The 2000s Housing Cycle With 2020 Hindsight: A Neo-Kindlebergerian View*

Gabriel Chodorow-Reich
Harvard University and NBER

Adam M. Guren
Boston University and NBER

Timothy J. McQuade
UC Berkeley

October 2022

Abstract

With "2020 hindsight," the 2000s housing cycle is not a boom-bust but a boom-bust-rebound. Using a spatial equilibrium regression in which house prices are determined by income, amenities, urbanization, and supply, we show that long-run city-level fundamentals predict not only 1997-2019 price and rent growth but also the amplitude of the boom-bust-rebound. This evidence motivates our model of a cycle rooted in fundamentals. Households learn about fundamentals by observing "dividends" but become over-optimistic in the boom due to diagnostic expectations. A bust ensues when beliefs start to correct, exacerbated by a price-foreclosure spiral that drives prices below their long-run level. The rebound follows as prices converge to a path commensurate with higher fundamental growth. The estimated model explains the boom-bust-rebound with a single shock and accounts quantitatively for the dynamics of prices, rents, and foreclosures in cities with the largest cycles. We draw implications for asset cycles more generally.

*Contact information: chodorowreich@fas.harvard.edu, guren@bu.edu, tmcquade@berkeley.edu.

We thank Elias Papaioannou (editor), four referees, and numerous seminar and conference participants and discussants for useful feedback. Sage Belz, Stefano Pica, and Lei Ma provided excellent research assistance. Chodorow-Reich thanks the Alfred P. Sloan Foundation and Harvard Ferrante Fund for generous financial support. Guren thanks the National Science Foundation (grant SES-1623801) for financial support. The computational work for this paper was performed on Boston University's Shared Computing Cluster. A replication package including details of the numerical computation is available from the authors' webpages.

1 Introduction

Asset price cycles are a recurring feature of financial history, yet our understanding of such cycles and the structural features that create fragility to them remains incomplete. Among the most consequential and studied in recent history is the 2000s housing cycle, both for its sharpness and its role in triggering the Great Recession. Real house prices in the United States rose by nearly 80% between 1997 and 2006 and then lost over two-thirds of their gain between 2006 and 2012. The boom-bust cycle was even more dramatic in some cities, with areas experiencing the most rapid price growth during the boom also having the largest price declines during the bust. A predominant view is that the boom-bust was the result of the emergence and popping of a house price "bubble" that was not rooted in economic fundamentals (Shiller, 2008; Charles et al., 2018).

We reevaluate the 2000s housing cycle from the perspective of 2020.[1] National real house prices grew steadily between 2012 and 2019, with the largest price growth in the same areas that had the largest booms between 1997 and 2006 and busts between 2006 and 2012. As a result, the areas that had the largest booms also had higher long-run price growth over the entire 1997-2019 period. With "2020 hindsight," the 2000s housing cycle is not a boom-bust but rather a boom-bust-rebound.

We argue that this pattern reflects a larger role for fundamentals than previously thought. In a first step, we use a standard spatial equilibrium framework to motivate fundamental determinants of location choice, land costs, housing supply, and house prices. We find that these determinants explain cross-city variation in long-run house price growth in reduced-form and structural regressions, as well as the amplitude of the boom-bust-rebound. In a second step, we introduce a model of a fundamentally-rooted house price cycle in which belief over-reaction amplifies the boom and a foreclosure spiral exacer-

[1] We consider developments until 2019 due to COVID-19's independent shock to housing markets.

bates the bust and discipline the model using simulated method of moments (SMM). The estimated model generates a boom-bust-rebound from a single fundamental shock and quantitatively matches the cross-city patterns in the metro areas most affected by the 2000s housing cycle. More broadly, our analysis formalizes the Kindleberger (1978) characterization of asset price cycles as initiated by fundamental improvements and amplified by overoptimism in the boom and fire sale dynamics in the bust and highlights the role of low interest rates in catalyzing these types of cycles.

Section 2 begins by establishing the strong cross-sectional correlation of price growth across the boom, bust, and rebound. We use the national time series to break the cycle into a boom (1997-2006), bust (2006-2012), and rebound (2012-2019). At the ZIP Code level, the boom correlates negatively with the bust but the bust correlates negatively with the rebound, each with an R^2 above 0.35. As a result, the boom correlates strongly with price growth between 1997 and 2019, with an R^2 of 0.62. We also show that rents experienced a trend break in the late 1990s, indicative of a fundamental improvement.

Section 3 introduces a long-run spatial equilibrium empirical framework to shed light on whether price growth over the boom-bust-rebound reflects fundamental forces. Building on Saiz (2010), we derive a structurally-interpretable supply regression of house price growth as a function of population growth, the land share of prices, housing supply regulation, and the premium to living downtown. We follow Saiz (2010), Diamond (2016), and others in selecting excluded instruments to address endogeneity, including shift-share predictors of local wages and employment, climate amenities, land unavailability, initial population density, taste-based determinants of regulation, and predictors of the movement of college-educated residents into land-scarce downtown areas.

Section 4 contains the empirical analysis of our framework. The excluded instruments strongly predict house price and rent growth over the boom-bust-rebound as well as the house price determinants of population growth, land share, regulation, and the downtown premium. In reduced form, they explain nearly 60% of the 1997-2019 cross-city variation in price growth and the structural IV relationship also has strong explanatory power. We define a city's long-run observed fundamental as the second stage fitted value, which depends only on the instruments and combines both demand and supply-side determinants of house prices using the structure of the spatial equilibrium framework. Areas with higher *long-run* fundamentals *also* have larger booms, deeper busts, and stronger rebounds.

Section 5 contains our quantitative model of a boom-bust-rebound. House prices equilibrate demand from potential entrants to a city and supply from the construction of new homes and foreclosures. An endogenous boom-bust-rebound cycle occurs in response to a single fundamental change, an increase in the growth rate of the income and amenities or "dividend" from living in the city.

Two model ingredients are instrumental to this result. First, agents learn about the true growth rate from observing dividends, but, in line with survey evidence (Shiller and Thompson, 2022), are over-optimistic in the boom. We formalize over reaction to news about fundamentals using *diagnostic* expectations (Bordalo et al., 2018, 2020b). Eventually, over-optimism peaks and beliefs start to converge toward the true growth rate, triggering the bust phase. On their own, however, a turn in beliefs cannot generate a bust in which prices fall below their full-information value or as steeply as in the data. This motivates the model's second key ingredient, mortgage borrowing and foreclosures. Consistent with empirical evidence, a foreclosure occurs when an under-water homeowner experiences a liquidity shock. Foreclosures add to housing supply, further depressing prices, putting more owners under-water, and leading to more foreclosures in a price-foreclosure spiral that pushes prices below their long-run level (Guren and McQuade, 2020). Finally, prices rebound as foreclosures recede and ongoing dividend growth causes new entry.

To quantitatively analyze the model, we use SMM to target long-run price movements, house price expectations, supply elasticities, and foreclosures, among other moments. We then assess the model's ability to explain untargeted short-run price dynamics and rent growth across four quartiles of cities grouped by their 1997-2019 price growth. The model fits the boom, bust, and rebound of prices and the growth of rents remarkably well for metro areas in the top half of the long-run growth distribution. It somewhat under-predicts the bust size in lower long-run growth areas, in part due to high supply elasticities in these cities and in part because we do not model spillovers across cities arising from either learning or the balance sheets of national lenders.

While several model features reflect our application to housing cycles, our results shed light on the economics of diagnostic asset price cycles more broadly. In particular, our model highlights how low interest rates create fragility. Intuitively, in a low interest rate environment current demand is more sensitive to expectations about dividends further in the future, making prices more responsive to over-optimism about dividend growth. Indeed, our quantitative analysis indicates the 2000s cycle would have been far milder with interest rates closer to their 1980s level.

Related literature. Our paper differs from other work on asset and housing cycles by presenting a model that generates a boom-bust-rebound from a single fundamental improvement and by using the rebound to infer the source of the cycle.

Most prior work relies on multiple exogenous changes in fundamentals or expectations for each turning point to deliver a boom-bust cycle (e.g., Kaplan et al. (2020) and Jacobson (2022)). Three notable exceptions generate endogenous cycles from a single initial shock as in our paper. First, Barberis et al. (1998) show that belief disagreement, over-optimism, or extrapolation in asset markets can generate a boom-bust. Second, Burnside et al.

(2016) present a model in which boom-busts occur through "social dynamics" similar to the epidemiological spread of disease. Third, in the context of housing Glaeser and Nathanson (2017) show that sufficiently extrapolative expectations lead to persistent price oscillations. Crucially, ours is the only theory that generates a boom-bust-rebound; the first two generate a boom-bust as over-optimism spreads and recedes about a fundamental improvement that is ultimately not realized, while the third leads to continual oscillations.

Empirically, prior work on housing has primarily considered the contribution of specific factors such as changes in credit conditions (Mian and Sufi, 2009; Favara and Imbs, 2015), speculators (Gao et al., 2020), over-optimism and expectations (Shiller and Thompson, 2022), or foreclosures (Guren and McQuade, 2020), although structural models have been used to disentangle the role of various factors (Kaplan et al., 2020; Favilukis et al., 2017; Greenwald and Guren, 2021). This literature has not focused on fundamentals, with a few notable exceptions.[2] Writing near the peak of the boom, Himmelberg et al. (2005) found "little evidence of a housing bubble" in part due to fundamental growth. Ferreira and Gyourko (2018) show that the timing of the boom start in each city was "fundamentally based to a significant extent" but that fundamentals revert in roughly three years.[3] Kaplan et al. (2020) and Glaeser and Nathanson (2017) microfound changes in expectations with shocks to future fundamentals but do not treat fundamentals as a measured and central feature of the cycle. Our focus on fundamentals accommodates a role in the boom-bust for forces such as subprime credit expansion or speculation, which we show are essentially uncorrelated with our measured fundamental in the cross-section of cities. Such transitory factors cannot, however, explain the rebound or why places with large booms also had

[2] For instance, Charles et al. (2018) describe a "consensus that much of the variation in housing prices during the boom and bust derived from a speculative 'bubble' and not from changes in standard determinants of housing values such as income, population, or construction costs."

[3] More recently, Howard and Liebersohn (2022, 2021) examine divergence in demand across areas as a fundamental and Schubert (2021) identifies spillovers of fundamentals across cities via migration networks.

higher growth over the full cycle.

Methodologically, our empirical analysis builds on an urban economics literature on the long-run determinants of housing costs and location choice, notably Saiz (2010). Our theoretical framework incorporates diagnostic expectations (Bordalo et al., 2019, 2020a) into a continuous time framework with asset supply and demand and contains methodological innovations that may prove useful in other contexts, including a new characterization of the impulse response of diagnostic beliefs.

Finally, our interpretation of the 2000s housing cycle echoes the seminal work of Kindleberger (1978). For instance, Kindleberger writes that "virtually every economic mania is associated with a robust economic expansion."[4] As in Kindleberger, our focus on long-run fundamentals does not imply that there was no "housing bubble." Our work expands on Kindleberger by emphasizing the role of the rebound in diagnosing the driving forces of the cycle and by estimating a quantitative model of a fundamentally-rooted cycle.

2 Boom, Bust, and Rebound

This section documents the boom, bust, and rebound phases of the 2000s housing cycle in both national and local data and the rise in rent growth around the start of the boom.

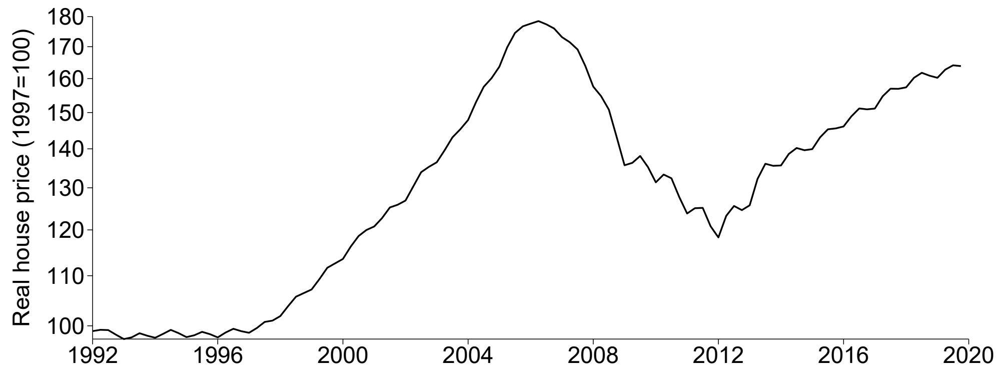
Figure 1 shows the national Case-Shiller house price index, deflated using the GDP price index. The series begins to rise in the late 1990s, peaks in 2006Q2, reaches a local trough in 2012Q1, and then grows again through the end of our sample in 2019. We use this timing to define the boom, bust, and rebound periods for the remainder of the paper.[5]

Figure 2 shows the correlation of the boom, bust, and rebound at the local level. Each

[4]Barberis et al. (2018) also emphasize this feature of Kindleberger, but they interpret the fundamental improvement as repeated good news about fundamentals rather than an actual fundamental improvement.

[5]Appendix Figure B.3 shows the timing of the boom start across cities using a procedure similar to Ferreira and Gyourko (2018). Because few booms start before 1997, we use 1997 as the boom start for our analysis. The figure shows much more uniformity in the timing of the price peak and trough across cities.

Figure 1: National Boom, Bust, and Rebound



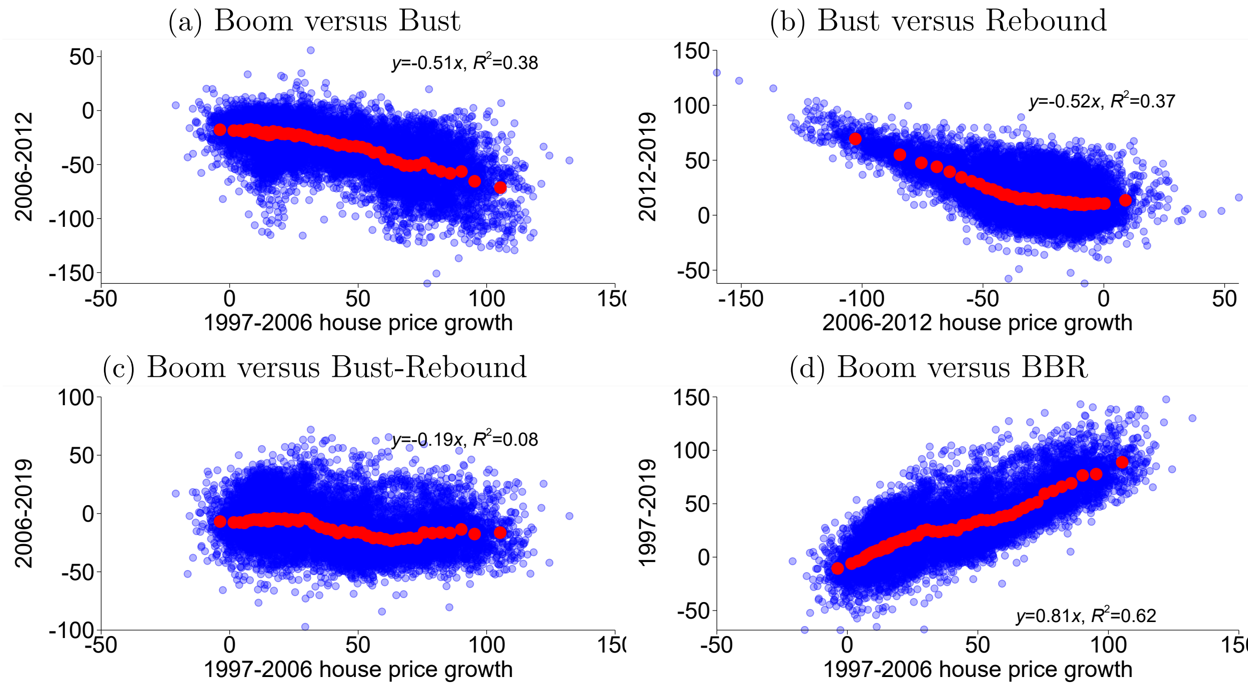
Notes: The figure shows the national Case-Shiller index deflated by the GDP price index.

blue circle represents one ZIP Code and the overlaid red circles show the mean value of the y-axis variable in each of 50 quantiles of the x-axis variable. Panel (a) shows the correlation of price growth in the boom and the bust. Each additional percentage point of house price appreciation in the boom is associated with an additional decline of 0.51 percentage point in the bust and the R^2 of this relationship is 0.38.

Panel (b) reveals an equally strong correlation between the bust and post-2012 price growth, with each additional percentage point decline during 2006-2012 associated with an additional 0.52 percentage point of growth during 2012-2019 and an R^2 of 0.37. The negative correlation between bust and rebound suggests an over-shooting of prices on the downside during the bust, just as the negative correlation between boom and bust points to a bubble in the boom. Putting the bust and rebound together in Panel (c), house price growth in the boom has a much weaker correlation with total price growth after 2006 than with the bust only. Panel (d) displays the corollary of this result: House price growth during the boom correlates strongly with growth over the entire 1997-2019 period (BBR for short), with a slope coefficient of 0.81 and R^2 of 0.62. The boom was not ephemeral.

Figure 3 shows the paths of rents for all cities and for averages of each of four quartiles of 1997-2019 house price growth based on an unbalanced sample of 27 CBSAs over this

Figure 2: Zip Code Boom, Bust, and Rebound



Notes: Each blue circle represents one ZIP Code. The red circles show the mean of the y-axis variable for 50 bins of the x-axis variable. Data from FHFA deflated using the national GDP price index.

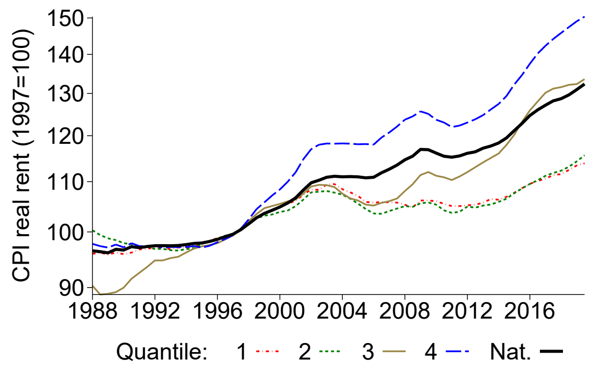
period. CBSAs with faster long-run price growth also had faster rent growth. Moreover, there is a visible trend break in the late 1990s that is stronger in higher-price-growth areas, which we confirm in Table B.2 using structural break tests. Appendix C.5 associates such rent growth acceleration with improving fundamentals. We nonetheless focus mostly on prices rather than rents because of their wider coverage and higher quality.[6]

3 Long-run Fundamentals: Framework

This section introduces a long-run supply-and-demand framework for house prices and describes our data. The framework allows us to move beyond correlations and associate areas with faster growth over the full 1997-2019 period with fundamental determinants

[6]Crone et al. (2010) document important methodological deficiencies prior to 1988 in the CPI's repeat-sampling methodology. The only other yearly city-level data set dating to the 1990s, HUD fair market rents, has imputations in most CBSAs prior to 2003 and has a break in 1995 when the reference percentile changes.

Figure 3: Rent Growth By House Price Quartile



Notes: The figure plots rent indexes from the Bureau of Labor Statistics' Consumer Price Index deflated by the national GDP price index for CBSAs grouped into quartiles based on 1997-2019 house price growth.

of house prices. As a by-product, it produces housing supply elasticities that we use to calibrate the model in Section 5.

3.1 Structural System

The supply block starts with an additive decomposition of house prices into the value of the structure and the land:

$$P_{i,t} = C_{i,t} + L_{i,t}, \tag{1}$$

where $P_{i,t}$ is the price of a house in area i at date t, $C_{i,t}$ is the replacement cost of the structure, and $L_{i,t}$ is the land cost.[7] Gyourko and Saiz (2006) argue that the construction sector is sufficiently competitive to justify this decomposition.

Both construction and land costs may increase with population in a city:

$$C_{i,t} = A_{i,t} H_{i,t}^{\alpha_i}, \tag{2}$$

$$L_{i,t} = B_{i,t} H_{i,t}^{\beta_i}. \tag{3}$$

[7]Other proportional costs such as broker fees drop out when we take log changes below. We abstract throughout from changes in average house size.

Here, $H_{i,t}$ is total city population (H for house), α_i and β_i are city-specific long-run elasticities, and $A_{i,t}$ and $B_{i,t}$ are cost shifters unrelated to population.[8] Letting $s_{i,t} = L_{i,t}/P_{i,t}$ denote the land share of prices, the overall long-run supply elasticity is $\eta_{i,t} \equiv [\alpha_i + s_{i,t}(\beta_i - \alpha_i)]^{-1}$. Importantly, these cost functions apply to long-run changes, as they omit short-run dynamics stemming from accelerating population growth or foreclosure dynamics that we introduce in Section 5. We apply equations (1) to (3) to price growth between 1997 and 2019, effectively assuming that *on average* across cities house prices at these two end points reflect long-run fundamentals.[9]

A spatial equilibrium demand equation relates population growth $\dot{H}_{i,t}/H_{i,t}$ to the present value of the income and amenities from living in city i, $V_{i,t}$, and the price:

$$\dot{H}_{i,t}/H_{i,t} = G_i \left(V_{i,t}/P_{i,t}\right)^{\gamma} \quad \text{where} \quad V_{i,t} = \mathbb{E}_t \int_t^{\infty} e^{-\rho s} D_{i,s} ds. \tag{4}$$

Here G_i is a constant, γ is the elasticity of demand, ρ is the discount rate, and $D_{i,t}$ is the "dividend" that captures the income/amenities from living in an area. Equation (4) is a dynamic spatial equilibrium condition in which households choose location to maximize their earning potential and amenity value net of housing costs. Section 5 provides a microfoundation for this functional form (see equation (17)); intuitively, $V_{i,t}/P_t$ is the Tobin's Q associated with an area.[10]

We next derive a regression specification that relates observable counterparts of the terms in equations (1) to (4) to the cross-section of 1997-2019 price growth.

