistent with speculative motives leading short-term buyers to enter and exit the local housing market in response to expected capital gains, it is also possible that some short-term sellers do not truly exit the market and instead choose to buy another house within the same MSA. Rather than speculation, such a pattern may reflect move-up purchases enabled by higher home equity during the boom, as in Stein (1995) and Ortalo-Magné and Rady (2006). Furthermore, these within-MSA movers complicate mapping the data to the model because in the model, sellers leave the city and expect to do so upon buying.

We take two approaches to exploring this alternative explanation. First, we follow the methodology of Anenberg and Bayer (2013) and construct a direct measure of within-MSA moves. We use the names of buyers and sellers to match transactions as being possibly linked in a joint buyer-seller event. For each sale transaction, we attempt to identify a purchase transaction in which the seller from the sale matches the buyer from the purchase. To allow the possibility that a purchase occurs before a sale or with a lag, we look for matches in a window of plus or minus one quarter around the quarter of the sale transaction. We only look for within-MSA matches, as purchases associated with cross-city moves are similar in spirit to our model.

Our match accounts for several anomalies that would lead a naive match strategy to understate the match rate.[12] Our approach is likely to overstate the number of true matches, because it does not use address information to restrict matches and it allows common names to match even if they represent different people. Because we find a low match rate even with this aggressive strategy, we do not make use of address information in our algorithm or otherwise attempt to refine matches.

We focus on transactions between 2002 and 2011 because the seller name fields are

---

[12]These include: inconsistent use of nicknames (e.g., Charles versus Charlie), initials in place of first names, the presence or absence of middle initials, transitions from a couples buyer to a single buyer via divorce, transitions from a single buyer to a couples buyer via cohabitation, and reversal of order in couples purchases.

incomplete in prior years for several cities. We also restrict sales transactions to those with human sellers, as indicated by the name being parsed and separated into first and last name fields by CoreLogic. The sample includes 16.3 million sales transactions. Of these, we are able to match 3.9 million to a linked buyer transaction, or 24%. Thus, three-quarters of transactions do not appear to be associated with joint buyer-seller decisions. Among sellers who had bought within the last three years, the match rate is slightly higher, equal to 31%, consistent with move-up purchase behavior. In addition, the match rates peak in 2005 at 29% and 38% for all transactions and short-term transactions, respectively. These patterns confirm and extend the findings in Anenberg and Bayer (2013), who conduct a similar match for the Los Angeles metro area. However, the evidence supports the notion that most of the short-term volume represents sellers not engaging in move-up purchases, even during the cycle's peak.

Our second approach provides evidence on whether down-payment constraints for move-up buyers can account for our results. In particular, we ask how much of the increase in short-term selling occurs among transactions when the original buyer used low leverage. To do this we focus on the subset of transactions where the seller originally purchased with a loan-to-value (LTV) ratio below 60 percent. Low LTV ratios suggest buyers are less likely to be down-payment-constrained when they subsequently choose to sell the property. Therefore, if we observe a significant increase in short-term selling among this subsample of buyers, we can be relatively certain that it is not due to a relaxation of the down-payment constraint on their next purchase.

We find that total volume by low-LTV sellers increased from 239,253 in 2000 to 466,650 in 2005. During the same period, total volume by low-LTV sellers in the 1-, 2-, and 3-year bins increased from 157,567 to 254,204. This low LTV growth accounts for approximately 22% $((254-158)/(929-485))$ of the total increase in volume in these bins. This pattern suggests that the compression in holding periods cannot be entirely due to down-payment constraints for move-up buyers, as a considerable share of the growth in short-term transactions occurred among low-LTV sellers. At the same time, the proportional growth in short-term buying was stronger among high-LTV sellers, which is consistent with the move-up channel being an important factor during the boom. This evidence is also consistent with high credit growth among speculative buyers during the boom, as documented by Haughwout, Lee, Tracy and van der Klaauw (2011) and Bhutta (2015).

3.3   The Composition of Buyers and Sellers in the Cross-section

Variation across MSAs

The shift in the composition of buyers and sellers toward shorter holding periods during the boom correlates highly with changes in total volume across local markets. This correlation can be seen clearly in Figure 5, which presents binned scatter plots of the percent change in total volume at the MSA-level from 2000–2005 versus the percent change in volume for short holding periods (< 3 years) in Panel (a) and long holding periods (≥ 3 years) in Panel (b).[13] Not only does the growth in volume of short-holding-period transactions correlate strongly with the increase in total volume across MSAs during this period, but this relationship is much stronger for short holding periods relative to long holding periods.

Panel (c) further shows that these cross-sectional differences in the growth rate of short-holding-period volume explain a significant portion of the differences in the growth in total volume across MSAs during this period. For each MSA, we plot the change in short-holding-period volume divided by initial total volume on the y-axis against the percent change in total volume on the x-axis. The slope of this line provides an estimate of how much of a given increase in total volume during this period came in the form of short-holding-period volume. The answer is 32%. Thus, shifts in the distribution of holding periods of buyers and sellers over the course of the cycle are a major determinant of changes in total transaction volume.

Proposition 1 suggests that expected capital gains increase demand through the extensive margin by lowering the threshold level of housing flow utility required to justify a purchase. A corollary of this result is that volume will increase more strongly in response to a change in expected capital gains among buyers with low housing utility. While we do not observe housing utility in our data, we do observe whether each purchased property is owner-occupied. Under the assumption that non-occupants receive less housing utility than occupants, we test this prediction by examining whether non-occupant purchases rose more than occupant purchases from 2000 to 2005.[14]

---

[13]For visual clarity, we group MSAs into 25 equal-sized bins based on their percent change in total volume during this period and calculate the average percent change in short- and long-holding-period volume in each of these bins.

[14]One interpretation of the housing utility received by investors, who represent one group of non-occupant buyers, is the rent they collect. This rent is less than the housing utility of many owner occupants because in a competitive market, rent equals the housing utility of the *marginal* occupant. Frictions that arise from the separation of ownership and control may further lower rent relative to occupant housing utility (Nathanson

To track non-occupant buyers in the market over time, we follow Chinco and Mayer (2016) by marking buyers as non-occupants when the transaction lists the buyer's mailing address as distinct from the property address. While this proxy may misclassify some non-occupants as living in the home if they choose to list the property's address for property-tax-collection purposes, we believe it to be a useful gauge of the level of non-occupant purchases. For the analysis of non-occupant purchases, we drop 13 MSAs for which the mailing address data are not consistently populated using a procedure specified in Appendix C.

Using this proxy, Figure 6 displays plots that are analogous to those in Figure 5 but use non-occupancy as the sorting variable rather than holding periods. We find that, like short-term volume, non-occupant volume is an important driver of total volume during the cycle. The top panels compare volume growth for non-occupant and occupant buyers; the relationship between total volume growth and non-occupant volume growth is much stronger. The bottom panel shows that non-occupant volume growth accounts for more than half of the growth in total volume across MSAs. We assess below the extent to which the short-term and non-occupant categories overlap.

Variation within MSAs

The MSA-level results from Figures 5 and 6 also hold across local neighborhoods *within* MSAs. To show this, in Table 1 we repeat the cross-MSA analysis using a ZIP-code-level panel data set constructed from the same underlying sample of transactions.[15] In column 1, we regress the percentage change in the number of short (< 3 years) holding period transactions in each ZIP code on the corresponding percentage change in total volume for that ZIP code. Column 2 runs the same regression, using the percentage change in the number of long (≥ 3 years) holding period transactions as the outcome. Both regressions include a full set of MSA fixed effects, so that the coefficient estimates reflect only within-MSA variation in transaction volume. Echoing the cross-MSA results from Figure 5, the change in short-holding-period volume is strongly correlated with changes in total volume and is much stronger than the corresponding relationship for long holding periods. Column 3, which is analogous to panel (c) of Figure 5, shows that these changes in short-holding-period volume are also quantitatively important for explaining cross-ZIP-code differences in

and Zwick, 2018).

[15]We assign each transaction to a Census ZIP-Code Tabulation Area (ZCTA) using the postal ZIP code of the property and a ZCTA-to-ZIP code crosswalk file provided by the Missouri Census Data Center.

the change in total volume within MSA. In this specification, we regress the change in short holding period volume measured as a fraction of the level of total volume in 2000 on the change in total volume for the ZIP code. The coefficient estimate implies that roughly 24% of the cross-ZIP-code variation within an MSA in the growth of volume over this period can be attributed to differences in the number of short-holding-period transactions.

The results for occupants versus non-occupants are similar. Columns 4–6 of Table 1 report estimates from similar regressions where we instead group borrowers by their occupancy status. As in Figure 6, we find an important role for non-occupants. The coefficient estimate in column 6 implies that roughly a third of the within-city variation in the growth rate of total volume between 2000 and 2005 can be attributed to changes in the number of purchases by non-occupants.

The Characteristics of Short-term Buyers

Our results thus far indicate that short-term buyers were a major driver of changes in transaction volume over time, across MSAs, and within MSAs during the boom years of the 2000s housing cycle. In this section, we explore three additional questions designed to shed further light on the nature of these speculative short-term purchases. The results point to the quantitative importance of a class of inexperienced speculative entrants into the housing market during the cycle.

First, we ask what share of short-term volume was due to firms or developers rather than individuals. To evaluate this question, we mark transactions as developer purchases when the buyer name is both not parsed as a person by CoreLogic and contains strings reflecting developer names.[16] Across the cities in our sample, these transactions account for 6% of total volume and 10% of the growth in volume between 2000 and 2005. Of the 4,015,261 transactions between 2000 and 2005 made by buyers with short-holding periods, 576,081 (14%) were developer buyers. Between 2000 and 2005, the number of short-term-buyer transactions increases from 510,309 to 951,516, while the number of short-term-developer-buyer transactions increases from 78,856 to 131,421, or 12% of the growth in short-term transactions. We conclude that, though developers were actively involved in the housing market, they did not contribute disproportionately to the growth in short-term volume during

---

[16]We identify developer names using CoreLogic's internal new construction flag as described in Appendix C.

the boom. A possible explanation is that developers were more likely to engage in speculation in the raw land market (Nathanson and Zwick, 2018).

Second, we ask what share of short-term volume was from sellers who were non-occupant buyers. The results above indicate that both short-term and non-occupant buyers were disproportionately active during the run-up in house prices from 2000 to 2005. However, there may be overlap between these two groups. Of the 2000–2005 short-term volume, we find that 848,933 (25%) were non-occupant buyers (excluding developer buyers). Between 2000 and 2005, the number of short-term-non-occupant-buyer transactions increases from 94,298 to 248,334, or 40% of the growth in short-term transactions overall. Non-occupant buyers thus account for an excess share of the growth in short-term buyers, further suggesting that speculative motives drive their trading behavior.

Finally, we ask what share of short-term buyers were experienced investors versus inexperienced speculators in merely one or two homes. To answer this question, we count the total number of transactions for each unique buyer name in an MSA and then ask what share of total transactions in that MSA are associated with buyers with few purchases during the entire sample period versus buyers with many purchases. We classify buyers with one or two purchases as inexperienced and those with three or more as experienced. Of the 2000–2005 short-term volume, 2,521,787 (73%) were inexperienced buyers (excluding developer buyers). Between 2000 and 2005, the number of inexperienced short-term-buyer transactions increases from 315,638 to 593,062, or 71% of the growth in short-term transactions. Consistent with the evidence in Bayer, Geissler, Mangum and Roberts (2011), who use a similar methodology for the Los Angeles metro area, entry of inexperienced buyers is critical for understanding the growth in aggregate volume. Bayer, Geissler, Mangum and Roberts (2011) show that inexperienced entrants pursue a momentum-trading strategy, consistent with the prediction of Proposition 1.

4 The Joint Dynamics of Prices and Volume

4.1 The Lead–Lag Relationship

One of the key features of our model is that it is able to produce dynamic patterns in prices and volume that are consistent with the patterns typical of housing cycles. In particular, it

can generate episodes in which both prices and volume go through a boom and bust cycle, with the volume cycle leading the price cycle. Panel (a) of Figure 1, which plots aggregate trends in prices and transaction volume between 2000–2011, shows that this pattern was clearly present in the 2000s housing cycle. Volume peaks before prices and there is a sustained period during which volume is falling rapidly on high prices. This period corresponds to the quiet in our model and is marked in the figure as the period between the peak of volume and the last peak of prices before their pronounced decline. In this section, we provide evidence that this lead–lag relation is a robust feature of housing cycles that holds not just in aggregate but also across MSAs and in a way that is directly related to the severity of the housing cycle itself.

Figure 7, Panel (a) presents evidence that the lead–lag relation holds on average across MSAs. We search for the horizon over which a given change in volume has the most predictive power for the contemporaneous change in prices at the MSA level. Changes in volume generally lead changes in prices if the correlation between prices and volume is maximized at a positive lag.

To implement this search, we build a monthly panel of log house prices and transaction volume at the MSA level running from January 2000 to December 2011. We normalize transaction volume in each MSA-month by dividing by the total housing stock for the MSA recorded in the 2000 Census. Using this panel, we run a series of simple regressions of the form:

$$p_{i,t} = \beta_\tau v_{i,t-\tau} + \eta_{i,t}, \tag{15}$$

where p is log price, v is volume, i indexes MSAs, and time is measured in months. To account for the seasonal adjustment in the CoreLogic price indices, for each regression we demean prices at the MSA level and demean volume at the MSA–calendar month level.[17]

The coefficient β_τ provides an estimate of how movements in volume around MSA–calendar month averages at a τ-month lag are correlated with contemporaneous movements in prices around MSA averages. We run these regressions separately for up to 4 years of lags ($\tau = 48$) and one year of leads ($\tau = -12$). Panel (a) of Figure 7 plots the implied correlation from each regression along with its 95% confidence interval.[18] The correlation is positive at

most leads and lags but reaches its maximum at a positive lag of 24 months. Thus, changes in volume generally lead changes in prices by about two years.

Figure 7, Panel (b) shows that this lead–lag relation is robust across MSAs and increasing in the size of the boom-bust cycle. We sort MSAs into 20 groups based on the increase in prices from January 2000 to peak and then plot average prices and volume relative to housing stock for each group. The results indicate that the strength of the lead–lag relation between prices and volume appears to increase in the size of the price boom, which provides further support that the patterns in volume through the cycle are linked to market prices.

In Figure 8, we further decompose volume in the cross-section of MSAs and ZIP codes and ask how volume growth in various categories correlates with cross-sectional variation in the size of the price boom. The goal is to connect the evidence on the composition of buyers and sellers during the volume boom to the rise of prices. We present binned scatter plots of growth for total volume, short-term volume, and non-occupant volume against the percent change in prices from January 2000 to peak. Panels (a), (c), and (e) plot the cross-MSA relationships, while Panels (b), (d), and (f) plot the cross-ZIP relationship after absorbing MSA fixed effects. We fix the axes so that the slopes are equal in each row and report the slope of the fitted line. The ZIP code plots weight observations by initial volume in 2000 to avoid undue influence from small ZIP codes.

Panels (a) and (b) confirm the facts in Figure 7 that the relationship between the volume and price booms is robustly positive, both across MSAs and across ZIP codes within MSAs. The slopes are quantitatively close and imply a 0.63-0.67% increase in volume per 1% increase in prices. Panels (c) and (d) show the relationship is stronger for short-term volume and robust across MSAs and ZIP codes within MSAs. Panels (e) and (f) reveal a somewhat stronger relationship between non-occupant purchases and the price boom at the ZIP level than across MSAs. Together this evidence suggests that the growth in both short-term and non-occupant volume documented in the aggregate above is not only robust across local markets but also directly related to the size of the boom in prices.

4.2   Listings over the Cycle

Because our model departs from Walrasian price discovery, it also has implications for the dynamics of listings and unsold inventory over the course of a housing cycle. In particular, Proposition 2 implies that when demand is sufficiently low relative to the stock of current

listings, an inventory of unsold listings will accumulate as extrapolative movers fail to cut prices quickly enough to clear the market. This can happen either because of an influx of listings from recent short-term buyers, or because of a fall in price growth that lowers speculative demand even holding current listings constant.

