

NBER WORKING PAPER SERIES

SPECULATIVE DYNAMICS OF PRICES AND VOLUME

Anthony A. DeFusco  
Charles G. Nathanson  
Eric Zwick

Working Paper 23449  
<http://www.nber.org/papers/w23449>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
May 2017

We thank Stefano Giglio, Edward Glaeser, Sam Hanson, Stijn Van Nieuwerburgh, Alp Simsek, Johannes Stroebel, Richard Thaler, and Rob Vishny for helpful comments. William Cassidy, Jessica Henderson, Saul Ioffie, Harshil Sahai, and Iris Song provided excellent research assistance. DeFusco and Nathanson thank the Guthrie Center for Real Estate Research for financial support, and Zwick gratefully acknowledges financial support from the Fama Miller Center and Booth School of Business at the University of Chicago. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Anthony A. DeFusco, Charles G. Nathanson, and Eric Zwick. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Speculative Dynamics of Prices and Volume

Anthony A. DeFusco, Charles G. Nathanson, and Eric Zwick

NBER Working Paper No. 23449

May 2017

JEL No. E32,G02,G12,R3

**ABSTRACT**

We present a dynamic theory of prices and volume in asset bubbles. In our framework, predictable price increases endogenously attract short-term investors more strongly than long-term investors. Short-term investors 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 US 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.

Anthony A. DeFusco  
Northwestern University  
Kellogg School of Management  
2211 Campus Drive  
Office 4463  
Evanston, Illinois 60208  
anthony.defusco@kellogg.northwestern.edu

Eric Zwick  
Booth School of Business  
University of Chicago  
5807 South Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
ezwick@chicagobooth.edu

Charles G. Nathanson  
Kellogg School of Management  
Northwestern University  
2211 Campus Drive, Room 4479  
Evanston, IL 60208-0898  
nathanson@kellogg.northwestern.edu

The role of speculation in driving asset prices has long been debated among economists (Keynes, 1936; Fama, 1970; Shiller, 1981; Black, 1986).<sup>1</sup> Modern empirical work on “speculative dynamics” begins with Cutler et al. (1991), who document short-run momentum and long-run reversals in the prices of many diverse assets. These patterns are especially strong during asset bubbles, which have drawn attention due to their frenzied activity and subsequent social costs (Kindleberger, 1978; Shiller, 2005; Glaeser, 2013). Several distinct theories have been offered to explain these asset pricing facts.<sup>2</sup>

This paper explores a less studied feature of asset bubbles—the speculative dynamics of volume. Large swings in transaction volume consistently accompany price cycles (Stein, 1995; Genesove and Mayer, 2001; Hong and Stein, 2007), yet many theories of bubbles ignore implications for volume. As Cochrane (2011) writes:

Every asset price “bubble”...has coincided with a similar trading frenzy, from Dutch tulips in 1620 to Miami condos in 2006. ... Is this a coincidence? Do prices rise and fall for other reasons, and large trading volume follows, with no effect on price? Or is the high price...explained at least in part by the huge volume? ... To make this a deep theory, we must answer why people trade so much.

Figure 1 plots time-series patterns in prices and volume for four distinct episodes: the 2000–2011 US housing market, the 1995–2005 market in technology stocks, the experimental bubbles studied by Smith et al. (1988), and the 1985–1995 Japanese stock market. During these episodes, prices and volume comove strongly. The figures also reveal a more nuanced feature of the data: in each case, volume peaks well before prices. Improving our understanding of bubbles requires focus on the complex, joint dynamics of prices and volume.

To take up this challenge, we first present a simple model of the joint speculative dynamics of prices and volume during bubbles. Following past work, the model features extrapolative expectations: investors expect prices to increase after past increases.<sup>3</sup> The model departs from past work in two ways. First, instead of featuring the standard dichotomy between feedback traders and rational arbitrageurs, the model focuses on investors who differ not in their beliefs but in their expected investment horizons. Some buyers plan to sell after one

---

<sup>1</sup>Harrison and Kreps (1978, p. 323) define speculation in the following way: “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.”

<sup>2</sup>These theories include Cutler et al. (1990), De Long et al. (1990), Barberis et al. (1998), Daniel et al. (1998), Hong and Stein (1999), Abreu and Brunnermeier (2003), Pástor and Veronesi (2006), Pástor and Veronesi (2009), Piazzesi and Schneider (2009), and Burnside et al. (2016).

<sup>3</sup>Key models with extrapolative expectations include Cutler et al. (1990), De Long et al. (1990), Hong and Stein (1999), Barberis and Shleifer (2003), Barberis et al. (2015), and Glaeser and Nathanson (2016).

year, while others plan to hold for many years. The second departure involves specifying a term structure for extrapolation. In particular, extrapolation declines with the forecast horizon, so that short-run expectations display more sensitivity than long-run expectations to past prices. When this term structure holds, past price growth disproportionately attracts short-horizon investors, who in turn generate excess volume when they sell.

These modeling innovations benefit from much empirical support. Prior work estimating extrapolative future expectations finds that short-run expectations display more sensitivity to past price changes than long-run expectations (Graham and Harvey, 2003; Vissing-Jorgensen, 2004; Armona et al., 2016). In survey evidence from the National Association of Realtors, expected holding times vary considerably across buyers in the housing market. Furthermore, the share of respondents reporting an expected holding time of less than 3 years comoves strongly with recent house price growth.

We model a housing market populated by extrapolative investors with heterogeneous horizons. Due to the generality of the stylized facts presented above, we abstract from several special features of the housing market, such as debt and new construction, in order to ease comparison with other asset markets. We focus on the housing market because data availability allows us to test directly the model’s predictions about the composition of buyers and sellers over the course of a bubble episode. In the model, potential buyers arrive each instant and decide whether to buy a house. If they buy, they must hold the house for some period, after which they are free to sell. The expected duration of this period and the flow utility received during it vary across potential buyers. Potential buyers expect to sell immediately upon the period’s expiration at the prevailing price. As in prior work (e.g., Hong and Stein, 1999), the asset price reacts sluggishly to changes in the number of potential buyers who wish to buy. This “price stickiness” combines with extrapolative expectations to generate positive feedback between price growth and demand that causes bubbles.

The model is analytically tractable, allowing us to characterize how prices, volume, and the composition of buyers respond to a one-time demand shock. We partition time following the shock into three epochs: a *boom* in which prices, volume, and the short-term buyer share rise; a *quiet* in which prices continue to rise while volume and the short-term buyer share fall; and a *bust* in which prices fall on low volume. This partition implies a lead–lag relationship between volume and prices that is consistent with the data in Figure 1 and absent from other

theories in which trading volume comoves with prices.<sup>4</sup> In our model, price growth during the boom attracts short-term buyers who raise volume by selling quickly. During the quiet, demand from short-term investors falls due to a slowdown in price growth. This decline in demand causes a drop in volume and makes it difficult for previous cohorts of short-term buyers to sell their houses.

We prove that the increases in prices and volume during this cycle are larger when the frequency of potential buyers with short horizons is greater. This finding ties together prices and volume during bubbles by demonstrating that a single factor is responsible for movements in both. Using the empirical literature on extrapolative expectations and the survey evidence on expected holding times, we calibrate our model and find that the marginal effect of short-term potential buyers on the price and volume booms is quantitatively large and first-order relevant for explaining aggregate price and volume dynamics.

The second part of the paper documents new facts on the composition of buyers and sellers during the recent US housing bubble using transaction-level data between 1995 and 2014 for 115 metropolitan statistical areas (MSAs) that represent 48% of the US housing stock. While we organize the analysis around the model’s key predictions, these facts provide a model-free depiction of the joint dynamics of prices and volume that highlights the role of speculation in driving the bubble.

We present three 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, as predicted by the model, the short-term buyer share falls during the 2005–2006 quiet and subsequent bust. Finally, 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.

We then present three facts about the the joint price–volume relationship. First, the lead–lag relationship between national prices and volume holds also within MSAs, with MSA-level prices correlated most strongly with a 24-month lag of volume. Second, the

---

<sup>4</sup>Alternative theories that generate comovement without a clear lead–lag relationship have focused on credit constraints (Stein, 1995; Ortalo-Magné and Rady, 2006), loss aversion (Genesove and Mayer, 2001), or disagreement (Scheinkman and Xiong, 2003; Piazzesi and Schneider, 2009).

2000–2005 rise in volume is driven entirely by new listings, whereas the 2005–2006 decline in volume is due primarily to a slowdown in the rate at which listings sell. This dichotomy matches the model’s predictions about the differential drivers of volume during the boom and quiet. Finally, the 2000–2011 house price cycle was larger in MSAs where the 2000 level of existing sales as a share of the housing stock was greater. As shown 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.

## Related Literature

Our theoretical approach builds on the strand of papers beginning with Cutler et al. (1990) and De Long et al. (1990) that explain asset price fluctuations using the interaction of fundamental and feedback traders, who play roles similar to long-run and short-run investors in our framework. The most closely related example is Hong and Stein (1999), who offer a model of price underreaction and overshooting in which price stickiness arises due to the slow diffusion of news about future dividends. Unlike our paper, this literature considers neither the joint dynamics of volume and prices nor the role of heterogeneous holding times.

A literature going back to Amihud and Mendelson (1986) emphasizes the importance of heterogeneous investment horizons for financial markets. Henderson and Ioannides (1989) and Kan (1999) establish this heterogeneity in the housing market by showing that demographics like age and education predict both the actual and self-reported expectation of the amount of time until moving. Edelstein and Qian (2014) demonstrate that the share of short-term buyers, as measured by such predictive characteristics, comoves with house prices between 1970 and 2005.<sup>5</sup> Our paper draws out the implications of this underlying heterogeneity for prices and volume in a bubble.

Our approach differs from but is complementary to models in which disagreement increases both prices and volume in financial markets (Hong and Stein, 2007; Simsek, 2013; Daniel and Hirshleifer, 2015).<sup>6</sup> The seminal disagreement papers (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003) link prices and volume through comparative statics and hence

---

<sup>5</sup>Like us, Edelstein and Qian (2014) present a model in which short-term investors are more attracted to expected capital gains than long-term investors. Unlike us, they do not solve for dynamics and do not derive explicit results about the effect of short-term investors on prices and volume.

<sup>6</sup>Bailey et al. (2016) present empirical evidence that house prices and volume are higher in US counties with larger disagreement about future house price growth. They measure an individual’s house price growth expectations using the past house price growth experienced by the people in that individual’s social network.

cannot explain the dynamics of volume over the course of a bubble episode in a single asset. Penasse and Renneboog (2016) show that volume and short-term investment rise and fall during a disagreement-induced bubble, but they assume that the extent of disagreement exogenously rises and falls during the bubble. In our model of a bubble, the only exogenous shift is a one-time initial demand shock.

That speculators anticipate wanting to sell in the near future distinguishes our framework from several recent dynamic models of disagreement in bubbles. In Burnside et al. (2016), the extent of disagreement varies over a bubble due to epidemiological forces governing random meetings between optimists and pessimists. In their framework, volume rises primarily due to an influx of optimistic buyers, which lowers the time that listings spend on the market. However, this increase in volume is not accompanied by the variation in holding times observed in the data. In Barberis et al. (2016), volume rises during a bubble through a process called “wavering” in which the beliefs of investors unexpectedly change.<sup>7</sup> The unanticipated nature of these belief changes implies that speculators do not have short investment horizons at the time of purchase, in contrast with our model and the empirical evidence we present in Section 1.

A large literature surveyed by Han and Strange (2015) uses search and matching models to understand some of the facts this paper presents concerning the joint dynamics of house prices and sales. The housing market in these models often is a closed system in which the joint buyer–seller problem determines equilibrium dynamics (Wheaton, 1990; Caplin and Leahy, 2011; Díaz and Jerez, 2013; Ngai and Tenreyro, 2014). This approach does not speak to the significant entry of non-occupant and short-term buyers we document. As the entry of such buyers explains 30% to 50% of the 2000–2005 rise in volume, our model explains an important share of volume dynamics that is missed by search and matching models.<sup>8</sup>

Volume growth during the boom in our model comes entirely from an increase in listings, whereas increases in the speed at which listings sell drive volume growth in several search and matching models (Díaz and Jerez, 2013; Head et al., 2014; Burnside et al., 2016; Hedlund, 2016). The 2000–2005 US housing boom matches our model well: as we show, listings

---

<sup>7</sup>Piazzesi and Schneider (2009) also present a model where investor beliefs unexpectedly change: optimistic homebuyers unexpectedly become pessimists upon buying a house, leading to short-term trading volume.

<sup>8</sup>Relative to our paper, search and matching models are often less concerned with price determination. In that literature, prices are usually set through Nash bargaining and react quickly to shocks rather than exhibiting the autocorrelation present in the data. Caplin and Leahy (2011), Díaz and Jerez (2013), Head et al. (2014), and Guren (2016) discuss this point and propose solutions.

increased sharply while selling speed remained flat. To establish this point we replicate the empirical analysis of Ngai and Sheedy (2016), who further present a theoretical explanation that stresses the moving decision of occupant homeowners. In contrast, our theory focuses on the entry of short-term buyers into the housing market when prices are rising.

Our empirical findings contribute to a growing literature documenting the significant role of speculators in the 2000s US housing cycle. Haughwout et al. (2011) and Bhutta (2015) demonstrate the importance of investors—defined as borrowers with simultaneous mortgages on multiple properties—for explaining mortgage credit growth during the boom.<sup>9</sup> Using deeds records, we confirm their mortgage-based findings using the alternate definition of non-occupant buyer proposed by Chinco and Mayer (2016). Bayer et al. (2011) and Bayer et al. (2016) estimate increased entry of short-term buyers into the Los Angeles, CA housing market during the boom, whereas Adelino et al. (2016) calculate a rise during this time in the national share of house purchases sold within one year. Our analysis extends these results by showing that the rise in short-term transactions accounts for much of the total increase in volume.

Last, our model can help explain the relationship between prices and volume in the stock market. Lee and Swaminathan (2000) and Jones (2002) show that higher volume predicts lower subsequent returns, and Statman et al. (2006) and Griffin et al. (2007) document a positive relationship between volume and recent returns. In our model these patterns arise as a result of changes in the activity of short-horizon investors and are particularly important during bubble episodes. In line with this mechanism, Cochrane (2002) and Ofek and Richardson (2003) argue that short-horizon trading became more prevalent during the technology bubble shown in Figure 1(b).

## 1 Motivating Evidence

### 1.1 The Term Structure of Extrapolative Expectations

Much of the early theoretical work on extrapolative expectations does not consider the forward term structure of extrapolation. The two papers we are aware of that explicitly model

---

<sup>9</sup>Gao et al. (2016) calculate an increase in investor activity using HMDA data, and Nathanson and Zwick (2017) show that developers increased their speculative land holdings during this time. Empirical studies of speculation in other markets include Fu and Qian (2014) and Penasse and Renneboog (2016).



how past price changes are extrapolated into expectations of future prices at varying horizons are Barberis et al. (2015) and Glaeser and Nathanson (2016).<sup>10</sup> In both papers, extrapolation is modeled in a way that leads short-term expectations to exhibit more sensitivity to recent price changes than long-term expectations. This approach, which we adopt in our model, is supported by a growing body of empirical evidence suggesting that past asset returns do indeed influence annualized expected capital gains more strongly over short versus long future horizons.

In the housing market, Armona et al. (2016) survey expected capital gains over 1- and 5-year horizons and relate those expectations to perceptions of recent local price changes. They find that 1-year-ahead expectations are nearly five times more sensitive to perceived past price changes than annualized 2-5-year-ahead expectations.<sup>11</sup> Furthermore, when provided with new information about local changes in house prices over the last year, respondents in the survey update their forecasts of 1-year price gains more strongly than their 2-5-year forecasts. Both of these facts suggests that short-run house price expectations display significantly more sensitivity to past returns than do long-run expectations.

Similar evidence exists for the US stock market. Vissing-Jorgensen (2004) reports the average expectation of annualized stock market returns over 1- and 10-year horizons among respondents to the UBS/Gallup Index of Investor Optimism survey between 1998 and 2002. Over this period, 1-year expectations moved closely with recent price changes—first rising from 10% to 16% as stock prices increased, and then falling to 6% as prices fell. In contrast, 10-year expectations remained relatively constant over this period and were uncorrelated with the large contemporaneous movements in the stock market.

These patterns persist even in a sample of more sophisticated survey respondents. The Duke CFO Global Business Outlook, which surveys chief financial officers of US firms, provides data on annualized 1- and 10-year stock return expectations. Graham and Harvey (2003) use data from the 2000–2003 waves of this survey and find that the 1-year expected

---

<sup>10</sup>Hong and Stein (1999) assume that past price changes are extrapolated linearly into a fixed future horizon. Their model does not deliver predictions about the effects of recent price changes on expectations at differing future horizons.

<sup>11</sup>The coefficient when regressing 1-year expectations on past 1-year price growth perceptions equals 0.262 (0.029), whereas the coefficient when regressing the implied 2-5-year annualized expectations on 1-year perceptions equals 0.058 (0.012). Case et al. (2012) conduct a similar survey and report the 1- and 10-year annualized capital gains expectations for the housing market at the city-year level. Using CoreLogic county house price indices and their survey data, we find a coefficient of 0.246 (0.025) when regressing 1-year expectations on the past year’s house price appreciation and a coefficient of 0.146 (0.038) when regressing the annualized 10-year expectations on the past year’s house price appreciation.

risk premium (expected return less treasury yield) is positively and significantly related to excess returns over the previous week, month, two months, and quarter, whereas the 10-year annualized expected risk premium is slightly negatively related to these past returns. In Appendix Table A1, we use the survey data from 2000 to 2011 and confirm that the 1-year expectations remain more sensitive to past returns than the 10-year expectations in the longer sample.<sup>12</sup> Thus, the available evidence all points toward a term structure for extrapolation in which short-run forecasts are more sensitive to recent prices changes than long-run forecasts.

## 1.2 Variation in Expected Holding Times

In the presence of a downward-sloping term structure for extrapolative expectations, 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 will depend on the degree of heterogeneity in the distribution of expected holding times among prospective investors. While not much 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.

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.

Figure 2(a) documents the substantial cross-sectional heterogeneity in expected holding times among respondents to the survey.<sup>13</sup> Each bar reports an equal-weighted average of the

---

<sup>12</sup>Although the data are available from 2000 to 2016, we use only 2000–2011 to match the window used by Greenwood and Shleifer (2014), who also find a coefficient of about 0.03 when regressing the 1-year return expectation on the prior year’s return. Interestingly, the coefficients decline considerably when the 2012–2016 sample is included, possibly because declines in interest rates over this time both increased lagged returns and decreased future return expectations.

<sup>13</sup>The bins in the figure are those used by the NAR in its data release (we do not have access to less aggregated data). We reclassify respondents who have already sold their properties as having an expected holding time in  $[0,1)$ .

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,” which is measured as the fraction of respondents who report an expected holding time of less than 3 years or had already sold their property by the time of the survey.<sup>14</sup> 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. As shown in Figure 2(b), the short-term buyer share moves closely with recent price appreciation in the housing market.<sup>15</sup> A simple regression of the pooled short-term buyer share on the equal-weighted average year-over-year change in the nominal quarterly FHFA US house price index during the survey year yields a statistically significant coefficient estimate of 0.82.<sup>16</sup> 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. The nominal house price appreciation in the US in 2005 was equal to 11% and was 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.

---

<sup>14</sup>In constructing this measure, we omit respondents who do not know their expected holding time.

<sup>15</sup>Appendix Figure A1 shows that internet search queries for “house flipping,” an explicitly short-term investment strategy, also move closely with the housing cycle in the US.

<sup>16</sup>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 Figures 2 and A1, which match house prices to other data over that time.

## 2 A Model of Investors with Heterogeneous Horizons

### 2.1 Primitives and Information Environment

We present an infinite-horizon, continuous-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. The flow utility of all agents equals the sum of composite consumption (whose price is normalized to 1) and housing utility received from owning a house in the city. Agents maximize the present value of this flow utility using an instantaneous discount rate of  $r$ .

In each instant, agents arrive. Each agent begins as a potential buyer and must decide between buying a home immediately and leaving the city forever. A potential buyer who buys a home becomes a stayer. Stayers receive housing utility  $\delta > 0$  from living in the city until receiving an idiosyncratic taste shock and becoming movers, at which point their housing utility drops to 0. This taste shock arrives with an instantaneous Poisson hazard  $\lambda > 0$ , which is distributed across potential buyers independently from  $\delta$  and according to a time-invariant probability density function  $f(\lambda)$ . For any  $\delta_0 > 0$ , the measure of agents arriving at  $t$  for whom  $\delta \geq \delta_0$  equals  $A_t \delta_0^{-\epsilon}$ , where  $\epsilon > 0$  and  $\int_0^\infty \lambda^\epsilon f(\lambda) d\lambda$  exists.<sup>17</sup>

Potential buyers and movers maximize the present value of flow utility by choosing whether to buy or list, respectively. Stayers do not sell their homes until becoming movers, at which point they choose whether to list their homes for sale at the current price  $P_t$ . Until selling, movers decide at each instant whether to list at the current price. All agents may borrow or lend at the common discount rate  $r$ .