[8] Appendix A.1 provides a microfoundation of equation (2) from the cost-minimization problem of a competitive construction sector. Appendix A.2 provides a microfoundation of equation (3) as in Saiz (2010) from an Alonso-Muth-Mills intra-city spatial equilibrium condition, wherein average land prices in a city grow with population because the premium to living in the city center (or equivalently most desirable neighborhoods) rises to induce new housing in less desirable locations.

[9] City-specific, idiosyncratic deviations from fundamentals are in the structural residual, as are city-specific changes in fundamentals not captured by our instruments. The latter make our conclusions about the role of fundamentals a likely *lower bound* for the total role of fundamental drivers.

[10] The assumption that V does not depend on P holds if utility from housing is independent of non-housing goods or both enter Cobb-Douglas.

3.2 Regression Specification

Taking logs of equation (1), differencing over time, and letting lower case p, c, ℓ, a, b denote the log differences of their respective upper case letters, we obtain:

$$p_{i,t} = (1 - s_{i,t})\, c_{i,t} + s_{i,t}\ell_{i,t} \tag{5}$$

$$= a_{i,t} + s_{i,t}\,(b_{i,t} - a_{i,t}) + (\alpha_i + s_{i,t}\,(\beta_i - \alpha_i))\, h_{i,t}. \tag{6}$$

Equation (5) decomposes house price growth into a weighted average of the growth of construction and land costs. Equation (6) imposes our functional forms. We further parameterize $\alpha_i = \alpha_0 + \alpha_1 m_i$ and $\beta_i = \beta_0 + \beta_1 m_i$, where m_i measures the regulatory burden of new construction in city i relative to the cross-city mean and α_0 and β_0 are the average elasticities. This parameterization allows construction and land costs to increase more steeply with the number of houses in places with stricter land-use regulations.

The component $b_{i,t}$, the growth in land prices unrelated to total population growth, contains the trend toward urbanization and gentrification that began in many cities around the turn of the millennium (Couture and Handbury, 2020). This movement of college-educated, high-income individuals into downtown neighborhoods where land is relatively scarce pushed up the cost premium to living in the city center, causing land and house prices to rise everywhere (Su, 2022). We capture the city-specific dimension of this trend by parameterizing $b_{i,t} = bu_{i,t} + \bar{b}_t + \hat{b}_{i,t}$, where $u_{i,t}$ denotes the log change in the house price premium for buying downtown relative to the rest of the city, $\bar{b}_t$ is the cross-city average secular land price growth, and $\hat{b}_{i,t} = b_{i,t} - bu_{i,t} - \bar{b}_t$. We similarly demean $a_{i,t}$ and let $\epsilon_{i,t} = \hat{a}_{i,t} + s_{i,t}\left(\hat{b}_{i,t} - \hat{a}_{i,t}\right)$ to arrive at the structural equation:

$$\begin{aligned} p_{i,t} = {} & \bar{a}_t + s_{i,t}\left(\bar{b}_t - \bar{a}_t\right) + \alpha_0 h_{i,t} + (\beta_0 - \alpha_0)\, s_{i,t}h_{i,t} \\ & + \alpha_1 m_i h_{i,t} + (\beta_1 - \alpha_1)\, m_i s_{i,t} h_{i,t} + s_{i,t} b u_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{7}$$

Equation (7) corresponds to the regression equation:

$$\begin{aligned} p_{i,t} &= c_0 + c_1 s_i + c_2 h_{i,t} + c_3 \left(s_i \times h_{i,t}\right) \\ &\quad + c_4 \left(m_i \times h_{i,t}\right) + c_5 \left(m_i \times s_i \times h_{i,t}\right) + c_6 \left(s_i \times u_{i,t}\right) + e_{i,t}, \end{aligned} \tag{8}$$

with $c_0 = \bar{a}_t, c_1 = \bar{b}_t - \bar{a}_t, c_2 = \alpha_0, c_3 = \beta_0 - \alpha_0, c_4 = \alpha_1, c_5 = \beta_1 - \alpha_1, c_6 = b$.

Equation (8) is a long-run supply equation.[11] The coefficient c_1 identifies average excess secular (not driven by population) increase in land prices over construction costs, c_2 the average long-run elasticity of construction costs to population, c_3 the difference in the average elasticities of land and construction costs to population growth, c_4 the increase in the construction cost elasticity from higher regulatory strictness, c_5 the difference in the increases in the land and construction cost elasticities, and c_6 the contribution of city-specific urbanization. The inverse supply elasticity is $c_2 + c_3 \times s_i + c_4 \times m_i + c_5 \times s_i \times m_i$.

3.3 Data, Measurement, and Excluded Instruments

Outcome and endogenous variables. We estimate equation (8) across 308 core-based statistical areas (CBSAs) over 1997-2019. House price data come from Freddie Mac, deflated by the national GDP price index. We measure quantity growth by aggregating Census intercensal estimates of county housing units. We obtain the CBSA land share in 2012 from Larson et al. (2021). We equate regulatory strictness with the 2006 Wharton Residential Land Use Regulatory Index (WRLURI) developed in Gyourko et al. (2008). We measure growth in the urban premium by defining a downtown as in Couture and Handbury (2020) and using ZIP code house price indexes to compute relative price growth,

[11] Equation (8) generalizes Saiz (2010) by treating the land share as observable, by allowing for construction costs to respond endogenously to population, and by explicitly modeling urbanization. Specifically, the Saiz (2010) model starts with equations (1) to (3) with $\alpha_i = 0\ \forall i$ and $b = 0$. In our notation, the final specification in Saiz (2010) is $p_{i,t} - c_{i,t} = k_1 (1 - \Lambda_i) \times h_{i,t} + k_2 \ln H_{i,0} \times (1 - \Lambda_i) \times h_{i,t} + k_3 \times m_i \times h_{i,t} + e_{i,t}$, where Λ_i denotes the share of land available for development. Our approach instead treats Λ_i and $\ln H_{i,0}$ as excluded instruments that help to identify the land share terms in the supply elasticity.

as described in Appendix B.4. Table B.4 presents summary statistics.

Excluded instruments. We treat population growth, land share, regulatory strictness, urbanization, and their interactions as potentially endogenous and estimate equation (8) using instrumental variables. A simultaneity problem arises because $\epsilon_{i,t}$ contains city-specific, secular (i.e. unrelated to population) changes in land or construction costs that increase house prices and reduce population growth or urbanization. Since we observe land share and zoning regulation only well into the cycle, if these variables respond to population or prices they are endogenous as well. We heuristically group excluded instruments by which endogenous variable they most closely affect and appropriately interact these groups as well to produce the complete excluded instrument set. We purposely follow the existing literature in choosing excluded instruments to make the point that a fundamentally-rooted interpretation of the 2000s cycle emerges from standard house price determinants.

Equation (4) motivates excluded instruments for population growth, which form the basis for the CBSA-level observable demand-side components of long-run fundamental growth. Specifically, labor demand and amenities that shift the growth of $V_{i,t}$ constitute valid demand shifters to identify the long-run supply elasticity. We follow Saiz (2010) and use shift-share predictors of employment and wage growth as labor demand shifters (see Appendix B.3 for details). We use January temperature and sunlight and July humidity as climate-related amenities that capture population movement toward the "sunbelt." We additionally follow Diamond (2016) and use the 1997 share of employment in restaurants.

We also follow Saiz (2010) in choosing instruments relevant for the land share of the price and regulatory strictness, which form the basis for the supply-side components of long-run fundamental growth. For land share, we use the fraction of land available for development (not water or steep slope) from Lutz and Sand (2019) and 1997 population density.[12] For regulatory strictness, we use the ratio of public expenditure on protective inspection to total tax revenue in the 1992 Census of Governments and the share of Christians in nontraditional denominations in the 1990 Census.[13]

Our instruments for the change in the downtown price premium draw on recent work that pinpoints changing tastes by college-educated residents for urban amenities such as bars and restaurants that started around the late 1990s (Couture and Handbury, 2020; Baum-Snow and Hartley, 2020). Appendix A formalizes this force by extending the Alonso-Muth-Mills model of intra-city spatial equilibrium to include college and non-college residents with time-varying preferences for living in a downtown core and motivates two excluded instruments (see equation (A.15)): (i) the interaction of the pre-boom (1990) share of college workers in the CBSA and pre-boom urban amenities, which we measure as the ratio of restaurant density in the downtown and non-downtown; and (ii) the interaction of the pre-boom relative likelihood of living downtown for college and non-college residents and the predicted change in the CBSA college share using a Bartik shift-share. Appendix B.4 contains details of the measurement.

Let $\mathcal{H}, \mathcal{L}, \mathcal{M}$, and $\mathcal{U}$ denote the sets of instruments heuristically assigned to population growth, land share, WRLURI, and urbanization, respectively. The full excluded instrument set consists of $\mathcal{H}, \mathcal{L}, \mathcal{M}, \mathcal{H} \times \mathcal{L}, \mathcal{H} \times \mathcal{M}, \mathcal{H} \times \mathcal{L} \times \mathcal{M}$, and $\mathcal{L} \times \mathcal{U}$, where $\times$ denotes element-wise cross-set multiplication. The linear combination of these instruments formed by the second stage fitted value is the long-run observable fundamental.

What is the Shock? Many of the excluded instruments are either persistent or time invariant. Consistent with this fact, areas with higher predicted house price growth over

[12] Appendix A.2 provides a formal motivation for these instruments in the context of the Alonso-Muth-Mills intra-city spatial equilibrium microfoundation of equation (3) (see equation (A.10).

[13] Saiz (2010) motivates protective expenditure as revealing an area's taste for regulation and the nontraditional Christian share because these denominations' ethos of individualism leads to reduced regulation.

1997-2019 based on these measures also had faster house price and population growth over 1975-1997. Persistence in instruments associated with land share and regulation occurs naturally, since supply heterogeneity is a persistent feature of areas. Nothing in the econometric setup precludes using persistent demand instruments to identify a long-run supply elasticity; instead, the framework simply requires demand instruments that shift population growth and are orthogonal to unobserved, location-specific supply shifters.

Persistence in some instruments, however, raises the question of what changed to trigger the cycle. Several of the instruments do embody changes in the late 1990s, notably educated workers moving to downtown neighborhoods and changes in industry growth.[14] Other, more persistent attributes such as climate may have coincided with changing preferences for amenities that accelerated in the late 1990s. More generally, rising tastes for city dwelling, due for example to the widespread decline in crime (Pope and Pope, 2012; Ellen et al., 2019) or changing preferences for urban amenities, increased urban land values nationwide, with the largest impact on house prices in areas with high land shares or already experiencing rapid growth.[15] These forces together comprise the shock.

4 Long-run Fundamentals: Results

This section contains our main empirical results. Section 4.1 shows that the excluded instruments strongly predict house price growth over 1997-2019. Section 4.2 reports the IV results. Section 4.3 defines observed long-run fundamental growth as a linear combination of the excluded instruments and shows that higher fundamental growth predicts a larger

[14] A regression of the predicted employment growth instrument on a shift-share for predicted employment growth over 1986-1996 has an R^2 of 0.35, indicating some but far from full persistence.

[15] Because the present value of dividends in high growth areas is dominated by dividends farther in the future, marginal increases in demand growth matter more, similar to convexity in bond pricing. Formally, if the present value of income/amenities $V_{i,t}$ has a Gordon growth representation with discount rate ρ and spot dividend growth μ_i, $V_{i,t} = D_{i,t}/(\rho - \mu_i)$, then a common increase in μ_i will increase V by more in places where μ_i was already large. The calibration in Section 5 accounts for differences in pre-boom house price growth rates across CBSAs, so this force is present in our model.

amplitude of the boom-bust-rebound cycle.

4.1 Reduced-Form Results

Appendix Table B.5 reports first-stage-type regressions for each endogenous variable separately, using only the excluded instruments motivated by that variable and also using the full set of uninteracted instruments. In brief, the instruments act as expected and strongly predict their respective endogenous variables.

Figure 4 plots the fitted values from regressing house price growth over 1997-2019 on all of the uninteracted instruments against actual house price growth in various subperiods. Panel (a) shows a strong reduced form fit with an R^2 of 0.58, illustrating that fundamental drivers of location choice, land share, regulation, and urbanization measured prior to the start of the boom explain a substantial amount of the variation in house price growth over 1997-2019. Panels (b)-(d) show that higher predicted long-run growth correlates positively with the magnitudes of the boom, bust, and rebound separately. Thus, the reduced-form evidence establishes a correlation between long-run fundamental determinants of house price growth and a boom-bust-rebound cycle without imposing the structure of equations (1) to (4) and the IV specification. Panel (e) shows that these fundamental determinants also predict 2000-2019 rent growth, with an R^2 of 0.39.[16]

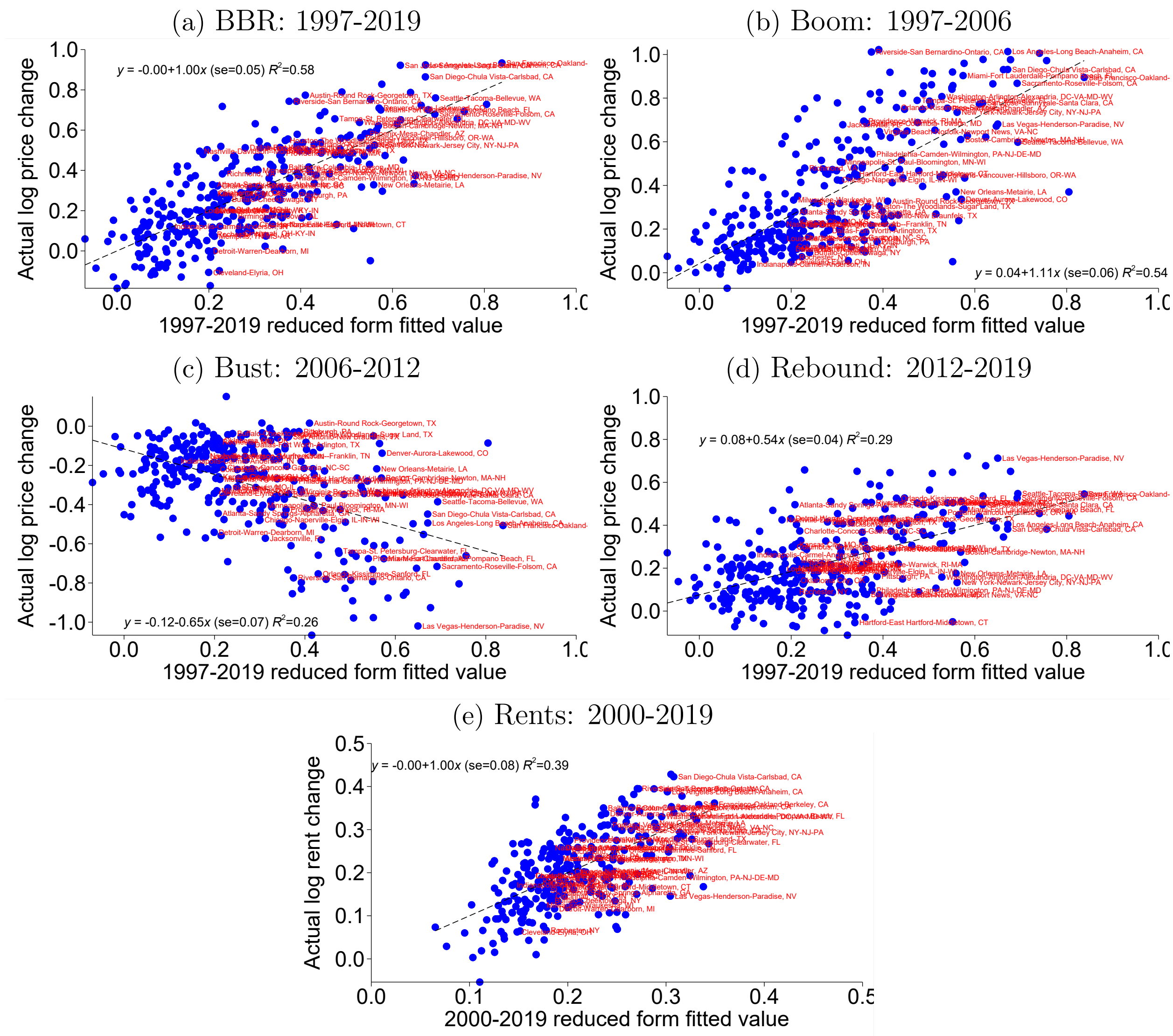
4.2 Structural IV Results

The full IV results impose additional restrictions by forcing the instruments to act through the endogenous variables in the model. They also yield structurally interpretable long-run housing supply elasticities that we use to calibrate the model in Section 5.

Table 1 presents the results from estimating equation (8). Column (1) shows OLS.

[16]Panel (e) uses Census median rents rather than the BLS data in Figure 3 to have a larger sample. The main drawbacks are that the Census data start in 2000 and there are no quality adjustments.

Figure 4: Reduced Form Correlations



Notes: In panels (a)-(d), each blue dot is the real house price growth in a CBSA over the period indicated on the vertical axis plotted against the predicted real house price growth over the period 1997-2019 based on column (7) of Table B.5. In panel (e), each blue dot is real rent growth between the 2000 Census and 2019 one-year ACS on the vertical axis plotted against the fitted value from a regression on the excluded instruments. CBSAs with more than one million people in 1997 are labeled in red.

CBSAs with higher land share and faster population growth had higher house price growth over the full BBR, and especially so in places with both high land share and high regulation. Evaluated at the (unweighted) mean land share and regulatory burden, the long-run inverse supply elasticity is 0.54 with a standard error of 0.26 using the delta method.

Table 1: Long-run OLS and IV Results

Dep. var.:	House Price Growth 1997-2019			Rent Growth 2000-19
	(1)	(2)	(3)	(4)
Land Share	0.64**	0.91*	0.78**	0.17*
	(0.24)	(0.37)	(0.20)	(0.08)
Units Growth	0.34	0.81	0.63**	0.16**
	(0.26)	(0.49)	(0.10)	(0.04)
Land Share × Units Growth	0.72	−0.64		
	(0.92)	(1.63)		
WRLURI × Units Growth	−0.01	0.32		
	(0.12)	(0.23)		
Land Share × WRLURI × Units Growth	0.79*	0.34	1.30**	0.34**
	(0.37)	(0.75)	(0.28)	(0.11)
Land share × Urbanization	1.22**	1.42**	1.40**	0.40**
	(0.19)	(0.37)	(0.38)	(0.14)
Constant	−0.05	−0.13	−0.11+	0.11**
	(0.06)	(0.11)	(0.06)	(0.02)
Estimator	OLS	2sls	2sls	2sls
Elasticity at $\bar{s}_j$	0.54	0.63	0.63	
Standard error of elasticity	0.26	0.50	0.10	
R^2	0.49	0.43	0.45	0.17
Observations	308	308	308	272

Notes: The table reports OLS (column 1) and IV (columns 2-4) regressions of real CBSA house price growth over 1997-2019 or rent growth over 2000-2010 on land share, housing unit growth over 1997-2019, their interactions with WRLURI and each other, and the interaction of land share and the change in the downtown price premium, as in equation (8). The standard error of the elasticity at the mean of land share is computed using the delta method. Heteroskedastic-robust standard errors in parentheses. $^{**}, ^{*}, ^{+}$ denote significance at the 1, 5, and 10 percent levels, respectively.

Columns (2) and (3) report our main IV results. Column (2) reports the IV specification using all of the excluded instruments. Given the large number of interaction terms and instruments, the data are unable to tightly identify each interaction in column (2). Column (3) consequently constrains the coefficients on land share × population growth and WRLURI × population growth to be zero.[17] Imposing these restrictions results in

[17] The supply framework interprets these restrictions as the intercepts of the elasticities of land and construction are the same, $\beta_0 = \alpha_0$, and the construction elasticity α_i does not vary with land-use regulation.

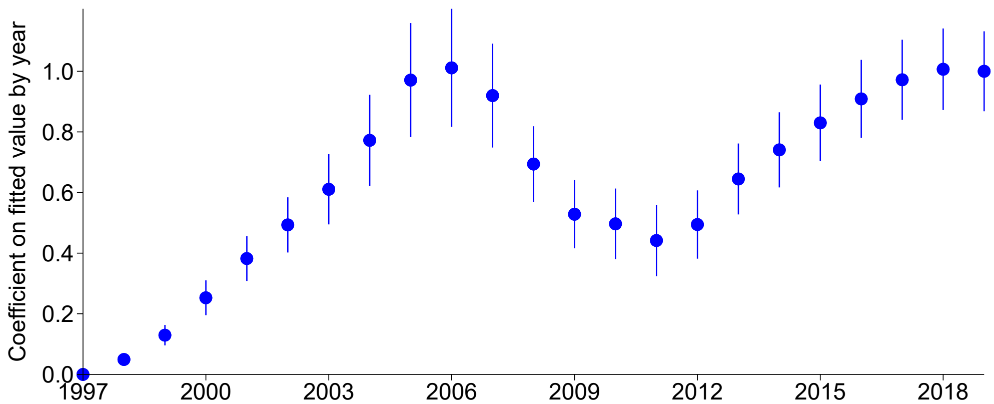
tightly identified coefficients without sacrificing fit: The remaining coefficients have much smaller standard errors, the R^2 *rises* slightly, and the mean inverse supply elasticity is unchanged. Column (3) is thus our preferred specification.[18]

Several features merit comment. The R^2 of 0.45 in column (3) reveals strong explanatory power of land share, population growth, WRLURI, and urbanization when imposing the IV coefficients, illustrating the central message that fundamentals can explain a substantial amount of the variation in house price growth over the entire BBR. The impact of population growth and the average inverse elasticity are larger with IV, consistent with the expected bias of OLS due to area-specific cost shocks. The coefficient on the triple interaction term indicates a larger inverse elasticity (price growth more sensitive to population) in areas with high land share and high regulation. The land share coefficient of 78 log points suggests a significant role for nationwide forces affecting land prices as discussed previously and is consistent with Nichols (2019). The large and positive coefficient on land share $\times$ urbanization indicates that a portion of the land share increase is due to increasing desirability of land-scarce downtown neighborhoods, underscoring the role of this particular channel in triggering the increase in fundamental growth in the late 1990s.