Figure 1, Panel (b) plots the empirical evolution of unsold listings and national prices between 2000 and 2011. Unsold listings equal the inventory of listed existing houses reported by the NAR (both volume and listings are seasonally adjusted using month-of-year fixed effects). As in Panel (a), we continue to shade the quiet period for reference. The existence of this quiet period confirms the prediction of Proposition 2, and the contemporaneous sharp increase in unsold listings matches the path of simulated unsold inventory in Figure 2, Panel (c).

Our model also has implications for the relation between listings and volume at different stages of the cycle. Proposition 2 predicts that during the boom, volume increases entirely because owners are more likely to list their houses for sale. The speed at which listings sell remains constant, with all listings selling instantaneously just as they do in steady state. Conversely, volume falls during the quiet because the speed at which listings sell begins to decline, as listings outstrip demand.

To determine the relative importance of listing and selling rates for explaining movements in volume, we use the methodology of Ngai and Sheedy (2016) to measure these rates during the cycle. In discrete time, the equations determining these rates are $V_t = s_t I_t$, $I_t = I_{t-1} - V_{t-1} + L_t$, and $L_t = n_t(K_t - I_t)$, where K denotes the housing stock, s denotes the rate at which listings sell, and n denotes the rate at which owners list unlisted houses. Using monthly data on V, I, and K, we calculate $s_t = V_t/I_t$ and $n_t = (I_t - I_{t-1} + V_{t-1})/(K_t - I_t)$ for each month. Figure 9 plots these probabilities averaged within each year.[19]

Figure 9 shows that the listing rate n_t rose 40% during the boom, whereas the rate s_t at which listings sold remained flat. In contrast, the decline in volume during the quiet is driven primarily by s_t, which contracts by about 30% while n_t remains high. As implied by the steady-state equation $V = Ksn/(s + n)$, these rates multiplicatively determine volume, so their proportional changes determine their relative importance. This evidence confirms the model's predictions about what drives volume during the boom and the quiet.

---

[19]In discrete time, V and L are integrals of their continuous time counterparts, whereas I equals the integral of continuous-time listings. In this exercise, we use existing sales from the NAR for V, because it is analogous to the data on I. Burnside, Eichenbaum and Rebelo (2016) plot the same series for s_t.

4.3   Initial Volume and the Size of the Cycle

An additional implication of the model is that an increase in the distribution of expected holding times $f_\lambda(\cdot)$ raises both the level of steady-state volume and the amplitude of the boom–bust cycle. All else equal, markets with a larger pool of potential short-term buyers should experience larger price swings during the cycle. To evaluate this prediction, we test whether steady-state volume and the amplitude of the price cycle are positively correlated across MSAs. Our measure of steady-state volume is the number of existing home sales in 2000 as a share of the housing stock.[20] The boom in each MSA is the percentage change in prices between January 2000 and the price peak, and the bust equals the percentage change between the month of the peak and the month in which prices reached their lowest level subsequent to the peak month.

Figure 10 plots the relationship between steady-state volume and the magnitude of the boom and bust in prices across MSAs. Panel (a) shows a clear positive relationship between initial volume and the magnitude of the price boom: MSAs with higher initial volume experienced significantly larger house price booms. As shown in Panel (b), these MSAs also experienced more drastic drops in prices following the boom.

Columns 1 and 3 of Table 2 quantify this relationship by reporting coefficient estimates from simple linear regressions of the price boom (Column 1) and bust (Column 3) on steady-state volume. A one percentage point increase in the share of the existing housing stock that turned over in 2000 is associated with a 15 percentage point higher increase in prices from January 2000 to peak and a 4 percentage point larger fall in prices from peak to trough. In Columns 2 and 4, we report analogous and nearly identical estimates from regressions that instead assume that the boom ended in January of 2006 for all MSAs. These results are strongly consistent with the prediction of our model that steady-state volume should be correlated with the magnitude of swings in house prices during boom–bust episodes.

---

[20]We omit new construction sales from this exercise in order to avoid conflating differences in supply elasticity or the rate of new construction with differences in steady-state existing sales volume. New construction sales are identified as described in Appendix C.

5 Comparison to Related Work

5.1 Empirics

Our empirical results build on a number of prior studies of the 2000s housing bubble. In particular, our focus on short-term investment relates our work to several other papers exploring the role of "flippers" in the U.S. housing market. Adelino, Schoar and Severino (2016) note that ZIP codes with faster house price growth between 2002 and 2006 witness more flipping activity, but do not directly explore the implications of this fact for aggregate transaction volume. Bayer, Geissler, Mangum and Roberts (2011) focus on Los Angeles and distinguish inexperienced flippers who trade only two or three houses from experienced flippers who trade many more. Inexperienced flippers sharply enter the housing market during the boom and, unlike experienced flippers, make money primarily from market-wide price growth. We find similarly high entry rates of inexperienced flippers in our larger sample of 115 MSAs, and our model explains why new short-term investors would disproportionately enter the market during a price boom. Because we use data from multiple MSAs, we are also able to show that the relationship between the entry of short-term investors and house price growth holds both within and across local markets. In related work, Bayer, Mangum and Roberts (2016) find evidence that contagion may have played a role in driving the entry of these flippers; contagion shares the spirit of extrapolation in that witnessing capital gains drives more people to pursue short-term strategies in the market.

Our results on non-occupant purchases also relate this paper to a growing empirical literature documenting the role of investors in the U.S. housing market. Haughwout, Lee, Tracy and van der Klaauw (2011) and Bhutta (2015) use credit registry data to show that borrowers with multiple first lien mortgages explain a disproportionate share of mortgage growth during the housing boom, especially in states with the largest price increases. Chinco and Mayer (2016) and Gao, Sockin and Xiong (2017) complement these findings by arguing that investors causally increased house price growth during this episode. We confirm the disproportionate role of non-occupant buyers with complete market-level information derived from deeds records across many MSAs. By using the entire universe of home sales, we account for the role of non-occupants for explaining aggregate volume.

Finally, several of our empirical findings relate our work to the large literature surveyed by Han and Strange (2015) that uses search and matching models to understand housing

market dynamics. As pointed out by Ngai and Sheedy (2016), volume rises during price booms in many search models because of an increase in selling speed (see, e.g., Díaz and Jerez (2013), Head, Lloyd-Ellis and Sun (2014), Burnside, Eichenbaum and Rebelo (2016), and Hedlund (2016)). Our results in Figure 9 show the opposite: volume rose from 2000 to 2005 because the number of listings grew while the speed at which listings sold remained flat. Our model explains this phenomenon via the entry of short-horizon buyers during the boom. Ngai and Sheedy (2016) offer a complementary theory based on homeowners who are more likely to move while prices are rising. Our finding that most sellers do not show up as repeat buyers within the same MSA and that the entry of short-term and non-occupant buyers are key to explaining aggregate volume also poses a challenge for many search models in which equilibrium dynamics are driven by the joint buyer–seller problem in a housing market that is modeled as a closed system (Wheaton, 1990; Caplin and Leahy, 2011; Díaz and Jerez, 2013; Ngai and Tenreyro, 2014).

In summary, our paper explains existing empirical results and establishes them more broadly in data that allow us to account for the role of short-term investors in driving total volume over the cycle. We connect these facts on speculation to the pronounced lead–lag relationship between prices and transaction volume at the aggregate level, which prior work has not emphasized. We document the robustness of this lead–lag relationship across MSAs and the cross-sectional relationship between the size of the price boom and the entry of speculative buyers.

5.2 Theory

Recent papers by Barberis, Greenwood, Jin and Shleifer (2017) and Liao and Peng (2018) are similar to ours: we all propose models of volume dynamics in extrapolative bubbles. All three papers predict that volume increases along with prices, but the mechanisms differ. Volume rises in our model due to the entry of short-horizon buyers into the market. In contrast, rising prices make an individual investor more likely to sell in the other two papers. Psychological mechanisms generate this channel: wavering in Barberis, Greenwood, Jin and Shleifer (2017) and realization utility (Barberis and Xiong, 2012) in Liao and Peng (2018). Other predictions are unique to our model.

First, only our model predicts a period of rising listings and falling volume at the end of the price boom. A divergence between listings and volume is not possible in models with

Walrasian price discovery such as Barberis, Greenwood, Jin and Shleifer (2017) and Liao and Peng (2018). While volume may decline before prices in a model with Walrasian price discovery, interest in selling must fall alongside volume. Exactly this phenomenon plays out in Barberis, Greenwood, Jin and Shleifer (2017) and Liao and Peng (2018). Figure 9 shows that the listing rate remained high during the 2006 quiet in the U.S. housing market.

Second, neither Barberis, Greenwood, Jin and Shleifer (2017) nor Liao and Peng (2018) predict that potential buyers with exogenously short horizons buy disproportionately in a boom. In the housing market, relatively exogenous demographics like age and education predict investment horizons (Henderson and Ioannides, 1989; Kan, 1999), and individuals with such characteristics are indeed more likely to buy a home when house prices have been rising (Edelstein and Qian, 2014). Because demographics provide a natural source of variation in average horizons across cities, only our model seems to predict that cities with more housing turnover in 2000 witness larger cycles thereafter.

Our model builds on the work of Frankel and Froot (1986), Cutler, Poterba and Summers (1990) and De Long, Shleifer, Summers and Waldmann (1990), who show that positive feedback trading causes booms and busts in asset prices. Cutler, Poterba and Summers (1990) and De Long, Shleifer, Summers and Waldmann (1990) feature three classes of agents: positive feedback traders who buy more stock when prices have been rising, fundamentalists who sell stock when prices rise above an eventual liquidation value, and arbitrageurs who rationally exploit the other two traders. As prices rise, their models generate short-term volume as arbitrageurs buy from the fundamentalists and then sell to the positive feedback traders. Thus, in each of our papers, extrapolation and short-run trading raise prices and volume after a shock.

Our contribution is to maintain the insight of their papers in a model that can be calibrated to the housing market. In particular, we endogenously derive asset demand as a function of past price changes (Proposition 1), whereas Cutler, Poterba and Summers (1990) and De Long, Shleifer, Summers and Waldmann (1990) assume that asset demand from positive feedback traders increases in past price changes. Our approach explicitly ties demand to investment horizons and beliefs in the housing market, allowing us to calibrate demand using the NAR data on horizons and the survey evidence on house price expectations from Armona, Fuster and Zafar (2016). Each agent from the earlier papers appears in some form in our work: the positive feedback traders emerge as the short-horizon buyers, the fundamen-

29

talists become the long-horizon buyers, and the arbitrageurs resemble the informed movers in that both determine market prices by rationally exploiting the extrapolative tendencies of other market participants. Our setup also permits an arbitrary number of periods and departs from Walrasian price discovery.

In emphasizing sluggish price adjustment and extrapolators with fixed horizons, our model is closer to the earlier work of Hong and Stein (1999). However, our models differ in key ways. Everyone extrapolates in our model, not just the short-horizon investors. We endogenously derive that short-horizon investors are more sensitive to expected capital gains. Another difference is our departure from Walrasian price discovery. Broadly speaking, our model takes extrapolation as a primitive and studies the implications for volume and market participation, whereas Hong and Stein (1999) start from inattention and derive extrapolation as a boundedly rational response for short-run arbitrageurs.

A related literature in behavioral finance pins overpricing and trading volume not on extrapolation but on disagreement (e.g., Scheinkman and Xiong (2003); see Hong and Stein (2007) and Daniel and Hirshleifer (2015) for surveys). While disagreement surely accounts for some of the average prices and volume in the housing market (as Bailey, Cao, Kuchler and Stroebel (2016) document empirically), most disagreement models do not characterize the dynamics of prices and volume over a single bubble episode. An exception is Burnside, Eichenbaum and Rebelo (2016). While that model produces an increase in volume during a house price boom, it does not predict an increase in short-term sales. More generally, models with heterogeneous beliefs explain short-term sales only when optimists are likely to become pessimists exogenously, as is the case in Piazzesi and Schneider (2009), Penasse and Renneboog (2016), and Barberis, Greenwood, Jin and Shleifer (2017). Disagreement also struggles to explain the widespread optimism about house price growth during the boom we study (Case, Shiller and Thompson, 2012; Foote, Gerardi and Willen, 2012; Cheng, Raina and Xiong, 2014), although it can account for the dispersion in these beliefs (Piazzesi and Schneider, 2009; Burnside, Eichenbaum and Rebelo, 2016).

6   Final Remarks

Our paper raises additional lines of inquiry within the housing market. We have argued, theoretically and empirically, that short-term investors play a crucial role in the housing cycle.

Do the expansions in credit that typically accompany housing booms appeal disproportionately to short-term investors? Barlevy and Fisher (2011) document a strong correlation across U.S. metropolitan areas between the size of the 2000s house price boom and the take-up of interest-only mortgages. These mortgages back-load payments by deferring principal repayment for some amount of time and thus might appeal especially to buyers who expect to resell quickly. The targeting of credit expansions to short-term buyers might explain the amplification effects of credit availability on real estate booms documented by Di Maggio and Kermani (2017), Favara and Imbs (2015), and Rajan and Ramcharan (2015).

A second line of inquiry within housing concerns tax policy. The capital gains tax discourages housing speculation by lowering expected after-tax capital gains. However, it discourages productive residential investment as well. Is this tax optimal, and if not, what type of tax policy would be better? Many economists have analyzed or proposed transaction taxes (Tobin, 1978; Stiglitz, 1989; Summers and Summers, 1989; Dávila, 2015). It is unclear whether these taxes would particularly discourage short-term investors, given that the incidence of this tax might fall more on buyers than sellers.

Although this paper focuses on the housing market, many of the patterns we study appear in other asset markets as well. Several famous bubbles involve large movements in transaction volume (Cochrane, 2011). The lead–lag relation between prices and volume holds in four other boom-bust episodes shown in Figure 11: the 1995–2005 market in technology stocks, the 1985–1995 Japanese stock market, the experimental bubbles studied by Smith, Suchanek and Williams (1988), and the 1985–1995 bubble in Postwar art. Short-horizon trading was prevalent during the technology boom (Cochrane, 2002; Ofek and Richardson, 2003). Even outside of bubbles, stock market volume increases following high returns and predicts negative subsequent returns (Lee and Swaminathan, 2000; Jones, 2002; Statman, Thorley and Vorkink, 2006; Griffin, Nardari and Stulz, 2007).

Cutler, Poterba and Summers (1991) document price dynamics such as momentum and mean reversion in many asset classes. They conclude the generality of these patterns suggests that inherent features of the speculative process likely explain them. Can our model of speculation explain the joint dynamics of prices and volume outside the housing market? The assumption most tailored to housing is the departure from Walrasian price discovery, which does not permit a direct analogy to more centralized markets. We hope that future work will investigate the striking similarity of volume dynamics in other markets.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2016. "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class." *Review of Financial Studies*, 29(7): 1635–1670.

Amromin, Gene, and Steve A. Sharpe. 2014. "From the Horse's Mouth: Economic Conditions and Investor Expectations of Risk and Return." *Management Science*, 60(4): 845–866.

Anenberg, Elliot, and Patrick Bayer. 2013. "Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem." National Bureau of Economic Research.

Armona, Luis, Andreas Fuster, and Basit Zafar. 2016. "Home Price Expectations and Behavior: Evidence from a Randomized Information Experiment." Working Paper.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. 2016. "Social Networks and Housing Markets." Working Paper.

Barberis, Nicholas, and Wei Xiong. 2012. "Realization Utility." *Journal of Financial Economics*, 104: 251–271.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015. "X-CAPM: An Extrapolative Capital Asset Pricing Model." *Journal of Financial Economics*, 115(1): 1–24.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2017. "Extrapolation and bubbles." Forthcoming, *Journal of Financial Economics*.

Barlevy, Gadi, and Jonas D.M. Fisher. 2011. "Mortgage Choices and Housing Speculation." Working Paper, Federal Reserve Bank of Chicago.