This environment 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 et al., 2016). In contrast to this literature, the Poisson hazard may differ across agents in our model. Another contrast is that movers depart the city after selling rather than simultaneously searching for another house. The absence of a joint buyer–seller problem simplifies the model and matches the fact discussed in Section 4

---

<sup>17</sup>This heterogeneity in  $\delta$  is necessary only to produce the demand curve in Lemma 3 that is a smoothly decreasing function of the current price. We believe that a similar demand curve would hold in a model with risk-averse agents and homogeneous  $\delta$ . This alternate model may better describe the stock market because dividends are the same for all owners of a stock.

that the median seller in our data does not reappear as a buyer in the same MSA within a year.

In addition to their individual types and the current price, all agents observe summary information about the complete history of prices. In particular, there exists a function  $\omega(\cdot) \rightarrow \mathbb{R}$  that maps the history of prices into a single factor observed by market participants at time  $t$ . We denote this factor by  $\omega_t \equiv \omega(\{P_{t'} \mid t' \leq t\})$ . Potential buyers and movers use  $\omega_t$  to form expectations of future prices, which govern their decisions of whether to buy or list, respectively. Agents form expectations regarding price growth between time  $t$  and  $t + \tau$  in a manner that is consistent with the following assumption:

**Assumption 1.** *There exist functions  $\gamma$  and  $g$  such that for all  $\delta, \lambda, \tau \geq 0$  and  $P_t, \omega_t \in \mathbb{R}$ ,*

$$\mathbb{E}[P_{t+\tau}/P_t \mid \delta, \lambda, P_t, \omega_t] = 1 + \gamma(\omega_t)g(\tau) \quad (1)$$

and the following properties hold:

- (a)  $(g(\tau)/\tau)' < 0$  for all  $\tau > 0$ ;
- (b)  $g(0) = 0$ ;
- (c)  $\gamma(\omega)g'(0) \leq r$  for all  $\omega \in \mathbb{R}$ ; and
- (d)  $\int_0^\infty e^{-r'\tau}g(\tau)d\tau > 0$  for all  $r' > r$ .

Assumption 1(a) endows agents with extrapolative expectations that satisfy the empirical evidence on the forward term structure presented in Section 1. The decrease of  $g(\tau)/\tau$  is necessary and sufficient for nontrivial increases in  $\omega_t$  to raise short-term expected capital gains more strongly than long-term expected capital gains:<sup>18</sup>

**Lemma 1.** *Given (1), Assumption 1(a) holds if and only if*

$$\frac{\partial^2}{\partial \tau \partial \omega_t} \mathbb{E} \left[ \frac{P_{t+\tau} - P_t}{\tau P_t} \mid \omega_t \right] < 0$$

for all  $\tau > 0$  and  $\omega_t \in \mathbb{R}$  such that  $\gamma'(\omega_t) > 0$ .

Assumption 1(b) imposes the weak constraint that  $\mathbb{E}[P_{t+\tau}/P_t] = 1$  for  $\tau = 0$ . Assumption 1(c) is necessary and sufficient for the expected growth rate of prices to always fall below  $r$ :

---

<sup>18</sup>Appendix A contains the proof of Lemma 1 as well as all other omitted proofs.

**Lemma 2.** *Given (1) and Assumptions 1(a) and 1(b),  $E[P_{t+\tau}/P_t \mid \omega_t] < e^{r\tau}$  for all  $\tau > 0$  and  $\omega_t \in \mathbb{R}$  if and only if Assumption 1(c) holds.*

In our risk-neutral framework, this condition is necessary for demand from arbitrarily short-term buyers to be finite.<sup>19</sup> Finally, Assumption 1(d) guarantees that an increase to  $\gamma(\omega_t)$  raises the present value of expected capital gains for all potential buyers.

## 2.2 Equilibrium Quantities

Solving the model requires knowing the number of agents of each type at each point in time, as well as the stock of previous listings that did not sell. In particular, we track the number of potential buyers who decide to buy  $D_t$  and the number of stayers  $S_t$ . These jointly determine the flow of listings  $L_t$ , the inventory of unsold listings  $I_t$ , and sales volume  $V_t$ .

Due to Lemma 2,  $P_t > e^{-r\tau}E[P_{t+\tau} \mid P_t, \omega_t]$  for all  $t$  and  $\tau > 0$ , so movers always choose to list their homes for sale. As a result, we may describe the evolution of listings using the stock of stayers of each type  $\lambda$ , or  $S_t(\lambda)$ . Because all new movers list, the flow of listings is simply the total number of stayers receiving idiosyncratic mover shocks:

$$L_t = \int_0^\infty \lambda S_t(\lambda) d\lambda. \quad (2)$$

Given  $A_t$ ,  $P_t$ , and  $\omega_t$ , a measure  $D_t$  of potential buyers decide to buy. If  $D_t$  is less than the number of homes listed for sale, then all interested potential buyers are able to buy. Otherwise, homes are rationed randomly among the set of interested potential buyers. Let  $D_t(\lambda)$  denote the measure of potential buyers of type  $\lambda$  who decide to buy. Then sales volume to potential buyers of type  $\lambda$  equals

$$V_t(\lambda) = \begin{cases} D_t(\lambda) & \text{if } L_t > D_t \text{ or } I_t > 0 \\ \frac{D_t(\lambda)}{D_t} L_t & \text{if } L_t \leq D_t \text{ and } I_t = 0, \end{cases} \quad (3)$$

where  $I_t$  is the inventory of unsold listings. The stayer stocks for each  $\lambda$  evolve according to the law of motion

$$\dot{S}_t(\lambda) = V_t(\lambda) - \lambda S_t(\lambda), \quad (4)$$

---

<sup>19</sup>As shown in the proof of Lemma 2, Assumption 1(c) implies the stronger condition that  $E[P_{t+\tau}/P_t \mid \omega_t] < 1 + r\tau$  for all  $\tau > 0$  and  $\omega_t \in \mathbb{R}$ , which guarantees that the transversality condition holds as appears to be the case empirically in housing markets (Giglio et al., 2016).

which is the number of new buyers of type  $\lambda$  less the number of stayers of type  $\lambda$  who become movers. Unsold listings follow the law of motion

$$\dot{I}_t = L_t - V_t, \tag{5}$$

where  $V_t = \int_0^\infty V_t(\lambda)d\lambda$  equals total sales.

These expressions have 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.

### 2.3 The Composition of Buyers

The interesting dynamics in the model concern how the composition of buyers—specifically, the composition of expected holding periods—varies over the cycle. This composition depends on the distribution of  $\lambda$  among buyers. The probability density function of  $\lambda$  among buyers at time  $t$  is given by the function  $V_t(\lambda)/V_t$ . By (3), this function coincides with  $D_t(\lambda)/D_t$ , the probability density function of demand across potential buyers at  $t$ . Thus to understand how the composition of buyers varies over time, we must calculate the distribution of demand across  $\lambda$ .

To derive  $D_t(\lambda)$ , we assume that potential buyers believe that they will sell their house as soon as they list it:<sup>20</sup>

**Assumption 2.** *Potential buyers believe that listing movers sell instantaneously.*

One consequence of Assumption 2 is that the expected holding time of a buyer of type  $\lambda$  equals  $1/\lambda$ . The other consequence is that a potential buyer of type  $(\delta, \lambda)$  tries to buy if the current price is below the present discounted value of the housing utility she would receive as a stayer plus the expected resale value of the house at the time she anticipates becoming a mover. That is, a potential buyer tries to buy if and only if

$$P_t \leq \int_0^\infty \lambda e^{-\lambda\tau} \left( \int_0^\tau e^{-r\tau'} \delta d\tau' + e^{-r\tau} E_t P_{t+\tau} \right) d\tau. \tag{6}$$

---

<sup>20</sup>Appendix B microfoundes Assumption 2 as a manifestation of limited attention.

The value of  $\delta$  at which a buyer of type  $\lambda$  is indifferent implies the equation for demand given in Lemma 3.

**Lemma 3.** *Demand from each  $\lambda$  type equals*

$$D_t(\lambda) = \underbrace{f(\lambda)A_t}_{\substack{\text{potential} \\ \text{buyer} \\ \text{measure}}} \times \underbrace{(rP_t)^{-\epsilon}}_{\substack{\text{fundamental} \\ \text{demand}}} \times \underbrace{\Sigma_\lambda(\omega_t)}_{\substack{\text{speculative} \\ \text{demand}}},$$

where

$$\Sigma_\lambda(\omega_t) = \left(1 - \frac{\lambda\gamma(\omega_t)}{r} \int_0^\infty e^{-(r+\lambda)\tau} g'(\tau) d\tau\right)^{-\epsilon}. \quad (7)$$

Demand for each  $\lambda$  is composed of three terms. The first term  $f(\lambda)A_t$  equals the relative measure of potential buyers of type  $\lambda$ . The second term, which we call “fundamental demand,” is a decreasing function of current prices with constant elasticity  $\epsilon$ . This term reflects the relationship between demand and prices were prices to remain permanently at their current level. The third term, which we call “speculative demand,” links current demand to expected capital gains. If buyers expect prices to remain constant, then  $\Sigma_\lambda(\omega_t) \equiv 1$  and speculative demand does not magnify total demand. Otherwise, speculative demand magnifies total demand when buyers expect capital gains and attenuates total demand when buyers expect capital losses, with the force of this multiplier depending on the buyer’s horizon.

Only speculative demand matters for variation in the composition of buyers over time, as the other two demand components are always proportional to each other across  $\lambda$  types. As a consequence, variation in the composition of buyers depends entirely on changes in expected capital gains. Proposition 1 formally states this relationship using the log of speculative demand, which we denote  $\sigma_\lambda \equiv \log \Sigma_\lambda$ :

**Proposition 1.** *At any  $\omega_t$  such that  $\gamma'(\omega_t) > 0$ , the following hold for all  $\lambda > 0$ .*

- (a) *Expected capital gains increase speculative demand from all types:  $\sigma'_\lambda(\omega_t) > 0$ .*
- (b) *Short-term buyers are more sensitive to expected capital gains:  $\partial\sigma'_\lambda(\omega_t)/\partial\lambda > 0$ .*
- (c) *Expected capital gains skew the composition of buyers shorter-term:*

$$\frac{\partial \int_0^\lambda V_t(\lambda') d\lambda' / V_t}{\partial\omega_t} \leq 0,$$

*with equality if and only if  $\mathbb{R}_{<\lambda} \cap \text{supp } f = \emptyset$  or  $\mathbb{R}_{>\lambda} \cap \text{supp } f = \emptyset$ .*



Part (a) formally states that greater expected capital gains increase demand. This effect appears in any user cost model of housing (e.g., Poterba, 1984). The focus in most user cost models is primarily the intensive margin demand for housing capital. Our model highlights the extensive margin instead, as the stimulative effect of expectations on demand operates through drawing new buyers into the market. Consistent with this mechanism, Agarwal et al. (2015) document increased participation in the owner-occupied housing market in response to rising prices.

Part (b) shows that the stimulative effect of capital gains is stronger for buyers with shorter expected holding times. Buyers looking to make a “quick buck” are drawn to rising prices more than those buying for the long run. The proof of Proposition 1 shows that this effect follows from the higher sensitivity of short-term expectations to  $\omega_t$  embodied by Assumption 1(a).

Part (c) links expected capital gains to the composition of buyers. Because short-horizon buyers are more sensitive to capital gains, an increase in expected capital gains skews the composition of buyers towards those with shorter holding times. This skewing occurs as long as heterogeneity exists in  $f$ —that is, as long as  $|\text{supp } f| > 1$ .<sup>21</sup> Proposition 1(c) explains the evidence presented in Section 1 that both expected capital gains and the short-term buyer share respond strongly to recent home price appreciation.

Proposition 1 illustrates the key mechanism generating time variation in volume in the model, namely, time variation in the composition of holding periods for buyers and sellers driven by time variation in expected capital gains. To close the model, we now specify how prices are determined.

## 2.4 Equilibrium Prices

As pointed out by Barberis et al. (2015), extrapolative expectations do not generate protracted booms and busts in asset prices without some additional source of sluggish adjustment. As the goal of this paper is to study volume dynamics during an extended price boom and bust, we specify a sluggish price adjustment process rather than assuming a Walrasian market in which prices equate the number of interested buyers and sellers at each instant.<sup>22</sup>

---

<sup>21</sup>If  $|\text{supp } f| > 1$ , then there always exists  $\lambda$  such that  $\mathbb{R}_{<\lambda} \cap \text{supp } f \neq \emptyset$  and  $\mathbb{R}_{>\lambda} \cap \text{supp } f \neq \emptyset$ . If  $|\text{supp } f| = 1$ , then such  $\lambda$  do not exist.

<sup>22</sup>An alternative source of sluggish adjustment is for agents to base extrapolative expectations on only lagged prices and not the current price (Glaeser and Nathanson, 2016; Barberis et al., 2016).

The log price  $p_t = \log P_t$  changes according to the rule

$$\dot{p}_t = c \log(D_t/\bar{D}), \tag{8}$$

where  $c > 0$  is a constant determining the speed of price adjustment, and  $\bar{D}$  is given by

$$\bar{D} = \left( \int_0^\infty \lambda^{-1} f(\lambda) d\lambda \right)^{-1}. \tag{9}$$

As shown in Appendix A, (9) gives the unique value of  $\bar{D}$  such that a steady state exists in which all listings sell instantaneously and prices are expected to and do remain constant. The rule in (8) brings prices towards such a steady state in a slow and predictable manner.

Appendix B microfounds (8) as the equilibrium outcome of a simple stock-flow matching model. In this microfoundation, some sellers are inattentive as in Hong and Stein (1999), Mankiw and Reis (2002), and Guren (2016) and do not adjust their listing prices in response to abnormal levels of demand, while other sellers observe current demand and choose their listing prices in response to it. The constant  $c$  is larger when the share of attentive sellers is higher and the demand elasticity  $\epsilon$  is lower. In equilibrium, the allocation of listed houses to potential buyers coincides with the rule in (3).

### 3 The Joint Dynamics of Prices and Volume

We provide a series of propositions that characterize the relationship between prices and volume over the course of a boom–bust cycle. We then present a numerical calibration, which allows us to study the potential quantitative relevance of the factors driving the model.

We focus on how volume and prices respond to a one-time permanent demand shock. In particular, we study an impulse response around the unique steady state that results from a single positive shock to the number of potential buyers,  $A_t$ , at a time we normalize to  $t = 0$ . Specifically,  $A_t$  follows the path where  $A_t = A^i$  for  $t < 0$  and  $A_t = A^f > A^i$  for  $t \geq 0$ .

To characterize how prices and volume respond to this shock, we impose additional structure on how agents form expectations. We specify both the information that agents have available to them when they form forecasts of future prices and how that information

influences these forecasts, which are governed in the model by  $\omega(\cdot)$  and  $\gamma(\cdot)$ , respectively.

Following Barberis et al. (2015), we assume that agents observe only a weighted average of past price changes:

$$\omega_t = \int_{-\infty}^t \mu e^{-\mu(t-\tau)} \dot{p}_\tau d\tau, \quad (10)$$

where the parameter  $\mu > 0$  measures the relative weight put on more recent price changes. To study dynamics, it is useful to know how this average changes in response to an instantaneous change in prices, which is simply the differential form of (10):

$$\dot{\omega} = \mu(\dot{p} - \omega). \quad (11)$$

Our specification of  $\omega$  is sufficiently general that the impulse response may not always result in a well-behaved boom and bust in prices. In some cases, prices may rise without falling or they may oscillate indefinitely. We restrict focus to a parameter region in which a boom is followed by a bust that asymptotes without indefinitely oscillating. In particular, we require  $\gamma(\cdot)$ , the function mapping past price information to future expected price changes, to satisfy the following assumption:

**Assumption 3.**  $\gamma(\omega) \equiv 0$  for  $\omega \leq 0$ ,  $\lim_{\omega \rightarrow 0^+} \gamma(\omega) = 0$ ,  $\gamma'(\omega) > 0$  for  $\omega > 0$ , and

$$\lim_{\omega \rightarrow 0^+} \gamma'(\omega) > \frac{r}{\min(c\epsilon, \mu)} \left( \int_0^\infty \int_0^\infty \lambda e^{-(r+\lambda)\tau} g'(\tau) d\tau f(\lambda) d\lambda \right)^{-1}.$$

The requirement that  $\gamma \equiv 0$  for  $\omega < 0$  rules out oscillations after the bust, as agents stop expecting capital gains once the historical average return becomes negative. The rest of Assumption 3 guarantees that price increases initially beget further increases and that prices eventually overshoot.

### 3.1 The Three Epochs

We present a series of propositions partitioning the boom–bust cycle into three epochs. To define these epochs, we provide a lemma that guarantees that prices do not rise indefinitely:

**Lemma 4.**  $t_2 \equiv \min\{t > 0 \mid \dot{p}_t = 0\}$  exists.

The time  $t_2$  marks the first time that prices stop rising. We set  $V^{max} \equiv \max\{V_t \mid t \in [0, t_2]\}$  as the peak of volume during this rise in prices and  $t_1 \equiv \max\{t \in [0, t_2] \mid V_t = V^{max}\}$  to be

the latest time this peak occurs ( $V^{max}$  and  $t_1$  exist due to the continuity of  $V_t$ ). The *boom* is the period  $(0, t_1]$ , the *quiet* is the period  $(t_1, t_2]$ , and the *bust* is the period  $(t_2, \infty)$ .

Proposition 2 characterizes the boom:

**Proposition 2.** *The boom exists ( $t_1 > 0$ ). Throughout the boom, prices rise and expected capital gains are positive ( $\omega > 0$ ). At least initially, prices are convex ( $\ddot{p} > 0$ ), expected capital gains rise ( $\dot{\omega} > 0$ ), no unsold listings accumulate ( $I = 0$ ), and if  $|\text{supp } f| > 1$ , volume rises ( $\dot{V} > 0$ ) and listings rise ( $\dot{L} > 0$ ). Volume remains constant throughout the boom if  $|\text{supp } f| = 1$ .*

As the proof shows, the shock to  $A$  at  $t = 0$  causes demand  $D$  to jump above available listings. As a result, the price of housing begins to increase, and this increase raises  $\omega$ . The rise in  $\omega$  stimulates demand, further increasing prices and leading to convexity in the price path. When  $|\text{supp } f| > 1$ , demand rises more sharply for potential buyers with higher values of  $\lambda$ , causing an increase in listings and hence volume. Volume rises solely due to the underlying heterogeneity in horizons among market participants—without this heterogeneity, volume and listings remain constant during the boom.

Proposition 2 fits the 2000–2005 US housing market remarkably well. As shown in Figure 1(a), prices and volume rose during this time, with prices rising at an increasing rate. According to the analysis of the Michigan Survey of Consumers presented by Piazzesi and Schneider (2009), during 2002–2003 an increasing share of respondents said now was a “good time to buy” a house, whereas during 2004–2005 an increasing share described housing as “too expensive” while the share of respondents expecting further house price increases doubled. This evidence is consistent with our characterization of the boom as an initial demand shock followed by a period of rising prices and expected capital gains.

Proposition 3 characterizes the quiet:

**Proposition 3.** *The quiet exists ( $t_1 < t_2$ ) if and only if  $|\text{supp } f| > 1$ . Prices increase and expected capital gains remain positive ( $\omega > 0$ ) throughout the quiet, and at least ultimately prices are concave ( $\ddot{p} < 0$ ) and expected capital gains decrease ( $\dot{\omega} < 0$ ). Volume is lower at the end of the quiet than at the beginning ( $V_{t_1} > V_{t_2}$ ), and unsold inventory may accumulate ( $I > 0$ ) during the quiet.*

Prices continue to increase during the quiet by the definition of  $t_2$ . Because  $\omega$  averages past price changes, it remains positive throughout the quiet. As time reaches  $t_2$ , prices

are concave as price growth slows, and this slowdown in price growth causes  $\omega$  to fall. The combination of rising prices and falling expected capital gains unambiguously lowers demand. It falls to the steady-state value  $\bar{D}$  at  $t_2$ , as this is the unique value at which price growth is flat according to the price adjustment rule (8). When  $|\text{supp } f| > 1$ , listings may remain above this steady-state value because past capital gains have lured short-term potential buyers disproportionately, so the decline in demand can trigger a buildup of unsold inventory and a fall in volume. When  $|\text{supp } f| = 1$ , listings equal  $\bar{D}$  for the entire time between 0 and  $t_2$ , so volume never falls as  $t_2$  is approached and the quiet fails to exist because  $t_1 = t_2$ .

Proposition 3 fits the US housing market between 2005 and 2007. As shown in Figure 1(a), in 2005 volume began to fall and prices began to grow at a slower rate. Piazzesi and Schneider (2009) report that the share of households in the Michigan Survey of Consumers expecting further house price growth sharply fell from 2005 to 2007, and expected house price growth in the four metropolitan areas surveyed by Case et al. (2012) also declined during this time. This survey evidence is consistent with the fall in  $\omega$  during the quiet.