The IV fitted value permits a decomposition of price growth into the part coming from CBSA-wide population growth, the fitted values of the terms including $h_{i,t}$ in equation (8), and the remainder which includes within-CBSA urbanization and the contribution of rising average land prices nationally. Applying this decomposition, CBSA-wide population growth at the estimated long-run elasticities explains 57% of average house price growth, but land price growth and urbanization explain slightly more of the cross-CBSA variance.

[18] Appendix B.5 collects several alternative specifications that address various potential concerns, including (i) alternative house price indexes; (ii) using the Saiz (2010) measure of land unavailability, or population in place of housing units; (iii) weighting by population or dropping areas with fewer than 150,000 people or shrinking population; (iv) changing the estimator to GMM, JIVE, or bias-adjusted 2SLS; (v) excluding groups of excluded instruments one-at-a-time; and (vi) controlling for lagged house price or units growth.

Figure 5: Fundamentals and House Prices Over Time



Notes: The figure plots the coefficients $\{\beta_{1,h}\}$ and 95% confidence intervals from regressions at each horizon h of house price growth between 1997 and 1997+h on the long-run fundamental using the specification $p_{i,t,t+h} = \beta_{0,h} + \beta_{1,h}\hat{p}_{i,t} + \nu_{i,h}$, where $\hat{p}_{i,t}$ denotes the second stage fitted value from column (3) of Table 1.

Figure B.5 provides a less structural decomposition and shows that the supply and demand instruments have approximately equal power for predicting 1997-2019 price growth and both predict more pronounced cycles. We conclude that all of the elements of the long-run framework matter to explaining house price growth in our sample.

Column (4) shows that long-run fundamentals also predict rent growth. The structural coefficients are smaller than for price growth, reflecting the fact that prices rose faster than rents over the full 1997-2019 period, a result we return to in Section 5, and the data limitation that we measure rents in the 2000 Census and 2019 ACS to maximize sample coverage and therefore lose the first few years of the boom.

4.3 Long-Run Fundamental and Boom-Bust-Rebound

We define the long-run observable fundamental as the second stage fitted value corresponding to column (3) of Table 1. Associating the fundamental with the second stage fitted value makes it a linear combination of the excluded instruments, none of which depends on local characteristics that evolve after the start of the boom.

Figure 5 reports the coefficients $\{\beta_{1,h}\}$ from a cross-sectional regression of log real price growth between 1997 and 1997+h on the long-run fundamental. Areas with higher long-run fundamentals have more pronounced booms, busts, and rebounds.[19] Appendix B.1 and Figure C.2 further show that fundamentals are essentially orthogonal to other forces such as speculation or subprime credit expansion. These results together motivate a quantitative model of a fundamentally driven boom-bust-rebound cycle.

5 A Quantitative Neo-Kindlebergerian Model

Areas experiencing faster price growth over 1997-2019 had higher fundamental growth and also experienced a larger boom-bust-rebound. We now offer a "neo-Kindlebergerian" interpretation of this finding using a structural model calibrated to the cross-section of cities. The model nests the empirical framework in the sense that it gives rise to the same long-run supply and demand system as in equations (1) to (4), and we use the estimated supply elasticities in the calibration. It enriches the prior framework by introducing an explicit stochastic process for income/amenities, non-rational beliefs, adjustment costs to housing, and mortgages with foreclosures that together give rise to the boom-bust rebound.

As in Kindleberger's *Manias, Panics, and Crashes*, a single change in the economy's fundamentals sets off the asset price cycle. In our urban setting, this change takes the form of an increase in the growth rate of the "dividend" from living in a city. Agents learn about the growth rate by observing the history of dividends but become overly optimistic due to *diagnostic* expectations, causing a boom. Eventually, beliefs correct, causing house prices to fall. As prices fall, some under-water homeowners default, triggering a price-foreclosure spiral in the bust. This "crash" causes prices to fall below their long-run level.

[19]This pattern is not driven only by cities that are historically more sensitive to aggregate dynamics. Controlling for the 1978-1996 cyclical sensitivity using the procedure in Guren et al. (2021a) results in yearly coefficients that never differ by more than 0.002 from those plotted in Figure 5.

Finally, ongoing dividend growth induces new buyers to enter and prices to rebound. Both over-optimism and credit with foreclosures are necessary to generate a realistic boom-bust-rebound from the single change in the economy's fundamentals.

5.1 Environment

Time is continuous. The economy consists of cities each with population $H_{i,t}$ and a residual "hinterland." We describe the determination of population, house prices, mortgage distribution, and foreclosures in each city and suppress the i subscript for convenience.

Dividend and beliefs. The driving force in the economy is the "dividend," D_t, from living in the city. As in equation (4), this unidimensional object captures the combination of income prospects and amenities from living in the city relative to the hinterland. The value of D_t evolves as a geometric Brownian motion with stochastic drift:

$$dD_t = \mu_t D_t dt + \sigma_D D_t dW_{D,t}, \tag{9}$$

where $dW_{D,t}$ is a standard Wiener process. The drift μ_t follows an Ornstein-Uhlenbeck process:

$$d\mu_t = \vartheta \left(\bar{\mu} - \mu_t\right) dt + \sigma_\mu dW_{\mu,t}, \tag{10}$$

where ϑ determines the rate of convergence to the unconditional mean $\bar{\mu}$ and $dW_{\mu,t}$ is a standard Wiener process uncorrelated with $dW_{D,t}$. Agents observe D_t but do not know the instantaneous drift rate μ_t. The single realized shock in the model will be an increase from $\bar{\mu}$ to $\mu_0 > \bar{\mu}$ at the start of the boom at date 0.

A Bayesian agent would form beliefs over μ_t from the path of observed dividends. Let $\mathcal{F}_t = \sigma \left\{D_s : -\infty \leq s \leq t\right\}$ denote the information set at time t. Applying the Kalman-Bucy filter, the Bayesian posterior belief has a normal distribution, $h_t\left(\mu_t | \mathcal{F}_t\right) \sim$

$\mathcal{N}\left(m_t, \sigma_m^2\right)$, where the current mean belief, m_t, follows the process:

$$dm_t = \vartheta\left(\bar{\mu} - m_t\right)dt + KdB_t, \tag{11}$$

and where the surprise innovation dB_t follows:

$$dB_t = \sigma_D^{-1}\left(dD_t/D_t - m_t dt\right). \tag{12}$$

A surprise dB_t causes the rational agent to update her mean belief according to the Kalman gain $K = \sigma_m^2/\sigma_D$, where the gain is increasing in the signal-to-noise ratio.[20]

As we show below, generating a boom-bust-rebound in prices following an increase from $\bar{\mu}$ to μ_0 requires over-optimism relative to the fully rational process described by equations (11) and (12). Bordalo et al. (2018) propose *diagnostic* expectations as one such departure that formalizes the representativeness heuristic of Tversky and Kahneman (1983). The representativeness heuristic causes agents to overweight the likelihood of a trait in a class when that trait has a higher likelihood in the class than in a reference population. Bordalo et al. (2018) give as an example the higher prevalence of red hair among Irish than non-Irish, which causes people to overestimate the share of Irish with red hair. In the asset price cycle context, the reference population is the full history of dividends and the class is recent dividends, with inference over the current drift rate.

We implement diagnostic expectations as follows. For a "look-back" parameter k, the background context at date t consists of information observed up to date $t-k$, $\mathcal{F}_{t-k}$. The diagnostic belief distribution of the drift rate is then:

$$h_t^\theta\left(\mu_t\right) = h_t\left(\mu_t|\mathcal{F}_t\right)\left[\frac{h_t\left(\mu_t|\mathcal{F}_t\right)}{h_t\left(\mu_t|\mathcal{F}_{t-k}\right)}\right]^\theta Z, \tag{13}$$

[20]We assume the asymptotic variance of the posterior drift rate, i.e. $\sigma_m^2 = \sigma_\mu^2/\left(2\vartheta + K/\sigma^D\right)$. Suppressing the city subscript i implicitly assumes that agents learn only from the dividends of a single city, even if the Wiener shocks across cities were correlated. Because learning from correlated shocks has an effect similar to a rescaling of the posterior variance, this assumption does not materially impact our results.

for a constant Z that ensures the density integrates to 1. The diagnostic distribution overweights states that have become relatively more likely in light of recent dividend news, that is, where $h_t(\mu_t|\mathcal{F}_t) > h_t(\mu_t|\mathcal{F}_{t-k})$. This "kernel-of-truth" property causes over-reaction to good news when the distribution of states satisfies the monotone-likelihood property – as with normally-distributed innovations – because higher growth states become more likely following positive news. The parameter θ controls the magnitude of departure from rational expectations and nests the rational case when $\theta = 0$. As in Bordalo et al. (2020b), the diagnostic posterior mean is simply shifted from the rational case by a term $\theta\mathcal{I}$:

$$h_t^\theta(\mu_t) \sim N\left(m_t^\theta, \sigma_m^2\right), \ m_t^\theta \equiv \mathbb{E}_t^\theta[\mu_t] = m_t + \theta\mathcal{I}_t, \tag{14}$$

$$\text{where: } \mathcal{I}_t \equiv m_t - \mathbb{E}_{t-k} m_t = K \int_{t-k}^{t} e^{-\vartheta(t-s)} dB_s. \tag{15}$$

$\mathcal{I}$ is the information the diagnostic agent neglects in forming her background context.[21]

The key driving force is the expected present value of dividends P^*. Denoting the discount rate by ρ, Appendix C.1 proves that P^* depends only on D_t, m_t^θ, and parameters:[22]

$$P^*\left(D_t, m_t^\theta\right) \equiv \int_{-\infty}^{\infty} \mathbb{E}_t\left[\int_t^\infty e^{-\rho(s-t)} D_s ds | \mu_t\right] h_t^\theta(\mu_t)\, d\mu_t. \tag{16}$$

The present value $P^*\left(D_t, m_t^\theta\right)$ encodes all relevant information coming from the evolution of the dividend and beliefs. Any belief process that produces the same path of P^* will produce the same results in our model. Diagnostic expectations offer an especially attractive microfoundation for beliefs about house prices because they immediately imply two key features of the cycle — independence of the lengths of the boom, bust, and

[21]Equations (14) and (15) extend the Bordalo et al. (2019) implementation of "slow-moving" information to our continuous time setting. Maxted (2022) proposes an alternative formulation that instead generalizes the decay parameter in equation (15) not to necessarily equal mean reversion ϑ. We prefer the formulation in equation (15) because it has a direct interpretation of the information received over a recent horizon and because the parameter k controls the length of the boom without directly impacting the length of the bust.

[22]Convergence of this present value requires $\bar{\mu} + \frac{1}{2}\frac{\sigma_\mu^2}{\vartheta^2} < \rho$. We set σ_μ^2 to ensure this inequality holds and verify that at the estimated parameters our results are not sensitive in the range of admissable values.

rebound from their amplitude in the cross-section and no over-shooting of house price expectations on the downside during the bust — in addition to their grounding in evidence from psychology, their tractability, and their success in non-housing settings as well. The non-overshooting of beliefs in the bust in particular distinguishes the path of P^* from models of learning from prices in which expectations overshoot on the downside.

To summarize, dividends follow a geometric Brownian motion with a drift rate that follows an Ornstein-Ulhenbeck process. When dividends rise unexpectedly, agents overweight the likelihood of high trend growth. Eventually, the positive surprises fall out of the diagnostic window and expected growth starts to converge towards the rational posterior.

Spatial equilibrium and housing demand. Each instant a mass $g_H H_t$ of potential entrants choose between purchasing in the city or in the hinterland. We normalize to 0 the dividend and hence the house price from living in the hinterland.[23] The total cost of purchasing in the city is the price of a house, P_t, plus the cost of a mortgage in up front origination fees or "points", W_t, as in Kaplan et al. (2020).

The value to a potential entrant from purchasing in the city has a common component, V_t, and an idiosyncratic component, ξ, which creates a downward-sloping demand curve. The common component comprises the expected dividends received while living in the city, $V_t = P^*\left(D_t, m_t^\theta\right)$.[24] The idiosyncratic component ξ is drawn from a Pareto distribution with $P\left(\xi > x\right) = \left(\frac{x_m}{x}\right)^\gamma$, such that γ governs the slope of the housing demand curve.

In spatial equilibrium, a potential entrant purchases a house in the city if ξV_t exceeds

[23] Potential entrants consider a single city, which arises naturally in continuous time as multiple "offers" never arrive in the same instant. More substantively, the normalization of hinterland house prices to 0 requires that potential entrants do not value the possibility of a future opportunity with a higher idiosyncratic component ξ (introduced below). We may justify this neglect either by assuming once-and-for-all decisions or by assuming a negative dividend to living in the hinterland that exactly offsets the option value.

[24] We assume that agents expect to live in the city forever for simplicity. One can motivate $V_t = P^*\left(D_t, m_t^\theta\right)$ as a 0^{th} order approximation in the moving probability to the value incorporating the possibility of leaving the city. In the data, the probability of moving across cities is approximately 2%.

the cost $P_t + W_t$. The total demand for houses from new entrants, Q_t, thus parallels equation (4) but augmented to account for points and foreclosures:

$$Q_t = g_H H_t P\left(\xi > \left(P_t + W_t\right)/V_t\right) = g_H H_t x_m^{\gamma}\left[V_t/\left(P_t + W_t\right)\right]^{\gamma}. \tag{17}$$

Construction. The cost of building an additional house, C_t, takes the form:

$$C_t = \left[A H_t^{1/\eta}\right]\left(I_t/\bar{I}_t\right)^{1/\chi}, \tag{18}$$

where $I_t \equiv dH_t/H_t$ denotes the construction rate and $\bar{I}_t \equiv \alpha I_{t-1} + (1-\alpha)\bar{I}_{t-1}$ is a trailing geometric average.[25] Equation (18) parallels the empirical specification for housing costs in equations (1) to (3), with two differences. First, for simplicity we do not separately model land and construction and instead directly set η as the overall long-run elasticity of supply. Second, the short-run adjustment cost $\left(I_t/\bar{I}_t\right)^{1/\chi}$ captures deviations from the existing capacity of the construction sector. The parameter χ governs the short-run inverse elasticity of supply, which is $d\ln C_t/d\ln H_t = \eta^{-1} + (\chi I_t)^{-1}$. The adjustment cost term disappears at long horizons as I_t approaches $\bar{I}_t$, leaving only the long-run supply equation.

Mortgages and foreclosures. A home-buyer at date t obtains an interest-only mortgage of $M_t = \phi P_t$, where the loan-to-value (LTV) ϕ is idiosyncratic for each buyer.[26] Mortgages end at the first date τ at which the mortgagee receives a "liquidity shock" and either refinances or defaults. Liquidity shocks arrive with Poisson intensity ι. A cash-out refinance occurs if the owner has positive equity, $M_t \leq RP_\tau$, where $R \sim N(1, \sigma_R^2)$ is an idiosyncratic house price shock. In a refinance, the owner pays off the old mortgage and obtains a new mortgage of $M_\tau = \phi P_\tau$. A default occurs if the owner receives the liquid-

[25] We set α to 1/30 years. The results are not sensitive to this value as long as $\bar{I}$ moves slowly.

[26] We consider interest-only mortgages in order to realistically capture the upper tail of the LTV distribution. In practice, the vast majority of defaulters took out a new loan relatively recently and principal pay-down is minimal at the beginning of a 30-year amortizing loan.

ity shock and has negative equity, $M_t > RP_\tau$. Thus, as in Guren and McQuade (2020), default requires a "double trigger."[27] Defaulters exit the housing market and the foreclosed homes enter supply in the instant after the default occurs.[28] The mortgage balance measure density, $g(M,t)$, evolves according to the Fokker-Planck equation:

$$\frac{\partial}{\partial t} g\left(M,t\right) = \underbrace{\left(I_t + \iota\right) H_t \phi\left(M/P_t\right)/P_t}_{\text{New originations}} - \underbrace{\iota g\left(M,t\right)}_{\substack{\text{Refis/}\\ \text{foreclosures}}}. \tag{19}$$

Competitive, risk-neutral lenders provide mortgages. These lenders have the same beliefs as buyers.[29] They set points W_t to make zero expected profits on each loan given that they only recover a fraction ψ of a house's value in a foreclosure:

$$W_t = \mathbb{E}^{\tau}\left\{\mathbb{E}_t^{\theta}\left[e^{-\rho(\tau-t)}\max\left\{M_t - \psi RP_\tau, 0\right\}\right]\right\}. \tag{20}$$

5.2 Equilibrium

An equilibrium in each city consists of paths for the prospective home buyers' common valuation V_t, house price P_t, city size H_t, mortgage points W_t, and mortgage balance measure density $g(M,t)$ such that:

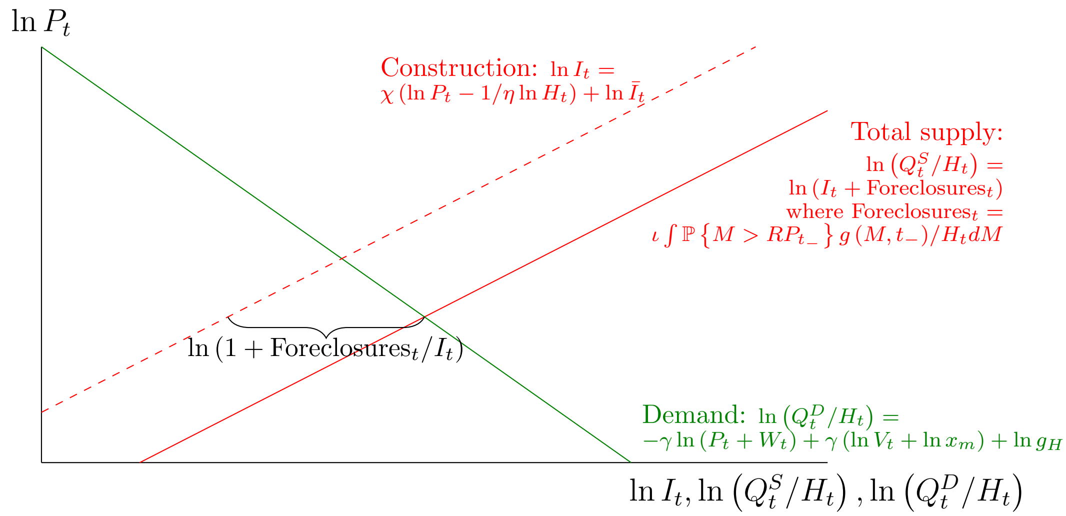
(i) Buyers' common valuation $V_t = P^*\left(D_t, m_t^\theta\right)$,reflects their beliefs $P^*\left(D_t, m_t^\theta\right)$ satisfies

[27]This is consistent with empirical evidence in Foote et al. (2008), Gerardi et al. (2018), Gupta and Hansman (2022), and Gupta et al. (2019). Gupta et al. in particular show that health shocks lead to foreclosure with negative equity and refinancing with positive equity as in our model. Ganong and Noel (2022) ascribe a larger role to liquidity shocks alone than to double trigger. If there were significant default unrelated to equity, our calibration would require a larger price impact of foreclosures or a higher rate of liquidity shocks during the bust. That being said, our model does have significant above water (to the econometrician) default, as the idiosyncratic house price shock R smooths the cliff function for default probability as a function of LTV as in Greenwald et al. (2021).

[28]Formally, a foreclosure at time t enters at time step $t + \Delta$ where in the limit $\Delta \to 0$. This timing assumption eliminates multiple equilibria that can arise if foreclosures and prices are determined jointly. The assumption that all defaulters exit (leave the city or move in with relatives) is made for simplicity; instead assuming a fraction of defaulters move into rental units would not change the qualitative conclusions but would impact the estimated values of parameters that govern the severity of the price-foreclosure spiral.

[29]Consistent with this assumption, Gerardi et al. (2008) show that lenders during the boom understood the consequences of falling house prices but put little weight on this possibility and Cheng et al. (2014) show that mortgage lenders behaved similar to the rest of the population in their own housing choices.

Figure 6: Short-Run Supply and Demand Diagrammatic Treatment of Equilibrium



equation (16).

(ii) Price equals the marginal cost of construction given by equation (18):

$$P_t = \left[AH_t^{1/\eta}\right] \left(I_t/\bar{I}_t\right)^{1/\chi}, \text{ where } I_t = \dot{H}_t/H_t. \tag{21}$$

(iii) Lenders make zero expected profits such that W_t satisfies equation (20).

(iv) Housing demand (equation (17)) equals new construction plus foreclosures:

$$g_H H_t x_m^\gamma \left[V_t/\left(P_t + W_t\right)\right]^\gamma = \dot{H}_t + \iota \int \Phi_R\left(M/P_{t-}\right) g\left(M, t_-\right) dM, \tag{22}$$

where $\Phi_R(.)$ denotes the cumulative density of a mean-one normal distribution with standard deviation σ_R.

(v) The mortgage density distribution $g(M, t)$ satisfies the Fokker-Planck equation (19).