Bayer, Patrick, Christopher Geissler, Kyle Mangum, and James W. Roberts. 2011. "Speculators and Middlemen: The Strategy and Performance of Investors in the Housing Market." NBER Working Paper 16784.

Bayer, Patrick, Kyle Mangum, and James W. Roberts. 2016. "Speculative Fever: Investor Contagion in the Housing Bubble." NBER Working Paper 22065.

Bhutta, Neil. 2015. "The Ins and Outs of Mortgage Debt during the Housing Boom and Bust." *Journal of Monetary Economics*, 76: 284–298.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2016. "Understanding Booms and Busts in Housing Markets." *Journal of Political Economy*, 124(4): 1088–1147.

Caplin, Andrew, and John Leahy. 2011. "Trading Frictions and House Price Dynamics." *Journal of Money, Credit, and Banking*, 43(2): 283–303.

Case, Karl E. 2008. "The Central Role of Home Prices in the Current Financial Crisis: How Will the Market Clear?" *Brookings Papers on Economic Activity*, 2: 161–193.

Case, Karl E., Robert J. Shiller, and Anne K. Thompson. 2012. "What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets." *Brookings Papers on Economic Activity*, Fall: 265–298.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. 2017. "Housing Booms and Busts, Labor Maket Opportunities, and College Attendance." Forthcoming, *American Economic Review*.

Cheng, Ing-Haw, Sahil Raina, and Wei Xiong. 2014. "Wall Street and the Housing Bubble." *American Economic Review*, 104(9): 2797–2829.

Chinco, Alex, and Christopher Mayer. 2016. "Misinformed Speculators and Mispricing in the Housing Market." *Review of Financial Studies*, 29(2): 486–522.

Clayton, Jim, Norman Miller, and Liang Peng. 2010. "Price-volume Correlation in the Housing Market: Causality and Co-movements." *Journal of Real Estate Finance and Economics*, 40(1): 14–40.

Cochrane, John H. 2002. "Stocks as Money: Convenience Yield and the Tech-Stock Bubble." Working Paper.

Cochrane, John H. 2011. "Discount Rates." *Journal of Finance*, 66: 1047–1108.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1990. "Speculative Dynamics and the Role of Feedback Traders." *American Economic Review, Papers and Proceedings*, 80(2): 63–68.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. "Speculative Dynamics." *Review of Economic Studies*, 58(3): 529–546.

Daniel, Kent, and David Hirshleifer. 2015. "Overconfident Investors, Predictable Returns, and Excessive Trading." *Journal of Economic Perspectives*, 29(4): 61–88.

Dávila, Eduardo. 2015. "Optimal Financial Transactions Taxes." Working Paper, NYU Stern.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, 45(2): 379–395.

Díaz, Antonia, and Belén Jerez. 2013. "House Prices, Sales, and Time on the Market: A Search-Theoretic Framework." *International Economic Review*, 54(3): 837–872.

Di Maggio, Marco, and Amir Kermani. 2017. "Credit-Induced Boom and Bust." *The Review of Financial Studies*, 30(11): 3711–3758.

Edelstein, Robert, and Wenlan Qian. 2014. "Short-Term Buyers and Housing Market Dynamics." *Journal of Real Estate Finance and Economics*, 49: 654–689.

Favara, Giovanni, and Jean Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review*, 105(3): 958–992.

Foote, Christopher L., Kristopher S. Gerardi, and Paul S. Willen. 2012. "Why Did So Many People Make So Many *Ex Post* Bad Decisions? The Causes of the Foreclosure Crisis." *Federal Reserve Bank of Boston Public Policy Discussion Papers*, 12(2).

Frankel, Jeffrey A., and Kenneth A. Froot. 1986. "Understanding the U.S. Dollar in the Eighties: The Expectations of Chartists and Fundamentalists." *Economic Record*, 24–38.

Frankel, Jeffrey A., and Kenneth A. Froot. 1990. "Chartists, Fundamentalists, and Trading in the Foreign Exchange Market." *American Economic Review*, 80(2): 181–185.

Gao, Zhenyu, Michael Sockin, and Wei Xiong. 2017. "Economic Consequences of Housing Speculation." Working Paper.

Genesove, David, and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics*, 116(4): 1233–1260.

Glaeser, Edward L. 2013. "A Nation of Gamblers: Real Estate Speculation and American History." *American Economic Review*, 103(3): 1–42.

Glaeser, Edward L., and Charles G. Nathanson. 2017. "An extrapolative model of house price dynamics." *Journal of Financial Economics*, 126(1): 147 – 170.

Graham, John R., and Campbell R. Harvey. 2003. "Expectations of Equity Risk Premia, Volatility and Asymmetry." Working Paper.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies*, 27(3): 714–746.

Griffin, John M., Federico Nardari, and René M. Stulz. 2007. "Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries." *Review of Financial Studies*, 20(3): 905–951.

Guren, Adam. 2018. "House Price Momentum and Strategic Complementarity." Forthcoming, *Journal of Political Economy*.

Guren, Adam M., and Timothy J. McQuade. 2015. "How do Foreclosures Exacerbate Housing Downturns?" Working Paper.

Han, Lu, and William C. Strange. 2015. "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage." In *Handbook of Regional and Urban Economics*. Vol. 5B, , ed. Gilles Duranton, J. Vernon Henderson and William C. Strange, Chapter 13, 813–886. Elsevier.

Hanushek, Eric A., and John M. Quigley. 1980. "What is the Price Elasticity of Housing Demand?" *Review of Economics and Statistics*, 62(3): 449–454.

Harrison, J. Michael, and David M. Kreps. 1978. "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations." *Quarterly Journal of Economics*, 92(2): 323–336.

Haughwout, Andrew, Donghoon Lee, Joseph Tracy, and Wilbert van der Klaauw. 2011. "Real Estate Investors, the Leverage Cycle, and the Housing Market Crisis." Federal Reserve Bank of New York Staff Report no. 514.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun. 2014. "Search, Liquidity, and the Dynamics of House Prices and Construction." *American Economic Review*, 104(4): 1172–1210.

Hedlund, Aaron. 2016. "The Cyclical Dynamics of Illiquid Housing, Debt, and Foreclosures." *Quantitative Economics*, 7: 289–328.

Henderson, J. Vernon, and Yannis M. Ioannides. 1989. "Dynamic Aspects of Consumer Decisions in Housing Markets." *Journal of Urban Economics*, 26: 212–230.

Hong, Harrison, and Jeremy C. Stein. 1999. "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets." *Journal of Finance*, 54(6): 2143–2184.

Hong, Harrison, and Jeremy C. Stein. 2007. "Disagreement and the Stock Market." *Journal of Economic Perspectives*, 21(2): 109–128.

Jones, Charles M. 2002. "A Century of Stock Market Liquidity and Trading Costs." Working Paper.

Kan, Kamhon. 1999. "Expected and Unexpected Residential Mobility." *Journal of Urban Economics*, 45: 72–96.

Lee, Charles, and Bhaskaran Swaminathan. 2000. "Price Momentum and Trading Volume." *Journal of Finance*, 55(5): 2017–2069.

Leung, Charles K. Y., Garrion C. K. Lau, and Youngman C. F. Leong. 2002. "Testing Alternative Theories of the Property Price-Trading Volume Correlation." *Journal of Real Estate Research*, 23(3): 253–264.

Liao, Jingchi, and Cameron Peng. 2018. "Price and Volume Dynamics in Bubbles." Working Paper.

Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics*, 117(4): 1295–1328.

Mayer, Christopher. 2011. "Housing Bubbles: A Survey." *Annual Review of Economics*, 3: 559–577.

Mayer, Christopher, Karen Pence, and Shane M. Sherlund. 2009. "The Rise in Mortgage Defaults." *Journal of Economic Perspectives*, 23(1): 27–50.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2015. "Foreclosures, House Prices, and the Real Economy." *Journal of Finance*, 70(6): 2587–2634.

Mian, Atif, and Amir Sufi. 2014. "What Explains the 2007–2009 Drop in Employment?" *Econometrica*, 82(6): 2197–2223.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. "Household Balance Sheets, Consumption, and the Economic Slump." *Quarterly Journal of Economics*, 128(4): 1687–1726.

Nathanson, Charles G., and Eric Zwick. 2018. "Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market." Forthcoming, *Journal of Finance*.

Ngai, L. Rachel, and Kevin D. Sheedy. 2016. "The Decision to Move House and Aggregate Housing-Market Dynamics." Working Paper.

Ngai, Rachel, and Silvana Tenreyro. 2014. "Hot and Cold Seasons in the Housing Market." *American Economic Review*, 104(12): 3991–4026.

Ofek, Eli, and Matthew Richardson. 2003. "Dotcom Mania: The Rise and Fall of Internet Stock Prices." *Journal of Finance*, 58(3): 1113–1137.

Ortalo-Magné, François, and Sven Rady. 2006. "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints." *Review of Economic Studies*, 73(2): 459–485.

Penasse, Julien, and Luc Renneboog. 2016. "Bubbles and Trading Frenzies: Evidence from the Art Market." Working Paper.

Piazzesi, Monika, and Martin Schneider. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review*, 99(2): 406–411.

Rajan, Raghuram, and Rodney Ramcharan. 2015. "The Anatomy of a Credit Crisis: The Boom and Bust in Farm Land Prices in The United States in the 1920s." *American Economic Review*, 105(4): 1439–1477.

Rognlie, Matthew, Andrei Shleifer, and Alp Simsek. 2017. "Investment Hangover and the Great Recession." Forthcoming, *American Economic Journal: Macroeconomics*.

Scheinkman, Jose A., and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy*, 111(6): 1183–1220.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams. 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica*, 56(5): 1119–1151.

Statman, Meir, Steven Thorley, and Keith Vorkink. 2006. "Investor Overconfidence and Trading Volume." *Review of Financial Studies*, 19(4): 1531–1565.

Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects." *Quarterly Journal of Economics*, 110(2): 379–406.

Stiglitz, Joseph E. 1989. "Using Tax Policy To Curb Speculative Short-Term Trading." *Journal of Financial Services Research*, 3: 101–115.

Summers, Lawrence H., and Victoria P. Summers. 1989. "When Financial Markets Work Too Well: A Cautious Case For a Securities Transaction Tax." *Journal of Financial Services Research*, 3: 261–286.

Tobin, James. 1978. "A Proposal for International Monetary Reform." *Eastern Economic Journal*, 4(3): 153–159.

Vissing-Jorgensen, Annette. 2004. "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions." In *NBER Macroeconomics Annual 2003, Volume 18*. 139–208. The MIT Press.

Wheaton, William C. 1990. "Vacancy, Search, and Prices in a Housing Market Matching Model." *Journal of Political Economy*, 98(6): 1270–1292.

FIGURE 1

The Dynamics of Prices, Volume, and Inventories

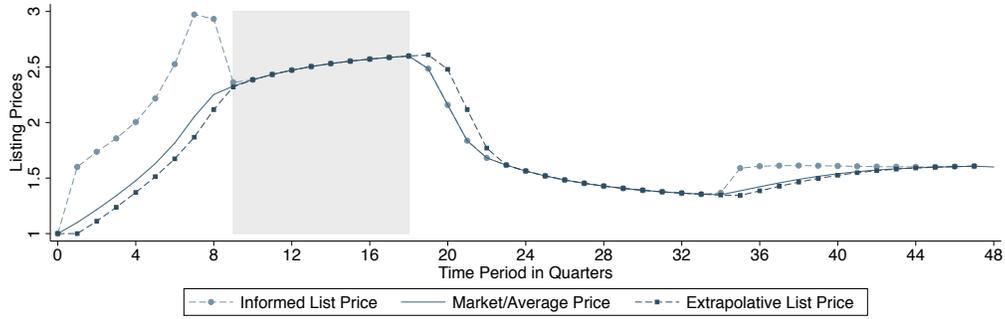(a) Prices and Volume
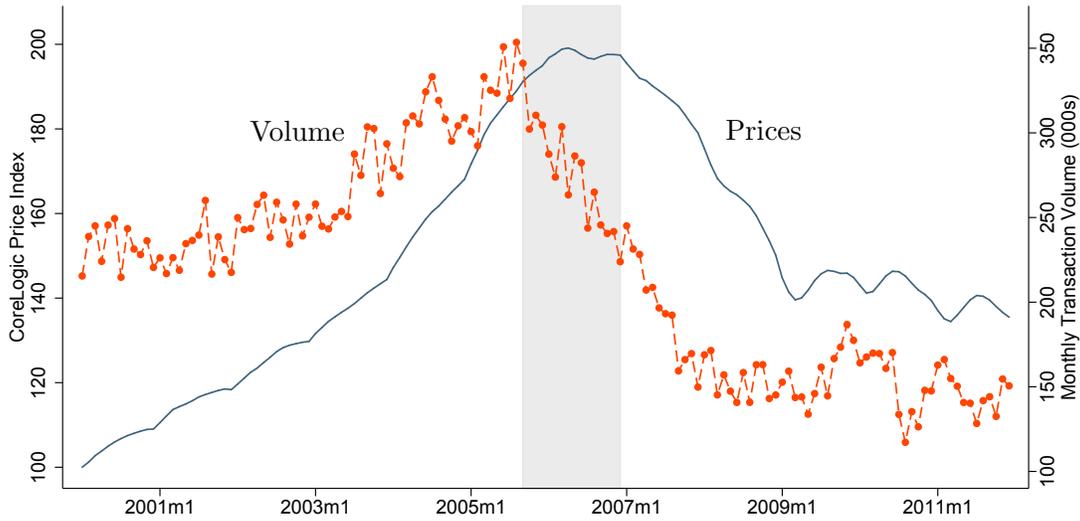
(b) Prices and Inventories
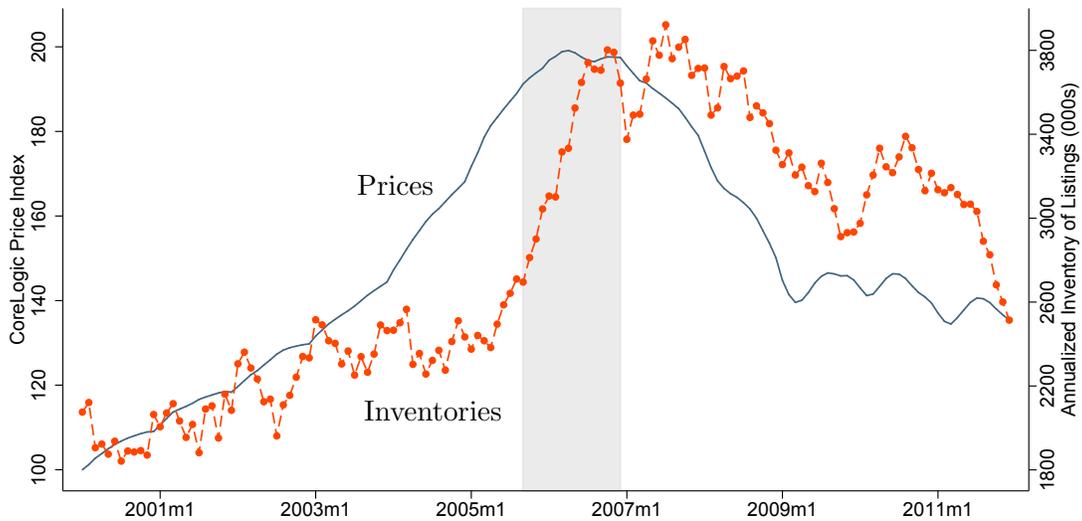
Notes: This figure displays the dynamic relationship between prices, volume, and the inventory of listings in the U.S. housing market between 2000 and 2011. Panel (a) plots monthly prices and sales volume, and panel (b) plots monthly prices and inventory. Monthly price index information comes from CoreLogic and monthly sales volume is based on aggregated transaction data from CoreLogic for 115 MSAs representing 48% of the U.S. housing stock. Inventory information comes from the National Association of Realtors. We apply a calendar-month seasonal adjustment for both volume and inventories.
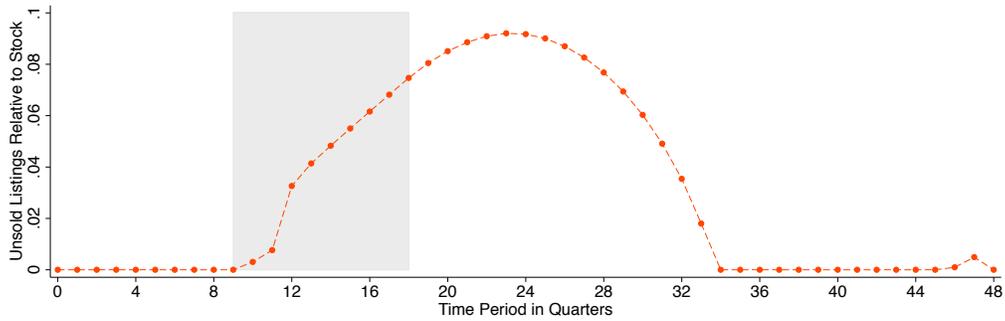
FIGURE 2
Simulation Results