Proposition 4 characterizes the bust:

**Proposition 4.** *The limit of prices is  $\lim_{t \rightarrow \infty} p_t = p_0 + \epsilon \log(A^f/A^i)$ . Prices decline, unsold inventory may exist ( $I > 0$ ), and volume is weakly below its steady-state value ( $V \leq \bar{D}$ )—strictly except possibly at isolated instants—during the bust until the limit is reached.*

At the end of the quiet, demand equals the steady-state value  $\bar{D}$  and is declining because expected capital gains are falling. As the bust begins, demand falls below  $\bar{D}$  and prices begin to fall. The negative price growth triggers further declines in expected capital gains and demand. This spiral stops once two conditions are met: expected capital gains reach their minimum value of 0, and prices reach their “fundamental” value of  $p_0 + \epsilon \log(A^f/A^i)$ .<sup>23</sup> At this fundamental price, demand returns to its steady-state value of  $\bar{D}$  and prices stabilize. Because demand remains low until this limit is reached, volume lies below  $\bar{D}$  and a stock of unsold inventory may persist during this phase of the bust.

The bust corresponds to the period after 2007 shown in Figure 1(a) during which prices

---

<sup>23</sup>Prices do not overshoot on the decline because Assumption 3 rules out negative expected capital gains. Without this assumption, prices would overshoot on the way down as is standard in models with extrapolative expectations (e.g., Glaeser and Nathanson, 2016). Negative overshooting may occur empirically for a variety of reasons, such as a continuing glut of listed houses (Shleifer and Vishny, 1992; Rognlie et al., 2016), foreclosures (Guren and McQuade, 2015), or a contraction in credit (Garriga and Hedlund, 2016).

fell and volume was lower than average. Low volume typically accompanies price declines in the housing market (Stein, 1995; Genesove and Mayer, 2001; Ortalo-Magné and Rady, 2006). Our explanation for this fact is that prices decline when housing has become overvalued relative to fundamental demand. This overvaluation, combined with the sluggish adjustment of prices, leads to an extended period during which prices decline towards their fundamental value on low transaction volume.

The three epochal propositions produce two empirical predictions about the joint dynamics of prices and volume. First, volume leads prices over their joint cycle (assuming heterogeneity in investment horizons). Second, the speed at which listings sell remains constant over the boom, declines during the quiet as a stock of unsold inventory appears, and remains low during the bust as the stock of unsold inventory persists on low volume.<sup>24</sup>

### 3.2 The Composition of Buyers over the Cycle

To illustrate the forces driving prices and volume over the three epochs, we now characterize the relative contribution of short-term and speculative potential buyers to demand over the cycle. A potential buyer is “speculative” if she would like to buy given  $\omega_t$  but would not want to buy if the current value of  $\omega$  were equal to 0. For a given  $\lambda$ , speculative potential buyers have lower flow utility  $\delta$  than nonspeculative potential buyers who would like to buy. By Lemma 3, the share of demand attributed to speculative potential buyers equals  $\int_0^\infty (1 - \Sigma_\lambda(\omega_t)^{-1}) f(\lambda) d\lambda$ , which we term the *speculative share*.

Proposition 5 characterizes the evolution of the short-term buyer and speculative shares:

**Proposition 5.** *For any  $\lambda'$  such that  $\mathbb{R}_{<\lambda'} \cap \text{supp } f \neq \emptyset$  and  $\mathbb{R}_{>\lambda'} \cap \text{supp } f \neq \emptyset$ , both the speculative share and the share of buyers for whom  $\lambda > \lambda'$  rise during the boom, fall but remain above steady-state during the quiet, and fall during the bust.*

Both the short-term buyer and speculative shares depend positively on  $\omega$ , so Proposition 5 follows from the facts about  $\omega$  established in Propositions 2–4.

Proposition 5 shows that speculative and short-term buyers begin to pull out of the market during the quiet after becoming more prevalent during the boom. The level of their

---

<sup>24</sup>The model’s prediction on listing speed is not sharp because Propositions 3 and 4 prove only that unsold listings may exist during the quiet and bust. In the calibration of the model below, we find that the stock of unsold listings appears precisely at the beginning of the quiet and persists into the bust.

activity during the quiet remains high, and it is this combination of high but decreasing speculative activity that generates rising prices and declining volume during the quiet.

This proposition generates two empirical predictions. First, the short-term buyer share—where the horizon is measured *ex post* as the amount of time between the purchase and the next listing—should grow over the boom and fall during the quiet. Second, the speculative share—measured as the share of buyers with low housing utility—should rise over the boom and fall during the quiet.

### 3.3 Short-Term Buyers and the Size of the Cycle

Propositions 2–4 characterize the relative timing of the price and volume responses. We now describe the relative *magnitude* of these responses. In particular, we show that a common factor—the distribution of expected holding times  $f(\cdot)$ —determines the magnitude of the price and volume responses as well as the level of steady-state volume. The following proposition characterizes the effects of increasing both the level of  $f(\cdot)$  (in the sense of first-order stochastic dominance) as well as its heterogeneity.

**Proposition 6.**

- (a) *An increase to  $f(\cdot)$  raises  $V_0$  and  $P^{max}$  while keeping  $P_0$  and  $P_\infty$  constant.*
- (b) *If  $1 = |\text{supp}(f_A)| < |\text{supp}(f_B)|$  and  $f_A < f_B$ , then  $P^{max}/P_0$ ,  $P^{max}/P_\infty$ ,  $V^{max}/V_0$ , and  $V_0$  are larger under  $f_B$  than  $f_A$ .*

$P_\infty$  denotes the limiting price specified in Proposition 4.

As shown in the proof of Proposition 6(a), a larger  $f(\cdot)$  raises steady-state volume  $V_0$  by increasing the average speed at which stayers choose to list. A larger  $f(\cdot)$  also raises the speculative demand function  $\Sigma(\cdot)$  because the demand of short-term buyers is more sensitive to expected capital gains; this increase to  $\Sigma(\cdot)$  raises the maximal price reached during the cycle without changing the steady states. Proposition 6(a) makes the empirical prediction that price cycles are larger in markets with higher steady-state volume.

Proposition 6(b) clarifies the importance of volume for understanding asset bubbles in a special case. In this example, all potential buyers have the same expected holding time under  $f_A$ , while potential buyers have shorter and more variable expected holding times under  $f_B$ . As shown by the proposition, the price *and* volume responses during the cycle are

larger under  $f_B$  than  $f_A$ . Volume is not a sideshow to prices, but rather a manifestation of the speculative forces responsible for price dynamics. These speculative forces are captured by the heterogeneity in  $f(\cdot)$ . We conjecture that a mean-preserving spread in  $f(\cdot)$  always increases  $P^{max}$  and  $V^{max}$  while leaving  $V_0$  unchanged. Although we are not able to prove this conjecture, we verify it in our calibration below.

### 3.4 Calibration

To study predictions of our model quantitatively, we simulate the impulse response characterized in Propositions 2–4. We choose parameter values using surveys and prior literature to discipline the calibration and ask whether, subject to this parameterization, the changes in volume and buyer composition are large.

#### Parameter Choices

Table 1 lists the model parameters and their sources. We calibrate  $f(\cdot)$  using the expected holding times reported in the NAR survey shown in Figure 2(a).<sup>25</sup> To select  $\mu$ , we rerun the regression shown in Figure 2(b) by replacing the lagged four-quarter house price change with  $\omega$  as given by (10). We choose  $\mu$  to maximize the  $R^2$  of this regression, leading us to  $\mu = 1.18$  (standard error 0.64) and  $R^2 = 68\%$ . This estimate is noisy due to the small amount of data used in the estimation, but it is close to the value of  $\mu = 0.5$  estimated by Barberis et al. (2015) in the context of the stock market.

To calibrate  $g(\cdot)$ , we adopt the functional form  $g(\tau) = \rho(1 - e^{-\tau/\rho})$ . For all  $\rho$ ,  $g'(0) = 1$ , so  $\rho$  controls the extent to which the initial gain is extrapolated into the future. The half-life of the cumulative gains implied by  $g(\cdot)$  equals  $\rho \log 2$ , so a larger  $\rho$  produces greater relative sensitivity of long-term expected gains to short-term expected gains.<sup>26</sup> In each year from 2014 to 2016, the New York Fed’s Survey of Consumer Expectations reports the median expectation of house price growth in the United States over the next 1 and 5 years (see Fuster and Zafar, 2015 and Kuchler and Zafar, 2016 for information on this survey). The ratio of these expectations equals  $(1 - e^{-5/\rho})/(1 - e^{-1/\rho})$ , so we use the sample ratios to obtain

<sup>25</sup>We map the expected holding time for each bin to its median, except for  $[11, \infty)$ , which we map to 20.  $\lambda$  equals the inverse of the expected holding time for each bin (e.g.,  $\lambda = 2$  for an expected holding time of 0.5 years). Using the share of sales to each property type in each year, we calculate  $f(\cdot)$  for each year and then take an equal-weighted average across years to obtain the distribution used in the calibration.

<sup>26</sup>The half-life is the value  $\tau_{hl}$  such that  $g(\tau_{hl}) = (1/2) \lim_{\tau \rightarrow \infty} g(\tau)$ .



a value of  $\rho = 20.9$  (half-life of 14.5 years). An alternative method is to use the coefficient estimates from Armona et al. (2016) of the relative sensitivities of 1-year and 5-year forward capital gains expectations to the prior year's house price appreciation. By equating these estimates to  $g(\tau) \lim_{\omega \rightarrow 0^+} \gamma'(\omega)$  for  $\tau = 1$  and  $\tau = 5$ , we obtain a much smaller number of  $\rho = 1.4$  (half-life of 1.0 year).<sup>27</sup> We use the average of these two numbers as our baseline, and return to each extreme in the sensitivity analysis.

To calibrate  $\gamma(\cdot)$ , we adopt the functional form  $\gamma(\omega) = r(1 - e^{-\phi\omega/r})$  for  $\omega > 0$ . Here,  $\phi > 0$  is a free parameter governing the sensitivity of expectations to small increases in  $\omega$ . As required by Assumption 1,  $\gamma(\omega)g'(0) \leq r$  for all  $\omega$ , and as required by Assumption 3,  $\gamma'(\omega) > 0$  for all  $\omega > 0$ . We use the results of the survey of homeowner expectations conducted by Case et al. (2012) to estimate  $\phi$ . Case et al. (2012) report the average expectation of the following year's price growth of homeowners in Alameda County (CA), Middlesex County (MA), Milwaukee County (WI), and Orange County (CA) in the spring of each year from 2003 to 2012. Using the CoreLogic monthly house price indices for these areas going back to 1976, we calculate  $\omega_t$  for each county and year with (10) and our estimate of  $\mu$  mentioned above. We then choose  $\phi$  and a constant to minimize the mean-squared error of  $g(1)\gamma(\omega)$  plus this constant versus the expectation reported by Case et al. (2012). The resulting value is  $\phi = 0.98$  (standard error 0.27).<sup>28</sup> Our specification explains 70% of the variance in 1-year expectations across counties and years in this sample.

We set  $r = 0.07$ , which corresponds to a steady-state price-rent ratio of about  $1/0.07 = 14$ . We choose  $c = 1$ , which implies a half-life of price adjustment of about 8 months. We set the elasticity of demand,  $\epsilon$ , equal to 0.6, a value in the range of estimates suggested by Hanushek and Quigley (1980). Finally, for round-number convenience, we choose a demand shock size to match a long-run price impact of 10%. This price impact equals  $(A^f/A^i)^{1/\epsilon}$ , so we set  $A^f/A^i = 1.06$ : a demand shock of 6%.

---

<sup>27</sup>Armona et al. (2016) calculate the expected 1-year gain, which equals  $\gamma(\omega)g(1)$ , and the annualized 2-5-year expected gain, which equals  $(1 + \gamma(\omega)(g(5) - g(1))/(1 + \gamma(\omega)g(1)))^{1/4} - 1$ . The right derivatives at  $\omega = 0$  are  $g(1) \lim_{\omega \rightarrow 0^+} \gamma'(\omega)$  and  $(1/4)(g(5) - g(1)) \lim_{\omega \rightarrow 0^+} \gamma'(\omega)$ . The ratio of these equals  $(1/4)(g(5)/g(1) - 1)$ , the value of which uniquely identifies  $\rho$ .

<sup>28</sup>The value of the constant is 0.022 (0.004).

## Results

The differential equations given by our model allow us to solve for the impulse response in continuous time. In order to quantify the marginal effects of heterogeneous holding times, we supplement the baseline model with one in which  $\lambda$  is the same for all potential homebuyers. We set this value to  $(\int_0^\infty \lambda^{-1} f(\lambda) d\lambda)^{-1}$ , the unique value at which steady-state volume remains unchanged.

Figure 3 displays the resulting impulse responses; the quiet is shaded. Panel (a) plots our two main objects of study: prices and volume. In the core model, prices significantly overshoot the long-run cumulative growth of 10%, more than doubling before decreasing to the new level. Prices initially are convex, with price changes begetting further changes. In contrast, prices display a much less pronounced boom and bust when expected holding times equal the average. In the baseline model, volume rises and then falls, beginning its decline 11 months before prices. This delay is close to the empirical delay of 24 months we document below. The total rise in volume in our simulation equals 22%, a substantial fraction of the 40% rise in total sales volume between 2000 and 2005 in the sample we use below. As shown on the right, volume remains constant as prices rise when  $\lambda$  is homogeneous.

Panel (b) documents the changing composition of buyers over the cycle. At each time, we calculate the share of purchases going to buyers whose expected holding time is less than 3 years. In steady state, this share equals 20% (as identified by the NAR survey), and it rises to a high of 39%. This rise occurs as prices increase, and it drives the concomitant and subsequent surge in volume. As predicted by Proposition 5, the share declines but remains high during the quiet. In the homogeneous simulation, no homebuyers have expected holding times of less than 3 years, as all holding times equal 10.5 years.

Finally, Panel (c) documents the evolution of unsold listings over the cycle. Until the quiet begins, all listings sell, so unsold listings equal 0. The stock grows as volume begins to decline. Quantitatively, it reaches 6% of the housing stock in the baseline model, but only 2% when horizons are homogeneous.

In sum, Figure 3 shows that our calibrated model can quantitatively generate large swings in prices and volume during the boom and quiet periods, a dramatic shift in the composition of buyers, and a sharp rise in inventories in the period leading to the bust. These features depend critically on heterogeneity in the expected holding times of potential homebuyers.

## Sensitivity Analysis

To provide intuition on how the parameters drive the results, we report key statistics of the simulation under parameters other than our baseline in Table 2. The three statistics we report are the excess price boom  $P^{max}/P_\infty - 1$ , the volume boom  $V^{max}/V_0 - 1$ , and the maximal inventory of unsold listings  $\max_t I_t$ . We vary each parameter to a low and high value while keeping the remaining parameters at the baseline values.

First we vary the degree of heterogeneity in  $f(\cdot)$ . In the “high” treatment, we keep steady-state volume  $(\int_0^\infty \lambda^{-1} f(\lambda) d\lambda)^{-1}$  constant but put all the mass in  $f(\cdot)$  on the most extreme values of  $\lambda$  in its support. The “low” treatment replicates the right panel of Figure 3, in which no heterogeneity exists. The booms in price and volume are much larger in the high treatment than in the baseline—prices more than quadruple, and volume almost doubles. Unsold inventories also rise to 18% of the housing stock. These results provide strong evidence tying together price booms, volume booms, and the distribution of expected holding times.

Varying the housing-demand elasticity produces similar effects. Our low treatment sets  $\epsilon = 0.3$  (half the baseline), whereas our high treatment sets  $\epsilon = 1.8$  (an average of values calculated by Diamond, 2016). Prices more than quadruple and volume nearly doubles under the high elasticity, whereas prices and volume are much more stable under the low treatment. Short-term buyers enter the market more aggressively when housing demand is more elastic, so increasing the elasticity achieves similar results to increasing the frequency of short-term buyers. The simulation results are less sensitive to variations in the other parameters.

## 4 Speculative Dynamics in the US Housing Bubble

While the calibration in the previous section allows for an assessment of the potential quantitative relevance of the factors that drive our model, their actual empirical relevance has yet to be established. In this section, we provide empirical evidence linking shifts in the distribution of realized holding periods over the course of the 2000–2011 US 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.

## 4.1 Data

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 US. Our main analysis sample spans the years 1995–2014 and includes data from 115 metropolitan statistical areas (MSAs), which together represent 48% of the US 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 our 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 US during our sample period.

We supplement these data with national and MSA-level housing stock counts from the US Census, national counts of sales and listings of existing homes from the NAR, and national and MSA-level nominal house-price indices from CoreLogic.

## 4.2 The Composition of Buyers

### Time-Series Evidence

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 was stated formally in Proposition 1 and implies that large swings in volume should be accompanied by equally large changes in the distribution of realized holding periods among those who choose to sell their homes at various points in the cycle.

As evidence for this prediction, Figure 4(a) presents a simple yet compelling illustration of the time variation in realized holding periods during the 2000–2011 US housing cycle. We define the holding period of each transaction 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 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.<sup>29</sup>

Figure 4(b) directly tests Proposition 5, which predicts that the short-term buyer share rises during the boom and falls during the quiet and bust. For each month, the short-term buyer share equals the fraction of purchases for which the expected time until listing conditional on the eventual sale date of property is less than 3 years.<sup>30</sup> We focus on the time until listing rather than sale for two reasons: (1) time until listing is what identifies horizons within our model, and (2) time until sale is confounded by the fact that the length of time between listing and sale varies empirically over the cycle. As predicted by the model, the resulting time series in Figure 4(b) first rises during the boom and then falls during the bust and quiet (using a procedure described in Section 4.3, we select a time interval as the quiet and shade this period in the figure). Figure 4(b) corresponds to the simulated series in Figure 3(b) that takes the same shape.

Some short-term sellers may stay within the MSA by buying another house. 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. To measure the extent of within-MSA moves, for each seller we search for buyers in the MSA within a quarter of the sale using the names in the deeds records. As documented in Appendix D, the match rates in this exercise fall

---

<sup>29</sup>This finding is in line with the evidence provided by Bayer et al. (2011), who document a similar increase in volume among short-holding-period buyers in the Los Angeles MSA during this period.

<sup>30</sup>Given purchase month  $t_b$  and sale month  $t_s$ , the expected listing month  $t_l$  assuming a uniform prior over such months is  $E[t_l|t_b, t_s] = \sum_{t=t_b}^{t_s} t s_t \prod_{t'=t}^{t_s-1} (1 - s_{t'}) / \sum_{t=t_b}^{t_s} s_t \prod_{t'=t}^{t_s-1} (1 - s_{t'})$ ;  $s_t$  is the probability we calculate in Section 4.3 that a house listed in month  $t$  sells during that month.

below 50% and are close to the rates found by Anenberg and Bayer (2013) when performing a similar match for the Los Angeles metro area. This evidence supports our claim that much of the short-term investment represents buyers entering and then exiting the local housing market.

### Cross-sectional Evidence

This shift in the composition of buyers and sellers toward shorter holding periods during the boom years correlates highly with changes in total volume across local markets. This correlation can be seen clearly in Figure 5, which presents 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 ( $\geq 3$  years) in Panel (b).<sup>31</sup> 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 33%. Thus, as predicted broadly by Proposition 1 and specifically by Proposition 5, shifts in the distribution of holding periods of buyers and sellers over the course of the cycle appear to be a major determinant of changes in total transaction volume.

Because expected capital gains increase demand through the extensive margin, Proposition 1 suggests that volume increases more strongly among buyers with low housing utility. Proposition 5 formalizes this intuition by showing that the speculative share of purchases increases during the boom. 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 these predictions by examining whether non-occupant purchases rose more than occupant purchases from 2000 to

---

<sup>31</sup>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.

2005.<sup>32</sup>

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.<sup>33</sup> 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. Similarly to the patterns we documented for short holding periods, we find that 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.

The MSA-level results from Figures 5 and 6 also hold across local neighborhoods *within* MSAs. Table 3 repeats the cross-MSA analysis using a ZIP-code-level panel data set constructed from the same underlying sample of transactions.<sup>34</sup> The first three columns report regression results that are directly analogous to Panels (a)-(c) of Figure 5; the latter three columns report results that are directly analogous to Panels (a)-(c) of Figure 6. All specifications include a full set of MSA fixed effects, so the coefficient estimates reflect only within-MSA variation in transaction volume. Echoing the earlier cross-MSA results, changes in short-holding-period and non-occupant volume are more strongly correlated with changes in total volume than are changes in long-holding-period and occupant volume. The third (respectively sixth) column indicates that 24% (respectively 36%) of the cross-ZIP-code variation within an MSA in the growth of volume over this period arises from differences in the

---

<sup>32</sup>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 and Zwick, 2017).

<sup>33</sup>In related work, Haughwout et al. (2011) use an alternative proxy for nonowner occupancy based on the number of first-lien mortgages present on an individual’s credit report. They also find a large increase in non-occupant purchases during this time period.

<sup>34</sup>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.

number of short-holding-period (respectively non-occupant) transactions.

### 4.3 The Joint Dynamics of Prices and Volume

Figure 7 plots prices, transaction volume, and unsold listings over the 2000–2011 US housing cycle; transaction volume equals the monthly count of all transactions in our data, and unsold listings equal the inventory of listed existing houses reported by the NAR (both series are seasonally adjusted using month-of-year fixed effects). We mark the quiet as the period between the peak of volume and the last peak of prices before its pronounced decline. The existence of this quiet period confirms the prediction of Proposition 3, and the contemporaneous sharp increase in unsold listings matches the path of simulated unsold inventory in Figure 3(c) that was raised as a possibility by Proposition 3. We now examine the joint behavior of prices, volume, and listings more carefully as they relate to the predictions of our model.