The equilibrium describes a supply-and-demand framework that determines the equilibrium house price at any instant, as illustrated in Figure 6. The demand curve for housing as a function of price P slopes down with elasticity determined by γ and shifts due to changes in beliefs about future dividends V or mortgage points W, as indicated by the downward-sloping green line. The supply curve slopes up due to construction with elasticity determined by χ , as indicated by the dashed upward-sloping red line, and shifts out due to foreclosures, as indicated by the shift from the dashed to solid red line. Over time the supply curve shifts in as population grows, according to the long-run elasticity η.

We solve the model globally by collocation on a Smolyak grid, simulating the expectation in equation (20) by Monte Carlo. We compute impulse responses to a one-time innovation $dW_{\mu,0}$ that increases the drift rate from $\bar{\mu}$ to μ_0. Appendix C.2 derives a novel, closed-form expression for the mean impulse response of beliefs m_t^θ to a one-time increase that we use to compute the path of $P^*\left(D_t, m_t^\theta\right)$, which facilitates the efficient calculation of impulse responses for prices and other model outcomes used in our estimation.

5.3 Calibration

We choose city-level targets for our calibration. Since we compute impulse responses of variables without the idiosyncratic Wiener shocks, we treat the model paths as representing groups of cities. Accordingly, we compute data moments after grouping the CBSAs in our sample into four (1997-population-weighted) quartiles of 1997-2019 house price growth. Crucially, we target long-run moments and expectations and evaluate how well the model can match untargeted short-term price dynamics.

We calibrate several parameters externally and set the remaining parameters by simulated method of moments (SMM). The top panel of Table 2 lists the model parameters, their value for each quartile of CBSAs, and the rationale for each parameter, while the bottom panel lists the moments (numbered) as well as the non-targeted short-run house price growth in the boom, bust, and rebound (unnumbered).[30]

We begin by describing the externally-calibrated parameters relating to beliefs. In a balanced growth path, prices grow at the rate μ (see Appendix C.3). Accordingly, we set

[30] We define the model boom as time zero to the price peak, the bust as the price peak to the price trough and the rebound as the seven years following the price trough given the lack of an end date for the rebound.

the long-run mean drift rate $\bar{\mu}$ in each quartile to the annualized average log change in real house prices over the 1977:Q1-1996:Q4 period.[31] We set the mean reversion rate ϑ to 0.005 so that the increase in the drift rate at date 0 is highly persistent.[32] We select a diagnostic window, k, of 8 to achieve a boom length of 8 years.

We next turn to externally-calibrated parameters relating to housing supply, mortgages, and foreclosures. We set the long-run supply elasticity, η, to the inverse of the empirically-estimated value from Table 1 obtained by multiplying the column (3) population growth and triple interaction coefficients by the respective endogenous variables and taking quartile averages. We set the short-run construction elasticity, χ, by imposing these long-run elasticities and estimating the analogous specification for short-run construction dynamics, as explained in Appendix B.6. We set the distribution of LTVs at mortgage origination, ϕ, to $N(0.87, 0.09)$ based on Adelino et al. (2018). We set the foreclosure recovery rate, ψ, to 64.5% as in Guren et al. (2021b). We set the growth rate of potential entrants, g_H, to 8.0% annually; this parameter has no impact on the impulse responses.

We set the remaining seven parameters – the drift rate at date 0, μ_0, the diagnostic over-shooting parameter, θ, the Kalman gain relative to the dividend noise, K/σ_D, the discount rate, ρ, the demand elasticity, γ, the liquidity shock, ι, and the idiosyncratic house price shock, σ_R^2 – to fit quartile-specific moments on long-run house price growth, long-run house price expectations in 2006 and 2019, the bust length, several moments relating to foreclosure dynamics, and idiosyncratic house price volatility. Because expectations data in the boom are only available for cities in the fourth quartile, we treat the learning parameters θ and K/σ_D and the discount rate ρ as deep parameters and fix their values

[31] The Freddie Mac house price index starts in 1975 but repeat sales indices are noisy in the first few years of a data set, so we allow for two years of burn in and start in 1977.

[32] Mean reversion in μ must be slow for a shock at the start of the boom to continue to drive house prices during the rebound. A value of 0.005 means that even 25 years after the boom start, μ_t has declined only $1 - e^{-.005\times 25} \approx 12\%$ of the distance from μ_0 to $\bar{\mu}$. Total mean reversion in prices or rents will be larger in any given city, reflecting the role of the idiosyncratic Wiener shocks.

Table 2: Model Parameterization and SMM Targeted and Untargeted Moments

Symbol	Description	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Rationale
Beliefs						
$\bar{\mu}$	Pre-boom drift rate	0.41%	0.58%	0.96%	1.51%	1977-96 growth rate
μ_0	New drift rate	0.63%	4.40%	4.90%	5.03%	SMM by quartile
θ	Diagnostic over-shooting	2.75	2.75	2.75	2.75	SMM in Q4
K/σ_D	Normalized Kalman gain	0.13	0.13	0.13	0.13	SMM in Q4
k	Diagnostic window	8.00	8.00	8.00	8.00	Boom length in data
ϑ	Mean reversion in drift	0.005	0.005	0.005	0.005	
Preferences						
ρ	Discount rate	7.75%	7.75%	7.75%	7.75%	SMM in Q4
γ	Demand elasticity	3.00	0.50	0.50	0.50	SMM by quartile
Construction and foreclosures						
η	Long-run supply elasticity	1.89	1.58	1.17	0.87	Empirical regressions
χ	Construction elasticity	4.50	3.51	2.40	1.69	Empirical regressions
g_H	Potential entrants	0.08	0.08	0.08	0.08	Normalization
ι	Liquidity shock	6.70%	4.55%	4.50%	5.30%	SMM by quartile
R	House price shock	$N(1, 0.0.13)$	$N(1, 0.11)$	$N(1, 0.08)$	$N(1, 0.09)$	SMM by quartile
ϕ	New origination LTV	$N(0.87, 0.09)$	$N(0.87, 0.09)$	$N(0.87, 0.09)$	$N(0.87, 0.09)$	Adelino et al. (2018)
ψ	Foreclosure recovery rate	0.645	0.645	0.645	0.645	Guren et al. (2021)

		Quartile 1		Quartile 2		Quartile 3		Quartile 4	
Moment	Source	Data	Achieved	Data	Achieved	Data	Achieved	Data	Achieved
Long run house price growth and expectations (annualized log points)									
1. 1997-2019 HPI	Freddie Mac	0.3	0.3	1.4	1.5	2.3	2.5	3.5	3.4
2. Expectations 2006	Shiller-Thompson							5	4.1
3. Expectations 2019	Shiller-Thompson							3.1	3.4
Phase length (years)									
4. Bust	Freddie Mac	6.00	3.67	5.75	7.29	6.00	6.67	6.00	5.58
Role of foreclosures									
5. Max Foreclosure Rate	CoreLogic	1.10	1.23	1.49	1.03	1.74	2.05	2.26	2.42
6. % Jan-07 Equity< 20%	Beraja et al.	32.4	47.6	25.6	21.8	17.8	19.8	18.2	19.9
7. % Jan-07 Equity< 10%	Beraja et al.	16.8	21.9	12.4	9.3	7.9	9.4	8.8	10.0
8. % Jan-07 Equity< 0%	Beraja et al.	4.9	4.4	3.0	2.0	2.1	2.5	3.5	2.9
9. Bust Speed (log points)	Freddie Mac	-9.2	-1.2	-9.4	-4.5	-13.1	-12.9	-23.4	-22.6
Standard deviation of house price shocks (log points)									
10. Repeat Sales Resid SD	DataQuick	9	13	9	11	9	8	9	9
Non-targeted short run house price growth (unannualized log points)									
Boom	Freddie Mac	20.0	4.2	35.3	43.1	59.4	67.1	84.9	85.6
Bust	Freddie Mac	-29.7	-1.8	-29.7	-21.2	-38.6	-48.1	-53.4	-60.4
Rebound	Freddie Mac	14.8	4.9	21.9	13.0	28.1	36.3	44.4	48.6

at the levels estimated in the fourth quartile, giving seven parameters to match 10 fourth quartile moments. We then set quartile-specific values of μ_0, ι, and σ_R^2, giving three parameters to fit eight quartile-specific moments in the remaining quartiles.

The first moment measures the growth in house prices over the full 1997-2019 BBR. The next two moments are average expected house price growth over the next 10 years at the peak in 2006 and in 2019 from Shiller and Thompson (2022) adjusted for inflation expectations using the Michigan Survey median. This is the only survey of house price expectations that covers the boom. Three of the four counties surveyed (Middlsex, MA, Alemeda, CA, and Orange, CA) are in quartile four, and we assume the average expectation for these counties is representative of the quartile. Heuristically, the size of the full boom-bust-rebound most directly informs μ_0, ρ, and γ as it reflects the actual path of dividends and their impact on the price. The over-shooting of the boom expectations relative to the long-run informs the learning parameters θ and K/σ_D.

The next moment is the length of the bust. In the model, the bust ends when beliefs begin to stabilize and actual dividends have risen enough to offset the earlier over-optimism. The speed of learning, K/σ_D, and the initial drift μ_0 influence this timing.

The next set of moments characterize the role of foreclosures. The fifth moment is the peak annualized foreclosure rate, which we compute using proprietary data from CoreLogic. The next three moments are the shares of properties with equity less than 20%, 10%, and underwater in January 2007 from the Beraja et al. (2019) data set. The equity distribution informs the liquidity shock ι, with the foreclosure moment additionally informing the supply and demand curve parameters χ and γ. Finally, we discipline the speed of the price-foreclosure spiral using the maximum four-quarter price decline in the bust.

The final moment closely corresponds to σ_R, the standard deviation of the house price shock. Using 1988-2013 DataQuick deeds data for non-distressed sales of single family homes and condominiums, we run a repeat sales regression of log price on house fixed effects and census tract-by-quarter fixed effects. The residuals of this regression reflect idiosyncratic variation in house prices. To mitigate outliers, the moment is the standard deviation of a normal distribution with the same inter-quartile range as the residuals.

We choose parameters to minimize the weighted sum of squared residuals between the model and the data moments. We calculate the model moments for the impulse response to a one-time increase from $\bar{\mu}$ to μ_0, using the analytic mean path of beliefs and starting from steady state values for all variables except the loan balance distribution, which is calibrated to match the 1995 Survey of Consumer Finances. We weight the three equity distribution moments together as much as each of the other moments. We then increase the weight on the long-run house price growth moment in each quartile to ensure this moment is within 10% of the data. This binds only for the lower two quartiles and enforces the spirit of the exercise of matching the overall BBR price growth.

5.4 Model Fit

Table 2 and Figure 7 show that the estimated model fits the targeted moments very well in quartile four, for which we estimate all parameters. All moments except the equity distribution and expectation moments are within ten percent of their target values. The equity distribution moments are close in percentage point terms. The model expectations are too low in 2006 and slightly too high in 2019. In fact, Laibson (2012) argues the survey expectations may be erroneously high in the boom because agents misunderstood the survey question in this period before it was rephrased.[33] The miss in 2019 reflects a tension between matching the expectation moment and average 1997-2019 price growth.

[33] Alternatively, agents may neglect that long-run supply is more elastic than short-run supply. A calibration where agents compute expectations as if on the balanced growth path matches the 2006 expectation better while still matching short-run price dynamics. The main change is a lower diagnosticity.

As shown at the bottom of Table 2 and in Figure 7, the model closely matches the short-run boom, bust, and rebound sizes in quartile four even though these moments are not targeted. The model also fits the third quartile targeted and untargeted short-run price moments well despite the deep parameters being held fixed in estimation, leading to many fewer degrees of freedom. In the second quartile, the model matches the boom quite well but does not generate enough of a price-foreclosure spiral to match the size of the bust and rebound. In the first quartile, the model generates very little price movement, in large part because this quartile has the most elastic short-run and long-run supply. One partial explanation for the model's failure in these lower quartiles may be that the price-foreclosure spiral in the real world works in part through a tightening of credit standards by nation-wide lenders, which spread the bust from the third and fourth quartiles to the first and second (Guren and McQuade, 2020). Another possible explanation is cross-quartile learning. In the model, we treat each quartile independently and shut down such linkages.

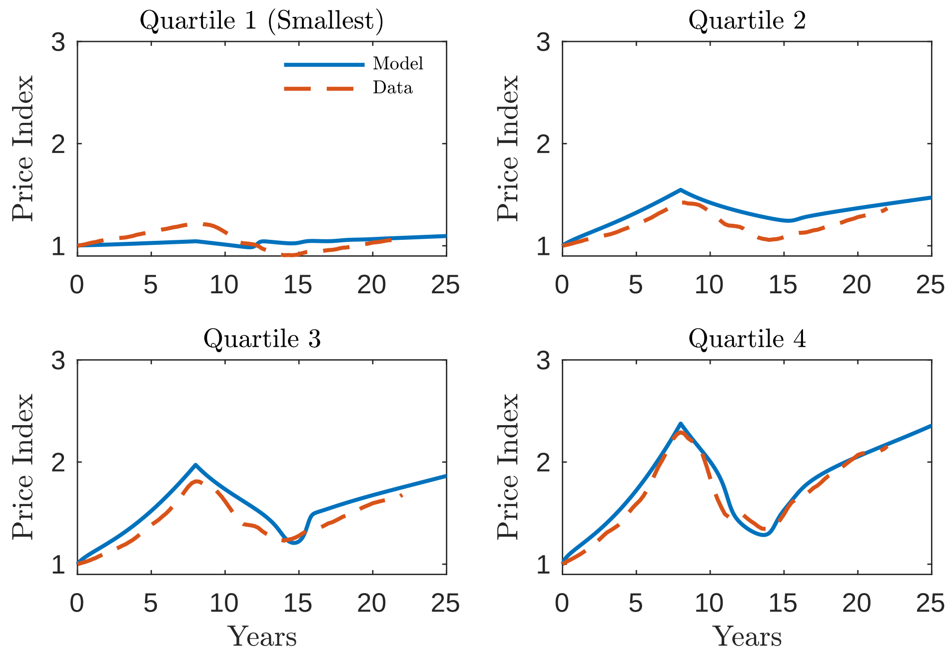
Several of the estimated parameters that deliver this fit merit comment. The new drift rate μ_0 is increasing across quartiles. The diagnosticity parameter θ of 2.75 is larger than the value of 0.9 in Bordalo et al. (2019) or the 0.3 to 1.5 range estimated in Bordalo et al. (2020b), perhaps reflecting the difference between the households in our setting and the financial analysts or professional forecasters in theirs or the subtleties described above in matching the expectations moments. The liquidity shock ι being in the neighborhood of 5% accords with mortgage pre-payment rates during the boom in Berger et al. (2021).

5.5 Implications for Rent Growth

We introduce rents in Appendix C.4 via a user cost indifference relationship.[34] Along the balanced growth path, this relationship takes the form $R_t^{\mathrm{BGP}} = (\rho - \mu) P_t + \zeta D_t$ (see

[34] In the background are deep-pocketed landlords willing to sell or to rent to households.

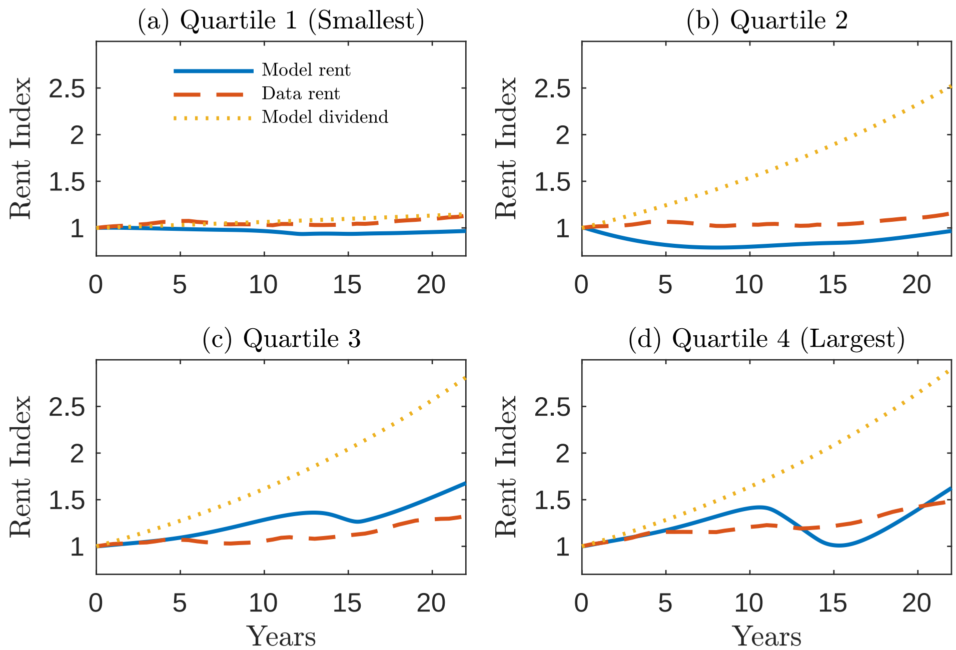
Figure 7: Quartile Price Paths in Model and Data



Notes: Each panel displays the path of house prices in the model following a change from μ_0 to $\bar{\mu}$ and in the data for a quartile of CBSAs, grouped by their their long-run 1997-2019 price growth.

equation (C.15)), where ζ encompasses the maintenance and property tax component of owning and is chosen to match an average price-rent ratio in quartile 4 of 19, which is what Begley et al. (2021) find for California. We make the following assumptions to calculate the transition path of rents: (i) households re-optimize their rent/own decision with Poisson intensity λ, where we parameterize $\lambda = 1/6$ to match tenure rates (ii) during a rental contract signed at date t, rents grow at a fixed rate m_t^θ, the nowcast of the dividend drift rate, (iii) households use the path of expectations for prices and dividends at each point in time. These assumptions capture the long-term nature of rent/own decisions, the stickiness of contract rents, and the volatility of expected house price appreciation off the balanced growth path. In Appendix C.4, we show that the path of average rent R_t is characterized by a "reset" rent $R_{t|t}$ for contracts signed at date t and laws of motion for R_t and the average within-contract growth rate g_t (see equations (C.12) to (C.14)).

Figure 8: Quartile Rent Paths in Model and Data



Notes: The solid lines plots model simulated rents for each quartile. The dashed lines plot indexes of annual average rent deflated by the national GDP price index for 22 CBSAs with CPI rent data available from 1988, grouped into the same quartiles. The dotted lines plot dividends.

Figure 8 shows rents in the model and data. Three features stand out. First, model rent growth substantially lags dividend growth over the full period, reflecting the decline in the rent-price ratio when expected price growth rises. Second, the model explains rent growth in each quartile over the entire 1997-2019 period reasonably well despite the fact that this moment is untargeted, with model rent growth a little low in quartiles 1 and 2 and high in quartiles 3 and 4. Third, the model features counterfactual rent drops in the bust in quartiles 3 and 4. These declines reflect the overshooting of prices due to foreclosures, as low prices and high expected price growth both drag down user-cost implied rent. Factors outside of the model such as downward stickiness of rents or a failure of rent-price arbitrage in the crisis could explain why rents in the data do not fall as much. Overall, the model's ability to explain rent growth across quartiles over the full BBR supports a fundamentals-based explanation of the path of prices.

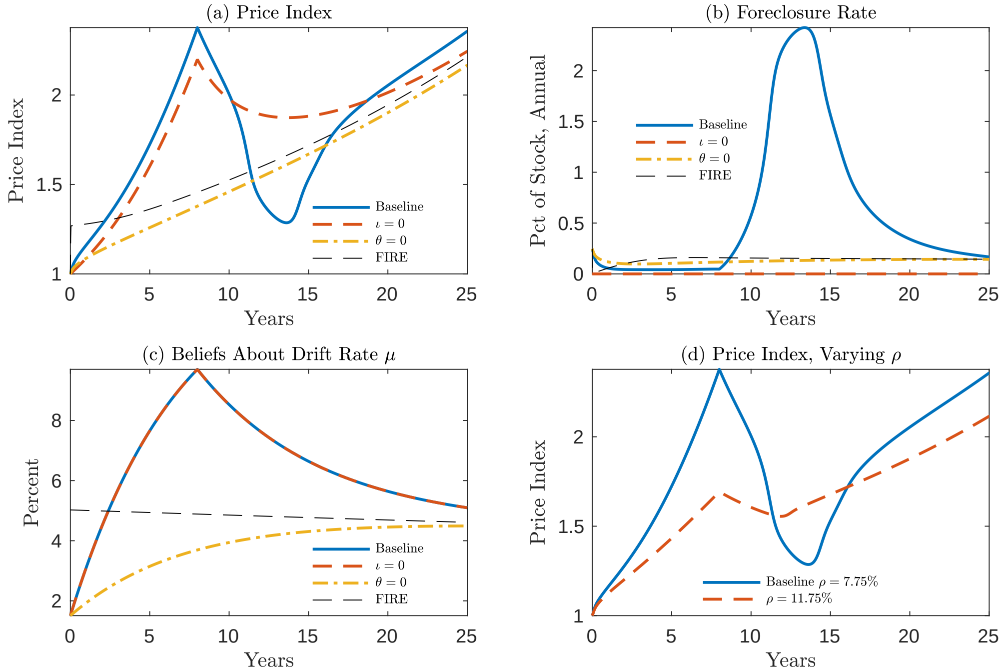
5.6 Unpacking the Mechanism

We now discuss the model features that generate a boom-bust-rebound. For parsimony, we use the parameters estimated for the fourth quartile. Panels (a)-(c) of Figure 9 show the evolution of prices, foreclosures, and beliefs following a change from $\bar{\mu}$ to μ_0. The solid blue lines show the paths at the estimated parameter values. Each remaining line shows a model permutation holding other parameters fixed: the dash-dot gold lines show the paths without diagnostic expectations ($\theta = 0$), the dashed red lines show the paths with no foreclosures ($\iota = 0$), and the thin dashed black line shows the full-information rational expectations (FIRE) model with immediate learning ($\theta = 0, K/\sigma_D \to \infty$).