(a) Prices and Volume

(b) List Prices

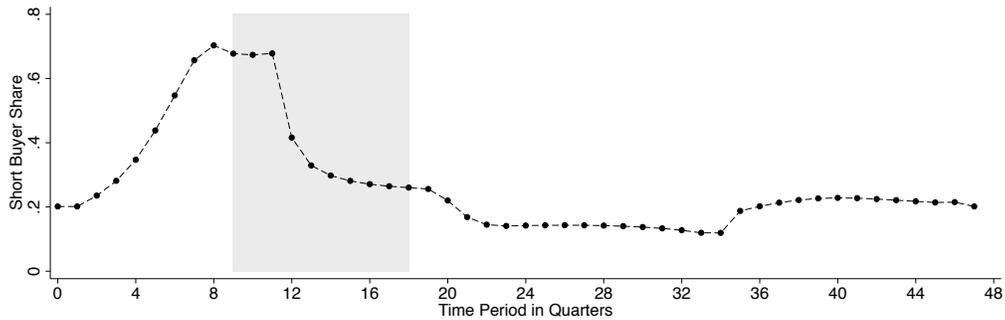Informed List Price — Market/Average Price — Extrapolative List Price
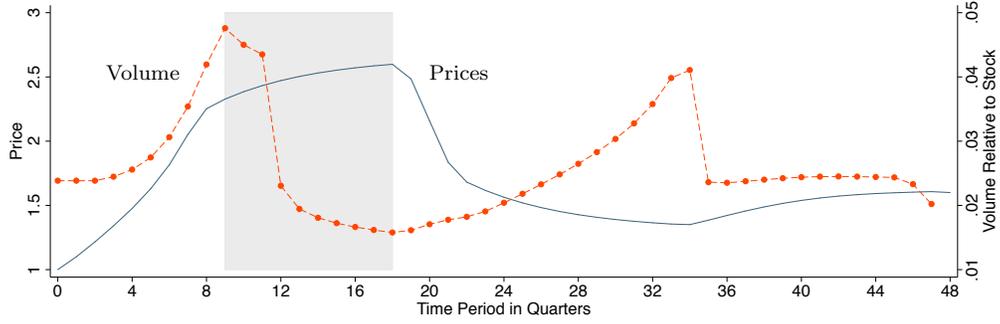
(c) Unsold Listings

(d) Short-term Buyer Share

Notes: These figures plot results from model simulation. The shaded area denotes the quiet. Panel (a) plots the joint pattern of prices and transaction volume relative to the total housing stock. Panel (b) plots the market price and the list prices of informed and extrapolative sellers. Panel (c) plots inventories relative to the total housing stock. Panel (d) plots the share of buyers with expected holding periods of less then 12 quarters.

39

FIGURE 3
Expected Holding Times of Homebuyers, 2008–2015

(a) Response Heterogeneity
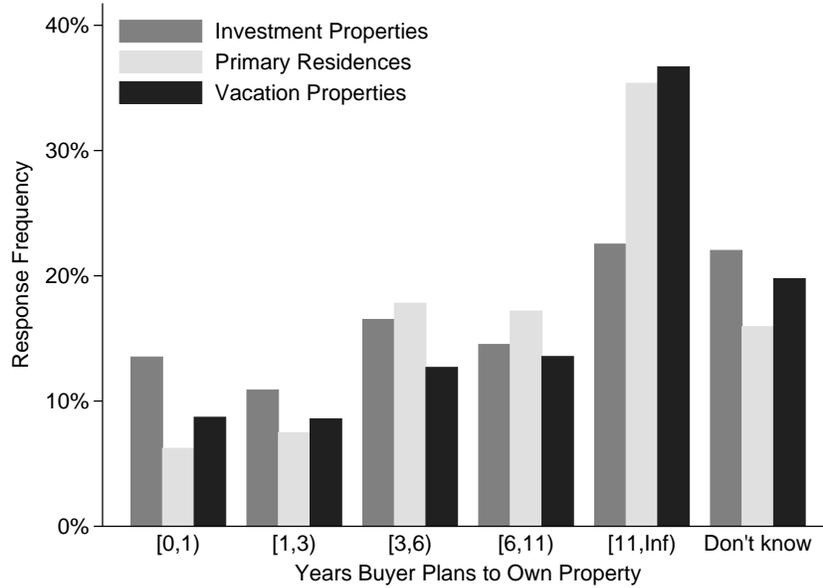
(b) Short-Term Buyers and Recent House Price Growth

Notes: Panel (a) plots the response frequency averaged equally over each year from 2008 to 2015. In Panel (b), "annual house price growth" equals the average across that year's four quarters of the log change in the all-transactions FHFA U.S. house price index from four quarters ago, and "short-term buyer share" equals the share of respondents other than those reporting "don't know" who report a horizon less than three years. Data come from the annual Investment and Vacation Home Buyers Survey conducted by the National Association of Realtors. We reclassify buyers who have already sold their properties by the time of the survey as having an expected holding time in [0,1).
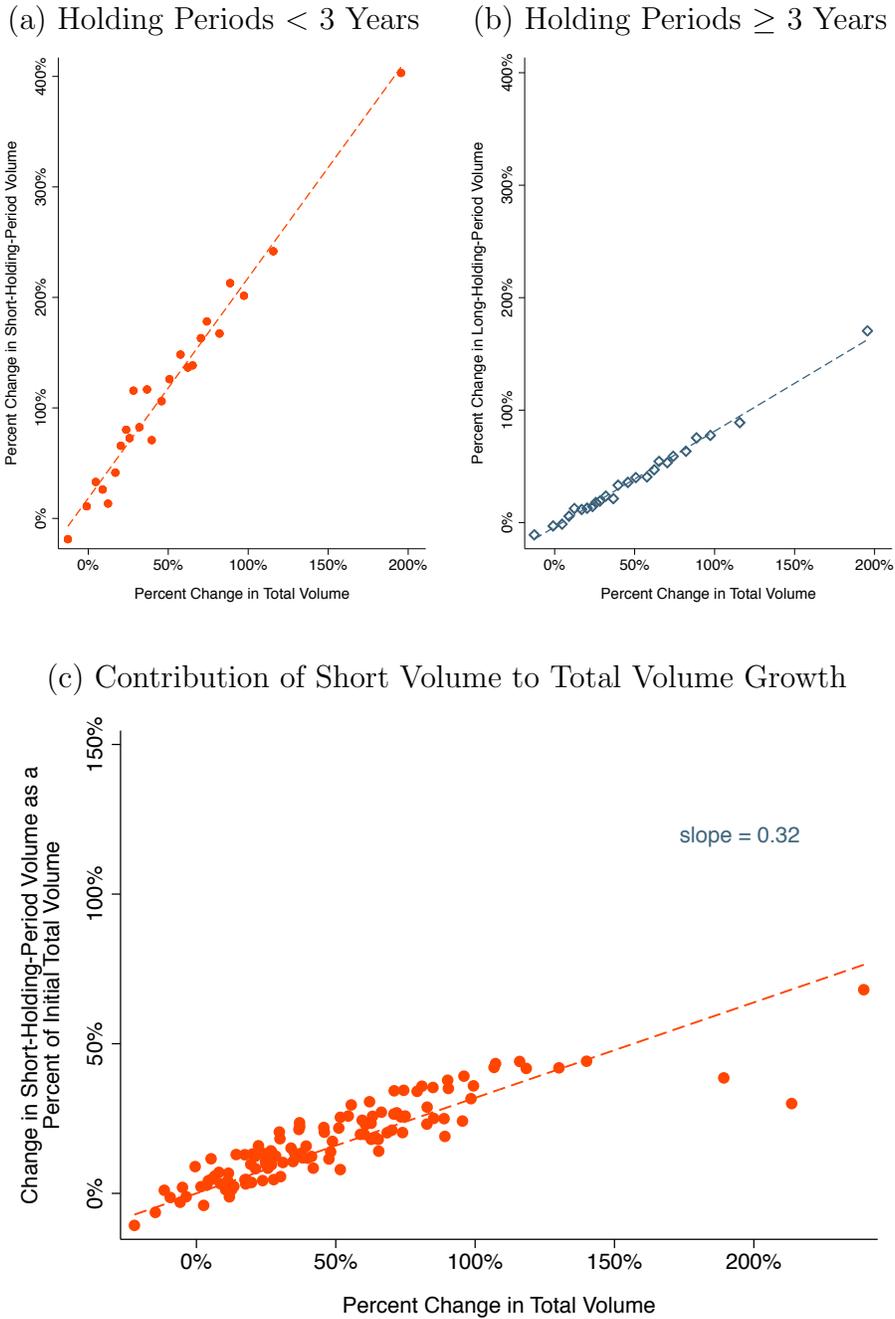
FIGURE 4
The Dynamics of Holding Times in the Housing Market
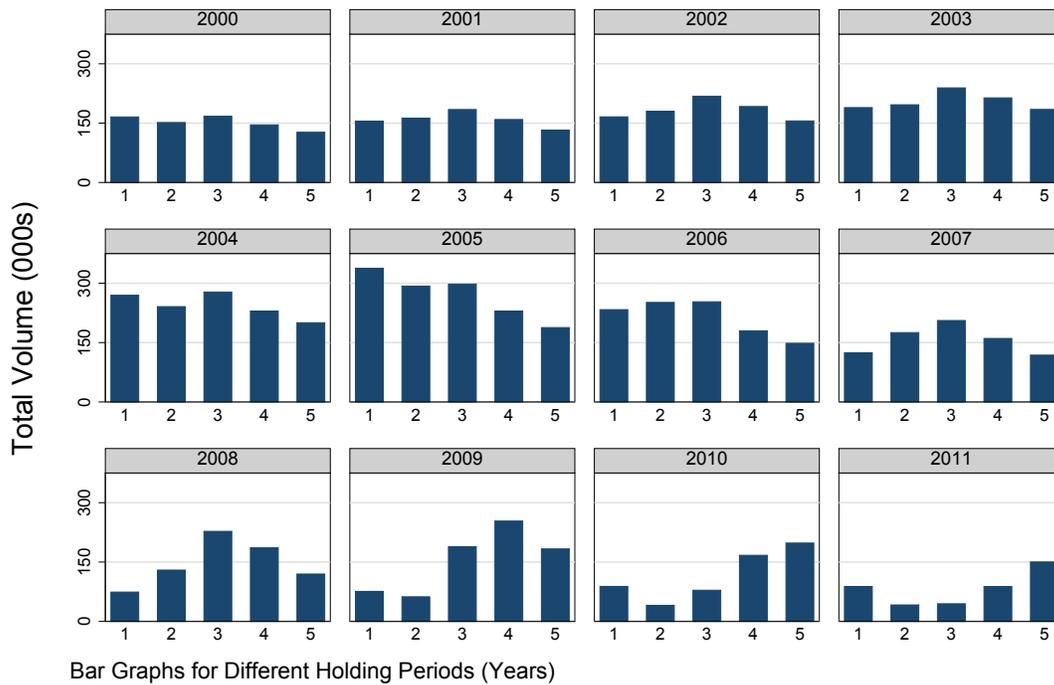
Notes: This figure illustrates the time variation in realized holding periods between 2000 and 2011 in the U.S. For each transaction, we define the holding period as the number of days since the last transaction of the same property. We then group all transactions with holding periods less than or equal to 5 years into bins of 1, 2, 3, 4, or 5 years, respectively. For each year between 2000 and 2011, we plot aggregate transaction counts in each of these five holding-period groups.
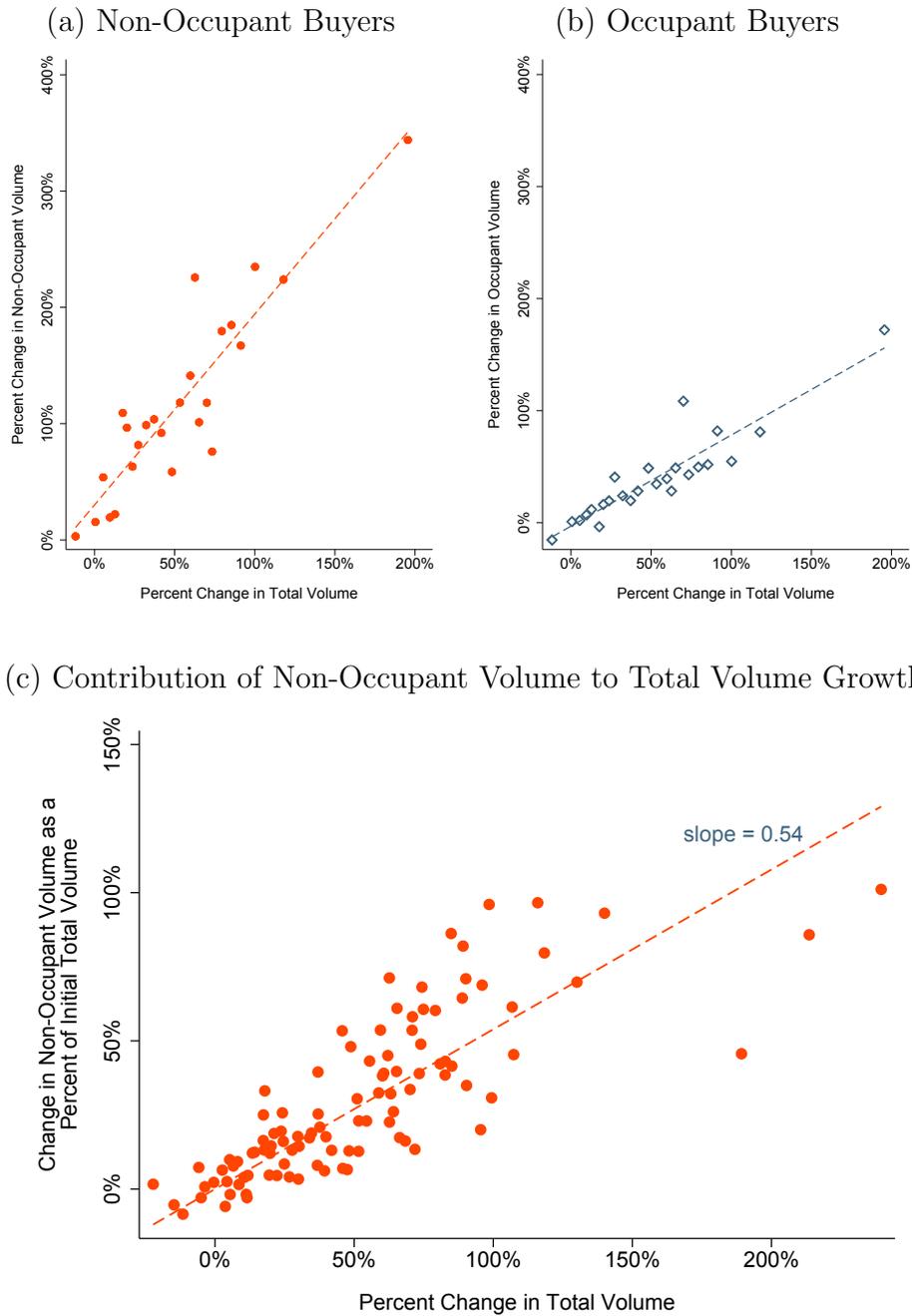
FIGURE 5

The Role of Short-Holding-Period Volume Growth for Total Volume Growth

(a) Holding Periods < 3 Years   (b) Holding Periods ≥ 3 Years

(c) Contribution of Short Volume to Total Volume Growth

Notes: This figure illustrates the quantitative importance of short-holding-period volume in accounting for the increase in total volume between 2000 and 2005. We present binned scatter plots ("binscatters") of the percent change in total volume from 2000 to 2005 versus the percent change in volume for short holding periods (< 3 years) in Panel (a) and long holding periods (≥ 3 years) in Panel (b). Panel (c) shows that the growth in short-holding-period volume is a quantitatively important component of the growth in total volume across MSAs. For each MSA, we plot the change in short-holding-time volume divided by initial total volume on the y-axis against the percent change in total volume on the x-axis.