#### The Lead–Lag Relationship

Propositions 2–4 predict that prices and volume both go through a boom and bust cycle, with the volume cycle leading the price cycle. In Figure 8, we present evidence that this relationship holds on average across MSAs in our sample. To do so, 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 construct a monthly panel of log house prices and transaction volume at the MSA level running from January 2000 to December 2011; volume equals the total number of transactions in our data in a given month and MSA divided by that MSA’s housing stock in the 2000 Census. Using this panel, we run a series of simple regressions of the form

$$p_{i,t} = \beta_{\tau} v_{i,t-\tau} + \eta_{i,t}, \tag{12}$$

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.

The coefficient  $\beta_{\tau}$  provides an estimate of how movements in volume around MSA–



calendar month averages at a  $\tau$ -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 ( $\tau = -12$ ). Figure 8 plots the implied correlation from each regression along with its 95% confidence interval.<sup>35</sup> 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.

## Listings over 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, Proposition 3 shows that 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 empirical movements in volume, we use the methodology presented by Ngai and Sheedy (2016) to measure each of these rates over the course of the housing 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  $L_t = n_t(K_t - I_t)$ ;  $K$  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. We average these probabilities within each year and plot the resulting averages in Figure 9.<sup>36</sup>

Figure 9 shows that the listing rate  $n_t$  rose 40% during the 2000–2005 boom, whereas the rate  $s_t$  at which listings sold remained flat. In contrast, the decline in volume during the 2005–2006 quiet is driven primarily by  $s_t$ , which contracts by about 30% while  $n_t$  remains high. As is suggested 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 the drivers of volume during the boom and the quiet.

<sup>35</sup>The implied correlation equals  $\beta_\tau \text{std}(v_{i,t-\tau})/\text{std}(p_{i,t})$ , where  $v_{i,t-\tau}$  and  $p_{i,t}$  are the demeaned regressors. Confidence intervals for each estimate come from standard errors clustered at the month level.

<sup>36</sup>In discrete time,  $V$  and  $L$  are the 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 et al. (2016) plot the same series for  $s_t$ .

## Initial Volume and the Size of the Cycle

Proposition 6 shows that an increase to the distribution of expected holding times  $f(\cdot)$  raises both the level of steady-state volume as well as the magnitude of the increase in prices during the boom and subsequent decreases in prices during the bust. Rather than measure  $f(\cdot)$  directly, we test whether steady-state volume and the amplitude of the price cycle are positively correlated across MSAs. Our measure of steady-state volume equals the number of existing home sales in 2000 as a share of the housing stock.<sup>37</sup> The boom in each MSA equals the percentage change in prices between January 2000 and the month in which prices peaked, 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.<sup>38</sup>

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 4 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.

---

<sup>37</sup>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.

<sup>38</sup>We restrict the price 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.

## 5 Conclusion

This paper develops a tractable model of asset bubbles that nevertheless generates a rich, joint dynamic of prices and transaction volume. Theoretical and empirical results on the composition of buyers and sellers during a bubble suggest that investigating the speculative dynamics of volume can help us understand what factors drive bubbles. In particular, short-term investors have the capacity to destabilize financial markets. We documented the importance of short-term investors in the 2000–2011 housing cycle. Studying this activity in other asset markets and historical episodes would be illuminating.

Our focus on short-term investors raises two additional lines of inquiry:

First: Do the expansions in credit that accompany asset price booms appeal disproportionately to short-term investors? Barlevy and Fisher (2011) document a strong correlation across US 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 asset price booms documented by Di Maggio and Kermani (2015), Favara and Imbs (2015), and Rajan and Ramcharan (2015).

Second: Do policies that aim to achieve financial-market stability work better if they discourage the participation of short-term investors? For instance, consider the financial transactions tax proposed by Tobin (1978), supported by Stiglitz (1989) and Summers and Summers (1989), and analyzed theoretically by Dávila (2015). If the incidence of this tax falls entirely on buyers, then the tax burden is independent of the investment horizon; if the incidence falls entirely on sellers, then the tax burden is larger in present-value terms for short-term investors who plan to resell quickly. Our model suggests that transaction taxes discourage bubbles more powerfully when their incidence falls more strongly on sellers. One policy that discourages short-term investors more directly is the short-term capital gains tax, and our model provides a rationale for this policy. Any tax that discourages short-term investors will also discourage the liquidity provision and, in the case of the housing market, the residential investment they provide. We hope that future work will weigh all of these effects to guide policy carefully.

# A Omitted Proofs of Mathematical Statements

## Lemma 1

By (1),  $E[(P_{t+\tau} - P_t)/(\tau P_t) \mid \omega_t] = \gamma(\omega_t)g(\tau)/\tau$ . The cross-partial in Lemma 1 equals  $\gamma'(\omega_t)(g(\tau)/\tau)'$ . Because  $\gamma'(\omega_t) > 0$ , this cross-partial is negative for all  $\tau > 0$  if and only if Assumption 1(a) holds.

## Lemma 2

If Assumption 1(c) fails, then we can find  $\omega_t$  such that  $\gamma(\omega_t)g'(0) > r$ . Then  $E[P_{t+\tau}/P_t \mid \omega_t] - e^{r\tau}$  equals 0 at  $\tau = 0$  and has a positive derivative with respect to  $\tau$  at  $\tau = 0$ , which means that  $E[P_{t+\tau}/P_t \mid \omega_t] > e^{r\tau}$  for some  $\tau > 0$ . Now suppose Assumption 1(c) holds. For  $\tau > 0$ ,  $g(\tau) = \int_0^\tau g'(\tau_0)d\tau_0 < \int_0^\tau g(\tau_0)/\tau_0 d\tau_0 < g'(0)\tau$ . As a result, for all  $\omega_t \in \mathbb{R}$   $E[P_{t+\tau}/P_t \mid \omega_t] = 1 + \gamma(\omega_t)g(\tau) < 1 + \gamma(\omega_t)g'(0)\tau \leq 1 + r\tau < e^{r\tau}$ . The last inequality follows because  $1 + r\tau$  and  $e^{r\tau}$  coincide for  $\tau = 0$  and the derivative of the latter exceeds that of the former for all  $\tau > 0$ .

## Lemma 3

By (6), a potential buyer buys if and only if

$$\delta \geq rP_t \left( 1 - \frac{\lambda}{r} \int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau} \left( E \left[ \frac{P_{t+\tau}}{P_t} \mid \omega_t \right] - 1 \right) d\tau \right).$$

Substituting (1) reduces the integral to  $\int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau} \gamma(\omega_t)g(\tau)d\tau$  and then integrating by parts further reduces it to  $\int_0^\infty e^{-(r+\lambda)\tau} \gamma(\omega_t)g'(\tau)d\tau$ . The measure of potential buyers at  $t$  of type  $\lambda$  whose flow utility exceeds some  $\delta_0 > 0$  equals  $f(\lambda)A_t\delta_0^{-\epsilon}$ , so we are done.

## Proposition 1

We prove the stronger statement of Proposition 1 in which each derivative with respect to  $\omega_t$  is replaced by the right-sided derivative  $\partial_+/\partial\omega_t$  or the left-sided derivative  $\partial_-/\partial\omega_t$  throughout. We write the proof in terms of  $\partial_+/\partial\omega_t$ ; the identical proof holds replacing those partials with  $\partial_-/\partial\omega_t$ . We use this more general form of the proposition in the proof of Proposition 2.

For notational ease, we define  $i(\lambda)$  to be the expression such that (7) produces  $\sigma_\lambda(\omega_t) = -\epsilon \log(1 - \gamma(\omega_t)i(\lambda))$ . From Lemma 2,  $1 + \gamma(\omega_t)g(\tau) < e^{r\tau}$  for  $\tau > 0$ , so  $\gamma(\omega_t)i(\lambda) < \lambda/r \int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau}(e^{r\tau} - 1)d\tau = 1$ . For  $\omega_t$  such that  $\partial_+\gamma(\omega_t)/\partial\omega_t > 0$ ,  $\partial_+\sigma_\lambda(\omega_t)/\partial\omega_t = \epsilon(\partial_+\gamma(\omega_t)/\partial\omega_t)i(\lambda)/(1 - i(\lambda)) > 0$  because  $i(\lambda) > 0$  by Assumption 1(d). Differentiating again yields  $\partial\partial_+\sigma_\lambda(\omega_t)/\partial\omega_t\partial\lambda = \epsilon(\partial_+\gamma(\omega_t)/\partial\omega_t)i'(\lambda)/(1 - i(\lambda))^2$ . Integrating by parts and applying Assumption 1(a) yields

$$ri'(\lambda) = \int_0^\infty e^{-(r+\lambda)\tau} g'(\tau)d\tau - \int_0^\infty \lambda e^{-(r+\lambda)\tau} \tau g'(\tau)d\tau > \frac{r}{r + \lambda} \int_0^\infty e^{-(r+\lambda)\tau} g'(\tau)d\tau > 0,$$

where the final inequality follows because the last integral equals  $r^2/(\lambda(r + \lambda))i(\lambda) > 0$ .

To prove (c), we first note that  $V_t(\lambda)/V_t = D_t(\lambda)/D_t$  by (3). By Lemma 3,  $D_t(\lambda)/D_t = f(\lambda)\Sigma_\lambda(\omega_t)/\int_0^\infty f(\lambda')\Sigma_{\lambda'}(\omega_t)d\lambda'$ . Thus, we must show that

$$\frac{\partial_+ \int_0^\lambda \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda'}{\partial\omega_t \int_0^\infty \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda'} \leq 0$$

for all  $\lambda$ . Differentiating, subtracting parts of the integral that appear on each side, and multiplying and dividing by  $Sigma_{\lambda'}$  reduces this inequality to

$$\begin{aligned} \int_\lambda^\infty \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \int_0^\lambda \frac{\partial_+\sigma_{\lambda'}(\omega_t)}{\partial\omega_t} \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \leq \\ \int_0^\lambda \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \int_\lambda^\infty \frac{\partial_+\sigma_{\lambda'}(\omega_t)}{\partial\omega_t} \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda'. \end{aligned}$$

The first part of the proof showed that  $\partial_+\sigma_\lambda(\omega_t)/\partial\omega_t > 0$  and  $\partial\partial_+\sigma_\lambda(\omega_t)/\partial\omega_t\partial\lambda > 0$  for all  $\lambda$ . Thus  $\partial_+\sigma_{\lambda'}(\omega_t)/\partial\omega_t$  increases in  $\lambda'$  and is positive, letting us reduce the inequality to

$$\begin{aligned} \int_\lambda^\infty \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \int_0^\lambda \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \leq \\ \int_0^\lambda \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda' \int_\lambda^\infty \Sigma_{\lambda'}(\omega_t)f(\lambda')d\lambda', \end{aligned}$$

which holds with equality. Strict inequality results if and only if  $\text{supp } f$  contains both a point above and a point below  $\lambda$ .

## Demand Target

Because prices remain constant,  $D_t = \bar{D}$  by (8). Because potential buyers expect prices to remain constant,  $\Sigma_\lambda(\omega_t) = 1$  for all  $\lambda$  and  $\omega_t$  by (7). Thus, by Lemma 3,  $D_t(\lambda)/D_t = f(\lambda)$  for all  $t$ . Because  $L_t = D_t$ ,  $V_t = L_t$  by (3), so  $V_t(\lambda) = f(\lambda)\bar{D}$ . In a steady state, each  $S_t(\lambda)$  remains constant, so  $S_t(\lambda) = V_t(\lambda)/\lambda = \bar{D}f(\lambda)/\lambda$ . Because  $V_t = L_t$ ,  $I_t = 0$ . As a result, the housing stock is comprised entirely of stayers, so  $1 = \int_0^\infty S_t(\lambda)d\lambda = \bar{D} \int_0^\infty f(\lambda)/\lambda d\lambda$ . Solving for  $\bar{D}$  provides (9).

## Lemma 4

By Lemma 3, total demand across all potential buyers equals

$$D_t = A_t(rP_t)^{-\epsilon}\Sigma(\omega_t), \tag{A1}$$

where  $\Sigma(\omega_t) \equiv \int_0^\infty \Sigma_\lambda(\omega_t)f(\lambda)d\lambda$  is aggregate speculative demand. Define  $\bar{p} \equiv \log(A^f)/\epsilon - \log(r) - \log(\bar{D})/\epsilon$  and  $\sigma(\omega) \equiv \log \Sigma(\omega)$ . Substituting (A1) into (8) yields

$$\dot{p} = c\epsilon(\bar{p} - p) + c\sigma(\omega). \tag{A2}$$

Substituting (A2) into (11) gives

$$\dot{\omega} = \mu c \epsilon (\bar{p} - p) + \mu c \sigma(\omega) - \mu \omega. \quad (\text{A3})$$

With the initial conditions  $p_0 = \bar{p} - \log(A^f/A^i)/\epsilon$  and  $\omega_0 = 0$ , (A2) and (A3) specify the joint dynamics of  $p$  and  $\omega$ . Figure A2 illustrates the corresponding phase diagram. The following two paragraphs justify the way it was drawn.

The  $\dot{p} = 0$  locus is given by  $p = \bar{p} + \sigma(\omega)/\epsilon$ . For a given  $\omega$ ,  $\dot{p} < 0$  for  $p$  above this locus and  $\dot{p} > 0$  for  $p$  below this locus. By Assumption 3,  $\sigma(\omega) = 0$  for  $\omega \leq 0$  and  $\sigma'(\omega) > 0$  for  $\omega > 0$ , so the  $\dot{p} = 0$  locus equals  $p = \bar{p}$  for  $\omega < 0$  and increases for  $\omega > 0$ . The  $\dot{\omega} = 0$  locus is given by  $p = \bar{p} + \sigma(\omega)/\epsilon - \omega/(c\epsilon)$ . For a given  $\omega$ ,  $\dot{\omega} < 0$  for  $p$  above this locus and  $\dot{\omega} > 0$  for  $p$  below this locus. For  $\omega < 0$ , this locus equals a decreasing line, and for  $\omega > 0$ , this locus lies beneath the  $\dot{p} = 0$  locus. The right slope of this locus at 0 equals  $\lim_{\omega \rightarrow 0^+} \sigma'(\omega)/\epsilon - 1/(c\epsilon)$ , which  $> 0$  as long as  $\lim_{\omega \rightarrow 0^+} \sigma'(\omega) > \epsilon/\min(c\epsilon, \mu)$ . This inequality holds because differentiating  $\sigma$  yields

$$\sigma'(\omega) = \frac{\epsilon \gamma'(\omega) \int_0^\infty (1 - \gamma(\omega) i(\lambda))^{-\epsilon-1} i(\lambda) f(\lambda) d\lambda}{\int_0^\infty (1 - \gamma(\omega) i(\lambda))^{-\epsilon} f(\lambda) d\lambda},$$

where  $i(\lambda)$  is as defined in the proof of Proposition 1. Taking limits yields  $\lim_{\omega \rightarrow 0^+} \sigma'(\omega) = (\lim_{\omega \rightarrow 0^+} \gamma'(\omega)) \epsilon \int_0^\infty i(\lambda) f(\lambda) d\lambda > \epsilon/\min(c\epsilon, \mu)$ , where Assumption 3 is used.

We have drawn the phase diagram such that the  $\dot{p} = 0$  locus is bounded for large  $\omega$  and that the  $\dot{\omega} = 0$  locus asymptotes to a decreasing linear function for large  $\omega$ . These features hold as long as  $\sigma(\omega)$  is bounded. As shown in the proof of Lemma 2,  $g(\tau) < g'(0)\tau$  for  $\tau > 0$ , so  $i(\lambda) < (\lambda g'(0)/r) \int_0^\infty (r + \lambda) e^{-(r+\lambda)\tau} \tau d\tau = (g'(0)/r) \lambda/(r + \lambda)$ . Therefore Assumption 1(c) implies that  $\sigma(\omega) < \log \int_0^\infty (r/(r + \lambda))^{-\epsilon} f(\lambda) d\lambda$ , which exists because  $\int_0^\infty \lambda^\epsilon f(\lambda) d\lambda$  exists.

Tracing the system from the initial point makes it clear that  $p$  must decrease in finite time. Because  $\bar{p} > p_0$  and  $\omega_0 = 0$ ,  $p$  and  $\omega$  increase at  $t = 0$  by (A2) and (A3). Eventually, the right  $\dot{\omega} = 0$  locus is reached because this locus goes to  $-\infty$  for large  $\omega$  due to its linearity. Next,  $\omega$  begins to increase while  $p$  continues increasing. The right  $\dot{p} = 0$  locus is then reached because it is bounded from above. After this point,  $\omega$  continues decreasing while  $p$  begins to decrease, as desired.

## Proposition 2

Because  $\bar{p} > p_0$  and  $\omega_0 = 0$ , (A2) implies that  $\dot{p}_0 > 0$ . By the definition of  $t_2$ ,  $\dot{p}_t > 0$  for  $t \in (0, t_2)$  and  $\dot{p}_{t_2} = 0$ . It follows that  $p_t$  increases on  $(0, t_2] \supset (0, t_1]$ . Because  $\dot{p}_t = 0$  for  $t < 0$ , the definition of  $\omega$  in (10) implies that  $\omega_t > 0$  for  $t \in (0, t_2] \supset (0, t_1]$ .

To show that  $\ddot{p} > 0$  at least initially, we show that  $\ddot{p}_0 > 0$ . Substituting (A1) into (8), differentiating, and using (11) yields  $\ddot{p}/c = -\epsilon \dot{p} + \sigma'(\omega) \mu (\dot{p} - \omega)$  for  $\omega > 0$ . Because  $\omega_0 = 0$  and  $\lim_{\omega \rightarrow 0^+} \sigma'(\omega) > \epsilon/\min(c\epsilon, \mu)$  as shown in the proof of Lemma 4,  $\ddot{p}_0 > 0$ . At least initially,  $\dot{\omega} > 0$  because  $\dot{\omega}_0 > 0$  given (A3) and the fact that  $\bar{p} > p_0$  and  $\omega_0 = 0$ .

To show that  $I = 0$  at least initially, we show that  $D > L$  at least initially and that  $D_t$  and  $L_t$  are continuous for  $t > 0$ . From these facts, (5) implies that  $\dot{I} = 0$  for some non-empty interval at the beginning of the boom, implying that  $I = 0$  during this interval. By (8) and (10),  $p_t$  and  $\omega_t$  are continuous, so by (A1)  $D_t$  is continuous for  $t > 0$ . The jump in  $A$  at

$t = 0$  implies that  $\lim_{t \rightarrow 0^+} D_t = A^f \bar{D} / A^i > \bar{D}$ . By (2),  $L_t$  is continuous if each  $S_t(\lambda)$  is, and by (4)  $S_t(\lambda) = \int_{-\infty}^t e^{-\lambda(t-\tau)} V_\tau(\lambda) d\tau$  so it is continuous. Therefore  $L_t$  is continuous, so in particular  $\lim_{t \rightarrow 0^+} L_t = \bar{D}$ . It follows that at least initially  $D$  and  $L$  are continuous and  $D > L$ , as claimed.

To prove that  $\dot{V} > 0$  at least initially when  $|\text{supp } f| > 1$ , we show that  $\dot{V}_0 = 0$  and  $\ddot{V}_0 > 0$  in this case. By (3),  $V_t = L_t$  when  $D_t > L_t$  and  $I_t = 0$ , both of which hold initially. As a result, we show that  $\dot{L}_0 = 0$  and  $\ddot{L}_0 > 0$ . We have  $\dot{L}_0 = \int_0^\infty \lambda \dot{S}_0(\lambda) d\lambda$ . But  $\dot{S}_0(\lambda) = \dot{V}_0(\lambda) - \lambda S_0(\lambda) = 0$  because  $V(\lambda)$  and  $S(\lambda)$  are continuous at  $t = 0$ . So  $\dot{L}_0 = 0$ . Taking another derivative yields  $\ddot{L}_0 = \int_0^\infty \lambda \ddot{S}_0(\lambda) d\lambda$ . We have  $\ddot{S}_0(\lambda) = \dot{V}_0(\lambda) - \lambda \dot{S}_0(\lambda) = \dot{V}_0(\lambda)$ . Thus  $\ddot{L}_0 = \int_0^\infty \lambda \dot{V}_0(\lambda) d\lambda$ . Because  $\dot{\omega}_0 > 0$ , by Proposition 1(c)  $\int_0^\lambda \dot{V}_0(\lambda') d\lambda' \leq 0$  for all  $\lambda > 0$  and strictly for some  $\lambda > 0$ , so  $V_0(\lambda) + \dot{V}_0(\lambda)$  strictly first-order stochastically dominates  $V_0(\lambda)$  as distributions. The former must have a larger mean as a result, so  $\int_0^\infty \lambda \dot{V}_0(\lambda) > 0$ , giving  $\ddot{L}_0 > 0$ .