The comparisons across lines in panels (a)-(c) illustrate the importance of both belief over-shooting and foreclosures to generating a realistic boom-bust-rebound. Under FIRE, beliefs equal the true μ, which jumps on impact and gradually mean reverts. Prices jump on impact due to the increase in the capitalization rate $\rho - \mu$ and then converge to the balanced growth path from above. Without diagnosticity but with learning ($\theta = 0$), beliefs monotonically and smoothly increase towards the truth and prices smoothly converge to the balanced growth path. The steady price appreciation reflects the increase in the capitalization rate and the increase over time in the dividend.

Diagnostic learning generates an over-shooting of beliefs about the drift rate μ, as shown in panel (c). These beliefs rise from their initial level of 1.5% up to above 9% before nearly converging to the true value of around 5% (shown by the FIRE thin dashed line) by the end of the sample window. The turning point coincides with the peak of the boom. Unlike in models of price extrapolation that generate overshooting of expectations on the downside, with diagnostic expectations the convergence is gradual and monotonic. The path of beliefs in the bust matches the time series of beliefs in the Shiller and Thompson (2022) survey, in which long-run expectations fall gradually from their boom peak.

Figure 9: Role of Structural Elements



Notes: This figure shows the simulated boom-bust-rebound in the model resulting from a single change from $\bar{\mu}$ to μ_0 at time $t = 0$. The solid blue line shows the baseline model with parameters estimated for the fourth quartile of CBSAs. In panels (a)-(c), the dashed red line shows the model without foreclosures, the dash-dot gold line shows a parameterization with foreclosures but no diagnostic expectations, and the thin dashed black line shows full information rational expectations. Panel (d) shows the effect of varying the discount rate ρ (held constant over the simulation) holding all other parameters unchanged. Panels (a) and (d) show the price index, panel (b) shows the foreclosure rate in annualized percent of the housing stock, and panel (c) shows beliefs about the drift μ in annualized percentage points.

Without foreclosures, however, the over-shooting of prices from diagnostic learning alone generates a much smaller boom and bust than in the data. The attenuation of the boom occurs for two reasons: credit expands in the baseline case as lenders perceive a decline in default risk and the increase in prices causes foreclosures to decline relative to the pre-boom period, contracting supply and further pushing up prices. The correction in beliefs on its own generates a counterfactually-small price dip in the bust, as prices converge smoothly toward their long-run path.

Foreclosures generate a much larger bust in prices, to below the level that would prevail with rational learning or FIRE.[35] The over-shooting occurs because of a price-foreclosure spiral (Guren and McQuade, 2020): foreclosures add to housing supply, which further depresses prices, putting more owners under-water, leading to more foreclosures. Reflecting this dynamic, the increase in foreclosures in panel (b) reaches a local extrema near the trough of the bust in prices. In sum, both elements — belief over-optimism and foreclosures — are required to generate the boom-bust-rebound.

An important implication of our analysis is that asset price cycles like the 2000s housing cycle are more likely when discount rates are low (or initial dividend growth is high). Under these conditions, the present value of dividends is dominated by terms further out in the future, so a shock to the expected dividend growth rate has a larger effect on $V_{i,t}$. This makes diagnostic cycles stronger. To illustrate, panel (d) of Figure 9 shows the baseline price path and an alternative where we feed in the same shock to μ but with the discount rate 4.0 percentage points higher, approximately the difference between the real interest rate during 1997-2019 and the average rate in the 1980s. A higher ρ substantially dampens the cycle, indicating the importance of the low *level* of discount rates throughout the 2000s.

This begs the question of the role of *time-varying* mortgage rates or credit conditions, which our analysis has so far omitted. In brief, while the decline in interest rates during the BBR constitutes another fundamental force affecting house prices, neither the timing of rate changes nor their magnitude suggest a pre-eminent role.[36] Changes in under-writing

[35]The price-foreclosure spiral in our model is quantitatively large: The quartile four bust size without foreclosures is 15.2 log points relative to 60.4 in the baseline model. The speed and magnitude of the bust and rebound explain why our exercise requires a substantial impact of foreclosures. For comparison, Guren and McQuade (2020) find a similarly large role for foreclosures, albeit split between supply-and-demand and bank balance sheet effects. Mian et al. (2015) conduct an empirical analysis with foreclosures instrumented by judicial requirements; extrapolating their local treatment effect elasticity to total foreclosures from 2007-2013 implies a 16 to 32 percentage-point decline in national house prices due to foreclosures.

[36]Figure C.3 shows that mortgage rates fell less than 1 p.p. during the boom, fell in the bust, and rose slightly during the rebound, a pattern inconsistent with declining mortgage costs explaining the periods of rising prices or rising mortgage costs explaining the bust. Using a richer mortgage structure than in our model, we calculate in Appendix C.6 that the 2.7 p.p. mortgage rate decline over the full BBR could on its own explain roughly 20 p.p. of house price growth, relative to the more than doubling in the top quartile.

costs, including credit spreads or screening costs, have relatively small effects on prices in our model simply because W_t is small relative to P_t.[37] Changes in credit that affect approval rates on the extensive margin offer greater potential for credit to impact prices, as they shift the number of entrants g_h and thus the demand curve directly. Such extensive margin shifts complement our focus on fundamentals and could operate independently or interact with changing fundamentals to the extent they reflect lender optimism or relax otherwise tightening payment-to-income constraints as prices rise.

6 Conclusion

We revisit the 2000s housing cycle with "2020 hindsight." At the city level, the areas with the largest price increases during the 1997-2006 boom had the largest busts but also the fastest growth after the 2012 trough, and as a result had the largest price appreciation over the full cycle. A long-run spatial equilibrium framework of house price growth determined by local income, amenities, and supply determinants fits the cross-section of city house price growth between 1997 and 2019. The implied long-run fundamental is correlated not only with long-run price growth but also with a strong boom-bust-rebound pattern.

Our neo-Kindlebergerian interpretation emphasizes the role of economic fundamentals in setting off asset price cycles. In our model, the boom results from over-optimism about an increase in the "dividend" growth rate, the bust ensues when beliefs correct, exacerbated by a price-foreclosure spiral that pushes prices below their full-information level, and eventually a rebound emerges as foreclosures recede and the economy converges to a price path commensurate with fundamental growth. We emphasize the low interest rate

[37] Figure C.1 confirms this intuition in a quantitative exercise where we replace lenders' diagnostic beliefs with perfect foresight, so that they perfectly anticipate the peak in buyers' beliefs and hence in prices. In this extreme case, W_t/P_t rises by 9.6 p.p., but the price path changes little. The reason W_t/P_t does not rise more is that lenders receive substantial cash flows even on mortgages made just prior to the price peak.

environment as a crucial catalyst of this cycle. We view our approach as complementary to work on the role of changes in credit availability and speculation, which we see as additional forces that may work in concert with overreaction to fundamentals.

Our conclusion about the fundamentally-driven roots of the 2000s housing cycle is important not only for understanding the cause of the cycle and asset bubbles more generally but also for macroprudential policy. If a boom-bust cycle were only due to exogenously-changing expectations or credit supply, a macroprudential policy maker might want to aggressively stamp out the boom phase. Our findings suggest a more delicate balance; while policy may want to temper over-optimism and aggressively mitigate fire sales, it is also important not to suffocate fundamentally-driven growth. Conversely, the consequences of over-optimism appear most dire when initial price growth is high and interest rates are low. Of course hindsight is 20-20; distinguishing fundamental growth from over-optimism in real time rather than after observing a full boom-bust-rebound poses a formidable task. Nonetheless, our findings imply that policy makers should heed Kindleberger's dictum that essentially all manias are to some degree grounded in fundamentals.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2018). "The Role of Housing and Mortgage Markets in the Financial Crisis". *Annual Review of Financial Economics* 10 (1): 25–41.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer (2018). "Extrapolation and Bubbles". *Journal of Financial Economics* 129 (2): 203–227.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny (1998). "A Model of Investor Sentiment". *Journal of Financial Economics* 49: 307–343.

Baum-Snow, Nathaniel and Daniel Hartley (2020). "Accounting For Central Neighborhood Change, 1980-2010". *Journal of Urban Economics* 117: 103228.

Begley, Jaclene, Lara Lowenstein, and Paul S. Willen (2021). "The Price-Rent Ratio in the 21st Century: Measurement and Implications". Working Paper.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra (2019). "Regional Heterogeneity and the Refinancing Channel of Monetary Policy". *The Quarterly Journal of Economics* 134 (1): 109–183.

Berger, David, Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra (2021). "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy". *American Economic Review* 111 (9): 2829–2878.

Bordalo, Pedro, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer (2020a). "Diagnostic Bubbles". *Journal of Financial Economics* 141 (3): 1060–1077.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic Expectations and Stock Returns". *The Journal of Finance* 74 (6): 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020b). "Overreaction in Macroeconomic Expectations". *American Economic Review* 110 (9): 2748–82.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018). "Diagnostic Expectations and Credit Cycles". *The Journal of Finance* 73 (1): 199–227.

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

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo (2018). "Housing Booms and Busts, Labor Market Opportunities, and College Attendance". *American Economic Review* 108 (10): 2947–2994.

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

Couture, Victor and Jessie Handbury (2020). "Urban revival in America". *Journal of Urban Economics* 119: 103267.

Crone, Theodore M, Leonard I Nakamura, and Richard Voith (2010). "Rents Have Been Rising, Not Falling, in the Postwar Period". *The Review of Economics and Statistics* 92 (3): 628–642.

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.

Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed (2019). "Has falling crime invited gentrification?" *Journal of Housing Economics* 46: 101636.

Favara, Giovanni and Jean Imbs (2015). "Credit Supply and the Price of Housing". *American Economic Review* 105 (3): 958–92.

Favilukis, Jack, Sydney C Ludvigson, and Stijn Van Nieuwerburgh (2017). "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk-Sharing in General Equilibrium". *Journal of Political Economy* 125 (1): 140–223.

Ferreira, Fernando and Joseph Gyourko (2018). "Anatomy of the Beginning of the Housing Boom Across U.S. Metropolitan Areas". NBER Working Paper No. 17374.

Foote, Christopher, Kristopher Gerardi, and Paul Willen (2008). "Negative Equity and Foreclosure: Theory and Evidence". *Journal of Urban Economics* 64 (2): 234–245.

Ganong, Peter and Pascal J Noel (2022). "Why Do Borrowers Default on Mortgages? A New Method For Causal Attribution". *The Quarterly Journal of Economics* Forthcoming.

Gao, Zhenyu, Michael Sockin, and Wei Xiong (2020). "Economic Consequences of Housing Speculation". *The Review of Financial Studies* 33 (11): 5248–5287.

Gerardi, Kristopher, Kyle F Herkenhoff, Lee E Ohanian, and Paul S. Willen (2018). "Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default". *The Review of Financial Studies* 31 (3): 1098–1131.

Gerardi, Kristopher, Andreas Lehnert, Shane M. Sherlund, and Paul Willen (2008). "Making Sense of the Subprime Crisis". *Brookings Papers on Economic Activity* 2008: 69–145.

Glaeser, Edward L and Charles G Nathanson (2017). "An Extrapolative Model of House Price Dynamics". *Journal of Financial Economics* 126 (1): 147–170.

Greenwald, Daniel and Adam M. Guren (2021). "Do Credit Conditions Move House Prices?" Working Paper, Boston University.

Greenwald, Daniel, Tim Landvoigt, and Stijn Van Nieuwerburgh (2021). "Financial Fragility With SAM?" *The Journal of Finance* 76 (2): 651–706.

Gupta, Arpit, Catherine Fedorenko Edward Morrison, and Scott Ramsey (2019). "Home Equity Mitigates the Financial and Mortality Consequences of Health Shocks: Evidence from Cancer Diagnoses". Working Paper, New York University.

Gupta, Arpit and Christopher Hansman (2022). "Selection, Leverage, and Default in the Mortgage Market". *The Review of Financial Studies* 35: 720–770.

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jon Steinsson (2021a). "Housing Wealth Effects: The Long View". *The Review of Economic Studies* 88 (2): 669–707.

Guren, Adam M., Arvind Krishnamurthy, and Timothy J. McQuade (2021b). "Mortgage Design in an Equilibrium Model of the Housing Market". *The Journal of Finance* 76 (1): 113–168.

Guren, Adam M. and Timothy J. McQuade (2020). "How Do Foreclosures Exacerbate Housing Downturns?" *The Review of Economic Studies* 87 (3): 1331–1364.

Gyourko, Joseph and Albert Saiz (2006). “Construction Costs and the Supply of Housing Structure”. *Journal of Regional Science* 46 (4): 661–680.

Gyourko, Joseph, Albert Saiz, and Anita Summers (2008). “A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index”. *Urban Studies* 45 (3): 693–729.

Himmelberg, Charles, Christopher Mayer, and Todd Sinai (2005). “Assessing High House Prices: Bubbles, Fundamentals and Misperceptions”. *Journal of Economic Perspectives* 19 (4): 67–92.

Howard, Greg and Jack Liebersohn (2021). “Why is the rent so darn high? The role of growing demand to live in housing-supply-inelastic cities”. *Journal of Urban Economics* 124: 103369.

— (2022). “Regional Divergence and House Prices”. Working Paper, University of Illinois.

Jacobson, Margaret M. (2022). “Beliefs, Aggregate Risk, and The U.S. Housing Boom”. Working Paper, Federal Reserve Board of Governors FEDS 2022-061.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante (2020). “The Housing Boom and Bust: Model Meets Evidence”. *Journal of Political Economy* 128 (9): 3285–3345.

Kindleberger, Charles (1978). *Manias, Panics and Crashes: A History of Financial Crises.* Palgrave Macmillan.

Laibson, David I. (2012). “Comment on What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets”. *Brookings Papers on Economic Activity*: 299–301.

Larson, William, Jessica Shui, Morris Davis, and Stephen Oliner (2021). “The Price of Residential Land for Counties, ZIP Codes, and Census Tracts in the United States”. *Journal of Monetary Economics* 118 (March): 413–431.

Lutz, Chandler and Ben Sand (2019). “Highly Disaggregated Topological Land Unavailability”. Working Paper, York University.

Maxted, Peter (2022). “A Macro-Finance Model with Sentiment”. *The Review of Economic Studies* Forthcoming.

Mian, Atif and Amir Sufi (2009). “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis”. *The Quarterly Journal of Economics* 124 (4): 1449–1496.

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

Nichols, Joseph (2019). “Regional Trends in Commercial and Residential Construction”. Working Paper, Federal Reserve Board.

Pope, Devin G. and Jaren C. Pope (2012). “Crime and Property Values: Evidence from the 1990s Crime Drop”. *Regional Science and Urban Economics* 42 (1): 177–188.

Saiz, Albert (2010). “The Geographic Determinants of Housing Supply”. *The Quarterly Journal of Economics* 125 (3): 1253–1296.

Schubert, Gregor (2021). “House Price Contagion and U.S. City Migration Networks”. Working Paper, Harvard University.

Shiller, Robert (2008). “Understanding Recent Trends in House Prices and Homeownership”. In: *Jackson Hole Conference Series*: pp. 85–123.

Shiller, Robert J. and Anne K. Thompson (2022). “What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets: A Ten-year Retrospect”. *Brookings Papers on Economic Activity* Forthcoming.

Su, Yichen (2022). “The Rising Value of Time and the Origin of Urban Gentrification”. *American Economic Journal: Economic Policy* 14 (1): 402–39.

Tversky, Amos and Daniel Kahneman (1983). “Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment”. *Psychological Review*: 293–315.

The 2000s Housing Cycle with 2020 Hindsight

Online Appendix

Gabriel Chodorow-Reich Adam Guren Tim McQuade

October 2022

This online appendix is split into three sections. In Section A, we microfound the construction and land costs in the empirical framework in Section 3. In Section B, we present additional empirical results on the role of investors and speculators, on the Bartik and urbanization instruments, on the robustness of our IV results, and a number of additional empirical results. In Section C, we present a number of model derivations and proofs.

A Microfounding Construction and Land Costs

A.1 Construction

Assume a construction function for producing houses out of materials $M_{i,t}$ and labor $N_{i,t}$:

$$\dot{H}_{i,t} = \tilde{A}_{i,t} \left(M_{i,t}^{\kappa} N_{i,t}^{1-\kappa} \right) H_{i,t}^{-\alpha_i}. \tag{A.1}$$

The term $H_{i,t}^{-\alpha_i}$ captures the possibility that construction becomes more difficult as easier-to-develop plots get built first. Competitive construction firms obtain materials at a price P_t^M on the national market and hire labor at local wage $W_{i,t}$. The FOC for cost minimization yields a cost-per-new-home $C_{i,t}$ of:

$$C_{i,t} = A_{i,t} H_{i,t}^{\alpha_i}, \tag{A.2}$$

$$\text{where:} \quad A_{i,t} = (P_t^M)^{\kappa} (W_{i,t})^{1-\kappa} / \tilde{A}_{i,t}. \tag{A.3}$$

The same result would arise if the local construction wage $W_{i,t}$ depended on population, with a re-definition of the exponent α_i.

A.2 Land

We extend Alonso (1964), Muth (1969), and Mills (1967) and Saiz (2010) to incorporate population growth, permitting restrictions, and an additional downtown premium.

We consider a city with population $H_{i,t}$ laid out on a disk with radius $\Phi_{i,t}$, of which a fraction $\Lambda_{i,t}$ is buildable land. A share $1-\varsigma_i$ of the population are spaced uniformly on the buildable part of the disk and the remaining share ς live in an "urban core" at the center of the disk, giving:

$$\Phi_{i,t} = \sqrt{\frac{(1-\varsigma)\,H_{i,t}}{\Lambda_{i,t}\pi}}, \tag{A.4}$$

where we have normalized lot size to 1. We allow the buildable share to increase with population with a semi-elasticity that depends on permitting restrictions:

$$\Lambda_{i,t} = \Lambda_{i,0} H_{i,t}^{g(m_i)}, \tag{A.5}$$

where m_i measures regulatory and permitting hurdles and $g'(m_i) < 0$.[1]

Outside of the urban core mass point, the rental cost of a plot of land $\nu_{i,t}(\tau)$ depends on its distance τ from the city center:

$$\nu_{i,t}(\tau) = \kappa_{i,t}\left(\Phi_{i,t}^{\chi} - \tau^{\chi}\right). \tag{A.6}$$

At the city's edge ($\tau = \Phi$), the rental value of land equals 0, a normalization that reflects a residual supply of unused land. The city-specific parameter $\kappa_{i,t} > 0$ shifts the value of all plots of land in a city proportionally. The parameter $\chi > 0$ is the elasticity of the premium to living one step outside the urban core, $\nu_{i,t}(0)$, relative to living 1% of the city radius outside of the center; denoting $\hat{\nu}_{i,t}(\tau) \equiv \frac{\nu_{i,t}(0)-\nu_{i,t}(\tau)}{\nu_{i,t}(0)} = \left(\frac{\tau}{\Phi_{i,t}}\right)^{\chi}$, $\chi = \frac{\partial \ln \hat{\nu}_{i,t}(\tau)}{\partial(\tau/\Phi_{i,t})}$. The term $\left(\Phi_{i,t}^{\chi} - \tau^{\chi}\right)$ has the literal interpretation of offsetting commuting costs to the city center;

[1] This functional form allows Λ to rise above 1, with the interpretation of allowing high-rises.

more generally it reflects the desirability of different neighborhoods. As a city grows, the premium to living closer to the city-center rises to preserve intra-city spatial equilibrium.

The urban core provides special amenities that create an additional premium (willingness-to-pay) above the value one step outside the core. This valuation is individual-specific and given by $U_{j,t}\nu_{i,t}(0)$, where $P(U > x) = 1 - F(U)$ for a CDF $F(.)$. Market clearing for living in the downtown core defines a cutoff $U^*_{i,t}$ such that $\varsigma_i = P(U > U^*_{i,t})$. Below we parameterize the CDF and characterize the cutoff $U^*_{i,t}$ as a function of resident types and preferences; for now we treat it as exogenous. Denoting the rental value in the urban core as $\nu^U_{i,t}(0)$ (superscript U for "urban"), we have:

$$\nu^U_{i,t}(0) = U^*_{i,t}\nu_{i,t}(0) = U^*_{i,t}\kappa_{i,t}\Phi^{\chi_i}_{i,t}. \tag{A.7}$$

Thus $U^*_{i,t}$ is the urban premium.

The price of a plot of land is the discounted future rents:

$$\text{Not in core:} \qquad L_{i,t}(\tau) = \int_t^\infty e^{-\rho(s-t)}\kappa_{i,s}\left(\Phi^{\chi_i}_{i,s} - \tau^{\chi_i}\right)ds, \tag{A.8}$$

$$\text{Urban core:} \qquad L^U_{i,t}(0) = \int_t^\infty e^{-r(s-t)}U^*_{i,s}\kappa_{i,s}\Phi^{\chi_i}_{i,s}ds. \tag{A.9}$$

We consider a balanced growth path with $\kappa_{i,s} = \kappa_i$, $U_{i,s} = U_i$, and population growth of I_i, giving $\Phi_{i,s} = \sqrt{\frac{(1-\varsigma_i)H_{i,s}}{\Lambda_i\pi}} = \sqrt{\frac{e^{I_i(s-t)}(1-\varsigma_i)H_{i,t}}{\Lambda_i\pi}} = e^{\frac{I_i}{2}(s-t)}\Phi_{i,t}$. Along this path:

$$L_{i,t}(0) = \int_t^\infty e^{-\rho(s-t)}\kappa_i e^{\frac{\chi_i I_i(1-g(m_i))}{2}(s-t)}\Phi^{\chi_i}_{i,t}ds = \frac{\kappa_i\Phi^{\chi_i}_{i,t}}{\rho - \chi_i I_i(1-g(m_i))/2},$$

$$L_{i,t}(\tau) = L_{i,t}(0) - \frac{\kappa_i\tau^{\chi_i}}{\rho}, \text{ and } L^U_{i,t}(0) = U^*_i L_{i,t}(0).$$

With positive population growth, the price $L_{i,t}(\Phi_{i,t})$ of land at the city boundary is strictly positive, reflecting the capitalization of future non-zero rents.