FIGURE 6
The Role of Non-Occupant Volume Growth for Total Volume Growth

(a) Non-Occupant Buyers

(b) Occupant Buyers

(c) Contribution of Non-Occupant Volume to Total Volume Growth

Notes: This figure illustrates the quantitative importance of non-occupant volume in accounting for the increase in total volume between 2000 and 2005. We present binned scatter plots ("binscatters") of the percent change in total volume from 2000 to 2005 versus the percent change in volume for non-occupant buyers (transactions with distinct mailing and property addresses) in Panel (a) and occcupant buyers (transactions with mailing address missing or matching the property address) in Panel (b). Panel (c) shows that the growth in non-occupant volume is a quantitatively important component of the growth in total volume across MSAs. For each MSA, we plot the increase in non-occupant volume divided by initial total volume on the y-axis against the percent change in total volume on the x-axis.
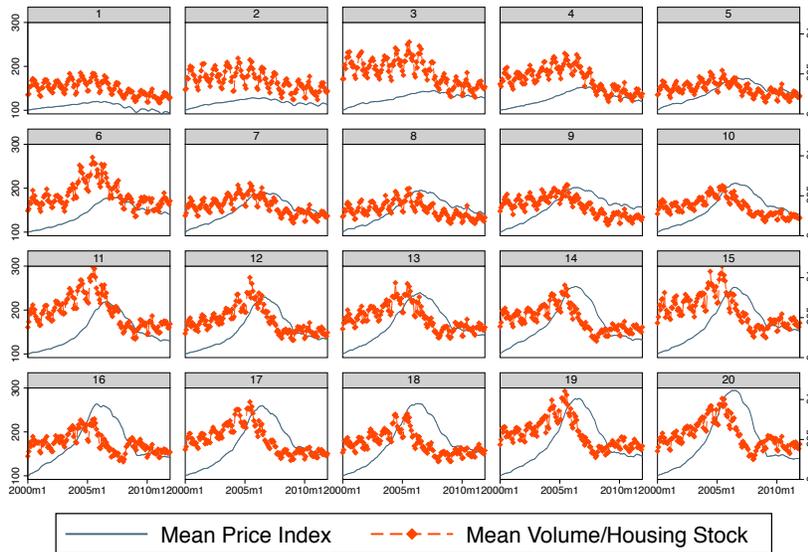
43

FIGURE 7

The Lead-Lag Relationship between Prices and Volume

(a) Correlation between Prices and Volume at Various Lags
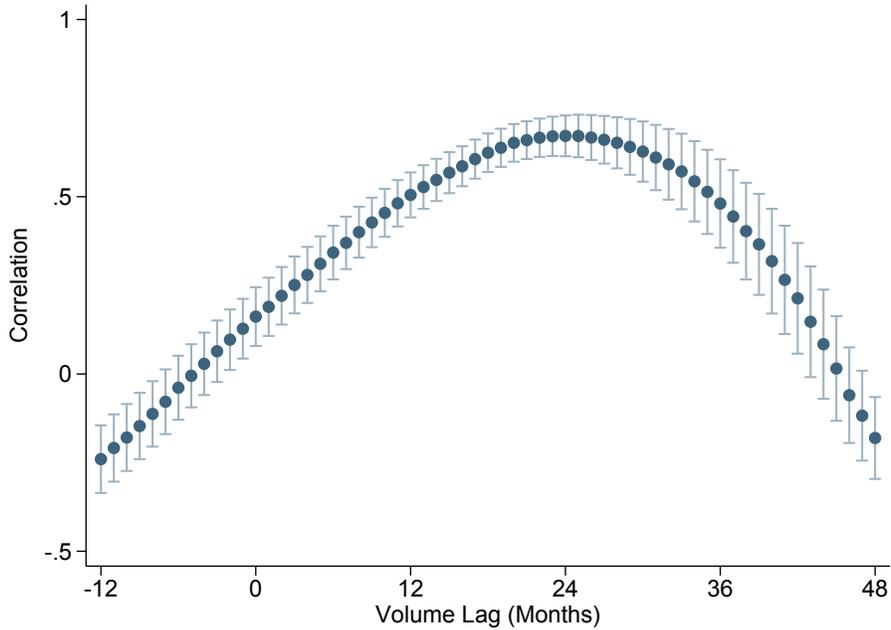
(b) Price Volume Correlation across MSAs with Different Price Booms


Graphs by 20 quantiles of boom

Notes: Panel (a) shows that the correlation between prices and lagged volume is robust across MSAs and maximized at a positive lag of 24 months. We regress the demeaned log of prices on seasonally adjusted lagged volume divided by the 2000 housing stock for each lag from -12 months to 48 months and plot the implied correlation and its 95% confidence interval calculated using standard errors that are clustered by month. Panel (b) shows the relationship is robust across MSAs and increasing in the size of the boom-bust cycle. We sort MSAs into 20 groups based on the increase in prices from January 2000 to peak and then plot average prices and volume relative to housing stock for each group.

FIGURE 8
The Cross-Sectional Relationship between Volume and the Price Run-up
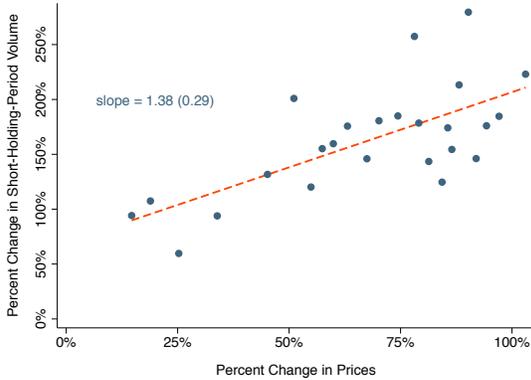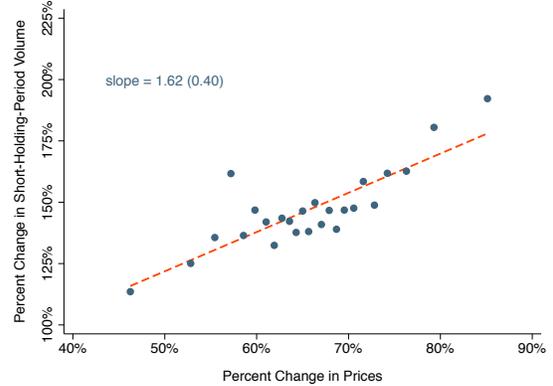
(a) Cross MSA Total Volume
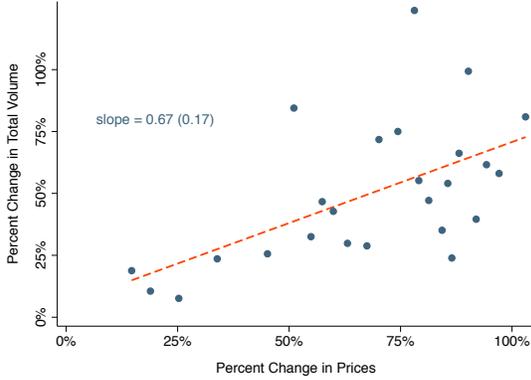
(b) Cross ZIP, Within MSA Total Volume
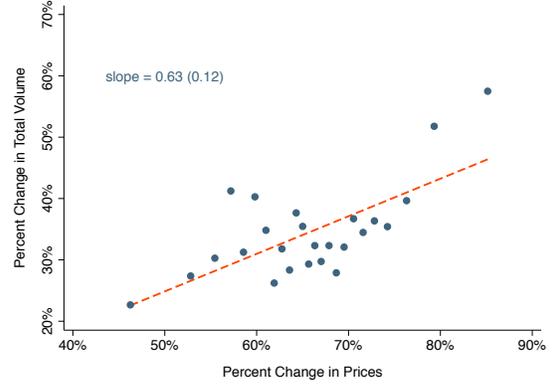
(c) Cross MSA Holding < 3 Years

(d) Cross ZIP, Within MSA Holding < 3 Years

(e) Cross MSA Non-Owner
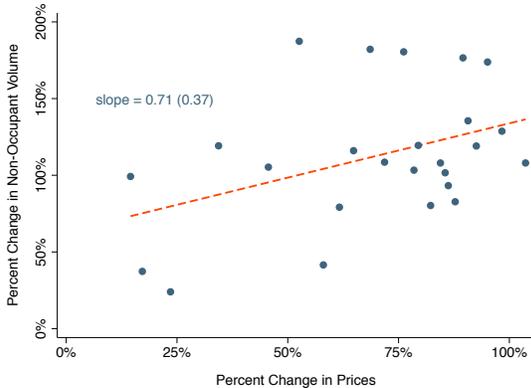
(f) Cross ZIP, Within MSA Non-Owner
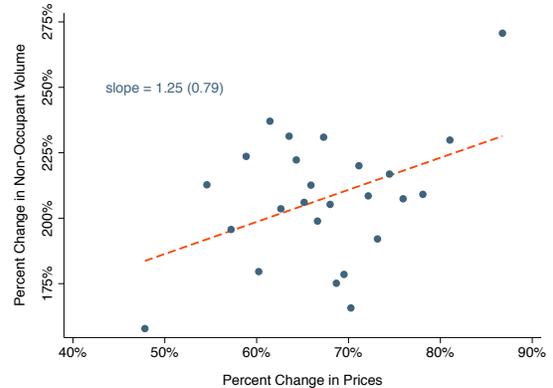
Notes: These figures demonstrate the relationship between prices and speculative volume across MSAs and across ZIP codes within MSAs. We present binned scatter plots ("binscatters") of growth in different categories of volume between 2000 and 2005 against the percent change in prices from January 2000 to peak. Binscatters for ZIP graphs plot residuals after absorbing MSA fixed effects. Standard errors for ZIP-linear fits are clustered at the MSA level.

45

FIGURE 9
Sale and Listing Rates for U.S. Existing Homes, 2000–2011

(a) Sale Rate
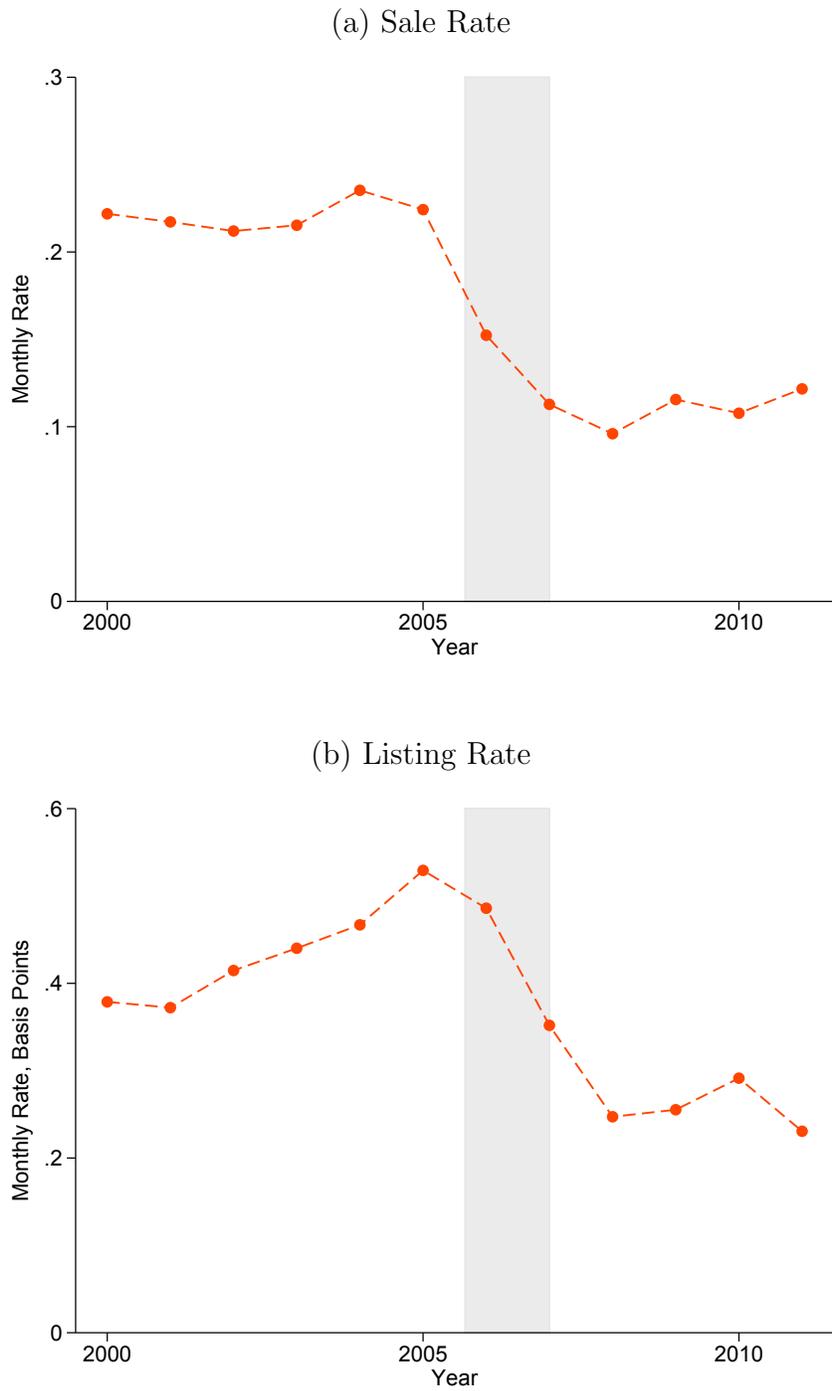
(b) Listing Rate

Notes: The figure displays the average monthly sale rate and listing rate for each year. For each month, the sale rate equals sales divided by listed inventory, and the listing probability equals new listings divided by the stock of unlisted houses. Monthly data on the U.S. housing stock are interpolated from quarterly estimates provided by the U.S. Census, and monthly sales and inventory numbers come from data provided by the National Association of Realtors; new listings are calculated using the inventory and sales data. The shaded region corresponds to the quiet as demarcated in Figure 1.