We now show that  $V_t = \bar{D}$  for  $t \in [0, t_2]$  when  $|\text{supp } f| = 1$ , which proves that volume remains constant during the boom in this case. By (9),  $f$  must be a single mass point on  $\lambda = \bar{D}$ . By (2),  $L_t = \bar{D} S_t(\bar{D}) \leq \bar{D}$  because  $S_t \leq 1$ , the total housing stock. Because  $\dot{p}_t > 0$  for  $t \in [0, t_2)$ ,  $D_t > \bar{D} \geq L_t$  for  $t \in [0, t_2)$  by (8). It follows from (3) that  $V_t = L_t$  for  $t \in [0, t_2]$  because  $D_t \geq L_t$  over this interval. As a result, (2) and (4) combine into the equation  $\dot{S}_t(\bar{D}) = 0$  for  $t \in [0, t_2)$ . It follows that  $S_t(\bar{D}) = S_0(\bar{D}) = 1$  and that  $V_t = L_t = \bar{D}$  for  $t \in [0, t_2]$ , as claimed.

### Proposition 3

Because  $\dot{p}_{t_2} = 0$ ,  $D_{t_2} = \bar{D}$  by (8). By (3),  $V_t \leq D_t$  for all  $t$ , so  $V_{t_2} \leq \bar{D}$ . As shown in the discussion of the demand target in this appendix,  $V = \bar{D}$  in steady state, so  $V_0 = \bar{D}$ . If  $|\text{supp } f| > 1$ , then Proposition 2 implies that  $V^{max} > \bar{D} = V_0$  because  $\dot{V}_0 > 0$  at least initially during the boom. It follows that  $V^{max} > V_{t_2}$  and that  $t_1 < t_2$ . If  $|\text{supp } f| = 1$ , then  $V_t = \bar{D}$  for all  $t \in [0, t_2]$  as shown at the end of the proof of Proposition 2. In this case  $t_1 = t_2$  and the quiet does not exist.

As shown at the beginning of the proof of Proposition 2,  $p_t$  increases and  $\omega_t > 0$  on  $(0, t_2]$ . From that proof,  $\ddot{p}/c = -\epsilon \dot{p} + \sigma'(\omega) \mu (\dot{p} - \omega)$ , so  $\ddot{p}_{t_2} = -\sigma'(\omega_{t_2}) \mu \omega_{t_2} < 0$  and  $\ddot{p} < 0$  at least ultimately during the quiet. From (11),  $\dot{\omega}_{t_2} = -\mu \omega_{t_2} < 0$ , so  $\dot{\omega} < 0$  at least ultimately during the quiet. By the definition of  $t_1$ ,  $V_{t_1} > V_{t_2}$  when  $t_1 < t_2$ . Figure 3 confirms the possibility that  $I$  may be positive during the quiet.

### Proposition 4

The proof of Lemma 4 showed that prices begin to decline when the right  $\dot{p} = 0$  locus shown in Figure A2 is reached. We prove that after this point,  $p$  limits to  $\bar{p}$  and that  $\dot{p} < 0$  or  $\dot{p} = 0$  for isolated instants until this limit is reached. Given that result, (8) implies that  $D_t < \bar{D}$  except at isolate instants at which  $D_t = \bar{D}$  until the limit is reached. Because (3) implies that  $V_t \leq D_t$ , the same statement applies to  $V_t$ .

Right after  $t_2$  when the  $\dot{p} = 0$  locus is first reached, the system shown in Figure A2 moves left so that  $\dot{p} < 0$ . The price decline continues until either the  $\omega = 0$  axis is reached or convergence to  $(0, \bar{p})$  occurs. Before either of these events, the system can never fall below

the right  $\dot{p} = 0$  locus, and any intersection with it occurs for but an instant. As a result,  $p$  continues to decrease until convergence to the steady-state or until  $\omega = 0$  but  $p > \bar{p}$ . In the first case, we are done. In the second case, we know that  $\omega < 0$  right after  $\omega = 0$  from examining the phase diagram. It is clear that  $\omega$  remains below 0 for the rest of time until steady-state is reached, as the system remains weakly above the left  $\dot{p} = 0$  locus. In this region,  $\sigma(\omega) = 0$ , so (A2) becomes  $\dot{p} = c\epsilon(\bar{p} - p)$ . Thus,  $\dot{p} < 0$  for the rest of time, as  $p > \bar{p}$  when  $\omega = 0$ .

Figure 3 confirms the possibility that  $I > 0$  during the bust when  $p > \bar{p}$ .

### Proposition 5

Propositions 2 and 3 showed that  $\omega > 0$  during the boom and quiet, initially rising during the boom and ultimately falling during the quiet. From Figure A2, it is clear that during the bust—that is, after the  $\dot{p} = 0$  is reached— $\omega$  monotonically decreases until reaching 0. To prove Proposition 5, it is sufficient to show that the short-term buyer share and speculative share equal their steady-state values for  $\omega \leq 0$  and strictly increase in  $\omega$  for  $\omega > 0$ .

We begin with the speculative share, which equals  $\int_0^\infty (1 - \Sigma_\lambda(\omega_t)^{-1})f(\lambda)d\lambda$  as explained in the text. For all  $\lambda > 0$ , Assumption 3 implies that  $\Sigma_\lambda(\omega) = 1$  for  $\omega \leq 0$  and  $\Sigma'_\lambda(\omega) > 0$  for all  $\omega > 0$ ; the second fact holds because that which  $\gamma(\omega_t)$  multiplies in (7) is positive as shown in the proof of Proposition 1. It follows that the speculative share equals its steady-state value (the value when  $\omega = 0$ ) of 0 when  $\omega \leq 0$  and increases in  $\omega$  for  $\omega > 0$ .

We turn now to the share of buyers for whom  $\lambda > \lambda'$ . For  $\omega > 0$ ,  $\gamma'(\omega) > 0$  by Assumption 3, so Proposition 1(c) implies that  $\int_0^{\lambda'} V_t(\lambda)d\lambda/V_t$  strictly decreases with  $\omega_t$ , so the share of buyers for whom  $\lambda > \lambda'$  strictly increases with  $\omega_t$ . When  $\omega_t \leq 0$ , Lemma 3 implies that  $V_t(\lambda)/V_t = D_t(\lambda)/D_t = f(\lambda)$ , its steady-state value.

### Proposition 6

We begin by proving part (a). Because  $V_0 = \bar{D} = (\int_0^\infty \lambda^{-1}f(\lambda)d\lambda)^{-1}$ , an increase to  $f(\cdot)$  increases  $V_0$  because  $1/\lambda$  decreases. We next show that  $\sigma(\cdot)$  increases pointwise if  $f(\cdot)$  increases. We have  $\partial\Sigma_\lambda/\partial\lambda = \epsilon\gamma(\omega)i'(\lambda)(1 - i(\lambda)\gamma(\omega))^{-\epsilon-1}$ , which equals 0 for  $\omega \leq 0$  and is positive otherwise. An increase to  $f(\cdot)$  thus increases  $\sigma = \log \int_0^\infty \Sigma_\lambda f(\lambda)d\lambda$  for  $\omega > 0$  and keeps it constant at  $\omega \leq 0$ .

We now show that  $P^{max}$  increases in  $\sigma$ .  $P_t = P^{max}$  when the system first intersects the  $\dot{p} = 0$  locus shown in Figure A2. Because this locus is given by  $p = \bar{p} + \sigma(\omega)/\epsilon$ , it shifts up for  $\omega > 0$  due to an increase in  $\sigma$ . If the path of  $(p, \omega)$  also shifts to the right, then the intersection with this locus must occur at a higher value of  $p$ . Thus we prove that the path does shift to the right. This shift occurs if  $dp/d\omega$  decreases as  $\sigma$  increases. This derivative equals  $\dot{p}/\dot{\omega} = (\mu(1 - \omega/\bar{p}))^{-1}$ . Because  $\omega > 0$  and  $\dot{p} > 0$  before the  $\dot{p} = 0$  locus is reached, we must show that  $\omega/\bar{p}$  decreases in  $\sigma$ . By (A2),  $\dot{p}$  increases in  $\sigma$ , so we are done.

As the final step in the proof of part (a), we show that  $P_0$  and  $P_\infty$  are independent of  $f$ . As shown in the proof of Proposition 4,  $\lim_{t \rightarrow \infty} p_t = \bar{p}$ , which does not depend on  $f$ . The initial condition  $p_0$  given in the proof of Lemma 4 also does not depend on  $f$ . Thus  $\sigma$  has no bearing on  $P_0$  or  $P_\infty$ .



We turn now to part (b). Because  $f_A < f_B$ , by Proposition 6  $V_0$  and  $P^{max}$  are larger under  $f_B$  while the steady-state prices  $P_0$  and  $P_\infty$  stay unchanged. Under  $f_A$ ,  $L_t = \lambda^* S_t(\lambda^*) \leq \lambda^*$ , where  $\lambda^*$  is the sole member of  $\text{supp } f_A$ , and  $V_0 = \lambda^*$ . Therefore  $V^{max}/V_0 = 1$  under  $f_A$ . By Proposition 2,  $V^{max}/V_0 > 1$  under  $f_B$  because  $|\text{supp } f_B| > 1$ .

## B Microfoundation of Price Adjustment Rule

This appendix presents one possible microfoundation for the price adjustment rule (8). This microfoundation is consistent with (3), which specifies how housing is allocated between sellers and interested buyers at each time.

### Market Mechanism

At each time, movers may post a listing price; we call listing movers “sellers.” A centralized mechanism matches potential buyers to sellers. We assume there exists a cutoff price  $P_t^c(\delta, \lambda)$  such that a potential buyer of type  $(\delta, \lambda)$  purchases from a matched seller if and only if the seller’s listing price is no greater than  $P_t^c(\delta, \lambda)$ . We verify the existence of  $P_t^c(\cdot, \cdot)$  below.

The matching mechanism works as follows. When the stock of sellers equals 0 (but the flow may still be positive), the potential buyer with the highest cutoff price matches to the seller with the lowest listing price, then the potential buyer with the second highest cutoff price matches to the seller with the second lowest listing price, et cetera until either the pool of potential buyers or the pool of sellers has been exhausted.<sup>1</sup> When the stock of sellers is positive, this mechanism is applied first to the oldest cohort of movers and then is applied successively to younger cohorts of movers until either the pool of potential buyers or the entire pool of sellers has been exhausted.<sup>2</sup>

Each matched pair trades at the seller’s listing price if and only if that price does not exceed the potential buyer’s cutoff price. If trade does not occur, the potential buyer permanently exits the market and the seller remains a mover.

### Information

Because trade may occur at different prices at the same time, some care is needed to define “the market price” at a given time. Letting  $i$  index sellers at  $t$ , we define the market price by  $P_t = e^{\int \log P_{i,t} di}$ , where  $P_{i,t}$  is seller  $i$ ’s price. Under this definition, the log of the market price equals the average of the logs of the listing prices.

Before the posting of list prices, movers and potential buyers at  $t$  receive two pieces of information: the price history  $\{P_{t'} \mid t' \leq t - dt\}$  and the number  $\tilde{D}_t$  of current potential buyers whose cutoff price is at least  $P_{t-dt}$ . We will take the limit as  $dt \rightarrow 0$  to describe the continuous-time model.

### Preferences

Movers act to maximize their utility after the matching process has occurred. We assume that they have lexicographic preferences over holding a house, preferring any outcome in which they do not own a house to all outcomes in which they do own a house. This extreme assumption strongly motivates movers to sell their houses at each time—selling a house is the only way to divest ownership of it (there is no free disposal)—and can be thought of as the limit as the “holding costs” assumed in housing search models (Piazzesi and Schneider,

---

<sup>1</sup>This mechanism presumes that  $\delta$  and  $\lambda$  are observable.

<sup>2</sup>The assumption that older cohorts of sellers are matched before younger cohorts is made in the stock-flow matching literature (Ebrahimi and Shimer, 2010).

2009; Guren and McQuade, 2015; Han and Strange, 2015) go to infinity. Within each class of outcomes, utility is linear in consumption.

Movers understand the market mechanism described above, but there does not exist common knowledge of the rationality of movers, leading to uncertainty over the prices that other movers will post. Following Gilboa and Schmeidler (1989), we assume that each mover displays “ambiguity aversion” over this uncertainty, acting to maximize her utility contingent on the pricing decisions made by other movers that would minimize her utility. In this setting, the mover’s decision is robust to the range of possible actions by other movers.

Potential buyers maximize the present value of utility discounted at a constant rate  $r$ , where utility is linear in consumption and the flow utility  $\delta$  received while a stayer.

## Rigidities

A fraction  $e^{-\beta dt}$  of movers are inattentive to  $\tilde{D}_t$  and assume that it equals its steady-state value  $\bar{D}$ . The remaining  $1 - e^{-\beta dt}$  of movers use the true value of  $\tilde{D}_t$  when setting prices. The allocation of movers to each group is independent of their prior histories, so in continuous time attentiveness arrives with Poisson intensity  $\beta$ . This attention rigidity is similar to the slow diffusion of news in Hong and Stein (1999) and the sticky information sets in Mankiw and Reis (2002).

Potential residents and movers hold rigid beliefs concerning the the number of movers at each moment. Instead of inferring how this number may vary over time, agents assume that the flow of new movers and the stock of existing movers remain constant over time at their values in the unique steady state in which the stock of movers equals 0. As shown in Appendix A, the flow of new movers in this steady state equals the value  $\bar{D}$  defined by (9).

As stated in Assumption 2, potential buyers assume that upon becoming a mover, they sell instantaneously at the prevailing market price. A potential buyer’s expectation of future market prices is given by  $E_t[P_{t+\tau} | \delta, \lambda, P_{i,t}, \omega_{t-dt}, \tilde{D}_t] = (1 + \gamma(\omega_{t-dt})g(\tau))P_{i,t}$  for  $\tau > 0$ ;  $P_{i,t}$  is the listing price to which the potential buyer is matched,  $\omega_{t-dt} = \omega(\{P_{t'} | t' \leq t - dt\})$ , and  $\gamma$  and  $g$  satisfy the conditions in Assumption 1.

## Cutoff Prices

The following lemma shows that, as assumed above, a potential buyer purchases the house to which she is matched if and only if the listing price does not exceed a time- and type-dependent cutoff.

**Lemma B1.** *A potential buyer of type  $(\delta, \lambda)$  buys if and only if  $P_{i,t} \leq \delta/r\Sigma_\lambda(\omega_{t-dt})^{-1}$ .*

*Proof.* The potential buyer anticipates selling at the market price with certainty when becoming a mover, so her expected utility of becoming a mover at  $t + \tau$  equals  $(1 + \gamma(\omega_{t-dt})g(\tau))P_{i,t}$ , where  $P_{i,t}$  is the listing price of the house to which she is matched at  $t$ . The expected utility of buying this house equals  $\int_0^\infty \lambda e^{-\lambda\tau} (\int_0^{\tau'} e^{-r\tau'} \delta d\tau' + e^{-r\tau} (1 + \gamma(\omega_{t-dt})g(\tau))P_{i,t}) d\tau$ . She buys if and only if this quantity is at least  $P_{i,t}$ , which occurs exactly when  $rP_{i,t}(1 - \lambda\gamma(\omega_{t-dt}))/r \int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau} g(\tau) d\tau \leq \delta$ . As shown in the proof of Lemma 2,  $\gamma(\omega_{t-dt})g(\tau) < e^{r\tau} - 1$  for all  $\tau > 0$  and  $\omega_{t-dt}$ . It follows that the term in parentheses exceeds  $1 - \lambda/r \int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau} (e^{r\tau} - 1) d\tau = 0$ , so it can be divided by to obtain the

decision rule  $P_{i,t} \leq \delta/r(1 - \lambda\gamma(\omega_{t-dt})/r \int_0^\infty (r + \lambda)e^{-(r+\lambda)\tau} g(\tau) d\tau)^{-1}$ . Differentiating by parts as in the proof of Lemma 3 and then applying (7) yields the cutoff rule in Lemma B1.  $\square$

Lemma B1 shows that the cutoff assumed above equals  $P_t^c(\delta, \lambda) = \delta/r\Sigma_\lambda(\omega_{t-dt})^{-1}$ . This cutoff leads to the following demand curve that is useful for proving the results below:

**Lemma B2.** *For any  $\tilde{P} > 0$ , the number of potential residents for whom  $P_t^c(\delta, \lambda) \geq \tilde{P}$  equals  $A_t(r\tilde{P})^{-\epsilon} \int_0^\infty \Sigma_\lambda(\omega_{t-dt})^{-\epsilon} f(\lambda) d\lambda$ .*

*Proof.*  $P_t^c(\delta, \lambda) \geq \tilde{P}$  if and only if  $\delta \geq r\tilde{P}\Sigma_\lambda(\omega_{t-dt})$ . For each  $\lambda$ , the measure of such potential buyers equals  $A_t f(\lambda)(r\tilde{P}\Sigma_\lambda(\omega_{t-dt}))^{-\epsilon}$ . Integrating over  $\lambda$  yields the result.  $\square$

## Listing Prices

The following lemma characterizes the price  $P_t^*$  chosen by attentive movers.

**Lemma B3.** *Each attentive mover chooses to list at the price  $P_t^* = P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}$ .*

*Proof.* Suppose the measure of listing movers equals  $L_t$  and the cumulative distribution function of their list prices equals  $F_t^p(\cdot)$ . Suppose a given mover  $i$  chooses a listing price  $P_{i,t}$ . Given Lemma B2, the number of potential buyers whose cutoff is at least  $P_{i,t}$  equals  $\tilde{D}_t(P_{i,t}/P_{t-dt})^{-\epsilon}$ . Under the mover's assumption that the stock of movers equals 0, the mover sells with certainty at  $P_{i,t}$  if and only if  $L_t F_t^p(P_{i,t}) \leq \tilde{D}_t(P_{i,t}/P_{t-dt})^{-\epsilon}$ . We define  $P_t^{sup}(L_t, F_t^p)$  to be the supremum of  $P_{i,t}$  satisfying this inequality. The listing mover can guarantee selling her house and attain a payoff arbitrarily close to  $P_t^{sup}(L_t, F_t^p)$  by choosing a price arbitrarily close to but below this value.

Due to ambiguity-averse preferences, the attentive mover sets  $P_t^* = \min_{L_t, F_t^p} P_t^{sup}(L_t, F_t^p)$ . Because this function decreases in  $L_t$ , the minimizing value of  $L_t$  is  $\bar{D}$ , the value that attains when all of the movers thought to exist decide to list. It follows that  $P_t^* = \min_{F_t^p} P_t^{sup}(\bar{D}, F_t^p)$ .

We claim that this minimum equals  $P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}$ . If  $F_t^p(P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}) = 1$ , then this assertion trivially holds. If  $F_t^p(P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}) < 1$ , then  $P_t^{sup}(\bar{D}, F_t^p) \geq P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}$ . It follows that  $P_t^* = \min_{F_t^p} P_t^{sup}(\bar{D}, F_t^p) = P_{t-dt}(\tilde{D}_t/\bar{D})^{1/\epsilon}$ , as claimed. Because the mover is certain to sell at this price, she always chooses to list.  $\square$

Lemma B3 allows us to characterize the evolution of the market price  $P_t$  over time. Because inattentive movers believe that  $\tilde{D}_t = \bar{D}$ , Lemma B3 implies that they list at  $P_{t-dt}$  at time  $t$ . With lower-case  $p$  denoting  $\log P$ , the log market price at  $t$  equals  $p_t = e^{-\beta dt} p_{t-dt} + (1 - e^{-\beta dt}) p_t^*$ . Substituting the expression for  $p_t^*$  from Lemma B3 yields  $p_t - p_{t-dt} = (1 - e^{-\beta dt}) \log(\tilde{D}_t/\bar{D})/\epsilon$ . Dividing by  $dt$  and taking the limit  $dt \rightarrow 0$  yields

$$\dot{p}_t = (\beta/\epsilon) \log(D_t/\bar{D}), \quad (\text{B1})$$

where we have used the definitional fact that  $\lim_{dt \rightarrow 0} \tilde{D}_t = D_t$ . In continuous time, movers are inattentive almost surely and post at  $P_t$ , so transactions take place at  $P_t$  almost surely. As a result, (B1) reproduces (8) with  $c = \beta/\epsilon$ .

The fact that inattentive movers post at  $P_t$  rationalizes Assumption 2. Because potential buyers believe that the stock of movers remains equal to 0 at all times, they must believe that all movers sell instantaneously, and all but a measure 0 of such movers sell at  $P_t$ . In this way, Assumption 2 follows from the assumed rigidities in attention to  $\tilde{D}_t$  and to the variation in the number of movers.

The price chosen by inattentive movers coincides with a rigidity assumed in Guren (2016) in which some sellers copy the most recent listing price plus an increase proportional to recent price growth. In continuous time, under this rigidity inattentive movers copy the latest market price, which is the behavior they follow in our microfoundation.

## Rationing

We now show that (3) holds given the market mechanism and the chosen prices of movers. In continuous time, all but a zero measure of matched potential buyers see a price of  $P_t$ .

If  $L_t > D_t$  or  $I_t > 0$ , all  $D_t$  potential buyers whose cutoff price is at least  $P_t$  are matched, and all of them buy. Volume  $V_t(\lambda)$  to potential buyers of type  $\lambda$  in this case equals  $D_t(\lambda)$ , the number of potential buyers of type  $\lambda$  whose cutoff price is at least exceeds  $P_t$  (according to the proof of Lemma B2).