We obtain the analog of equation (3) in the main text by integrating over the available land at each distance τ to arrive at the average price of a plot of land in the city:

$$
\begin{aligned}
L_{i,t} &= \frac{1}{H_{i,t}} \int_0^{\Phi_{i,t}} L_{i,t}(\tau)\, \Lambda_{i,t} 2\pi\tau d\tau + \varsigma_i L_{i,t}^U(0) \\
&= B_{i,t} H_{i,t}^{\beta_i}, \qquad \text{(A.10)}
\end{aligned}
$$

$$
\text{where:} \quad B_{i,t} = \kappa_i \left(\frac{1 + \varsigma_i \left(U_i^* - 1\right)}{\rho - \chi_i I_i \left(1 - g\left(m_i\right)\right)/2} - \frac{\left(1 - \varsigma_i\right)}{\rho\left(\chi_i/2 + 1\right)} \right) \left(\frac{\left(1 - \varsigma_i\right)}{\Lambda_{i,0}\pi} \right)^{\frac{\chi_i}{2}}, \qquad \text{(A.11)}
$$

$$
\beta_i = \left(1 - g\left(m_i\right)\right) \chi/2. \qquad \text{(A.12)}
$$

Setting $g(m_i) = \zeta_0 - \zeta_1 m_i$, we have $\beta_i = \beta_0 + \beta_1 m_i$, where $\beta_0 = (1 - \zeta_0)\chi_i/2$ and $\beta_1 = \zeta_1 \chi_i/2$, as in the main text. Furthermore, as in Saiz (2010) the average price of land is higher in places with a smaller share $\Lambda_{i,0}$ of land available for development. In addition, equation (A.11) illustrates how a reduction in the discount rate ρ or increase in the city center premium $U_{i,t}^*$ both raise average land prices in a city.

We next use this framework to motivate instruments for a changing city center premium, $u_{i,t}^* = d \ln U_{i,t}^*$. We partition the city's residents into two types, college graduate (E) and non-college graduate ($\not{E}$), with a population share $\omega_{i,t}$ of college graduates. An individual's willingness-to-pay to live in the urban core, $U_{j,t}\nu_{i,t}(0)$, is $U_{j,t} = U_{i,t}^E \exp(\epsilon_j)$ if a college graduate and $U_{j,t} = U_{i,t}^{\not{E}} \exp(\epsilon_j)$ if not, where ϵ_j is a random variable drawn from a uniform distribution with support $[0, \bar{\epsilon}]$. Let $\varsigma_{i,t}^E = P\left(U > U_{i,t}^* | E\right)$ and $\varsigma_{i,t}^{\not{E}} = P\left(U > U_{i,t}^* | \not{E}\right)$ denote the share of college graduates and of non-college graduates who live in the downtown core, respectively. For $k \in \{E, \not{E}\}$, we have:

$$
\varsigma_{i,t}^k = P\left(U > U_{i,t}^* | k\right) = 1 - (1/\bar{\epsilon})\left(\ln U_{i,t}^* - \ln U_{i,t}^k\right). \qquad \text{(A.13)}
$$

Substituting equation (A.13) into the market clearing condition $\varsigma_i = \omega_{i,t}\varsigma_{i,t}^E + (1 - \omega_{i,t})\varsigma_{i,t}^{\not{E}}$

and rearranging gives:

$$\ln U_{i,t}^{*} = \omega_{i,t}\left(\ln U_{i,t}^{E} - \ln U_{i,t}^{\not{E}}\right) + (1-\varsigma_{i,t})\,\overline{\epsilon} + \ln U_{i,t}^{\not{E}}, \tag{A.14}$$

$$u_{i,t}^{*} = \omega_{i,t}\left(d\ln U_{i,t}^{E} - d\ln U_{i,t}^{\not{E}}\right) + d\omega_{i,t}\left(\ln U_{i,t}^{E} - \ln U_{i,t}^{\not{E}}\right) + \text{other}. \tag{A.15}$$

Equation (A.15) motivates two instruments for the change in the downtown premium: (i) from the first term, the initial share of the CBSA that are college graduates interacted with the relative change in valuation of downtown amenities by college graduates, which we measure using the initial relative restaurant density downtown; and (ii) from the second term and using $\ln U_{i,t}^{E} - \ln U_{i,t}^{\not{E}} = \overline{\epsilon}\left(\varsigma_{i,t}^{E} - \varsigma_{i,t}^{\not{E}}\right)$, the predicted change in the CBSA college share interacted with the initial difference in the share of each type who live downtown.[2]

B Empirical Appendix

Appendix B.1 characterizes the relationship between our fundamental and investors/speculators. Appendix B.2 reports time series break tests for rent growth. Appendix B.3 provides details on the Bartik instruments. Appendix B.4 describes the data and measurement underlying the downtown housing premium and associated excluded instruments. Appendix B.5 demonstrates robustness of our main IV results. Appendix B.6 explains how we estimate the short-run elasticity χ. Appendix B.7 contains additional empirical results.

B.1 Investors/Speculators

In this appendix we explain how our long-run fundamental determinant of house prices relates to work emphasizing the role of speculation in the boom. We make five observa-

[2] This framework also accommodates forces discussed in the urbanization literature that do not suggest obvious city-specific instruments, including: (i) a rising skill-premium and convexity in the wage-hours worked profile, which increased the value of time for high-skill workers and hence their demand for short commutes to the city business center (Edlund et al., 2015; Su, 2022); (ii) exogenous improvement in downtown amenities, such as the decline in crime (Pope and Pope, 2012; Ellen et al., 2019); and (iii) endogenous improvement in downtown amenities in response to rising average incomes (Diamond, 2016; Su, 2022).

tions: (i) speculative activity appears potentially important to house price growth late in the boom in some places such as Las Vegas; (ii) long-run fundamentals also explain price growth late in the boom; (iii) speculative activity has much less explanatory power for price growth in the boom up to 2004 or for the full 1997-2019 period, especially compared with the explanatory power of long-run fundamentals; (iv) the degree of speculative activity in the late boom is uncorrelated with the long-run fundamental in an area; and (v) speculators did not contribute disproportionately to selling pressure during the bust. These observations suggest forces orthogonal to fundamentals that made some areas prone to speculation late in the boom, rather than systematic price return chasers who reversed course and added to selling pressure during the bust.

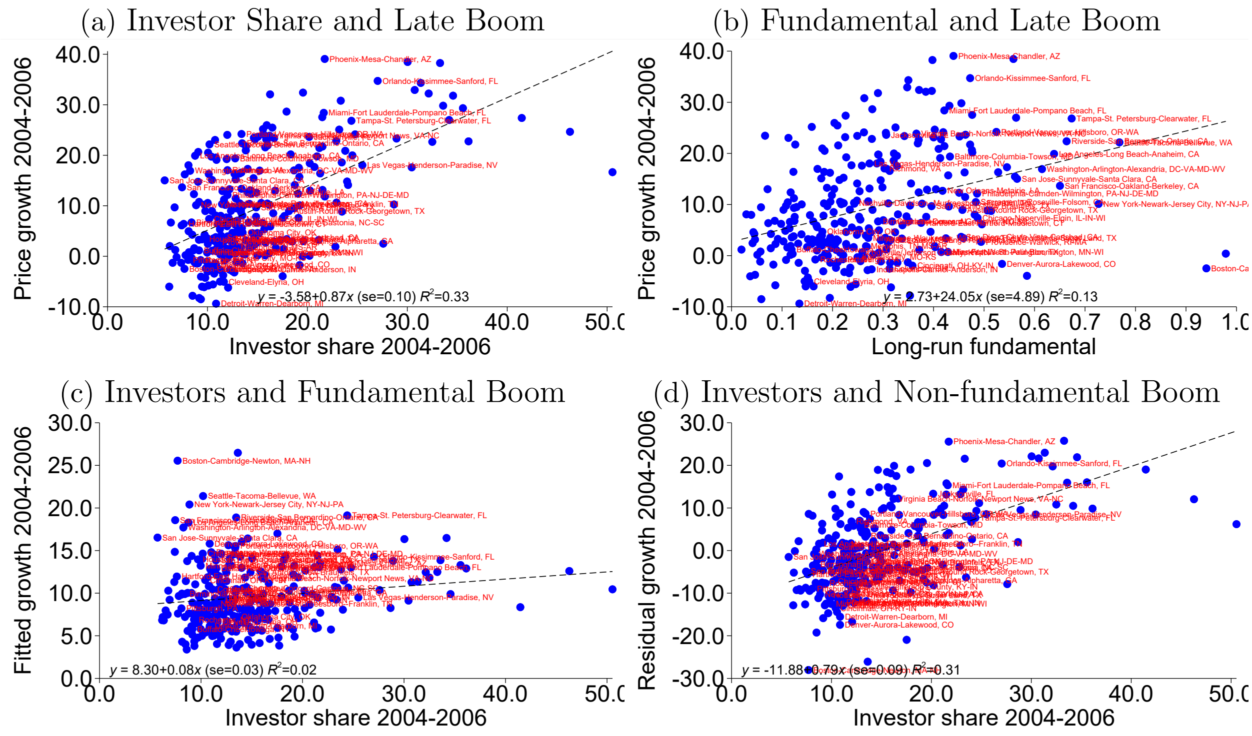
We follow Gao et al. (2020) and associate speculative activity by investors with the non-owner occupier share of purchase mortgages, measured using Home Mortgage Disclosure Act (HMDA) data.[3] Gao et al. (2020) show that the investor share in 2004-06 predicts house price growth in the same period, including when instrumented with state tax treatment of capital gains. Panel (a) of Figure B.1 replicates in our sample of CBSAs the OLS result found in Gao et al. (2020) and shows a strong positive relationship.[4]

Panel (b) shows that house price growth during 2004-2006 also correlates strongly with

[3]DeFusco et al. (2017) show that non-occupant buyers are somewhat more likely to pay in cash, suggesting the HMDA-based measure may understate the investor share. On the other hand, some non-occupant buyers likely purchase for reasons other than speculation, such as for the utility of a vacation home. Any such differences should have minimal impact on the conclusions that follow, because uniform level differences between the HMDA-based measure and the actual share of speculators rescale the investor measure and the comparisons across periods in Table B.1 hold fixed the investor share.

[4]For readability, the figure omits the seven CBSAs with pre-boom, 1994-1996 average share above 20%: Barnstable Town, MA (Cape Cod, 28.3%); Cape Coral-Fort Myers, FL (26.6%); Daphne-Fairhope-Foley, AL (25.6%); Hilton Head Island-Bluffton, SC (27.8%); Myrtle Beach-Conway-North Myrtle Beach, SC-NC (37.6%); Napes-Marco Island, FL (34.4%); and Ocean City, NJ (45.1%). The investor share in 1994-1996 correlates strongly with the share in 2004-2006 (correlation coefficient of 0.81), reflecting persistence in tax treatment of capital gains and that some areas have high non-owner occupier shares because they are common vacation destinations, as suggested by the list of areas with the highest shares in 1994-1996. The patterns shown in Figure B.1 and Table B.1 continue to hold if we replace the 2004-2006 level of the investor share with the change in the share from 1994-1996.

Figure B.1: Investors' Role in the Late Boom and Relation to Fundamentals



Notes: Panel (a) plots the HMDA investor share of purchases averaged over 2004-2006 against house price growth in 2004-2006. Panel (b) plots the long-run fundamental against house price growth in 2004-2006. Panel (c) plots the investor share against the fitted value from a regression of 2004-2006 house price growth on the long-run fundamental. Panel (d) plots the investor share against the residual from this regression. All panels exclude seven CBSAs with a 1994-96 share above 20%: Barnstable, Town, MA (Cape Cod, 28.3%); Cape Coral-Fort Myers, FL (26.6%); Daphne-Fairhope-Foley, AL (25.6%); Hilton Head Island-Bluffton, SC (27.8%); Myrtle Beach-Conway-North Myrtle Beach, SC-NC (37.6%); Napes-Marco Island, FL (34.4%); and Ocean City, NJ (45.1%). CBSAs with more than 1 million persons in 1997 are labeled in red.

the long-run fundamental, measured as usual as the fitted value from column (3) of Table 1. Thus, even at the end of the boom when speculative activity was plausibly most rampant, long-run fundamentals continue to explain house price growth.

Table B.1 summarizes the relationship among investors, fundamentals, and house price growth for several periods. Columns (1)-(2) reproduce the positive correlations shown in panels (a) and (b) of Figure B.1 of investor share and fundamentals with 2004-2006 house price growth. Column (3) shows that both variables contain predictive power when entered into a joint regression. Consistent with speculative activity peaking in the late boom,

Table B.1: House Price Growth, Investors, and Fundamentals

House price growth:	2004-2006			1997-2004			1997-2019		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Investor share	0.87**		0.81**	0.57**		0.32*	0.92**		0.58**
	(0.10)		(0.09)	(0.16)		(0.15)	(0.16)		(0.14)
Long-run fundamental		0.24**	0.19**		0.73**	0.71**		1.01**	0.97**
		(0.05)	(0.04)		(0.08)	(0.08)		(0.07)	(0.08)
Standard deviation of explanatory variables:									
Investor share	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6
Fundamental ($\times$100)	15.6	15.6	15.6	15.6	15.6	15.6	15.6	15.6	15.6
R^2	0.330	0.131	0.407	0.034	0.290	0.300	0.068	0.429	0.455
Observations	301	301	301	301	301	301	301	301	301

Notes: The table reports the coefficients from regressions of real house price growth by CBSA on the investor share, measured as the 2004-2006 average share of purchase mortgages to non-owner occupiers in HMDA, and the long-run fundamental, measured as the fitted value of column (3) of Table 1. Robust standard errors in parentheses. **,$^{+}$ denote significance at the 1, and 10 percent levels, respectively.

the investor share has much less explanatory power for 1997-2004 house price growth ($R^2 = 0.03$, column (4)), especially compared to the explanatory power of the long-run fundamental for the early boom ($R^2 = 0.29$, column (5)). The R^2 of the long-run fundamental for house price growth over the full 1997-2019 period of 0.43 (column (8)) substantially exceeds the R^2 of 0.07 for the investor share (column (7)).

Panels (c) and (d) of Figure B.1 decompose the correlation from Panel (a) into the correlation of investor share with the part of house price growth explained by long-run fundamentals and a residual, respectively. We measure the part explained by fundamentals as the fitted value from the relationship plotted in Panel (b) and the non-fundamental part as the residual from this regression. Panel (c) displays a small positive correlation between the investor share and the part of 2004-2006 price growth correlated with fundamentals, but the explanatory power is weak ($R^2 = 0.02$) and the positive sign does not survive weighting by population. In other words, the 2004-2006 investor share is essentially uncorrelated with the long-run fundamental that is the focus of our paper. Las Vegas provides an example of a CBSA with a high investor share but a relatively low long-run fundamental and hence a small predicted value for 2004-2006 house price growth.

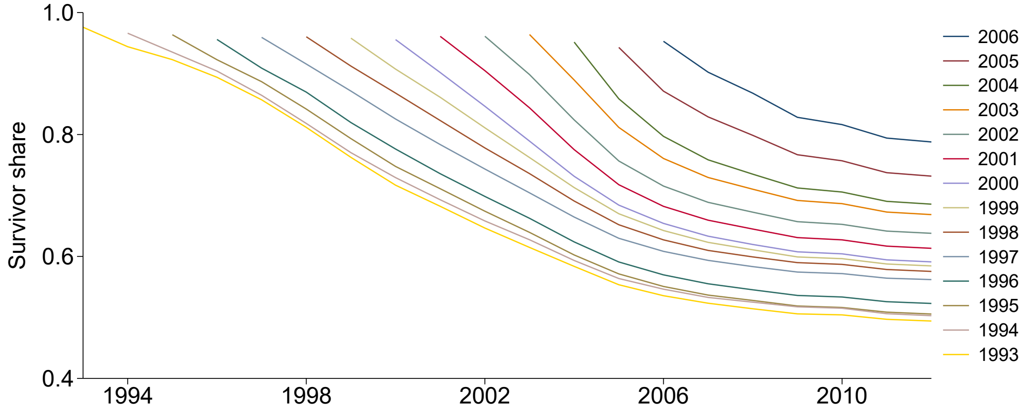
Panel (d) plots the investor share against the non-fundamental component of house price growth in 2004-2006. Areas with late price booms not explained by their long-run fundamental had higher investor shares of purchases, explaining essentially all of the overall correlation shown in Panel (a). Las Vegas again provides a leading example, with faster house price growth than its fundamental would predict and a high investor share.

Finally, we show empirically that investors do not contribute to significant selling pressure in the bust. Using a merge of HMDA and the DataQuick deeds data from Diamond et al. (2019) and the HMDA-based investor measure, we identify the next arms-length transaction on a property previously purchased by an investor. Figure B.2 shows that the survival functions for continuing to own the property flatten around 2006 for all investor cohorts; rather than dump their properties *en masse* in the bust, investors became *less* likely to sell. Consequently, while the emergence and receding of investor demand can potentially also explain a rise and fall in prices, it cannot substitute for the role of foreclosures in explaining why prices fell far below their long-run level in the bust.[5]

Overall, these results are consistent with speculation playing a role late in the house price boom in areas such as Las Vegas. However, they also suggest that the role of investors was mostly or wholly orthogonal to the role of fundamentals, less important than fundamentals to explaining the entirety of the boom or the full 1997-2019 period, and mostly unrelated to the over-shooting of prices in the bust.

[5] Las Vegas again illustrates the exception that proves the rule: It had a larger-than-predicted bust given the size of its fundamental and a relatively muted rebound given the magnitude of the boom, indicating a larger role for speculation in driving the boom-bust cycle than in the typical area. This pattern is representative; a regression of 2006-2019 house price growth on the 2004-2006 investor share indicates 5 percentage points lower house price growth in the bust-rebound for each 10 percentage points higher investor share (Table B.1 column (7) less sum of columns (1) and (4)).

Figure B.2: Survival Function For Investors From Matched HMDA-DQ Data



Notes: This figure shows survival functions for cohorts of investors in the matched HMDA-DataQuick data. Investors are defined as non-owner-occupiers in the HMDA data. For each cohort of investors that purchased in a given year, we compute the fraction of investors who have yet to sell at each year. The figure plots this survival function for each cohort.

B.2 Rent Break Tests

Table B.2 investigates the acceleration in rent growth around the start of the housing boom using a structural break test. The upper panel shows that the national break in rent growth in 1997Q3 closely coincides with the break in price growth in 1998Q2. The lower panel shows that among the 17 CBSAs with a statistically significant price break between 1994Q1 and 2000Q4 and with CPI rent data, 10 have statistically significant breaks in rent growth within two years of the price break. Moreover, these cities have larger and more statistically significant mean jumps in price at the timing of their breaks.

B.3 Bartik Instrument Details

We combine the CBP files provided by the Census with the files from Eckert et al. (2020) that optimally impute suppressed employment cells and provide a consistent correspondence to NAICS 2012. We use 1998 rather than 1997 as the initial year because the NAICS version of the data start in that year. The final year of data available is 2018. We implement "leave-one-out" shift shares: defining $E_{i,j,0}$ as employment in area i and industry j as

Table B.2: Rent Break Timing

National price growth break:			1998Q2	
National CPI rent growth break:			1997Q3	
Price-rent break gap	Number of CBSAs	Mean price jump (p.p.)	Mean rent jump (p.p.)	Mean price break test statistic
0-1 year	5	2.5	1.7	98.5
1-2 years	5	2.1	1.3	109.3
3-4 years	1	0.8	1.1	54.7
No rent break	6	0.9	.	55.5

Notes: The top panel reports the quarter (price) or half-year (rent) date of the Bai and Perron (1998) test for a series break between 1992 and 2006 as implemented in Ditzen et al. (2021). Both breaks are statistically significant at the 1% level. The bottom panel reports statistics grouped by the gap in years between the break date for prices and rents for the 17 CBSAs with a price break significant at the 5% level and CPI rent data. The last column reports the mean of the double maximum test statistic for the price break.

a share of total date 0 employment in area i, $g_{-i,j}$ as the growth rate of employment in industry j in all other areas between dates 0 and 1, $w_{-i,j,t}$ as the wage (payroll per employee) in industry j in all other areas at date t, and $\hat{E}_{i,j,1} \equiv E_{i,j,0} \times g_{-i,j} / \left[\sum_k E_{i,k,0} \times g_{-i,k}\right]$ as the predicted date 1 area i employment share in industry j, the shift-share for the growth of employment is $\sum_j E_{i,j,0} g_{-i,j}$ and the shift-share for the growth of the average wage is $\left[\sum_j \hat{E}_{i,j,1} \times w_{-i,j,1}\right] / \left[\sum_j E_{i,j,0} \times w_{-i,j,0}\right] - 1$, where date 0 is 1998 and date 1 is 2018. For area-industries with suppressed wage data, we replace $w_{-i,j,t}$ with $w_{j,t}$, where $w_{j,t}$ is the national wage in industry j at date t.