46

FIGURE 10

Initial Volume and the Magnitude of the Housing Boom and Bust

(a) Boom
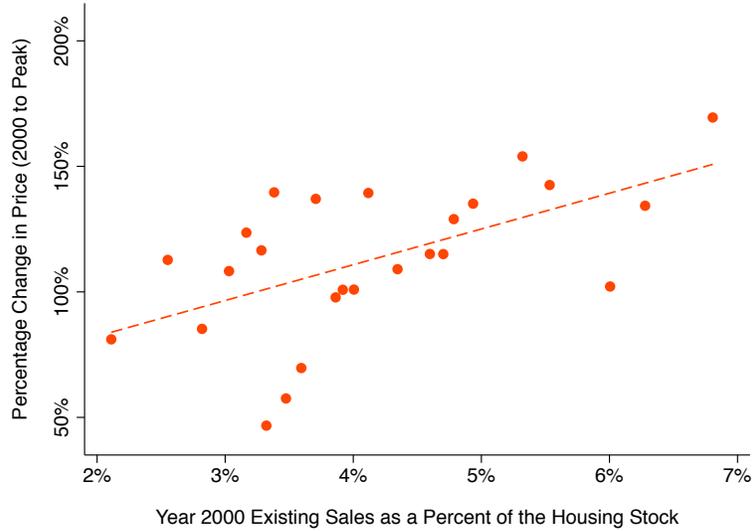
(b) Bust
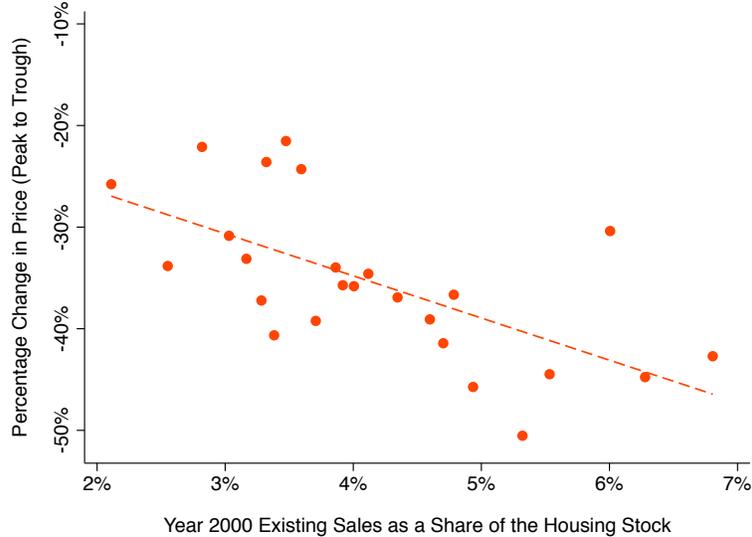
Notes: This figure provides empirical support for the cross-sectional prediction that the magnitude of price swings during boom–bust episodes should be correlated with the level of steady-state transaction volume across markets. We present binned scatter plots ("binscatters") of the percent change in prices from January 2000 to peak (Panel (a)) and from peak to trough (Panel (b)) versus total existing homes in 2000. To facilitate comparisons across cities of different sizes, we normalize existing sales by the size of the housing stock in 2000 for each city. House prices are measured using the monthly CoreLogic repeat-sales house price indicies. The price peak for each MSA is measured as the highest price recorded for that MSA prior to January, 2012. We restrict the peak to occur prior to January 2012, since prices in some markets had already recovered to levels higher than those experienced during the boom by the end of our sample. The trough is the lowest price subsequent to the month of the peak.

FIGURE 11
The Joint Dynamics of Prices and Volume

(a) US Equities, Tech (1995–2005)

(b) Japan Equities (1985–1995)

(c) Experimental Markets, SSW (1988)

(d) Art Market (1985–1995)
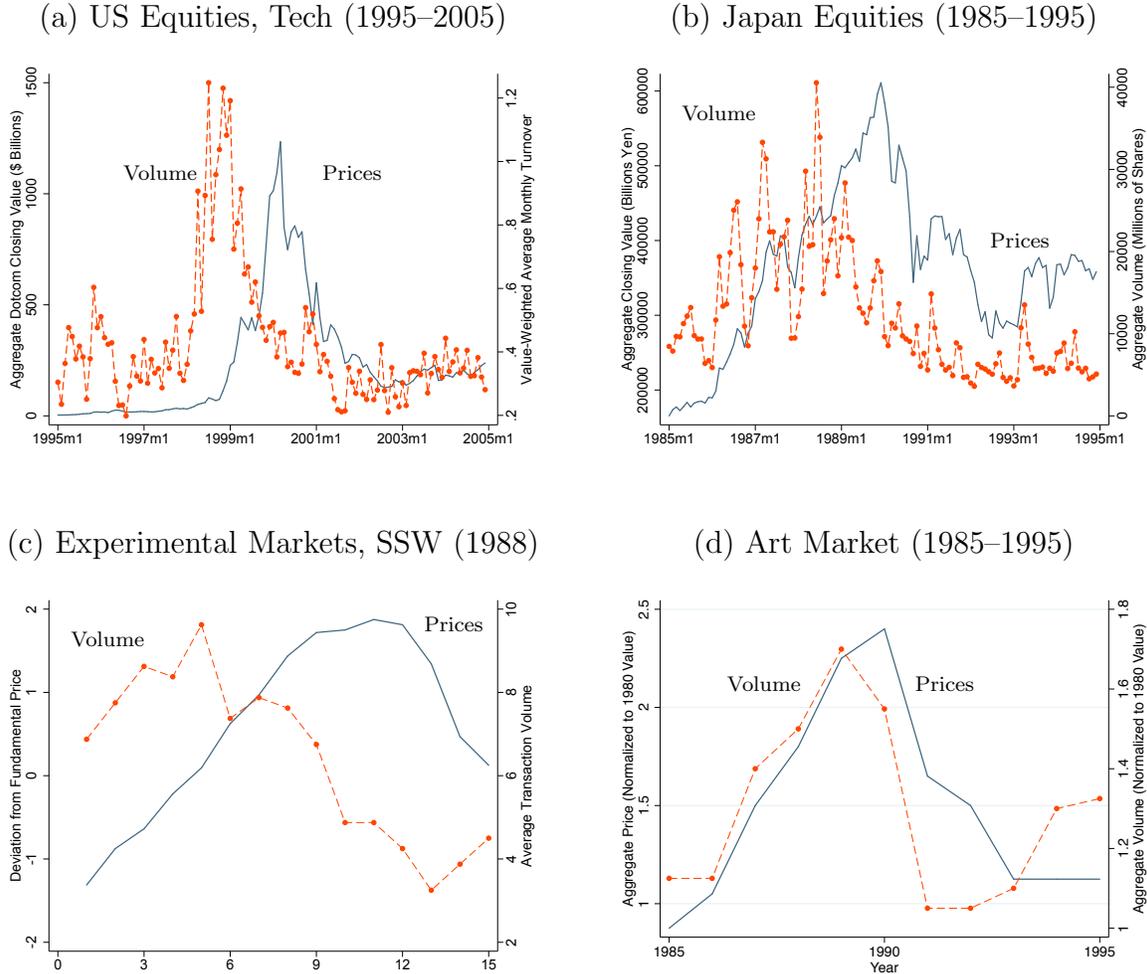
Notes: These figures display the dynamic relationship between prices and transaction volume for four distinct bubble episodes: (a) the 1995–2005 market in technology stocks, (b) the 1985–1995 Japanese stock market, (c) the bubbles in experimental asset markets, and (d) the 1985–1995 bubble in the Postwar art market. Panel (a) data come from CRSP and cover the Dotcom sample in Ofek and Richardson (2003). For prices, we plot aggregate Dotcom market capitalization. For volume, we plot average monthly turnover (shares traded/shares outstanding), weighted by market cap. Panel (b) data come from the Tokyo stock exchange online archive and cover all first- and second-tier (i.e., large and micro-cap) stocks. For volume, we plot total shares traded per month (shares-outstanding data are not available). For prices, we plot aggregate market capitalization. Panel (c) data were manually entered from the published Smith, Suchanek and Williams (1988) manuscript and cover all eight experiments that include a price boom and bust (IDs are 16, 17, 18, 26, 124xxf, 39xsf, 41f, 36xx). For prices, we plot average deviations from fundamental value. For volume, we plot the average number of trades. Panel (d) data come from Figure 1 of the working paper version of Penasse and Renneboog (2016) and cover aggregate art prices and transaction volume from auction houses for paintings and works on paper for more than 10,000 artists.

TABLE 1

Within-MSA Role of Short-Holding-Period and
Non-Occupant Volume Growth for Total Volume Growth (2000–2005)

	Holding Period			Occupancy Status		
	%Δ Short Volume	%Δ Long Volume	Δ Short/ Total Volume	%Δ Non-Occupant Volume	%Δ Occupant Volume	Δ Non-Occupant/ Total Volume
%Δ Total Volume	1.517***	0.936***	0.241***	1.525***	1.044***	0.356***
	(0.104)	(0.020)	(0.018)	(0.094)	(0.093)	(0.027)
R^2	0.57	0.95	0.77	0.17	0.48	0.75
Observations	6,763	6,763	6,763	5,597	5,597	5,597

Notes: This table reports estimates of the quantitative importance of short-holding-period and non-occupant volume in accounting for the increase in total volume between 2000 and 2005 at the ZIP-code level. We assign each transaction to a Census ZIP Code Tabulation Area (ZCTA) using the postal ZIP code of the property and a ZCTA-to-ZIP code crosswalk file provided by the Missouri Census Data Center. Each column reports estimates from a separate regression of the change in a given component of volume on the percent change in total volume in the ZCTA. All specifications include MSA fixed effects. A short holding period is defined as any holding period less than three years. Occupancy of the buyer is identified using information on the mailing address of the property as described in the text. In columns 3 and 6, we divide the level change in short-holding-period and non-occupant volume by total volume in the ZCTA in 2000. All other changes are expressed as percent changes from the 2000 level for the indicated type of transaction. ZCTAs are weighted according to their total volume in 2000. To eliminate the influence of outliers, all specifications drop the 99.9th and 0.1st percentile of the left- and right-hand-side variables. Standard errors are reported in parentheses and are clustered at the MSA level. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

TABLE 2

Initial Volume and the Magnitude of the Housing Boom and Bust

	Percentage Change in Prices			
	2000–Peak	2000–2006	Peak–Trough	2006–Trough
Existing Sales/Stock (2000)	14.626***	14.477***	−4.325***	−4.677***
	(3.813)	(3.739)	(1.172)	(1.259)
R^2	0.11	0.10	0.12	0.12
Observations	115	115	115	115

Notes: This table reports estimates of the cross-sectional relationship between the magnitude of the housing boom and bust and initial transaction volume at the MSA level. Each column reports estimates from a separate regression where the dependent variable is the percentage change in prices measured over the indicated horizon. Initial transaction volume is measured as total year 2000 existing home sales in each MSA scaled by the total number of housing units in the MSA as reported in the 2000 Census. House prices are measured using the monthly CoreLogic repeat-sales house price indicies. The price peak for each MSA is measured as the highest price recorded for that MSA prior to January, 2012. The trough is measured as the lowest price subsequent to either the month in which the peak occurred (column 3) or January, 2006 (column 4). In columns 1, 2, and 4, price changes are calculated using the January price level in 2000 and 2006. Heteroskedasticity robust standard errors are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

A Omitted Proofs of Mathematical Statements

A.1 Lemma 1

Lemma 1. *Given* (5), *the forward term structure of expectations is*

$$E(P_{t+\tau}) = P_t + \frac{(1 - m^\tau)\gamma}{1 - m}\omega_t, \tag{16}$$

where $m = \mu + (1 - \mu)\gamma < 1$.

Proof. For $t > 0$, the definition of ω_t may be written iteratively as $\omega_t = \mu\omega_{t-1} + (1 - \mu)(P_t - P_{t-1})$. Thus, for any $k > 0$ we have $E(\omega_{t+k}) = \mu E(\omega_{t+k-1}) + (1 - \mu)E(P_{t+k} - P_{t+k-1})$. From (5), this equation reduces to $E(\omega_{t+k}) = (\mu + (1 - \mu)\gamma)E(\omega_{t+k-1}) = mE(\omega_{t+k-1})$. Iterating backward gives $E(\omega_{t+k}) = m^k E(\omega_t) = m^k \omega_t$. Applying this formula and (5) gives

$$E(P_{t+\tau}) = P_t + \sum_{k=0}^{\tau-1} E(P_{t+k+1} - P_{t+k}) = P_t + \sum_{k=0}^{\tau-1} E(\gamma\omega_{t+k}) = P_t + \sum_{k=0}^{\tau-1} m^k \gamma\omega_0, \tag{17}$$

which reduces to (16). □

A.2 Proposition 1

Proof. Because potential buyers expect to sell with certainty at their expected list price upon becoming movers, $E(\mathcal{V}_{i,t+\tau}^m \mid \mathcal{F}_{i,t}) = E(P_{i,t+\tau} \mid \mathcal{F}_{i,t}) = E(P_{t+\tau} \mid \mathcal{F}_{i,t})$. Iterating (2) forward gives

$$\mathcal{V}_{i,t}^s = \frac{\rho\delta_i}{1 - \rho(1 - \lambda_i)} + \rho\lambda_i \sum_{\tau=1}^{\infty} \rho^{\tau-1}(1 - \lambda_i)^{\tau-1} E\left(P_{t+\tau} \mid \mathcal{F}_{i,t}\right). \tag{18}$$

The potential buyer's information set is $\mathcal{F}_{i,t} = \{P_{i,t}, \omega_{t-1}\}$. Because ω_{t-1} conveys no information about the price level, $E(P_t \mid \mathcal{F}_{i,t}) = P_{i,t}$. As shown in the proof of Lemma 1, $E(\omega_t \mid \omega_{t-1}) = m\omega_{t-1}$. Substituting this equation and (16) into (18) gives

$$\mathcal{V}_{i,t}^s = \frac{\rho\delta_i}{1 - \rho(1 - \lambda_i)} + \frac{\rho\lambda_i P_{i,t}}{1 - \rho(1 - \lambda_i)} + \frac{m\gamma\lambda_i\rho\omega_{t-1}}{(1 - \rho(1 - \lambda_i))(1 - m\rho(1 - \lambda_i))}. \tag{19}$$

The potential buyer buys when $\mathcal{V}_{i,t}^s \geq P_{i,t}$. With the substitution $\rho = 1/(1 + r)$, this inequality reduces to

$$\delta_i \geq rP_{i,t} - \frac{m\gamma\lambda_i(1 + r)\omega_{t-1}}{1 + r - m(1 - \lambda_i)}, \tag{20}$$

which matches (7). The coefficient $\phi(\lambda_i; \mu, \gamma, r)$ is positive because $\gamma, \lambda_i > 0$ and increasing in λ_i because $\partial\phi/\partial\lambda_i = m(1 + r)(1 + r - m)\gamma/(1 + r - m(1 - \lambda_i))^2 > 0$. □

A.3 Proposition 2

Proof. Define $U_t(P)$ to be the measure of movers listing below P plus the measure of informed movers listing at P. In equilibrium, all informed movers post the highest price that

51

guarantees sale, which is

$$P_t^n = \max\{P \mid U_t(P) \leq D_t(P)\}. \tag{21}$$

Hence, all informed movers post the same price in equilibrium. We solve for equilibrium in the three cases delineated in the proposition.

First, consider the boom, in which $D_t(P_t^x) > L_t$. If $P_t^n \leq P_t^x$, then $U_t(P) = L_t$ for $P \geq P_t^x$. Because $D_t(\cdot)$ is continuous and $L_t < D_t(P_t^x)$, there exists $P > P_t^x$ such that $U_t(P) = L_t \leq D_t(P)$. The existence of such P contradicts (21). Thus, the only possible equilibria in this case have $P_t^n > P_t^x$. In this case, $U_t(P) = (1-\beta)L_t$ for $P_t^x < P < P_t^n$, and $U_t(P) = L_t$ for $P \geq P_t^n$. Because $D_t(\cdot)$ is continuous and weakly decreasing, (21) holds if and only if $D_t(P_t^n) = L_t$ and $D_t(P)$ is strictly maximized for $P \geq P_t^n$ at $P = P_t^n$. From (8) and (9), $D_t(P)$ strictly decreases over the domain of P such that $D_t(P) < A_t$, which is the maximum of D_t. Thus, because $D_t(P_t^n) = L_t < D_t(P_t^x)$, $D_t(P_t^n) < A_t$, and $D_t(P)$ strictly decreases past P_t^n. In summary, the only equilibrium when $D_t(P_t^x) > L_t$ has $P_t^n = D_t^{-1}(L_t)$. Because a share β of movers post P_t^n and a share $1-\beta$ post P_t^x, $P_t = (1-\beta)P_t^x + \beta D_t^{-1}(L_t)$. The informed movers all sell, and because the extrapolative movers post a lower price, they all sell as well. As a result, $V_t = L_t$.