If  $L_t \leq D_t$  and  $I_t = 0$ , then only the  $L_t$  potential buyers with the highest cutoff prices buy. As shown in the proof of Lemma B2, the number of potential buyers of type  $\lambda$  whose cutoff price is at least  $\tilde{P}_t$  equals  $D_t(\lambda)(\tilde{P}_t/P_t)^{-\epsilon}$ . The market-clearing cutoff for  $\tilde{P}_t$  solves  $L_t = \int_{\tilde{P}_t}^{\infty} D_t(\lambda)(\tilde{P}_t/P_t)^{-\epsilon} f(\lambda) d\lambda = D_t(\tilde{P}_t/P_t)^{-\epsilon}$ , so  $\tilde{P}_t = P_t(D_t/L_t)^{1/\epsilon}$ . It follows that  $V_t(\lambda) = D_t(\lambda)(\tilde{P}_t/P_t)^{-\epsilon} = L_t D_t(\lambda)/D_t$ , as given by (3).

## 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.<sup>3</sup> 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 A2 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.
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.

---

<sup>3</sup>We rely on CoreLogic’s internal transaction-type categorization to determine whether a transaction occurred at arms length.

- Drop sales of vacant land parcels in MSAs where the CoreLogic data includes such sales.<sup>4</sup> We define a vacant land sale to be any transaction where the sale occurs a year or more before the property was built.

Table A3 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.<sup>5</sup> The MSAs that remain after these two filters are marked with an “x” in Table A2.

### **Identifying New Construction Sales**

In Table 4 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, 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

---

<sup>4</sup>MSAs are flagged as including vacant land sales if more than 5 percent of the sales in the MSA occur more than two years before the year in which the property was built.

<sup>5</sup>This step drops only Chicago-Naperville-Elgin, IL-IN-WI.

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.



## D Matching Sellers and Buyers

Our model abstracts from the situation in which sellers become buyers and this joint buyer-seller problem—anticipating that selling in the future will necessitate buying another house—affects their initial decision to buy. To explore how restrictive this assumption is, we conduct an empirical analysis of the frequency of move-up buyers using the sample of MSAs defined in Appendix C. This appendix describes our algorithm and results in more detail.

The deeds data are well suited for tracking properties over time, but information about buyers and sellers is limited to names and addresses. As a first pass, 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 a quarter around the quarter of the sale transaction. We only look for matches within MSA, as purchases associated with cross-city moves are more similar in spirit to our model’s departure from the standard search framework.

Our match accounts for several anomalies, which would lead a naive match strategy to understate the match rate. 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.

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 2014 because the seller name fields are 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 18.9 million sales transactions. Of these, we are able to match 6.9 million to a linked buyer transaction, or 37%. Thus, nearly two-thirds of transactions do not appear to be associated with joint buyer-seller decisions.

The mean match rate also varies across cities, ranging from 16% at the 10th percentile to 40% at the 90th percentile. The match rate increases with the cycle, reaching nearly 50% at the peak of the volume cycle in mid-2005 and falling to approximately 25% during the bust years of 2008 through 2011. These patterns are largely consistent with the findings in Anenberg and Bayer (2013), who conduct a similar match for the Los Angeles metro area.

We have not attempted to decompose the covariance in match rates with the cycle into components due to true joint buyer-seller decisions versus false matches from having a larger haystack to search through. Thus, the time series patterns of match rates should be judged with caution. We leave to future work a comprehensive investigation of how the joint buyer-seller problem interacts with the forces at play in our model.

## References

- Abreu, Dilip and Brunnermeier, Markus K.** (2003). ‘Bubbles and Crashes’, *Econometrica* 71(1), 173–204.
- Adelino, Manuel, Schoar, Antoinette and Severino, Felipe.** (2016). ‘Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class’, *Review of Financial Studies* 29(7), 1635–1670.
- Agarwal, Sumit, Hu, Luojia and Huang, Xing.** (2015). ‘Rushing into the American Dream? House Prices Growth and the Timing of Homeownership’. Forthcoming, *Review of Finance*.
- Amihud, Yakov and Mendelson, Haim.** (1986). ‘Asset Pricing and the Bid-Ask Spread’, *Journal of Financial Economics* 17(2), 223–249.
- Anenberg, Elliot and Bayer, Patrick.** (2013), Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem, Technical report, National Bureau of Economic Research.
- Armona, Luis, Fuster, Andreas and Zafar, Basit.** (2016). ‘Home Price Expectations and Behavior: Evidence from a Randomized Information Experiment’. Working Paper.
- Bailey, Michael, Cao, Ruiqing, Kuchler, Theresa and Stroebel, Johannes.** (2016), Social Networks and Housing Markets. Working Paper.
- Barberis, Nicholas, Greenwood, Robin, Jin, Lawrence and Shleifer, Andrei.** (2015). ‘X-CAPM: An Extrapolative Capital Asset Pricing Model’, *Journal of Financial Economics* 115(1), 1–24.
- Barberis, Nicholas, Greenwood, Robin, Jin, Lawrence and Shleifer, Andrei.** (2016). ‘An Extrapolative Model of Bubbles’. Working Paper.
- Barberis, Nicholas and Shleifer, Andrei.** (2003). ‘Style Investing’, *Journal of Financial Economics* 68, 161–199.
- Barberis, Nicholas, Shleifer, Andrei and Vishny, Robert.** (1998). ‘A Model of Investor Sentiment’, *Journal of Financial Economics* 49(3), 307–343.
- Barlevy, Gadi and Fisher, Jonas D.M.** (2011). ‘Mortgage Choices and Housing Speculation’. Working Paper, Federal Reserve Bank of Chicago.
- Bayer, Patrick, Geissler, Christopher, Mangum, Kyle and Roberts, James W.** (2011), Speculators and Middlemen: The Strategy and Performance of Investors in the Housing Market. NBER Working Paper 16784.
- Bayer, Patrick, Mangum, Kyle and Roberts, James W.** (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.
- Black, Fischer.** (1986). ‘Noise’, *Journal of Finance* 41(3), 528–543.
- Burnside, Craig, Eichenbaum, Martin and Rebelo, Sergio.** (2016). ‘Understanding Booms and Busts in Housing Markets’, *Journal of Political Economy* 124(4), 1088–1147.
- Caplin, Andrew and Leahy, John.** (2011). ‘Trading Frictions and House Price Dynamics’, *Journal of Money, Credit, and Banking* 43(2), 283–303.
- Case, Karl E., Shiller, Robert J. and Thompson, Anne K.** (2012). ‘What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets’, *Brookings Papers on Economic Activity* Fall, 265–298.
- Chinco, Alex and Mayer, Christopher.** (2016). ‘Misinformed Speculators and Mispricing in the Housing Market’, *Review of Financial Studies* 29(2), 486–522.
- 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., Poterba, James M. and Summers, Lawrence H.** (1990). ‘Speculative Dynamics and the Role of Feedback Traders’, *American Economic Review, Papers and Proceedings* 80(2), 63–68.
- Cutler, David M., Poterba, James M. and Summers, Lawrence H.** (1991). ‘Speculative Dynamics’, *Review of Economic Studies* 58(3), 529–546.
- Daniel, Kent and Hirshleifer, David.** (2015). ‘Overconfident Investors, Predictable Returns, and Excessive Trading’, *Journal of Economic Perspectives* 29(4), 61–88.
- Daniel, Kent, Hirshleifer, David and Subrahmanyam, Avanidhar.** (1998). ‘Investor Psychology and Security Market Under- and Overreactions’, *Journal of Finance* 53(6), 1839–1885.
- Dávila, Eduardo.** (2015). ‘Optimal Financial Transactions Taxes’. Working Paper, NYU Stern.
- De Long, J. Bradford, Shleifer, Andrei, Summers, Lawrence H. and Waldmann, Robert J.** (1990). ‘Positive Feedback Investment Strategies and Destabilizing Rational Speculation’, *Journal of Finance* 45(2), 379–395.
- Diamond, Rebecca.** (2016). ‘The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000’, *American Economic Review* 106(3), 479–524.
- Díaz, Antonia and Jerez, Belén.** (2013). ‘House Prices, Sales, and Time on the Market: A Search-Theoretic Framework’, *International Economic Review* 54(3), 837–872.

- Di Maggio, Marco and Kermani, Amir.** (2015). ‘Credit-Induced Boom and Bust’. Working Paper, Columbia Business School.
- Ebrahimi, Ehsan and Shimer, Robert.** (2010). ‘Stock-Flow Matching’, *Journal of Economic Theory* 145, 1325–1353.
- Edelstein, Robert and Qian, Wenlan.** (2014). ‘Short-Term Buyers and Housing Market Dynamics’, *Journal of Real Estate Finance and Economics* 49, 654–689.
- Fama, Eugene F.** (1970). ‘Efficient Capital Markets: A Review of Theory and Empirical Work’, *Journal of Finance* 25(2), 383–417.
- Favara, Giovanni and Imbs, Jean.** (2015). ‘Credit Supply and the Price of Housing’, *American Economic Review* 105(3), 958–992.
- Fuster, Andreas and Zafar, Basit.** (2015). ‘The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey’. Working Paper.
- Fu, Yuming and Qian, Wenlan.** (2014). ‘Speculators and Price Overreaction in the Housing Market’, *Real Estate Economics* 42(3), 977–1007.
- Gao, Zhenyu, Sockin, Michael and Xiong, Wei.** (2016). ‘Economic Consequences of Housing Speculation’. Working Paper.
- Garriga, Carlos and Hedlund, Aaron.** (2016), Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession. Working Paper.
- Genesove, David and Mayer, Christopher.** (2001). ‘Loss Aversion and Seller Behavior: Evidence from the Housing Market’, *Quarterly Journal of Economics* 116(4), 1233–1260.
- Giglio, Stefano, Maggiori, Matteo and Stroebel, Johannes.** (2016). ‘No-Bubble Condition: Model-Free Tests in Housing Markets’, *Econometrica* 84(3), 1047–1091.
- Gilboa, Itzhak and Schmeidler, David.** (1989). ‘Maxmin Expected Utility with Non-Unique Prior’, *Journal of Mathematical Economics* 18(2), 141–153.
- 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 Nathanson, Charles G.** (2016). ‘An Extrapolative Model of House Price Dynamics’. Working Paper.
- Graham, John R. and Harvey, Campbell R.** (2003). ‘Expectations of Equity Risk Premia, Volatility and Asymmetry’. Working Paper.
- Greenwood, Robin and Shleifer, Andrei.** (2014). ‘Expectations of Returns and Expected Returns’, *Review of Financial Studies* 27(3), 714–746.
- Griffin, John M., Nardari, Federico and Stulz, René M.** (2007). ‘Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries’, *Review of Financial Studies* 20(3), 905–951.

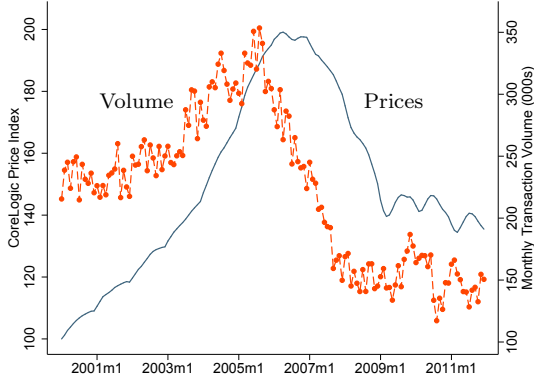
- Guren, Adam M.** (2016). ‘House Price Momentum and Strategic Complementarity’. Working Paper.
- Guren, Adam M. and McQuade, Timothy J.** (2015). ‘How do Foreclosures Exacerbate Housing Downturns?’. Working Paper.
- Han, Lu and Strange, William C.** (2015), The Microstructure of Housing Markets: Search, Bargaining, and Brokerage, *in Gilles Duranton, J. Vernon Henderson and William C. Strange.*, eds, ‘Handbook of Regional and Urban Economics’, Vol. 5B, Elsevier, chapter 13, pp. 813–886.
- Hanushek, Eric A. and Quigley, John M.** (1980). ‘What is the Price Elasticity of Housing Demand?’, *Review of Economics and Statistics* 62(3), 449–454.
- Harrison, J. Michael and Kreps, David M.** (1978). ‘Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations’, *Quarterly Journal of Economics* 92(2), 323–336.
- Haughwout, Andrew, Lee, Donghoon, Tracy, Joseph and van der Klaauw, Wilbert.** (2011). ‘Real Estate Investors, the Leverage Cycle, and the Housing Market Crisis’. Federal Reserve Bank of New York Staff Report no. 514.
- Head, Allen, Lloyd-Ellis, Huw and Sun, Hongfei.** (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 Ioannides, Yannis M.** (1989). ‘Dynamic Aspects of Consumer Decisions in Housing Markets’, *Journal of Urban Economics* 26, 212–230.
- Hong, Harrison and Stein, Jeremy C.** (1999). ‘A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets’, *Journal of Finance* 54(6), 2143–2184.
- Hong, Harrison and Stein, Jeremy C.** (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.
- Keynes, John Maynard.** (1936), *The General Theory of Employment, Interest and Money*, Macmillan, New York.
- Kindleberger, Charles P.** (1978), *Manias, Panics, and Crashes: A History of Financial Crises*, Wiley, 4th Edition.

- Kuchler, Theresa and Zafar, Basit.** (2016). ‘Personal Experiences and Expectations About Aggregate Outcomes’. Working Paper.
- Lee, Charles and Swaminathan, Bhaskaran.** (2000). ‘Price Momentum and Trading Volume’, *Journal of Finance* 55(5), 2017–2069.
- Mankiw, N. Gregory and Reis, Ricardo.** (2002). ‘Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve’, *Quarterly Journal of Economics* 117(4), 1295–1328.
- Nathanson, Charles G. and Zwick, Eric.** (2017), Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market. NBER Working Paper 23030.
- Ngai, L. Rachel and Sheedy, Kevin D.** (2016). ‘The Decision to Move House and Aggregate Housing-Market Dynamics’. Working Paper.
- Ngai, Rachel and Tenreyro, Silvana.** (2014). ‘Hot and Cold Seasons in the Housing Market’, *American Economic Review* 104(12), 3991–4026.
- Ofek, Eli and Richardson, Matthew.** (2003). ‘Dotcom Mania: The Rise and Fall of Internet Stock Prices’, *Journal of Finance* 58(3), 1113–1137.
- Ortalo-Magné, François and Rady, Sven.** (2006). ‘Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints’, *Review of Economic Studies* 73(2), 459–485.
- Pástor, L’uboš and Veronesi, Pietro.** (2006). ‘Was There a Nasdaq Bubble in the Late 1990s?’, *Journal of Financial Economics* 81(1), 61–100.
- Pástor, L’uboš and Veronesi, Pietro.** (2009). ‘Technological Revolutions and Stock Prices’, *American Economic Review* 99(4), 1451–1483.
- Penasse, Julien and Renneboog, Luc.** (2016), Bubbles and Trading Frenzies: Evidence from the Art Market. Working Paper.
- Piazzesi, Monika and Schneider, Martin.** (2009). ‘Momentum Traders in the Housing Market: Survey Evidence and a Search Model’, *American Economic Review* 99(2), 406–411.
- Poterba, James M.** (1984). ‘Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach’, *Quarterly Journal of Economics* 99(4), 729–752.
- Rajan, Raghuram and Ramcharan, Rodney.** (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, Shleifer, Andrei and Simsek, Alp.** (2016), Investment Hangover and the Great Recession. Working Paper.

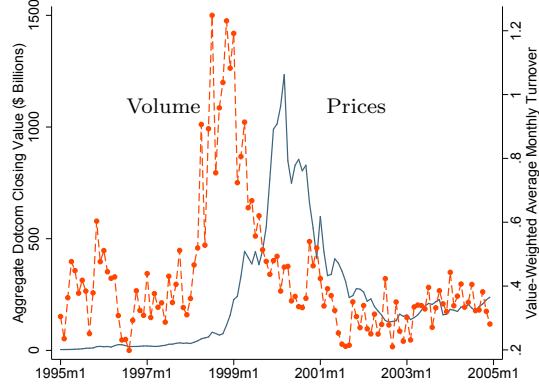
- Scheinkman, Jose A. and Xiong, Wei.** (2003). ‘Overconfidence and Speculative Bubbles’, *Journal of Political Economy* 111(6), 1183–1220.
- Shiller, Robert J.** (1981). ‘Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?’, *American Economic Review* 71(3), 421–436.
- Shiller, Robert J.** (2005), *Irrational Exuberance*, Princeton University Press.
- Shleifer, Andrei and Vishny, Robert W.** (1992). ‘Liquidation Values and Debt Capacity: A Market Equilibrium Approach’, *Journal of Finance* 47(4), 1343–1366.
- Simsek, Alp.** (2013). ‘Speculation and Risk Sharing with New Financial Assets’, *Quarterly Journal of Economics* 128(3), 1365–1396.
- Smith, Vernon L., Suchanek, Gerry L. and Williams, Arlington W.** (1988). ‘Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets’, *Econometrica* 56(5), 1119–1151.
- Statman, Meir, Thorley, Steven and Vorkink, Keith.** (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 Summers, Victoria P.** (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’, The MIT Press, pp. 139–208.
- 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 Joint Dynamics of Prices and Volume

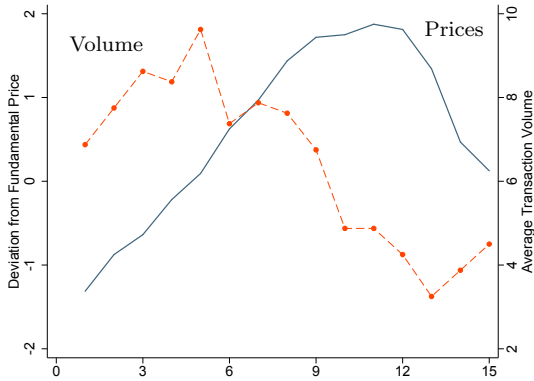
(a) US Housing Market (2000–2011)



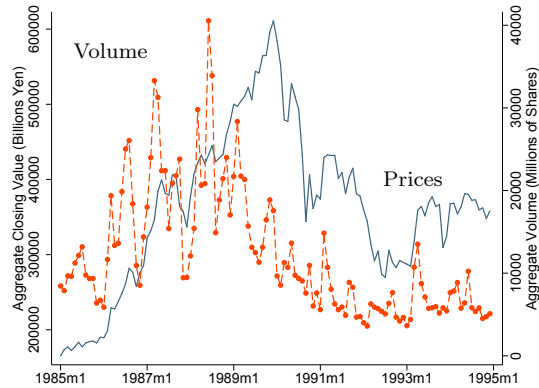
(b) US Equities, Tech (1995–2005)



(c) Experimental Markets, SSW (1988)



(d) Japan Equities (1985–1995)

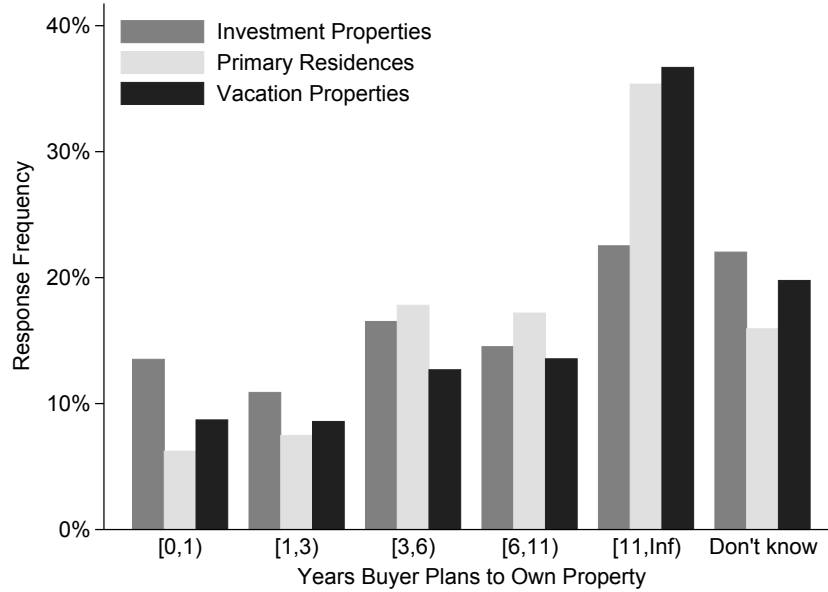


*Notes:* These figures display the dynamic relationship between prices and transaction volume for four distinct bubble episodes: (a) the 2000–2011 US housing market, (b) the 1995–2005 market in technology stocks, (c) the bubbles in experimental asset markets, and (d) the 1985–1995 Japanese stock market. Panel (a) data come from CoreLogic and cover 115 cities. For prices, we use CoreLogic’s single-family home price index, which is based on repeat sales. For volume, we plot a seasonally adjusted monthly count of transactions in the data set used in Section 4; we seasonally adjust volume by removing calendar-month fixed effects. Panel (b) 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 (c) data were manually entered from the published Smith et al. (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 average number of trades. Panel (d) 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.