B.4 Urbanization Measurement and Instrument Details

We measure the downtown price premium and associated excluded instruments as follows. We follow Holian and Kahn (2012) and define the center of the downtown of each CBSA as the coordinates returned from inputting the largest city in the CBSA into Google Earth. As in Couture and Handbury (2020), we then rank all Census tracts in the CBSA by their distance to the downtown center and define the downtown as those tracts covering the closest 5% of population using 2010 tract definitions and 1990 Census population counts apportioned to 2010 tract definitions by the US2010 Project webpage.[6]

We compute the downtown price premium using ZIP code-level house price data from Zillow and FHFA and tract-level data from the Census. We map tracts into ZIP codes using the 2014 crosswalk from the Missouri Census Data Center, assigning partial downtown shares to ZIP codes covering tracts in and out of the downtown. We compute population-weighted house price growth for ZIP codes or tracts in the downtown and not in the downtown for each CBSA. Both Zillow and FHFA have incomplete coverage of ZIP codes in 1997. We use the log change in the downtown premium from Zillow if data exist and from FHFA otherwise. If neither Zillow nor FHFA cover ZIP codes in the downtown and remainder of a CBSA in 1997 (30 of 308 and CBSAs), we use the fitted value from a regression of the log change in the Zillow/FHFA premium from 1997 to the end year on the log change in the premium from the 2000 Census to the 2015-2019 ACS.

We measure the 1990 share of the CBSA that are college graduates using the decennial Census as compiled in the US2010 Project webpage. We measure the relative density of restaurants (SIC code 5812) in the downtown using the 1997 County Business Pattern ZIP code files and the mapping described above from ZIP codes to downtown. We measure the 1990 shares of college and non-college residents who live in downtown tracts using the Census data from the US2010 Project webpage. Finally, we predict for each CBSA the change in the share of the CBSA with a college degree by combining the actual 1990 CBSA industry distribution from the County Business Patterns, the national share of workers in each industry with a college degree in the 1990 Census and 2019 ACS, and the predicted 2019 industry distribution obtained by applying the national industry growth rates between the 1990 Census and 2019 ACS to the actual 1990 industry distribution.

[6] http://www.s4.brown.edu/us2010/Researcher/Bridging.htm.

Table B.3: Robustness of Long-run Regression

Regressor:	s		h		$s \times m \times h$		$s \times u$		Obs.
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	
Specification:									
1. Baseline	**0.78**	0.20	**0.63**	0.10	**1.30**	0.28	**1.40**	0.38	308
2. FHFA HPI	**0.70**	0.21	**0.55**	0.09	**1.22**	0.27	**1.26**	0.37	306
3. CoreLogic HPI	**0.80**	0.21	**0.46**	0.10	**1.20**	0.29	**1.35**	0.38	308
4. Zillow HPI	**0.81**	0.24	**0.37**	0.10	**1.01**	0.25	**1.07**	0.34	225
5. Saiz unavail.	**0.92**	0.21	**0.76**	0.10	**1.10**	0.28	**1.21**	0.35	260
6. Alt. pop.	**0.87**	0.18	**0.61**	0.08	**1.70**	0.29	**1.14**	0.37	308
7. Pop. weighted	**1.71**	0.28	**0.87**	0.14	**1.22**	0.39	0.07	0.21	308
8. Drop pop.$< 150K$	**0.80**	0.23	**0.60**	0.10	**0.97**	0.23	**1.42**	0.35	219
9. Drop shrinking	**1.00**	0.22	**0.55**	0.10	**1.12**	0.26	**0.98**	0.36	277
10. GMM	**0.57**	0.14	**0.79**	0.08	**1.46**	0.22	**1.60**	0.32	308
11. Bias-adjusted 2SLS	**0.69**	0.28	**0.71**	0.14	**1.83**	0.62	1.34	0.71	308
12. JIVE	0.31	0.42	**0.70**	0.18	**2.25**	0.84	1.82	1.81	308
13. No climate instr.	**1.01**	0.24	**0.76**	0.13	**1.45**	0.40	**1.09**	0.41	308
14. No lifestyle instr.	**0.78**	0.20	**0.63**	0.10	**1.30**	0.28	**1.40**	0.38	308
15. No Bartik instr.	**0.72**	0.22	**0.66**	0.10	**1.41**	0.32	**1.55**	0.47	308
16. No land avail. instr.	**0.77**	0.25	**0.83**	0.12	**1.45**	0.35	**1.09**	0.37	308
17. No density instr.	**0.89**	0.22	**0.59**	0.11	**1.52**	0.34	**1.65**	0.59	308
18. Control lag units	**0.75**	0.20	**0.61**	0.15	**1.29**	0.28	**1.39**	0.39	308
19. Control lag HPI	**0.76**	0.21	**0.64**	0.09	**1.33**	0.29	**1.32**	0.38	308

Notes: Each row reports coefficients and standard errors from a separate modification of the specification in column (3) of Table 1. In the table header, s denotes the land share, h units growth, m the 2006 WRLURI, and u the growth of the price premium in downtown neighborhoods. Coefficients in bold font are statistically different from 0 at the 5% level.

We construct this shift-share using the ind1990 variable from IPUMS adjusted to give a balanced panel by updating the file from David Dorn and a Census crosswalk from the 1987 SIC variable in the County Business Patterns to ind1990.[7]

B.5 IV Robustness

Table B.3 collects several specifications that address potential concerns with the baseline IV regression. Each row reports the coefficients and standard errors from a separate specification. The first row reproduces the baseline coefficients from column (3) of Table 1.

[7] See `https://www.ddorn.net/data/subfile_ind1990dd.zip` and `https://www.census.gov/content/dam/Census/library/working-papers/2003/demo/techpaper2000.pdf`.

Rows (2)-(4) show robustness to alternative house price indexes from FHFA, CoreLogic, and Zillow. Although these indexes vary in their samples and methodologies, all yield similar results.[8] Row (5) replaces the land unavailability instrument with the measure from Saiz (2010).[9] Because Saiz (2010) developed his measure for 1999 MSA definitions, we lose 16% of the sample, but the coefficients change little. Row (6) replaces housing units growth with the growth of population with little change.

Rows (7) to (9) explore robustness to the sample, in row (7) by weighting by population, in row (8) by excluding 89 CBSAs with 1997 population below 150,000, and in row (9) by excluding CBSAs with declining population. The only notable difference is that the weighted specification has a higher loading on the main effect on land share and a smaller loading on the urbanization term, reflecting the fact that much of the variation in urbanization occurs across large and small CBSAs (Couture and Handbury, 2020).

Rows (10) to (12) explore robustness to the estimator, in row (10) by replacing two-stage least squares with GMM, in row (11) with the JIVE estimator of Angrist et al. (1999), and in row (12) with the biased-adjusted estimator of Donald and Newey (2001). The JIVE and bias-adjusted estimators address a particular concern that many instruments over-fit the first stage, biasing the second stage toward OLS (Bekker, 1994; Bound et al., 1995).[10] Unlike in the canonical many weak instrument case of Angrist and Krueger (1991), however, Table B.5 shows that the instruments are generally strong predictors of the endogenous variables, and these specifications produce qualitatively similar coefficients.

[8] Like Freddie Mac, FHFA uses a repeat-sales methodology in a sample of loans purchased by Fannie Mae or Freddie Mac, but weights the sales differently. CoreLogic also uses a repeat sales methodology but includes sales not associated with mortgages purchased by a GSE. Zillow combines sales and other data in order to estimate the average price of a home in the middle tercile of each market regardless of whether it transacts in a period. Row (4) contains all CBSAs with non-missing Zillow data in 1997.

[9] Lutz and Sand define the CBSA boundary as the polygon containing the CBSA plus a 5% buffer. They argue that this improves on the Saiz (2010) measure of the 50km radius around each metropolitan city.

[10] JIVE avoids overfitting by obtaining the fitted value for each observation using a first-stage coefficient vector estimated by excluding that observation from the sample. The Donald and Newey (2001) bias adjustment is a K-class estimator that exactly corrects the IV bias when residuals are homoskedastic.

Rows (13)-(17) remove groups of excluded instruments. The estimation does not critically depend on any particular instrument for population or land, with similar results omitting the climate variables (row 13), CBSA restaurant employment (row 14), shift-shares (row 15), land unavailability (row 16), and population density (row 17). Rows (18) and (19) show that the results change little after controlling for lagged growth.

B.6 Estimation of Short-run Elasticity χ

We augment equations (2) and (3) in the main text to include short-run adjustment costs:

$$\begin{aligned} C_{i,t} &= A_{i,t} H_{i,t}^{\alpha_i} \left(I_{i,t}/\bar{I}_i\right)^{1/\chi_i^c}, \\ L_{i,t} &= B_{i,t} H_{i,t}^{\beta_i} \left(I_{i,t}/\bar{I}_i\right)^{1/\chi_i^\ell}, \\ \text{where: } 1/\chi_i^c &= \alpha_0^\chi + \alpha_1^\chi m_i, \quad 1/\chi_i^\ell = \beta_0^\chi + \beta_1^\chi m_i. \end{aligned}$$

Then applying the same transformations and zero restrictions as in the main text and letting $i_{i,t} = d\ln I_{i,t}$, we have the analogous regression to equation (8):

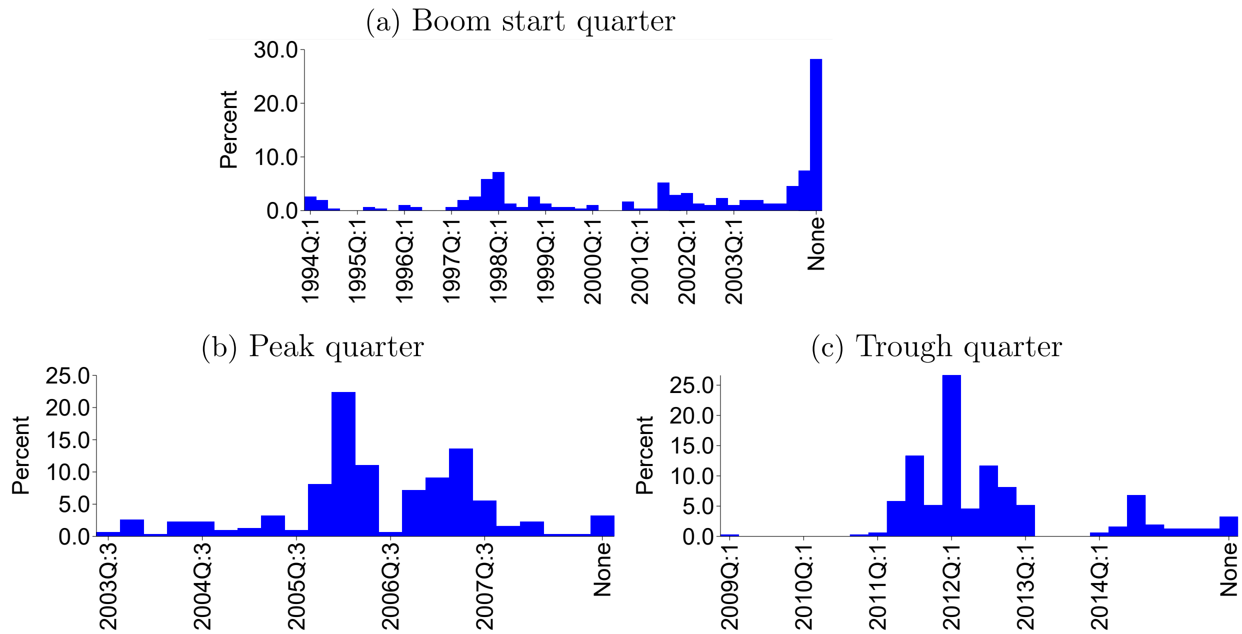
$$p_{i,t} - (1/\eta_i)h_{i,t} = c_0 + c_1 s_{i,t} + c_2 i_{i,t} + c_5\left(s_{i,t} \times m_i \times i_{i,t}\right) + c_6\left(s_{i,t} \times u_{i,t}\right) + e_{i,t}, \tag{B.1}$$

where we form the left hand side using the estimated values of η_i implied by column (3) of Table 1. The inverse short-run supply elasticity is $c_2 + c_5 \times s_i \times m_i$. We estimate equation (B.1) over the period 2000-2005 (to avoid the censal break in 2000) using the same excluded instruments as in the long-run regression and obtain $c_1 = 1.02$ ($s.e.$ 0.23), $c_2 = 0.29$ ($s.e.$ 0.05), $c_5 = 0.78$ ($s.e.$ 0.19), $c_6 = 0.65$ ($s.e.$ 1.10), from which we form $1/\chi_i$.

B.7 Additional Empirical Results

Figure B.3 plots the timing of boom starts, peaks, and bust troughs across CBSAs. Figure B.4 replicates Figure 2 across CBSAs, as discussed in Footnote 5. Table B.4 reports

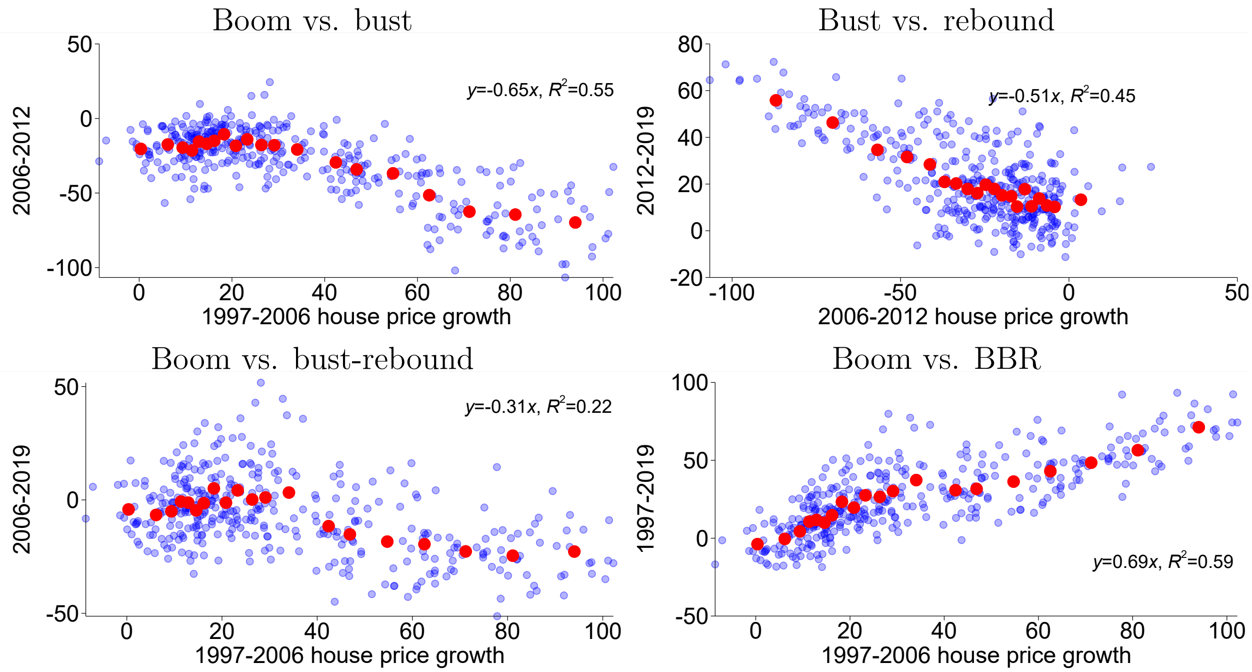
Figure B.3: CBSA Boom, Peak, and Trough Timing



Notes: Panel (a) reports a histogram of the first quarter between 1992:Q1 and 2006:Q2 with a positive structural break in the growth rate of real house prices, using the structural break test of Bai and Perron (1998, 2003) as implemented in Ditzen et al. (2021). Areas without a break identified at the 95% confidence level are shown in the bar labeled "None." Panels (b) and (c) report histograms of the quarter with the peak in real house prices between 2003:Q2 and 2009:Q2 and trough in prices between 2007:Q1 and 2015:Q4, respectively, with areas without an interior extremum shown in the bar labeled "None."

summary statistics. Table B.5 reports first-stage-type regressions for each endogenous variable separately, using only the excluded instruments motivated by that variable and also using the full set of uninteracted instruments to establish the explanatory power of the instruments without broaching many-instrument asymptotics. As discussed in Section 4.1, the instruments are strong and enter with the expected sign. The final column shows the reduced form of house price growth over the BBR on the (uninteracted) instruments, which corresponds to Figure 4. Figure B.5 repeats Figure 4 using only the supply or demand instruments separately and shows that both contribute to higher long-run growth and to the boom-bust-rebound pattern, as discussed in Section 3.3.

Figure B.4: CBSA Boom, Bust, and Rebound



Notes: Each blue circle represents one CBSA. The red circles show the mean value of the y-axis variable for 20 bins of the x-axis variable. Data from Freddie Mac deflated using the national GDP price index.

Table B.4: Summary Statistics

Variable	Mean	SD	P10	P50	P90	Obs.
House price growth 1997-2019	28.0	23.4	−0.6	25.5	59.9	308
Population growth 1997-2019	19.9	17.2	−0.1	17.7	42.3	308
Units growth 1997-2019	25.2	14.8	8.1	23.0	45.0	308
Land share	28.0	9.4	17.8	26.2	40.8	308
WRLURI 2006	−11.8	81.9	−104.5	−23.0	89.7	308
Log change in downtown premium	1.4	15.7	−14.3	0.8	18.5	308
Bartik employment 1998-2018	22.8	6.3	15.6	22.9	30.4	308
Bartik wage 1998-2018	84.5	9.1	71.5	85.4	94.1	308
January temperature	35.6	12.1	21.4	34.3	51.9	308
January sunlight hours	151.1	39.0	104.0	150.9	210.0	308
July humidity	56.4	16.4	26.0	60.3	73.6	308
Land unavailable	30.8	20.4	6.7	26.8	62.7	308
1997 population density	26.6	30.7	5.7	17.3	52.7	308
Non-traditional Christian share	41.5	23.0	10.4	39.5	73.8	308
Inspection/tax revenue	0.8	0.8	0.2	0.5	1.7	308
Col. share $\times$ restaurants	27.1	17.5	9.6	22.9	47.9	308
Downtown diff. $\times \Delta$ col. share	−1.3	25.8	−32.9	−3.1	30.8	308

Notes: This table shows summary statistics for our cross-section of CBSAs.

Table B.5: Pseudo First-stage and Reduced Form Regressions

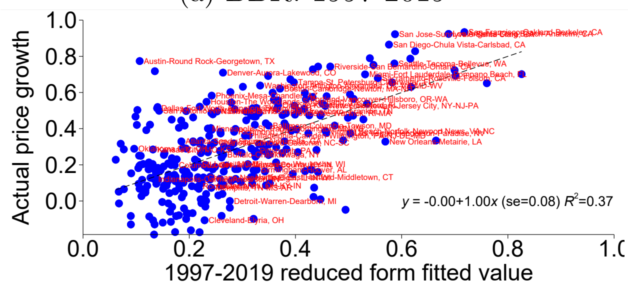
Dep. var.:	s		h		m		u		p
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Unavailability	3.0**	3.0**		−0.6		22.6**		1.3^{+}	10.0**
	(0.5)	(0.5)		(0.9)		(4.5)		(0.8)	(1.1)
Pop. density	4.8**	4.4**		−2.3**		18.3**		2.5*	2.2^{+}
	(0.5)	(0.5)		(0.5)		(5.7)		(1.2)	(1.2)
Bartik wage		0.3	1.3	1.1		3.8		−0.3	1.8^{+}
		(0.5)	(0.8)	(0.7)		(4.4)		(0.7)	(1.0)
Bartik emp.		−0.3	−0.1	1.5^{+}		7.7^{+}		1.8*	2.2*
		(0.5)	(0.8)	(0.8)		(4.3)		(0.8)	(1.0)
January temp.		0.4	4.5**	3.7**		−2.1		−0.2	3.2*
		(0.6)	(0.9)	(1.3)		(6.4)		(1.1)	(1.5)
January sunlight		0.3	2.2*	2.7**		9.2^{+}		1.7*	3.2**
		(0.5)	(0.9)	(0.9)		(5.1)		(0.8)	(1.2)
July humidity		−0.2	−4.6**	−3.7**		−14.3**		−2.0*	−7.0**
		(0.5)	(0.9)	(0.9)		(5.0)		(0.8)	(1.2)
Restaurants		1.9**	4.5**	3.1**		−2.9		−1.0	0.9
		(0.4)	(0.8)	(0.9)		(5.2)		(0.9)	(1.0)
Nontrad. Christ.		−1.2*		1.5	−26.7**	−23.1**		−0.7	−3.1**
		(0.6)		(1.1)	(4.3)	(5.1)		(1.0)	(1.1)
Inspection/tax		0.9^{+}		−1.4	21.9**	8.1*		−0.1	3.7**
		(0.5)		(0.9)	(3.4)	(3.7)		(0.7)	(1.1)
Col. share × rest.		1.1*		0.8		2.7	5.6**	4.5**	4.8**
		(0.5)		(0.8)		(4.9)	(0.9)	(1.0)	(1.2)
Downtown diff. ×		−0.4		2.9**		−1.2	2.0*	2.8**	0.3
$\widehat{\Delta\text{col. share}}$		(0.4)		(0.8)		(4.0)	(0.9)	(0.8)	(0.9)
Effective F	64.5	19.2	19.1	13.9	39.3	12.3	24.1	8.8	
R^2	0.366	0.444	0.295	0.382	0.179	0.341	0.158	0.241	0.575
Observations	308	308	308	308	308	308	308	308	308

Notes: Columns (1), (3), (5), and (7) report regressions of an endogenous variable on the group of excluded instruments associated with that variable. Columns (2), (4), (6), and (8) report regressions of an endogenous variable on all excluded instrument main effects. Column (9) reports the reduced form regression of house price growth on all excluded instrument main effects. In the table header, s denotes the land share, h units growth from 1997 to 2019, m the 2006 WRLURI, u the growth of the price premium in downtown neighborhoods from 2007 to 2019, and p the log change in the house price from 2007 to 2019. All independent variables normalized to have unit variance. Heteroskedastic-robust standard errors in parentheses. The effective F-statistic is computed as in Montiel Olea and Pflueger (2013). **,* ,$^{+}$ denote significance at the 1, 5, and 10 percent levels, respectively.