We next consider the quiet, in which $D_t(P_t^x) \in [\beta L_t, L_t]$. If $P_t^n < P_t^x$, then $U_t(P_t^x) = \beta L_t \leq D_t(P_t^x)$, contradicting (21). Therefore, $P_t^n \geq P_t^x$. We now consider two cases of the quiet.

In the first case, $D_t(P_t^x) = L_t = A_t$. This is the special case referenced in footnote 9. Define $P_t^{max} = \max\{P \mid D_t(P) = A_t\}$, which exists by the continuity of $D_t(P)$. If $P_t^n < P_t^{max}$, then $U_t(P_t^{max}) = L_t = A_t = D_t(P_t^{max})$, contradicting (21). If $P_t^n > P_t^{max}$, then $U_t(P_t^n) = L_t = A_t > D_t(P_t^n)$, contradicting (21). If $P_t^n = P_t^{max}$, then $U_t(P_t^n) = L_t = A_t = D_t(P_t^n)$, and $U_t(P) = L_t > D_t(P)$ for all $P > P_t^n = P_t^{max}$. Therefore, $P_t^n = P_t^{max}$ is the only solution to (21). In this case,

$$P_t = (1-\beta)P_t^x + \beta \max\{P \mid D_t(P) = A_t\}. \tag{22}$$

Because demand at the high price equals the number of movers, $D_t(P_t^n) = A_t = L_t$, all movers sell, giving $V_t = L_t$. This special case of the quiet resembles the boom in that all listings sell and informed movers may post a higher price than extrapolative movers. Unlike the boom, there is not any excess demand at P_t^x.

In the other quiet case, either $D_t(P_t^x) < L_t$ or $D_t(P_t^x) < A_t$. Given these conditions, $D_t(P) < L_t$ for all $P > P_t^x$, either because $D_t(P) \leq D_t(P_t^x) < L_t$ or because $D_t(P_t^x) < A_t$, which implies that D_t strictly decreases past P_t^x so that $D_t(P) < D_t(P_t^x) \leq L_t$. Suppose that $P_t^n > P_t^x$. Then $D_t(P_t^n) < L_t$, so $U_t(P_t^n) = L_t > D_t(P_t^n)$, contradicting (21). The only possible solution to (21) is then $P_t^n = P_t^x$, and this solution is valid because $U_t(P_t^x) = \beta L_t \leq D_t(P_t^x)$ while $U_t(P) = L_t > D_t(P)$ for all $P > P_t^x$. In summary, $P_t^n = P_t^x$ is the unique equilibrium. As a result, $P_t = P_t^x$. The total demand at this price equals $D_t(P_t^x)$, and as this demand is at most the number of listings L_t, all of the demand is served. Volume thus equals demand, $V_t = D_t(P_t^x)$, as claimed.

Finally, we consider the bust: $D_t(P_t^x) < \beta L_t$. Because $\beta L_t \leq \beta < A_t$, $D_t^{-1}(\beta L_t)$ exists. Suppose that $P_t^n < D_t^{-1}(\beta L_t)$. Then $U_t(D_t^{-1}(\beta L_t)) = \beta L_t = D_t(D_t^{-1}(\beta L_t))$, contradicting (21). Suppose that $P_t^n > D_t^{-1}(\beta L_t)$. Then $U_t(P_t^n) \geq \beta L_t > D_t(P_t^n)$, contradicting (21). If

$P_t^n = D_t^{-1}(\beta L_t)$, then $U_t(P_t^n) = \beta L_t = D_t(P_t^n)$, and $U_t(P) \geq \beta L_t > D_t(P)$ for all $P > P_t^n$, validating (21). Thus, $P_t^n = D_t^{-1}(\beta L_t)$ is the only equilibrium. In this case, demand at P_t^n equals supply, so there are no further transactions at P_t^x. As a result, $P_t = D_t^{-1}(\beta L_t)$ and $V_t = \beta L_t$. $\qquad\square$

B Calibration Details

B.1 Efficient steady state

Before solving for the efficient steady state, we first update Proposition 1 in light of the finite horizon now assumed.

Lemma 2. *In the finite horizon model, potential buyer i matched to price $P_{i,t}$ buys when*

$$\delta_i \geq rP_{i,t} - \phi_1(\lambda_i; \mu, \gamma, r, T-t)\omega_{t-1} - \phi_2(\lambda_i; \mu, \gamma, r, T-t)(P^T - P_{i,t}), \tag{23}$$

where

$$\phi_1(\lambda_i; \mu, \gamma, r, T-t) = \frac{\frac{\lambda_i m\gamma}{1-m}\left(1 - \left(\frac{1-\lambda_i}{1+r}\right)^{T-t-1} - \frac{m(r+\lambda_i)\left(1 - \left(\frac{m(1-\lambda_i)}{1+r}\right)^{T-t-1}\right)}{1+r-m(1-\lambda_i)}\right)}{1 - \left(\frac{1-\lambda_i}{1+r}\right)^{T-t}} \tag{24}$$

and

$$\phi_2(\lambda_i; \mu, \gamma, r, T-t) = \frac{(r+\lambda_i)(1-\lambda_i)^{T-t-1}}{(1+r)^{T-t} - (1-\lambda_i)^{T-t}}, \tag{25}$$

with $m = \mu + (1-\mu)\gamma < 1$.

Proof. Because all owners of housing at T receive P^T, $\mathcal{V}_{i,T}^s = \mathcal{V}_{i,T}^m = P^T$. For $t < T$, (2) continues to hold, so we substitute $\rho = 1/(1+r)$ and iterate forward to obtain

$$\mathcal{V}_{i,t}^s = \frac{\left(1 - \left(\frac{1-\lambda_i}{1+r}\right)^{T-t}\right)\delta_i}{r+\lambda_i} + \frac{(1-\lambda_i)^{T-t-1}P^T}{(1+r)^{T-t}} + \sum_{\tau=1}^{T-t-1} \frac{\lambda_i(1-\lambda_i)^{\tau-1}\mathrm{E}(P_{t+\tau} \mid \mathcal{F}_{i,t})}{(1+r)^\tau}. \tag{26}$$

As in the proof of Proposition 1, we substitute $\mathrm{E}(\omega_t \mid \omega_{t-1}) = m\omega_{t-1}$ and (16) into (26) to obtain

$$\mathcal{V}_{i,t}^s = \frac{\left(1 - \left(\frac{1-\lambda_i}{1+r}\right)^{T-t}\right)\delta_i}{r+\lambda_i} + \frac{(1-\lambda_i)^{T-t-1}(P^T - P_{i,t})}{(1+r)^{T-t}} + \frac{\left(\left(\frac{1-\lambda_i}{1+r}\right)^{T-t}r + \lambda_i\right)P_{i,t}}{r+\lambda_i}$$
$$+ \left(\frac{1 - \left(\frac{1-\lambda_i}{1+r}\right)^{T-t-1}}{r+\lambda_i} - \frac{m\left(1 - \left(\frac{m(1-\lambda_i)}{1+r}\right)^{T-t-1}\right)}{1+r-(1-\lambda_i)m}\right)\frac{\lambda_i m\gamma\omega_{t-1}}{1-m}. \tag{27}$$

Simplifying $\mathcal{V}_{i,t}^s \geq P_{i,t}$ yields Lemma 2. $\qquad\square$

As Lemma 2 shows, the finite horizon introduces two new features to the potential buyer decision rule. First, the rule now depends on $P^T - P_{i,t}$, the difference between the current list price and the eventual liquidation value. In the event that the potential buyer remains a stayer until T, this difference affects the expected cost of holding the housing. Second, the coefficients on ω_{t-1} and $P^T - P_{i,t}$ now depend on $T - t$, the remaining time until the final period. As $T \to \infty$, $\phi_1 \to \phi$ and $\phi_2 \to 0$, meaning that Lemma 2 reduces to Proposition 1.

Proposition 3 proves the existence of a unique efficient steady state and characterizes it.

Proposition 3. *If $A^i > (\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda)^{-1}$, then a unique efficient steady state exists, which is given by the liquidation value*

$$ P^T = \left(A^i \int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda \right)^{\frac{1}{\epsilon}} \frac{\delta_0}{r} \tag{28} $$

and initial conditions

$$ P_0 = \left(A^i \int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda \right)^{\frac{1}{\epsilon}} \frac{\delta_0}{r} \tag{29} $$

$$ \omega_0 = 0 \tag{30} $$

$$ I_0 = 0 \tag{31} $$

$$ S_0(\lambda) = \frac{\lambda^{-1} f_\lambda(\lambda)}{\int_0^1 (\lambda')^{-1} f_\lambda(\lambda') d\lambda'}. \tag{32} $$

In this steady state, volume and listings for $0 < t < T$ both equal

$$ L_t = V_t = \left(\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda \right)^{-1}. \tag{33} $$

Proof. We solve for variable values in an efficient steady state. Because variables do not depend on time, we sometimes drop time subscripts. The recursive equation for ω implies that $\omega_t = \mu \omega_{t-1} + (1 - \mu)(P_t - P_{t-1})$, or $\omega = \mu \omega$. Because $\mu < 1$, the only solution to this equation is $\omega = 0$. Eq. (5) implies that $P^x = P + \gamma \omega = P$. From 2, $P_t^x = P_t$ only when $P_t^x = P_t^n = P$. Hence, $P^n = P$.

We claim that $D(P^n) < A^i$. For a contradiction, assume not. Then $D(P^n) = A^i$, as the maximal value of demand is the number of potential buyers, A^i. In this case, demand from each λ type must also equal its maximum: $D(P^n, \lambda) = A^i f_\lambda(\lambda)$. Eq. (14), which implies that $S(\lambda) = (1 - \lambda)S(\lambda) + D(P^n, \lambda)$, thus gives $S(\lambda) = \lambda^{-1} A^i f_\lambda(\lambda)$. Because $I = 0$, end-of-period stayers comprise the entire housing stock, which has measure 1. Therefore, $1 = \int_0^1 S(\lambda) d\lambda = A^i \int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda$, implying that $A^i = (\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda)^{-1}$, a contradiction of the condition that $A^i > (\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda)^{-1}$.

Because $D(P^n) < A^i$, Lemma 2 implies that for some λ,

$$ D(P^n, \lambda) = A^i f_\lambda(\lambda) \delta_0^\epsilon \left(r P^n - \phi_2(\lambda; \mu, \gamma, r, T - t)(P^T - P^n) \right)^{-\epsilon}. \tag{34} $$

Because $\phi_2(\lambda; \mu, \gamma, r, T-t)$ depends on t, we must have $P^T = P^n$. It follows that $D(P^n, \lambda) = A^i f_\lambda(\lambda) \delta_0^\epsilon (r P^n)^{-\epsilon}$ for all λ. Substituting this equation for $D(P^n, \lambda)$ into (14) yields $S(\lambda) =$

$(1-\lambda)S(\lambda)+A^i f_\lambda(\lambda)\delta_0^\epsilon(rP^n)^{-\epsilon}$, or $S(\lambda)=A^i\delta_0^\epsilon(rP^n)^{-\epsilon}\lambda^{-1}f_\lambda(\lambda)$. Because $I=0$, $\int_0^1 S(\lambda)d\lambda=1$, the size of the housing stock. We may therefore solve for P^n to get

$$P^n=\left(A^i\int_0^1 \lambda^{-1}f_\lambda(\lambda)d\lambda\right)^{1/\epsilon}\frac{\delta_0}{r}, \tag{35}$$

which equals P^T and P as well, as we have already shown that $P=P^n=P^T$. We substitute this equation for P^n into $S(\lambda)=A\delta_m^\epsilon(rP^n)^{-\epsilon}\lambda^{-1}f_\lambda(\lambda)$ to obtain

$$S(\lambda)=\frac{\lambda^{-1}f_\lambda(\lambda)}{\int_0^1 (\lambda')^{-1}f_\lambda(\lambda')d\lambda'}.$$

We have proven that unique values of the initial and terminal conditions can give rise to an efficient steady state. To prove that the efficient steady state exists, we verify that P^n is optimal. By Proposition 2, $P^n=P^x$ if $D(P^x)\in[\beta L,L]$ and $D(P^x)<A^i$ (see footnote 9). By (10), the equation for listings, $L=\int_0^1 \lambda S(\lambda)d\lambda$, so

$$L=\left(\int_0^1 \lambda^{-1}f_\lambda(\lambda)d\lambda\right)^{-1}.$$

Plugging the formula for $P^x=P^n=P^T$ into (23) and integrating over λ yields

$$D(P^x)=\begin{cases}\left(\int_0^1 \lambda^{-1}f_\lambda(\lambda)d\lambda\right)^{-1} & \text{if } A^i>\left(\int_0^1 (\lambda)^{-1}f_\lambda(\lambda)d\lambda'\right)^{-1}\\ A^i & \text{if } A^i\leq\left(\int_0^1 (\lambda)^{-1}f_\lambda(\lambda)d\lambda\right)^{-1}.\end{cases}$$

Because $A^i>(\int_0^1 \lambda^{-1}f_\lambda(\lambda)d\lambda)^{-1}$, $D(P^x)=L<A^i$, verifying seller optimality and the existence of the efficient steady state. Because $I=L-V=0$, $L=V$, as desired. □

B.2 Parameter values

f_λ

Each March, as part of the Investment and Vacation Home Buyers Survey, the National Association of Realtors (NAR) surveys a nationally representative sample of around 2,000 individuals who purchased a home in the previous year. The survey asks respondents to report the type of home purchased (investment property, primary residence, or vacation property) as well as the "length of time [the] buyer plans to own [the] property." Data on expected holding times and the share of purchases of each type are available for 2008–2015.

The data on expected holding times are binned. For each property type, we see the share of recent buyers whose expected holding time is in each of $[0,1)$, $[1,3)$, $[3,6)$, $[6,11)$, and $[11,\infty)$ years. We also see the share of recent buyers of each type who have already sold the property or do not know their expected holding time. We drop recent buyers who do not know their expected holding time and reclassify buyers who have already sold the property as having an expected holding time in $[0,1)$. Using the share of sales to each property type

in each year, we calculate the distribution of expected holding times over the bins for each year and then take an equal-weighted average across years to obtain a single distribution across the bins. The resulting probability weights across the bins are (rounded) 0.1, 0.1, 0.2, 0.2, and 0.4.

We then map each bin to a value of λ as follows. When $T = \infty$, the expected holding time of potential buyer i, conditional on the expectation of selling immediately upon becoming a mover, equals $\sum_{\tau=1}^{\infty} \tau \lambda_i (1-\lambda_i)^{\tau-1} = 1/\lambda_i$. Given this relation, we map an empirical expected holding time to a value of λ by taking its inverse. For each bin in the NAR data, we select an expected holding time within it, giving us expected holding times of 0.25, 1.5, 4, 8, and 20 years, or 1, 6, 20, 32, and 80 quarterly periods. The resulting values of λ are (rounded) 1, 0.17, 0.06, 0.03, and 0.01.