FIGURE 2  
 Expected Holding Times of Homebuyers, 2008–2015

(a) Response Heterogeneity

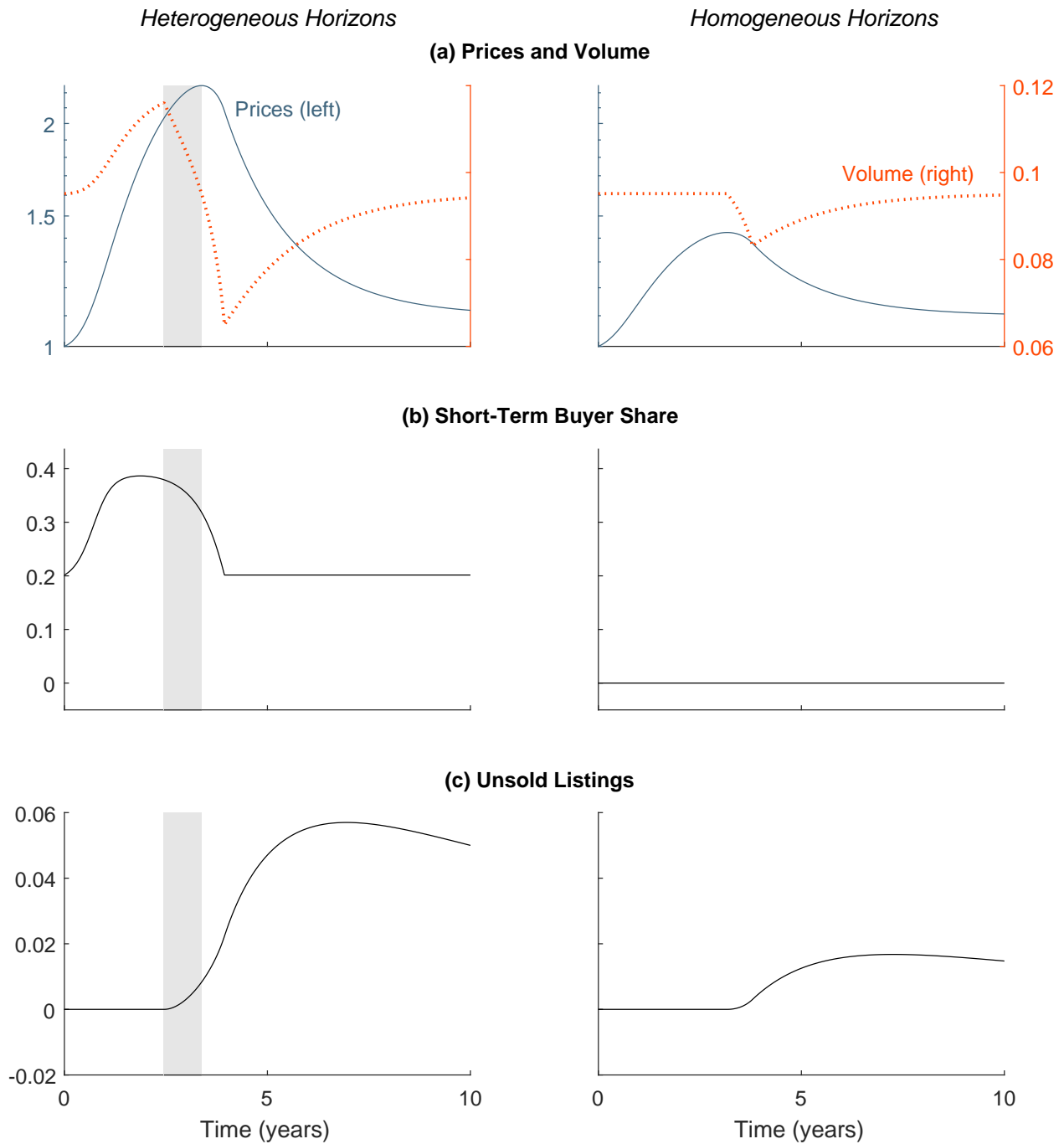


(b) Short-Term Buyers and Recent House Price Growth



*Notes:* 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). 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 US 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.

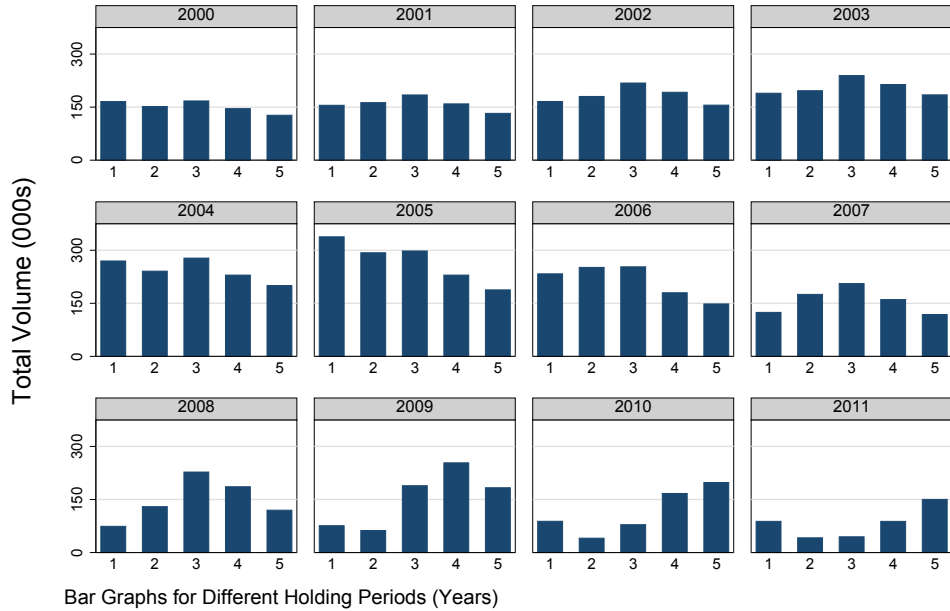
FIGURE 3  
Simulation Results



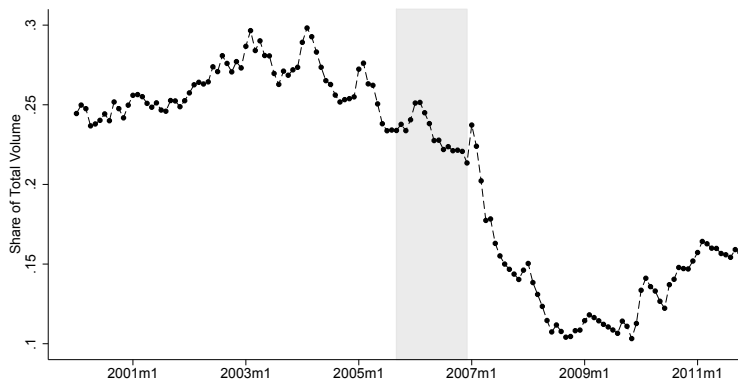
*Notes:* The shaded area denotes the quiet. (a) The price when the demand shock occurs is normalized to 1. Volume is in units of share of the housing stock. (b) “Short-term buyer” is defined as one whose expected holding time is less than 3 years. The panel plots the share of such individuals among all buyers. (c) Unsold listings are in units of share of the housing stock.

FIGURE 4  
The Dynamics of Holding Times in the Housing Market

(a) Realized Holding Times of Sales

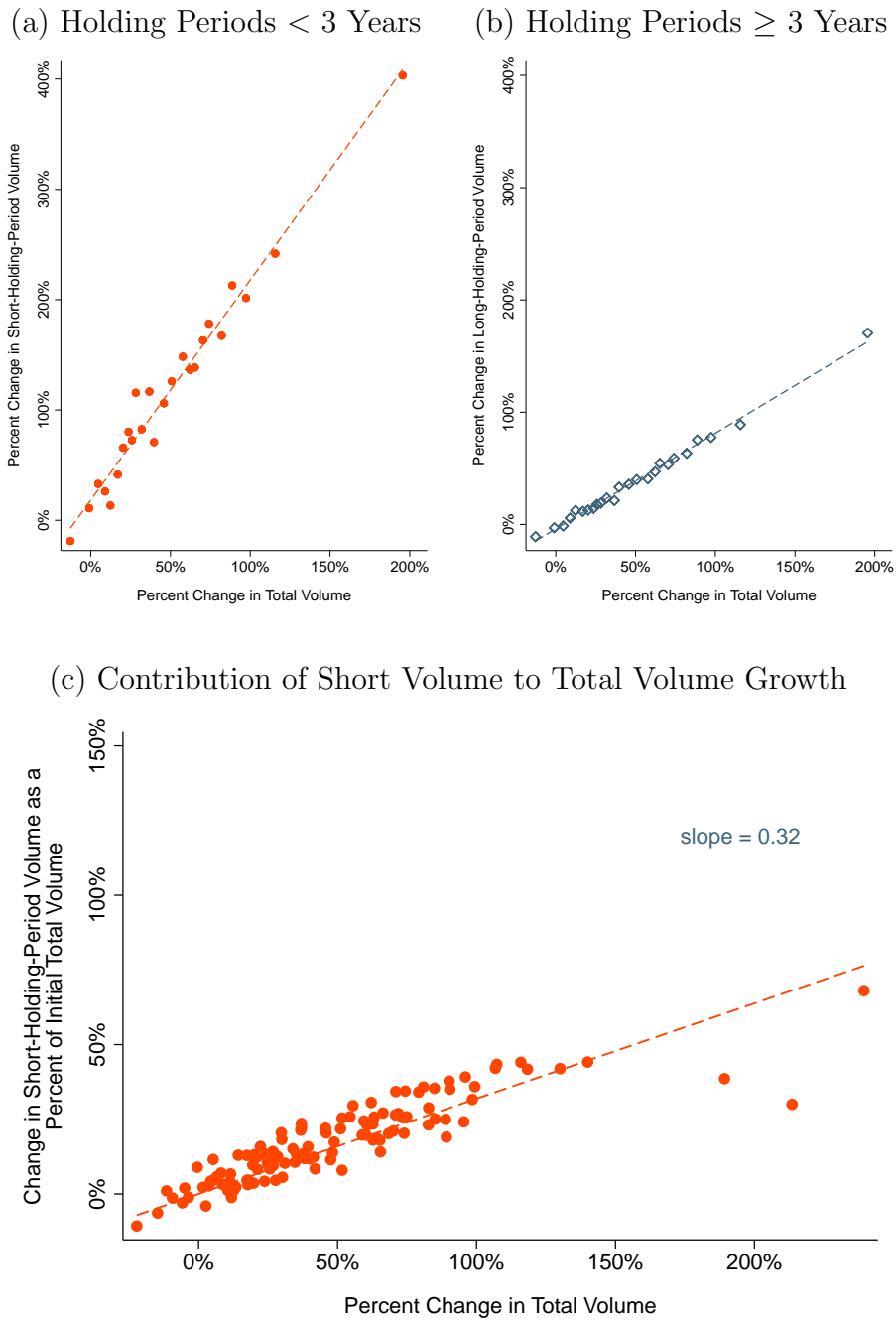


(b) Imputed Share of Buyers with Horizons < 3 Years



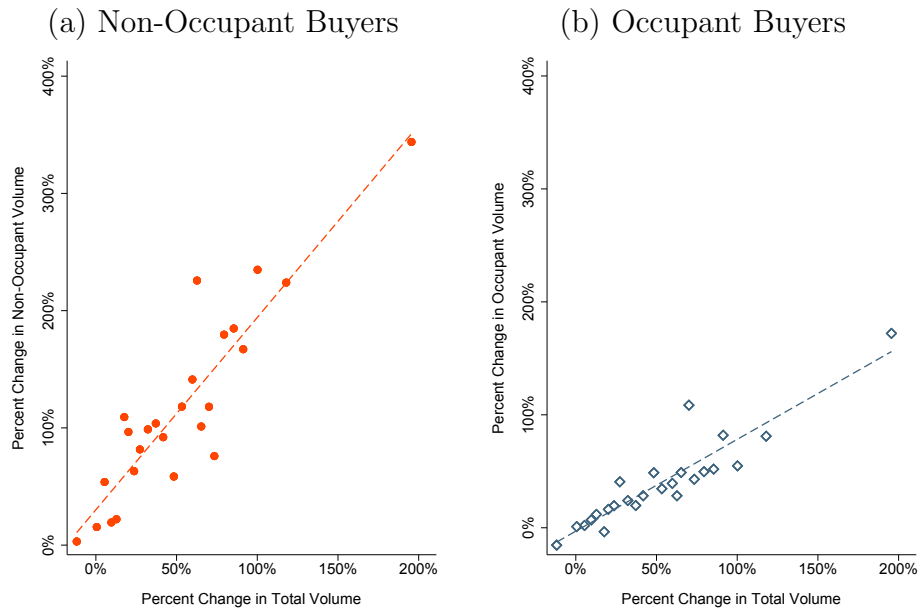
*Notes:* Panel (a) illustrates the time variation in holding periods between 2000 and 2011 in the US. For each transaction, we define the holding period as the number of days since the last transaction of the same property. We then group all transactions with holding periods less than or equal to 5 years 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. Panel (b) plots the share of total volume accounted for by transactions for which the buyer eventually sells in two years or sooner. As described in the text, we adjust short buyer counts to account for the changing probability of selling conditional on listing over time.

FIGURE 5  
 The Role of Short-Holding-Period Volume Growth for Total Volume Growth

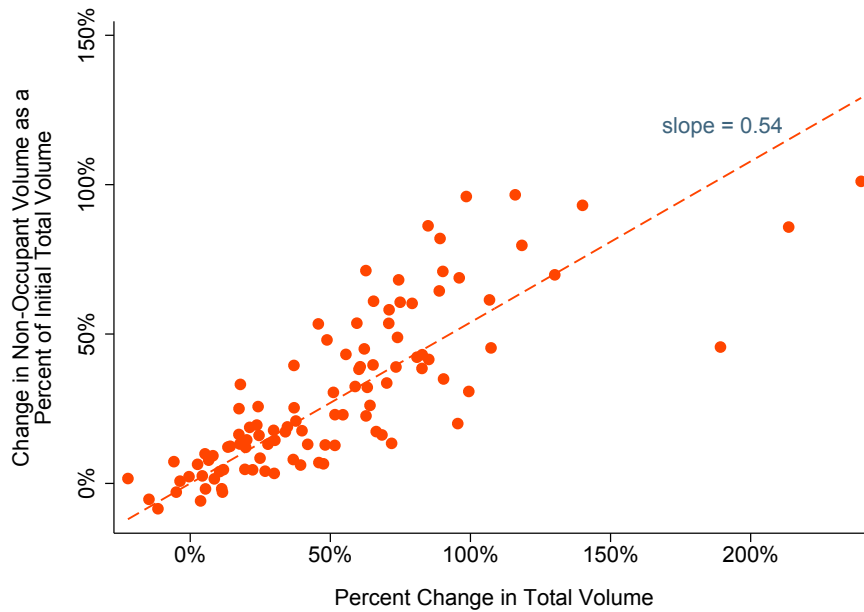


*Notes:* This figure illustrates the quantitative importance of short-holding-period volume in accounting for the increase in total volume between 2000 and 2005. We present binned scatter plots (“binscatters”) of the percent change in total volume from 2000 to 2005 versus the percent change in volume for short holding periods (< 3 years) in Panel (a) and long holding periods ( $\geq 3$  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



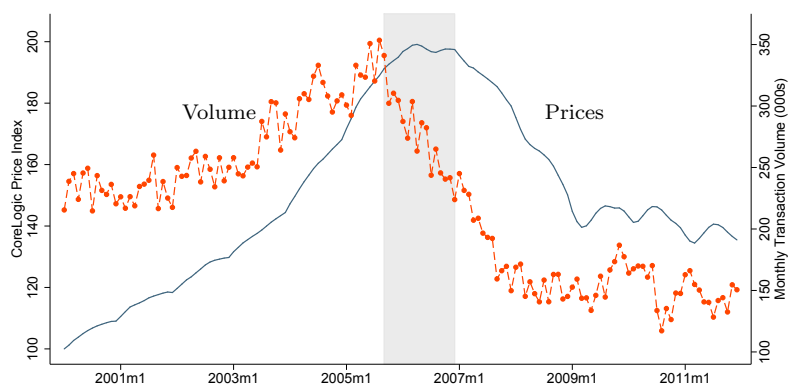
(c) Contribution of Non-Occupant Volume to Total Volume Growth



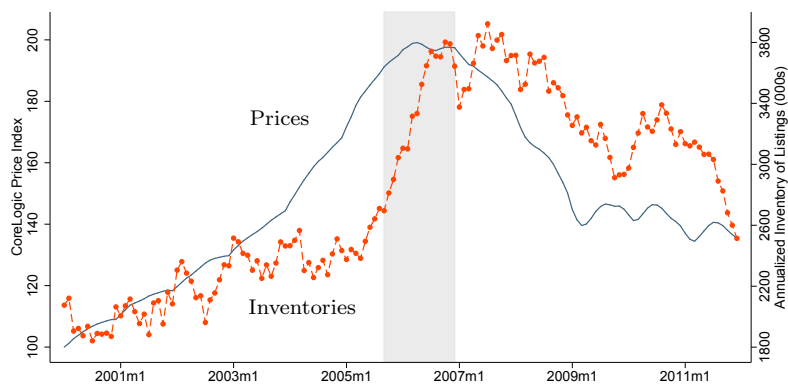
*Notes:* This figure illustrates of the quantitative importance of non-occupant volume in accounting for the increase in total volume between 2000 and 2005. We present binned scatter plots (“binscatters”) of the percent change in total volume from 2000 to 2005 versus the percent change in volume for non-occupant buyers (transactions with distinct mailing and property addresses) in Panel (a) and occupant 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.

FIGURE 7  
The Dynamics of Prices, Volume, and Inventories

(a) Prices and Volume

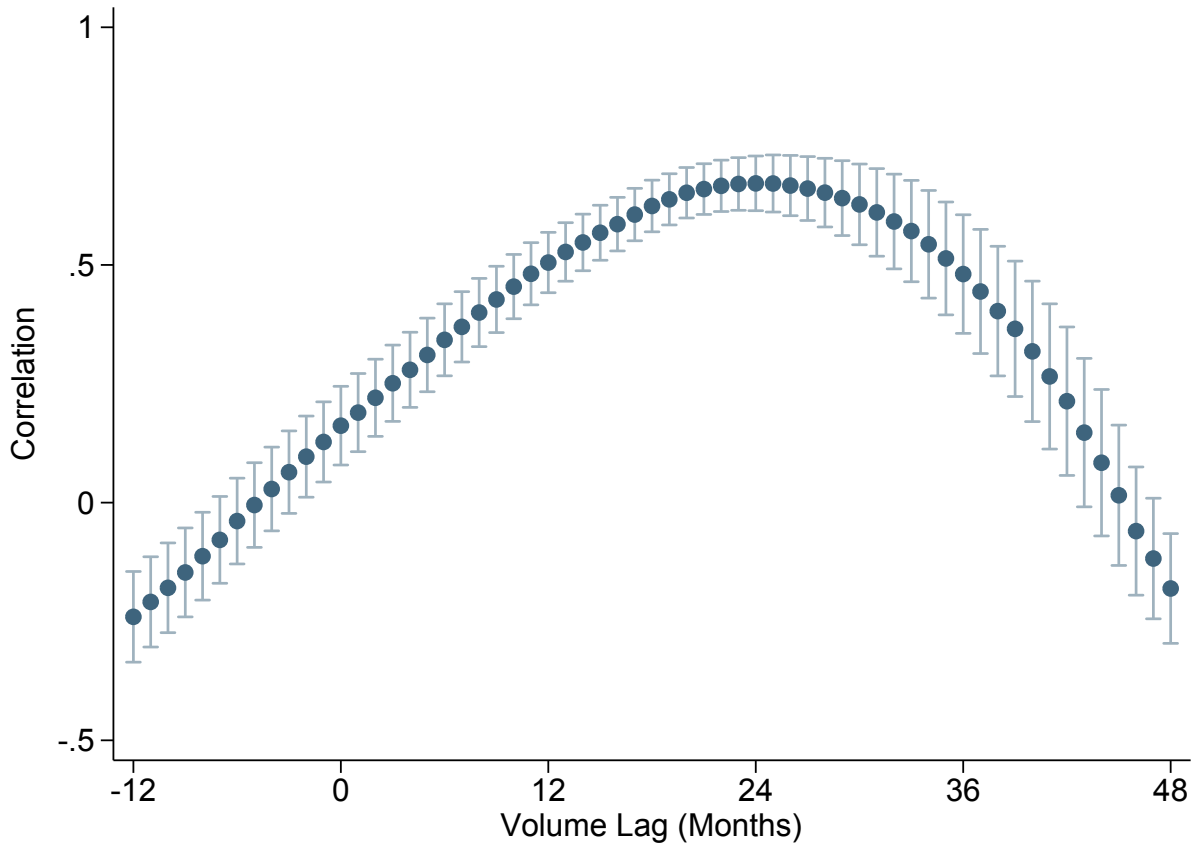


(b) Prices and Inventories



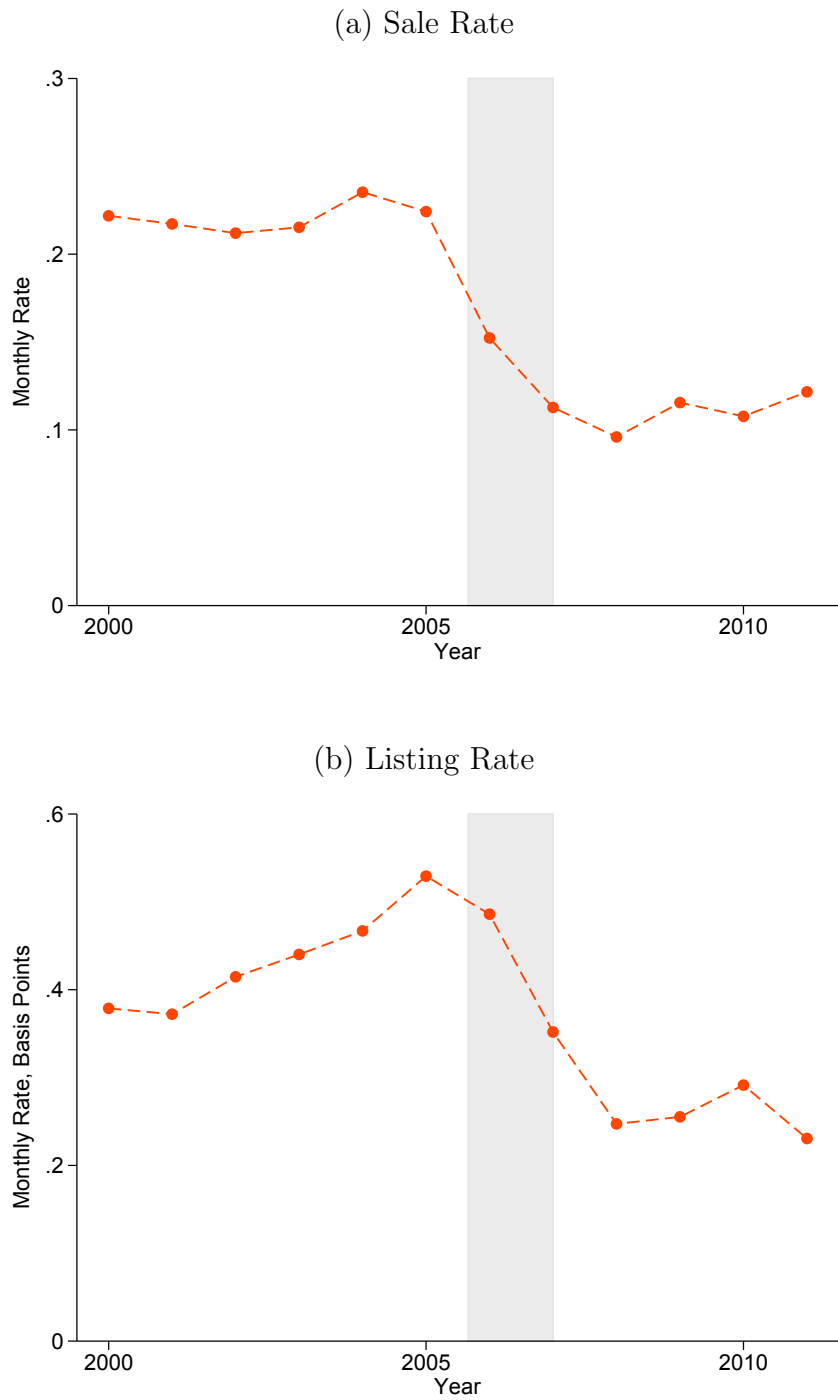
*Notes:* These figures display the dynamic relationship between prices, volume, and the inventory of listings in the US housing market between 2000 and 2011. Panel (a) plots monthly prices and sales volume, and panel (b) plots monthly prices and inventory. Inventory information comes from the National Association of Realtors. We apply a calendar-month seasonal adjustment for both volume and inventories.

FIGURE 8  
The Correlation between Prices and Volume at Various Lags



*Notes:* This figure shows that the correlation between prices and lagged volume is robust across cities 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.

FIGURE 9  
 Sale and Listing Rates for US Existing Homes, 2000–2011

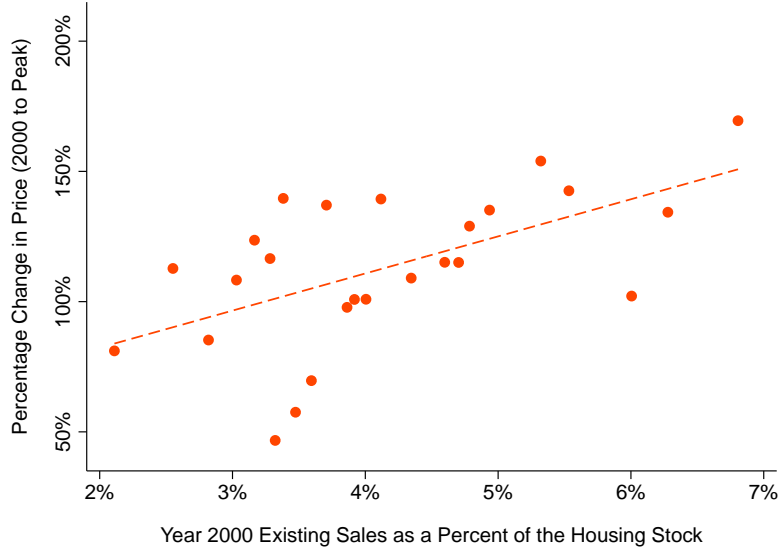


*Notes:* The figure displays the average monthly sale rate and listing rate for each year. For each month, the sale rate equals sales divided by listed inventory, and the listing probability equals new listings divided by the stock of unlisted houses. Monthly data on the US housing stock are interpolated from quarterly estimates provided by the US 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 7.

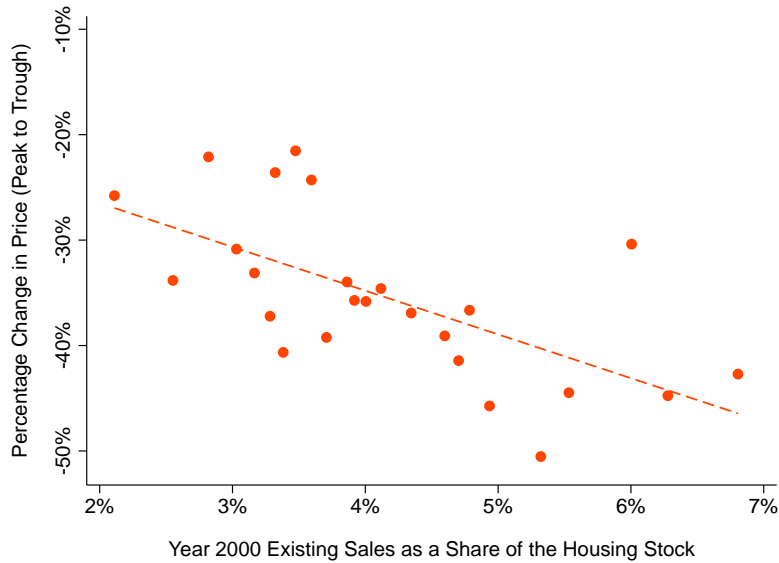


FIGURE 10  
Initial Volume and the Magnitude of the Housing Boom and Bust

(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 indices. 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 the month in which the peak occurred.

TABLE 1  
Calibration Sources

| Calibrated Quantity | Role in Model  | Source or Assumed Value                               |
|---------------------|--|---|
| $f(\cdot)$          | Distribution of expected holding times   | NAR Investment and Vacation Home Buyers Survey        |
| $\mu$               | Relative weight in expectations on recent price changes versus those in distant past | Estimation using NAR survey                           |
| $g(\cdot)$          | Forward term structure of expectations   | Survey of Consumer Expectations; Armona et al. (2016) |
| $\gamma(\cdot)$     | Extrapolation function   | Estimation using Case et al. (2012) survey            |
| $r$                 | Discount rate  | 0.07  |
| $c$                 | Price stickiness   | 1   |
| $\epsilon$          | Housing demand elasticity  | 0.6   |
| $A^f/A^i$           | Size of demand shock   | 1.06  |

*Notes:* This table lists the model quantities we calibrate to produce Figure 3. Further details are provided in Section 3.4.

TABLE 2  
Sensitivity of Simulation Results to Parameters

|                            | $f$  | $\rho$ | $\mu$ | $c$  | $\phi$ | $\epsilon$ | $A^f/A^i$ | $r$  |
|----------------------------|------|--------|-------|------|--------|------------|-----------|------|
| A. Excess Price Boom       |      |        |       |      |        |            |           |      |
| Low                        | 0.30 | 0.18   | 0.83  | 0.40 | 0.75   | 0.36       | 0.92      | 1.79 |
| Baseline                   | 1.05 | 1.05   | 1.05  | 1.05 | 1.05   | 1.05       | 1.05      | 1.05 |
| High                       | 3.31 | 1.46   | 0.99  | 1.48 | 1.41   | 3.46       | 1.17      | 0.26 |
| B. Volume Boom             |      |        |       |      |        |            |           |      |
| Low                        | 0.00 | 0.09   | 0.15  | 0.11 | 0.39   | 0.05       | 0.21      | 0.28 |
| Baseline                   | 0.22 | 0.22   | 0.22  | 0.22 | 0.22   | 0.22       | 0.22      | 0.22 |
| High                       | 0.94 | 0.26   | 0.23  | 0.19 | 0.25   | 0.86       | 0.23      | 0.14 |
| C. Maximal Unsold Listings |      |        |       |      |        |            |           |      |
| Low                        | 0.02 | 0.01   | 0.05  | 0.04 | 0.02   | 0.02       | 0.06      | 0.09 |
| Baseline                   | 0.06 | 0.06   | 0.06  | 0.06 | 0.06   | 0.06       | 0.06      | 0.06 |
| High                       | 0.18 | 0.07   | 0.05  | 0.05 | 0.08   | 0.16       | 0.06      | 0.01 |

*Notes:* The excess price boom equals  $P^{max}/P_\infty - 1$ , the volume boom equals  $V^{max}/V_0 - 1$ , and maximal unsold listings equal  $\max_t I_t$ . The alternate values for  $f$ ,  $\rho$ , and  $\epsilon$  are described in the text. The low and high values for the remaining parameters are half and double their baseline values (we half and double  $\log(A^f/A^i)$ ).

TABLE 3  
 Within-MSA Role of Short-Holding-Period and  
 Non-Occupant Volume Growth for Total Volume Growth (2000–2005)

|                         | Holding Period             |                           |                                 | Occupancy Status                  |                               |  |
|-------------------------|----------------------------|---------------------------|---------------------------------|-----------------------------------|-------------------------------|--|
|                         | % $\Delta$ Short<br>Volume | % $\Delta$ Long<br>Volume | $\Delta$ Short/<br>Total Volume | % $\Delta$ Non-Occupant<br>Volume | % $\Delta$ Occupant<br>Volume | $\Delta$ Non-Occupant/<br>Total Volume |
| % $\Delta$ Total Volume | 1.517***<br>(0.104)        | 0.936***<br>(0.020)       | 0.241***<br>(0.018)             | 1.525***<br>(0.094)               | 1.044***<br>(0.093)           | 0.356***<br>(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                                  |

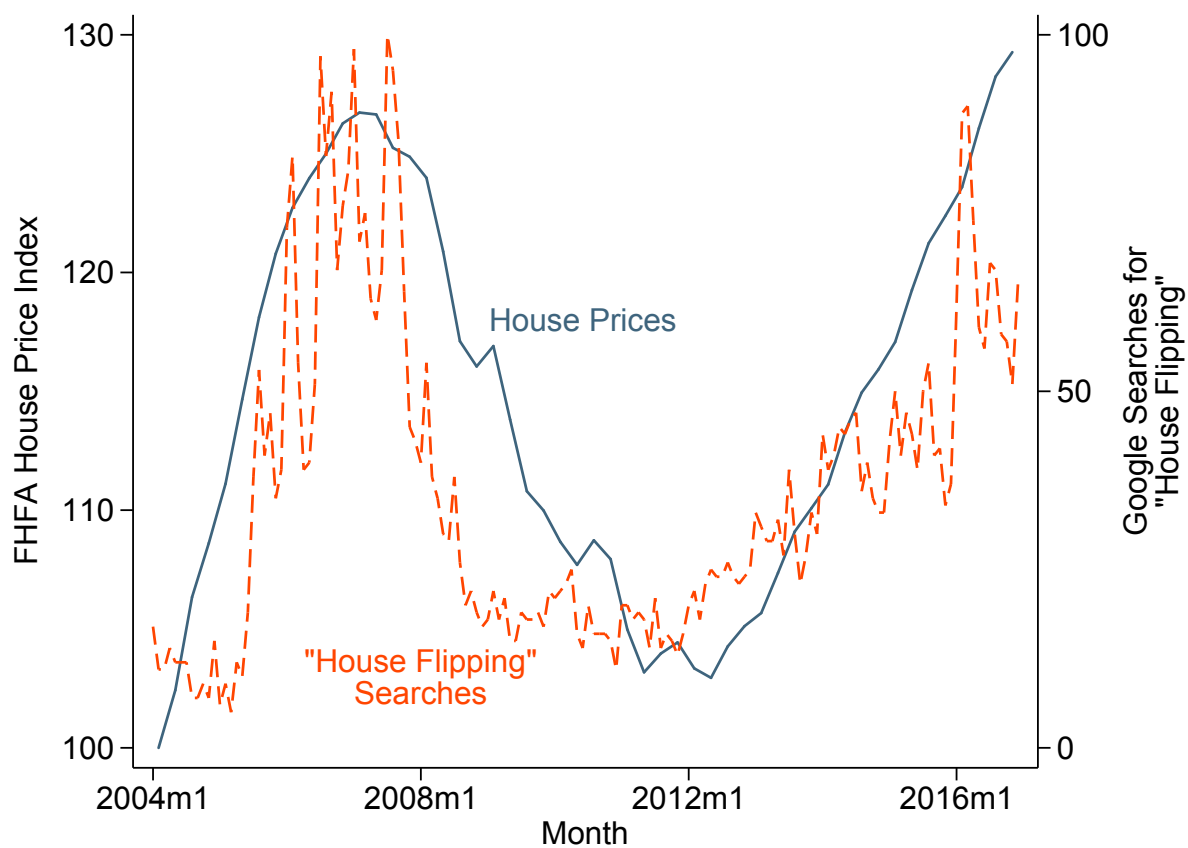
*Notes:* This table reports estimates of the quantitative importance of short-holding-period and non-occupant volume in accounting for the increase in total volume between 2000 and 2005 at the ZIP-code level. We assign each transaction to a Census ZIP Code Tabulation Area (ZCTA) using the postal ZIP 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 4  
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***<br>(3.813)        | 14.477***<br>(3.739) | −4.325***<br>(1.172) | −4.677***<br>(1.259) |
| $R^2$                       | 0.11                        | 0.10                 | 0.12                 | 0.12                 |
| Observations                | 115                         | 115                  | 115                  | 115                  |

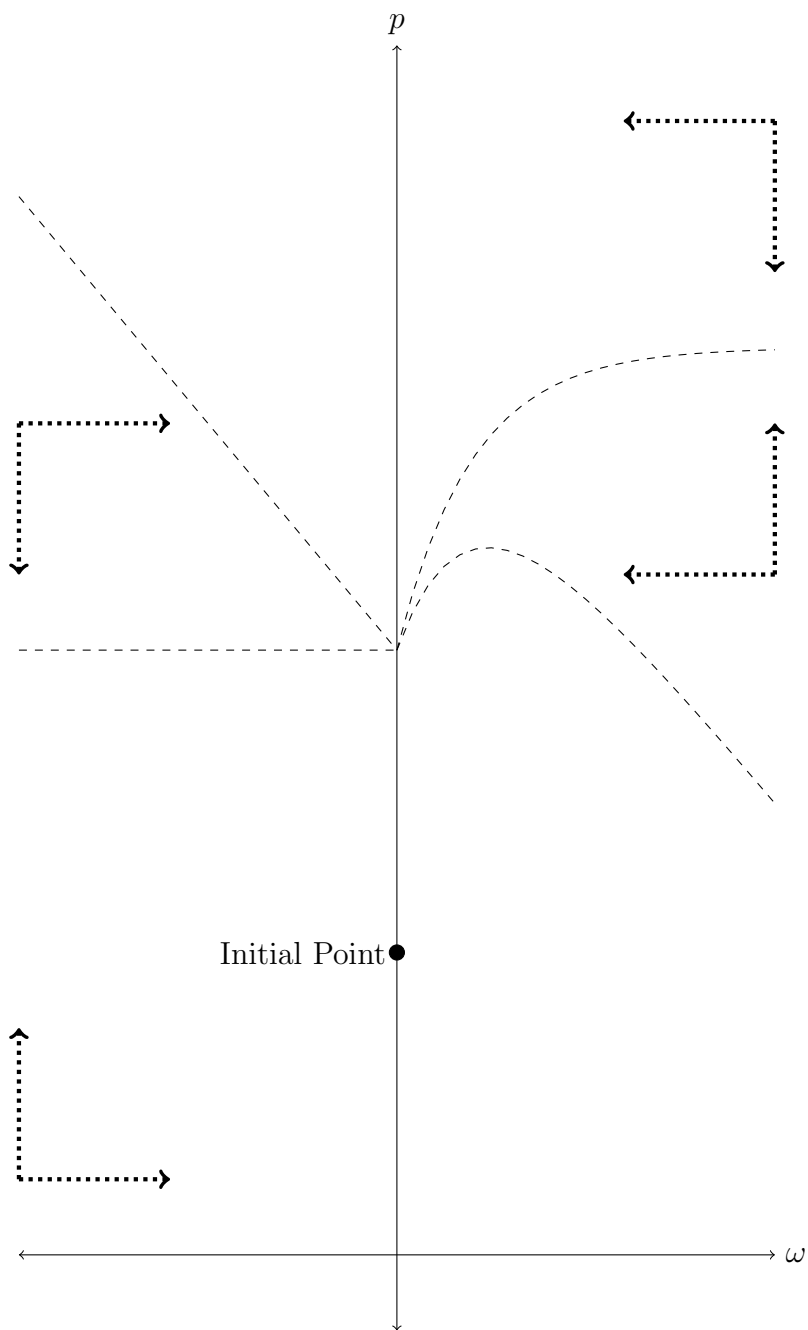
*Notes:* This table reports estimates of the cross-sectional relationship between the magnitude of the housing boom and bust and initial transaction volume at the MSA level. Each column reports estimates from a separate regression where the dependent variable is the percentage change in prices measured 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 indices. 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.

FIGURE A1  
Interest in “House Flipping” over the Housing Cycle



*Notes:* Search data were downloaded from Google Trends on January 31, 2017. House price index is the FHFA’s “all-transactions” series, which is nominal and quarterly. We normalize it to 100 in 2004q1 and map quarters to months using the midpoints.

FIGURE A2  
Phase Diagram



*Notes:* This figure illustrates the phase diagram for the  $(p, \omega)$  system specified by equations (A2) and (A3);  $p$  denotes the log house price, and  $\omega$  denotes the historical average log price change given by equation (10). The dashed loci indicate points at which either  $\dot{p} = 0$  or  $\dot{\omega} = 0$ . The dotted arrows indicate the directions  $p$  and  $\omega$  move in each of the four areas demarcated by the dashed loci. The system begins at the marked point on the  $p$ -axis.

TABLE A1  
Sensitivity of S&P 500 Return Forecasts to  
Historical Returns, 2000Q3–2011Q4

|                   | Lagged Annual Return |                                      | 75-Year Weighted Average Return |                                      |
|-------------------|----------------------|--------------------------------------|---------------------------------|--------------------------------------|
|                   | One-Year<br>Forecast | Ten-Year<br>Forecast<br>(Annualized) | One-Year<br>Forecast            | Ten-Year<br>Forecast<br>(Annualized) |
| Historical Return | 0.029**<br>(0.014)   | −0.013*<br>(0.007)                   | 0.045***<br>(0.010)             | 0.011<br>(0.015)                     |
| R-squared         | 0.12                 | 0.06                                 | 0.17                            | 0.03                                 |
| Observations      | 44                   | 44                                   | 44                              | 44                                   |

*Notes:* Return forecasts come from the Duke CFO Global Business Outlook, a quarterly survey of chief financial officers of U.S. firms. Historical returns on the S&P 500 come from CRSP’s daily dividend-inclusive value-weighted return series `vwretd`. The historical return equals  $P_t/P_{t-1} - 1$  in the first two columns and  $\mu(1 - e^{-\mu T})^{-1} \int_0^T e^{-\mu\tau} \dot{p}_{t-\tau} d\tau$  in the latter two columns, with  $p = \log P$ ,  $\mu = 0.5$ , and  $T = 75$ . The sample period is chosen to match that used by Greenwood and Shleifer (2014). Observations for 2001Q3 and 2002Q3 are dropped due to errors or gaps in the Duke CFO Global Business Outlook. Newey-West standard errors are reported in parentheses. Significance levels 10%, 5%, and 1% are denoted by \*, \*\*, and \*\*\*, respectively.



TABLE A2  
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 A3  
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.