μ and γ

We solve for the effects estimated by Armona, Fuster and Zafar (2016) within our expectations framework, giving us two equations that pin down μ and γ. Suppose that house price growth perceived in the prior year increases by Δ. We interpret this change as a uniform increase of $\Delta/4$ over each of the prior four quarterly periods. Then ω_t, which can be written as $\omega_t = (1-\mu)((P_t-P_{t-1})+\mu(P_{t-1}-P_{t-2})+\mu^2(P_{t-2}-P_{t-3})+\mu^3(P_{t-3}-P_{t-4}))+\mu^4\omega_{t-4}$, increases by $\Delta(1-\mu^4)/4$. By (6), the resulting change in expectations for the cumulative increase over the next year is $g(\mu,\gamma,4)\Delta(1-\mu^4)/4$, and the resulting change in expectations for the annualized increase over the next two-to-five years is $(1/4)(g(\mu,\gamma,20) - g(\mu,\gamma,4))\Delta(1-\mu^4)/4$. Substituting the expression for $g(\cdot,\cdot,\cdot)$ and equating the results with the empirical estimates from Armona, Fuster and Zafar (2016) gives the following two equations:

$$0.226 = \frac{(1 - m^4)(1 - \mu^4)\gamma}{4(1 - m)} \tag{36}$$

$$0.047 = \frac{m^4(1 - m^{16})(1 - \mu^4)\gamma}{16(1 - m)}, \tag{37}$$

where $m = \mu + (1 - \mu)\gamma$. To solve these equations, we first take the quotient, which gives us a single equation for m that can be written as $0.83 = m^4(1 + m^4)(1 + m^8)$. The right side strictly increases from 0 to 1 for $0 < m < 1$, so the equation has a unique solution for m, which we calculate numerically to be $m = 0.83$ (rounded). We substitute this solution and $\gamma = (m - \mu)/(1 - \mu)$ into (36) to obtain $0.83(1 - m)(1 - \mu)/(m - \mu) = (1 - m^4)(1 - \mu^4)$. For $0 < \mu < m$, the left increases while the right decreases, so the equation has at most one solution in μ. We numerically calculate this solution to be $\mu = 0.71$. We then use m and μ to calculate $\gamma = 0.40$.

δ_0, A^i, and A^f

From Proposition 3,

$$1 = P_0 = \left(A^i \int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda \right)^{\frac{1}{\epsilon}} \frac{\delta_0}{r} \tag{38}$$

and

$$1.6 = P^T = \left(A^f \int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda \right)^{\frac{1}{\epsilon}} \frac{\delta_0}{r}. \tag{39}$$

Dividing these equations gives $A^f/A^i = 1.6^\epsilon$. Because we have already calculated or are assuming certain values for f_λ, ϵ, and r, we have two equations in the three unknowns A^f, A^i, and δ_0 that have a solution once one is specified. We experiment with different values of A^i to find a value producing a simulation in which prices and volume converge over the finite horizon. Doing so gives $A^i = 0.2$, meaning that in the initial steady state, arriving potential buyers have a measure equal to 20% of the city's housing stock. This value implies that $A^f = 0.27$ and $\delta_0 = 0.00042$. The requirements that A^i and A^f exceed β and $(\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda)^{-1}$ are satisfied because $\beta = 0.17$ and $(\int_0^1 \lambda^{-1} f_\lambda(\lambda) d\lambda)^{-1} = 0.02$.

B.3   Holding cost bound

Here we calculate the bound on k given in Section 1 in equilibrium and calculate the bound in the calibration. For $1 \leq t < T$, $\pi_t^{-1}(1)$ equals the highest price that guarantees sale. As the proof of Proposition 2 shows, this price is P_t^n. The proof further shows that informed movers sell at this price. At $t = T$, informed movers receive the liquidation value P^T. Thus, $\mathcal{V}_{i,t}^{nm} = P_t^n$, where we have formally defined $P_T^n = P^T$. Because the model is deterministic, informed movers foresee the future perfectly, meaning that $\mathrm{E}(\mathcal{V}_{i,t+1}^{nm} \mid \mathcal{F}_{i,t}) = \mathcal{V}_{i,t+1}^{nm} = P_{t+1}^n$. The parameter restriction on k in Section 1 becomes

$$k > P_{t+1}^n - \rho^{-1} P_t^n \tag{40}$$

for $1 \leq t < T$. Thus we require that k be above the maximum value of this difference over the possible values of t. Using the values of P_t^n we calculate in the simulation (that appear in panel b of figure 2), we compute the maximal value of this difference to be 0.41.

For Online Publication

C   Data Appendix

To conduct our empirical analysis we make use of a transaction-level data set containing detailed information on individual home sales taking place throughout the US between 1995–2014. The raw data was purchased from CoreLogic and is sourced from publicly available tax assessment and deeds records maintained by local county governments.

Selecting Geographies

To select our sample, we first focus on a set of counties that have consistent data coverage going back to 1995 and which, together, constitute a majority of the housing stock in their respective MSAs. In particular, to be included in our sample a county must have at least one "arms length" transaction with a non-negative price and non-missing date in each quarter from 1995q1 to 2014q4.[21] Starting with this subset of counties, we then further drop any MSA for which the counties in this list make up less than 75 percent of the total owner-occupied housing stock for the MSA as measured by the 2010 Census. This leaves us with a final set of 250 counties belonging to a total of 115 MSAs. These MSAs are listed below in Table A1 along with the percentage of the housing stock that is represented by the 250 counties for which we have good coverage. Throughout the paper, when we refer to counts of transactions in an MSA we are referring to the portion of the MSA that is accounted for by these counties.

Selecting Transactions

Within this set of MSAs, we start with the full sample of all arms length transactions of single family, condo, or duplex properties and impose the following set of filters to ensure that our final set of transactions provides an accurate measure of aggregate transaction volume over the course of the sample period:

1. Drop transactions that are not uniquely identified using CoreLogic's transaction ID.

2. Drop transactions with non-positive prices.

3. Drop transactions that are recorded by CoreLogic as nominal transfers between banks or other financial institutions as part of a foreclosure process.

4. Drop transactions that appear to be clear duplicates, identified as follows:

    (a) If a set of transactions has an identical buyer, seller, and transaction price but are recorded on different dates, keep only the earliest recorded transaction in the set.

    (b) If the same property transacts multiple times on the same day at the same price keep only one transaction in the set.

---

[21]We rely on CoreLogic's internal transaction-type categorization to determine whether a transaction occurred at arms length.

5. If more than 10 transactions between the same buyer and seller at the same price are recorded on the same day, drop all such transactions. These transactions appear to be sales of large subdivided plots of vacant land where a separate transaction is recorded for each individual parcel but the recorded price represents the price of the entire subdivision.

6. Drop sales of vacant land parcels in MSAs where the CoreLogic data includes such sales.[22] We define a vacant land sale to be any transaction where the sale occurs a year or more before the property was built.

Table A2 shows the number of transactions that are dropped from our sample at each stage of this process as well as the final number of transactions included in our full analysis sample.

Identifying Occupant and Non-Occupant Buyers

We identify non-occupant buyers using differences between the mailing addresses listed by the buyer on the purchase deed and the actual physical address of the property itself. In most cases, these differences are identified using the house numbers from each address. In particular, if both the mailing address and the property address have a non-missing house number then we tag any instance in which these numbers are not equal as a non-occupant purchase and any instance in which they are equal as occupant purchases. In cases where the mailing address property number is missing we also tag buyers as non-occupants if both the mailing address and property address street names are non-missing and differ from one another. Typically, this will pick up cases where the mailing address provided by the buyer is a PO Box. In all other cases, we tag the transaction as having an unknown occupancy status.

Restricting the Sample for the Non-Occupant Analysis

Our analysis of non-occupant buyers focuses on the growth of the number of purchases by these individuals between 2000 and 2005. To be sure that this growth is not due to changes in the way mailing addresses are coded by the counties comprising the MSAs in our sample, for the non-occupant buyer analysis we keep only MSAs for which we are confident such changes do not occur between 2000 and 2005. In particular, we first drop any MSA in which the share of transactions in any one year between 2000 and 2005 with unknown occupancy status exceeds 0.5. Of the remaining MSAs, we then drop those for which the increase in the number of non-occupant purchases between any year and the next exceeds 150%, with the possible base years being those between 2000 and 2005.[23] The MSAs that remain after these two filters are marked with an "x" in Table A1.

Identifying New Construction Sales

In Table 2 and Figure 10 we correlate the size of the 2000–2005 house price boom with the level of initial volume relative to the total housing stock in 2000. In performing this exercise,

---

[22]MSAs are flagged as including vacant land sales if more then 5 percent of the sales in the MSA occur more then two years before the year in which the property was built.

[23]This step drops only Chicago-Naperville-Elgin, IL-IN-WI.

we omit new construction sales from the calculation of transaction volume. This is done to ensure that our calculations are not simply picking up cross-sectional differences in supply elasticity or new construction rates across markets.

To identify sales of newly constructed homes, we start with the internal CoreLogic new construction flag and make several modifications to pick up transactions that may not be captured by this flag. CoreLogic identifies new construction sales primarily using the name of the seller on the transaction (e.g. "PULTE HOMES" or "ROCKPORT DEV CORP"), but it is unclear whether their list of home builders is updated dynamically or maintained consistently across local markets. To ensure consistency, we begin by pulling the complete list of all seller names that are ever identified with a new construction sale as defined by CoreLogic. Starting with this list of sellers, we tag any transaction for which the seller is in this list, the buyer is a human being, and the transaction is not coded as a foreclosure sale by CoreLogic as a new construction sale. We use the parsing of the buyer name field to distinguish between human and non-human buyers (e.g. LLCs or financial institutions). Human buyers have a fully parsed name that is separated into individual first and name fields whereas non-human buyer's names are contained entirely within the first name field.

This approach will identify all new construction sales provided that the seller name is recognized by CoreLogic as the name of a homebuilder. However, many new construction sales may be hard to identify simply using the name of the seller. We therefore augment this definition using information on the date of the transaction and the year that the property was built. In particular, if a property was not already assigned a new construction sale using the builder name, then we search for sales of that property that occur within one year of the year that the property was built and record the earliest of such transactions as a new construction sale.

Finally, for properties that are not assigned a new construction sale using either of the two above methods, we also look to see if there were any construction loans recorded against the property in the deeds records. If so, we assign the earliest transaction to have occurred within three years of the earliest construction loan as a new construction sale. We use a three-year window to allow for a time lag between the origination of the construction loan and the actual date that the property was sold. Construction loans are identified using CoreLogic's internal deed and mortgage type codes.

List of Metropolitan Statistical Areas Included in the Analysis Sample

Metropolitan Statistical Area	Share of Housing Stock Represented	Included in Non-Occupant Analysis	Metropolitan Statistical Area	Share of Housing Stock Represented	Included in Non-Occupant Analysis
Akron, OH	1.00	x	New York-Newark-Jersey City, NY-NJ-PA	0.97	x
Ann Arbor, MI	1.00	x	North Port-Sarasota-Bradenton, FL	1.00	x
Atlanta-Sandy Springs-Roswell, GA	0.80		Norwich-New London, CT	1.00	
Atlantic City-Hammonton, NJ	1.00	x	Ocala, FL	1.00	x
Bakersfield, CA	1.00	x	Ocean City, NJ	1.00	x
Baltimore-Columbia-Towson, MD	1.00	x	Olympia-Tumwater, WA	1.00	x
Barnstable Town, MA	1.00		Orlando-Kissimmee-Sanford, FL	1.00	x
Bellingham, WA	1.00	x	Oxnard-Thousand Oaks-Ventura, CA	1.00	x
Bend-Redmond, OR	1.00	x	Palm Bay-Melbourne-Titusville, FL	1.00	x
Boston-Cambridge-Newton, MA-NH	0.89		Pensacola-Ferry Pass-Brent, FL	1.00	x
Boulder, CO	1.00	x	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1.00	x
Bremerton-Silverdale, WA	1.00	x	Phoenix-Mesa-Scottsdale, AZ	1.00	x
Bridgeport-Stamford-Norwalk, CT	1.00		Pittsfield, MA	1.00	
Buffalo-Cheektowaga-Niagara Falls, NY	0.80	x	Portland-Vancouver-Hillsboro, OR-WA	0.97	x
California-Lexington Park, MD	1.00	x	Port St. Lucie, FL	1.00	x
Canton-Massillon, OH	0.92	x	Prescott, AZ	1.00	x
Cape Coral-Fort Myers, FL	1.00	x	Providence-Warwick, RI-MA	0.78	
Champaign-Urbana, IL	0.82	x	Punta Gorda, FL	1.00	x
Charleston-North Charleston, SC	0.79	x	Raleigh, NC	0.78	x
Chicago-Naperville-Elgin, IL-IN-WI	0.90		Reading, PA	1.00	x
Chico, CA	1.00	x	Redding, CA	1.00	x
Cincinnati, OH-KY-IN	0.78	x	Reno, NV	0.99	x
Cleveland-Elyria, OH	1.00	x	Riverside-San Bernardino-Ontario, CA	1.00	x
Colorado Springs, CO	0.95	x	Rockford, IL	0.84	x
Crestview-Fort Walton Beach-Destin, FL	1.00	x	Sacramento–Roseville–Arden-Arcade, CA	1.00	x
Dallas-Fort Worth-Arlington, TX	0.85	x	Salem, OR	0.79	x
Dayton, OH	0.86	x	Salinas, CA	1.00	x
Deltona-Daytona Beach-Ormond Beach, FL	1.00	x	San Diego-Carlsbad, CA	1.00	x
Denver-Aurora-Lakewood, CO	0.95	x	San Francisco-Oakland-Hayward, CA	1.00	x
El Centro, CA	1.00	x	San Jose-Sunnyvale-Santa Clara, CA	1.00	x
El Paso, TX	0.99	x	Santa Cruz-Watsonville, CA	1.00	x
Elmira, NY	1.00	x	San Luis Obispo-Paso Robles-Arroyo Grande, CA	1.00	x
Erie, PA	1.00	x	Santa Maria-Santa Barbara, CA	1.00	x
Eugene, OR	1.00	x	Santa Rosa, CA	1.00	x
Flagstaff, AZ	1.00	x	Seattle-Tacoma-Bellevue, WA	1.00	x
Fort Collins, CO	1.00	x	Sebastian-Vero Beach, FL	1.00	x
Fresno, CA	1.00	x	Sebring, FL	1.00	x
Gainesville, FL	0.91	x	Sierra Vista-Douglas, AZ	1.00	x
Gainesville, GA	1.00		Spokane-Spokane Valley, WA	0.87	x
Hanford-Corcoran, CA	1.00	x	Springfield, IL	0.93	x
Hartford-West Hartford-East Hartford, CT	1.00		Springfield, MA	1.00	
Homosassa Springs, FL	1.00	x	Springfield, OH	1.00	x
Ithaca, NY	1.00	x	Stockton-Lodi, CA	1.00	x
Jacksonville, FL	0.98	x	Tampa-St. Petersburg-Clearwater, FL	1.00	x
Kahului-Wailuku-Lahaina, HI	1.00	x	The Villages, FL	1.00	x
Kingston, NY	1.00	x	Toledo, OH	0.92	x
Lake Havasu City-Kingman, AZ	1.00	x	Trenton, NJ	1.00	x
Lakeland-Winter Haven, FL	1.00	x	Tucson, AZ	1.00	x
Lancaster, PA	1.00	x	Urban Honolulu, HI	1.00	x
Las Vegas-Henderson-Paradise, NV	1.00	x	Vallejo-Fairfield, CA	1.00	x
Los Angeles-Long Beach-Anaheim, CA	1.00	x	Vineland-Bridgeton, NJ	1.00	x
Madera, CA	1.00	x	Visalia-Porterville, CA	1.00	x
Merced, CA	1.00	x	Washington-Arlington-Alexandria, DC-VA-MD-WV	0.95	x
Miami-Fort Lauderdale-West Palm Beach, FL	1.00	x	Worcester, MA-CT	1.00	
Modesto, CA	1.00	x	Youngstown-Warren-Boardman, OH-PA	0.80	x
Napa, CA	1.00	x	Yuba City, CA	1.00	x
Naples-Immokalee-Marco Island, FL	1.00	x	Yuma, AZ	1.00	x
New Haven-Milford, CT	1.00				

Notes: This table lists the Metropolitan Statistical Areas that are included in the final analysis sample along with the share of the total 2010 owner-occupied housing stock for each MSA that is represented by the subset of counties for which CoreLogic has consistent data coverage back to 1995.

TABLE A2

Number of Transactions Dropped During Sample Selection

Original Number of Transactions	57,668,026
Dropped: Non-unique CoreLogic ID	50
Dropped: Non-positive price	3,309,100
Dropped: Nominal foreclosure transfer	531,786
Dropped: Duplicate transaction	609,756
Dropped: Subdivision sale	1,304,920
Dropped: Vacant lot	831,774
Final Number of Transactions	51,080,640

Notes: This table shows the number of transactions dropped at each stage of our sample-selection procedure.