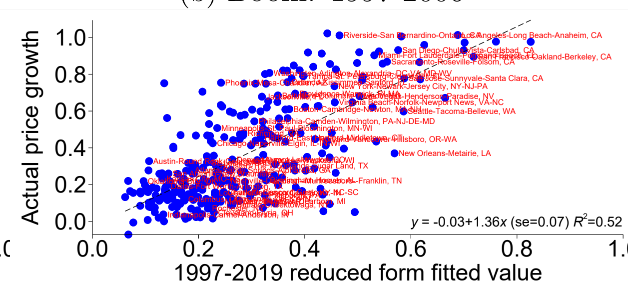
Figure B.5: Predicted and Actual Price Growth

Land Share and Regulation Instruments Only

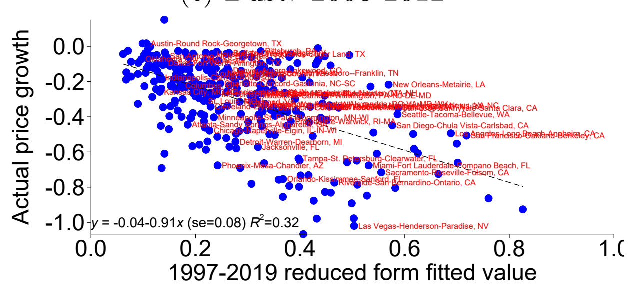
(a) BBR: 1997-2019



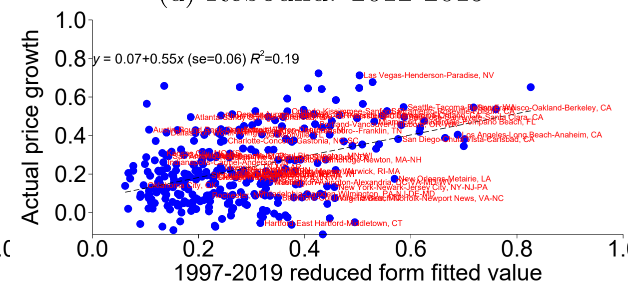
(b) Boom: 1997-2006



(c) Bust: 2006-2012

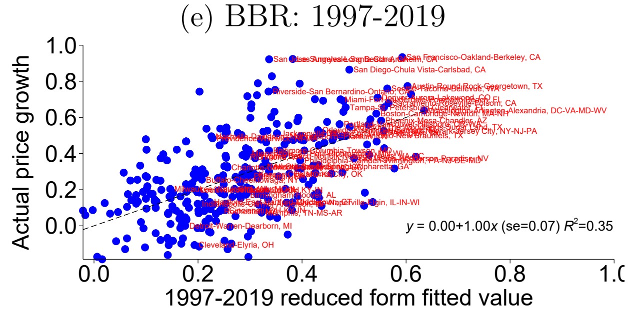


(d) Rebound: 2012-2019

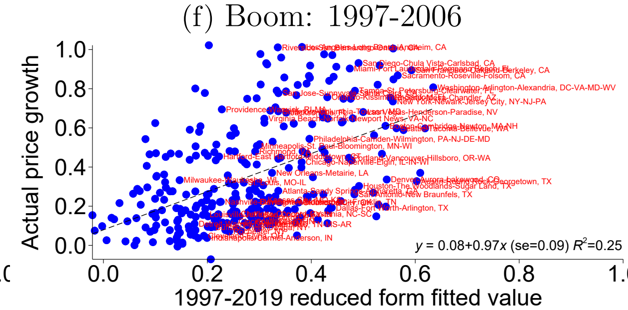


Population and Urbanization Instruments Only

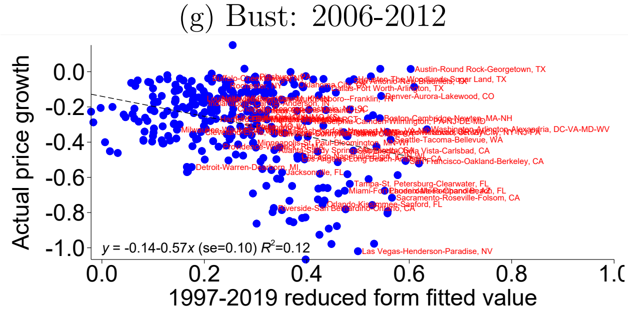
(e) BBR: 1997-2019



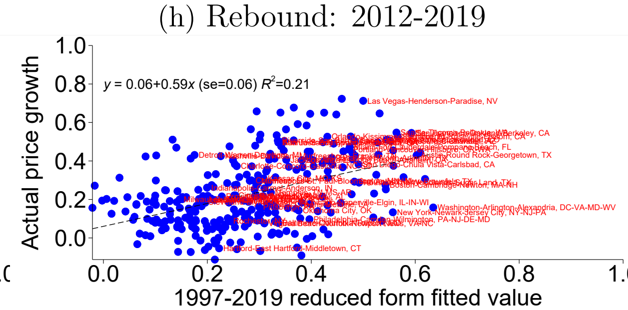
(f) Boom: 1997-2006



(g) Bust: 2006-2012



(h) Rebound: 2012-2019



Notes: In each panel, each blue dot is the real house price growth in a CBSA over the period indicated on the vertical axis plotted against the predicted real house price growth over the period 1997-2019. Panels (a)-(d) predict price growth using only the excluded instruments associated with land share and WRLURI (land unavailability, 1997 population density, non-traditional Christian share, and public expenditure on protective inspection). Panels (e)-(h) predict price growth using only the excluded instruments associated with population growth and urbanization (employment and wage shift-shares, climate variables, CBSA restaurant density, relative restaurant density in the downtown × the 1990 college share, and the 1990 difference in the college/non-college likelihood of living downtown × the CBSA predicted change in college share). CBSAs with more than 1 million persons in 1997 are labeled in red.

C Model Appendix

This appendix provides derivations and proofs for the model in Section 5. Appendix C.1 derives the present value of dividends. Appendix C.2 provides an analytic path of beliefs that we use to calculate impulse responses. Appendix C.3 defines the balanced growth path. Appendix C.4 derives the user cost for rents. Appendix C.5 associates rent acceleration with dividend growth. Appendix C.6 shows impulse responses when lenders have perfect foresight rather than diagnostic beliefs.

C.1 Present Value of Dividends

We restate equation (16) for convenience:

$$P_t^* = \int_{-\infty}^{\infty} \mathbb{E}_t \left[\int_t^{\infty} e^{-\rho(s-t)} D_s ds | \mu_t \right] h_t^{\theta} \left(\mu_t \right) d\mu_t.$$

We want to prove that this integral depends only on D_t, m_t^{θ}, and parameters. Start by fixing μ_t and D_t. Since D_t is a geometric Brownian motion, we have $e^{-\rho(s-t)} D_s = D_t \exp\left(-\rho\left(s-t\right) - \frac{1}{2}\sigma_D^2 \left(s-t\right) + \int_t^s \mu_\tau d\tau + \sigma_D \int_t^s dW_{D,\tau}\right)$. Taking an expectation:

$$\mathbb{E}_t \left[e^{-\rho(s-t)} D_s | \mu_t \right] = D_t \exp\left[-\rho\left(s-t\right) + \mathbb{E}_t\left[\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_{\mathrm{t}} \right] + \frac{1}{2} \mathrm{Var}\left(\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_{\mathrm{t}} \right) \right].$$

One can show: $\mathbb{E}_t \left[\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_t \right] = \bar{\mu}\left(s-t\right) + \frac{1}{\vartheta}\left(1 - e^{-\vartheta(s-t)}\right)\left(\mu_t - \bar{\mu}\right),$

$$\mathrm{Var}\left(\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_t \right) = \frac{\sigma_\mu^2}{\vartheta^2}\left(s-t\right) - \frac{3\sigma_\mu^2}{2\vartheta^3} + \frac{\sigma_\mu^2}{2\vartheta^3}\left[4e^{-\vartheta(s-t)} - e^{-2\vartheta(s-t)}\right],$$

giving: $$\mathbb{E}_t \left[e^{-\rho(s-t)} D_s | \mu_t \right] = D_t \exp\left[-\left(\rho - \bar{\mu}\right)\left(s-t\right) + \frac{1}{\vartheta}\left(1 - e^{-\vartheta(s-t)}\right)\left(\mu_t - \bar{\mu}\right) + G\left(s-t\right) \right],$$

where: $$G\left(s-t\right) = \frac{\sigma_\mu^2}{2\vartheta^2}\left(s-t\right) - \frac{3\sigma_\mu^2}{4\vartheta^3} + \frac{\sigma_\mu^2}{4\vartheta^3}\left[4e^{-\vartheta(s-t)} - e^{-2\vartheta(s-t)}\right].$$

Substituting these expressions into equation (16) gives the desired result:

$$\begin{aligned} P_t^*/D_t &= \int_0^\infty \mathbb{E}_{\mu_t}^{\theta} \exp\left[-(\rho-\bar{\mu})\tau + \frac{1}{\vartheta}\left(1-e^{-\vartheta\tau}\right)(\mu_t-\bar{\mu}) + G(\tau)\right] d\tau \\ &= \int_0^\infty \exp\left[-(\rho-\bar{\mu})\tau + \frac{1}{\vartheta}\left(1-e^{-\vartheta\tau}\right)\left(m_t^{\theta}-\bar{\mu}\right) + G(\tau)\right] \exp\left[\frac{\sigma_m^2}{2\vartheta^2}\left(1-e^{-\vartheta\tau}\right)^2\right] d\tau. \end{aligned} \tag{C.1}$$

C.2 Analytic Path of Beliefs

We solve for the mean path of beliefs m_t^{θ} starting from the initial condition $m_0 = \bar{\mu}$ and the initial drift rate μ_0. That is, we solve for m_t^{θ} if all subsequent Wiener shocks are equal to 0. From equation (14), we have:

$$m_t^{\theta} = m_t + \theta \mathcal{I}_t. \tag{C.2}$$

We first characterize the path of m_t, and then the path of $\theta \mathcal{I}_t$.

We first solve the SDE for m_t. Substituting equations (9) and (12) into equation (11):

$$\begin{aligned} dm_t &= \vartheta(\bar{\mu}-m_t)\,dt + K dB_t \\ &= \vartheta(\bar{\mu}-m_t)\,dt + K\sigma_D^{-1}\left(\mu_t dt + \sigma_D dW_{D,t} - m_t dt\right) \\ &= \left(\vartheta\bar{\mu} + \kappa\mu_t - (\kappa+\vartheta)\,m_t\right) dt + K dW_{D,t}, \end{aligned} \tag{C.3}$$

where $\kappa \equiv K/\sigma_D$. The solution to this SDE is:

$$m_t = m_0 e^{-(\kappa+\vartheta)t} + \vartheta\bar{\mu}\int_0^t e^{-(\kappa+\vartheta)(t-s)} ds + \kappa\int_0^t e^{-(\kappa+\vartheta)(t-s)}\mu_s ds + K\int_0^t e^{-(\kappa+\vartheta)(t-s)} dW_{D,s}. \tag{C.4}$$

Note that equation (10) implies:

$$\mathbb{E}_0\left[\mu_t|\mu_0\right] = e^{-\vartheta t}\mu_0 + \left(1-e^{-\vartheta t}\right)\bar{\mu}. \tag{C.5}$$

Taking a conditional expectation of (C.4), using (C.5), and simplifying terms:

$$\mathbb{E}_0\left[m_t|\mu_0, m_0\right] = \bar{\mu} + (\mu_0 - \bar{\mu})\, e^{-\vartheta t} - (\mu_0 - m_0)\, e^{-(\kappa+\vartheta)t}. \tag{C.6}$$

We next solve for the mean path of $\theta\mathcal{I}_t$. Using equations (9), (12) and (15), we have:

$$\theta\mathcal{I}_t = K\theta \int_{t-k}^{t} e^{-\vartheta(t-s)} dB_s \quad = \theta\kappa \int_{t-k}^{t} e^{-\vartheta(t-s)} \left(\mu_s - m_s\right) ds + \theta K \int_{t-k}^{t} e^{-\vartheta(t-s)} dW_{D,s}. \tag{C.7}$$

Note that equations (C.5) and (C.6) together imply that for any $s \geq 0$, $\mathbb{E}_0\left[\mu_s - m_s|\mu_0, m_0\right] = (\mu_0 - m_0)\, e^{-(\kappa+\vartheta)s}$. Therefore:

$$\begin{aligned}\mathbb{E}_0\left[\theta\mathcal{I}_t|\mu_0, m_0\right] &= \theta\kappa \int_{\max\{t-k,0\}}^{t} e^{-\vartheta(t-s)} \left(\mu_0 - m_0\right) e^{-(\kappa+\vartheta)s} ds \\ &= \theta\left(\mu_0 - m_0\right)\left(e^{-\kappa \max\{t-k,0\} - \vartheta t} - e^{-(\kappa+\vartheta)t}\right).\end{aligned} \tag{C.8}$$

Equations (C.6) and (C.8) together characterize the mean path of diagnostic beliefs $\mathbb{E}_0\left[m_t^\theta|\mu_0, m_0\right]$ that we use to solve for the path of $P^*\left(D_t, m_t^\theta\right)$.

C.3 Balanced Growth Path

A balanced growth path (BGP) consists of a fixed μ, constant rate of construction and foreclosures, and constant ratio of points to price. Let δ^f denote the BGP ratio of foreclosures to population and ω denote the ratio of points to price. Substituting this notation into the market-clearing condition (22) gives:

$$g_H H_t x_m^\gamma \left[V_t / \left(1+\omega\right) P_t\right]^\gamma = \dot{H}_t + \delta^f H_t. \tag{C.9}$$

Dividing through by H_t and recognizing that a constant construction rate means $dH_t/H_t = I_t$ is constant on the BGP, it is apparent that V_t/P_t is also constant. Since V_t grows at rate μ (see the valuation function (C.1)), BGP prices also grow at μ.

It only remains to verify that the foreclosure rate is constant on the BGP. With a constant liquidity shock ι, this will be true if the loan-to-value (LTV) distribution remains stable. Let $M(s,t)$ denote the balance in period s of a buyer who bought in period t and $m(s,t) = M(s,t)/P_s$ the current LTV of that buyer. With initial LTV of ϕ and allowing for generality for a mortgage pay down rate of ς, we have that $m(s,t) = \phi e^{-(\varsigma+\mu)(s-t)}$. This expression can be inverted to find the date t at which someone with LTV m at date s must have bought: $t(m,s) = s - \frac{\ln(\phi/m)}{\varsigma+\mu}$. Let $F(s,t)$ be the cumulative share of mortgages outstanding that bought before date t. Since new mortgages are written at a rate of $I+\iota$ each period, $F(s,t) = e^{-(I+\iota)(s-t)}$. Let $G(m,s)$ denote the share of mortgages outstanding at date s with LTV of less than m. Then:

$$G(m,s) = \left(\frac{\phi}{m}\right)^{-\frac{I+\iota}{\varsigma+\mu}}, \tag{C.10}$$

confirming that the LTV distribution is stable on the BGP.

C.4 Rent Details

We first derive a general user cost expression. To clarify notation, throughout this section we suppress the i subscript for an individual city and let τ index time elapsed since period t, $j = \tau/\Delta$ be the number of periods of length Δ that have elapsed after τ time units (e.g., if Δ is 1 week and τ is in years then at the end of one year when $\tau = 1$ there have been 52 periods of length 1 week), and $t+T$ be the end of time.

With Poisson intensity λ a household may re-optimize the rent-own decision. Let $R_{t+\tau|t}$ denote the rent paid in period $t+\tau$ for a contract signed in period t. An agent must be indifferent between renting and owning, where owning involves an outlay of the down payment $(1-\phi)P_t$ at date t plus the discounted sum of interest payments $i_t\phi P_t$, other costs of owning such as maintenance or property taxes that are assumed to be proportional to the dividend, ζD_t, and expected cash-flow at sale $\mathbb{E}_t^\theta [P_{t+j}] - \phi P_t$, where the notation $\mathbb{E}_t^\theta$ indicates that expectations are taken using the diagnostic measure.[11] We have:

$$\sum_{j=0}^{T/\Delta} \left(\frac{1-\lambda\Delta}{1+\rho\Delta}\right)^{\tau/\Delta} \Delta \mathbb{E}_t^\theta \left[R_{t+\Delta j|t}\right]$$
$$= (1-\phi) P_t + \sum_{j=0}^{T/\Delta} \left(\frac{1-\lambda\Delta}{1+\rho\Delta}\right)^{\tau/\Delta} \left[\phi P_t i_t \Delta + \zeta \Delta \mathbb{E}_t^\theta [D_{t+\Delta j}] - \lambda\Delta \left(\mathbb{E}_t^\theta [P_{t+\Delta j}] - \phi P_t\right)\right].$$

Taking the limit as $\Delta \to 0, T \to \infty$ and solving the integral multiplying P_t gives:

$$\int_0^\infty e^{-(\rho+\lambda)\tau} \mathbb{E}_t^\theta \left[R_{t+\tau|t}\right] d\tau = \left(1 + \phi \left(\frac{i_t + \lambda}{\rho + \lambda} - 1\right)\right) P_t$$
$$- \lambda \int_0^\infty e^{-(\rho+\lambda)\tau} \mathbb{E}_t^\theta [P_{t+\tau}] d\tau + c_t \zeta D_t. \tag{C.11}$$

In the continuous time representation, the left hand side is the expected present value of rents paid until the next rent/own decision. On the right hand side, the first term is the price gross of expected discounted interest costs (the interest rate i_t is locked in when the mortgage is signed). The second term is the expected discounted cash flow at sale, which can be written as $\int_0^\infty \lambda e^{-\lambda\tau} \mathbb{E}_t^\theta [e^{-\rho\tau} P_{t+\tau}] d\tau$ to make clear that $\lambda e^{-\lambda\tau}$ is the probability of selling at date $t+\tau$ and $\mathbb{E}_t^\theta [e^{-\rho\tau} P_{t+\tau}]$ is the expected discounted cash flow if the sale occurs at date $t+\tau$. The third term is the expected present value of maintenance costs of owning and is written as $\int_0^\infty e^{-(\rho+\lambda)\tau} \mathbb{E}_t^\theta \zeta D_{t+\tau} d\tau = c_t \zeta D_t$ for a scalar c_t that depends on beliefs of the drift rate and parameters and is given by the right hand side of equation (C.1) with ρ replaced by $\rho + \lambda$. Without uncertainty, $c_t = 1/(\rho + \lambda - \mu)$.

Equation (C.11) involves an indeterminacy since only the expected discounted present value of rents is pinned down by the expected cost of owning. We resolve this indeterminacy by assuming a contract where rents grow at a rate $g_{t|t}$ until the λ reoptimization shock

[11] This derivation involves a slight abuse of notation, as the present value V should subtract these other owning costs. This simply involves a redefinition of ζ.

hits: $R_{t+\tau|t} = e^{g_{t|t}\tau} R_{t|t}$. This assumption captures the empirical regularity that contract rents are sticky but grow at some trend rate. We assume in particular that $g_{t|t} = m_t^{\phi}$, the nowcast of the drift rate, but provide the derivation for a general $g_{t|t}$. Imposing this assumption on equation (C.11), the "reset" rent can then be written as:

$$R_{t|t} = \left(\frac{\rho + \lambda - g_{t|t}}{\rho + \lambda}\right) \left(\left((1-\phi)\,\rho + \phi i_t \right) P_t - \lambda \int_0^{\infty} (\rho + \lambda)\, e^{-(\rho+\lambda)\tau} \left(\mathbb{E}_t^{\theta} P_{t+\tau} - P_t \right) d\tau \right)$$
$$+ \left(\rho + \lambda - g_{t|t}\right) c_t \zeta D_t. \tag{C.12}$$

One can show that in the limit as $\lambda \to \infty$, so that the rent-own decision is re-optimized each instant, this expression collapses to $R_{t|t} = \left((1-\phi)\,\rho + \phi i_t\right) P_t - \mathbb{E}_t^{\theta}\left[\dot{P}_t\right] + \zeta D_t$.

The average rent paid at date t, denoted R_t, evolves according to:

$$\dot{R}_t = g_t R_t + \lambda \left(R_{t|t} - R_t\right), \tag{C.13}$$

$$\text{where: } \dot{g}_t = \left(R_{t|t}/R_t\right) \lambda \left(g_{t|t} - g_t\right). \tag{C.14}$$

Given initial conditions for R_t and g_t, equations (C.12) to (C.14) characterize the path of rents. The initial condition for g_t is $g_0 = \bar{\mu}$, the pre-boom growth rate of house prices and rents. The initial condition for R_0 is chosen to match an average price-rent ratio of 13 over the sample, which we achieve by choosing ζ.

C.5 Rent Growth Acceleration and Fundamentals

Along a BGP with dividend growth of μ, $g_t = \mu$, and $i = \rho$, (C.12) to (C.14) imply:

$$R_t^{\mathrm{BGP}} = (\rho - \mu)\, P_t + \zeta D_t. \tag{C.15}$$

The following lemma shows that a BGP-to-BGP acceleration of rent growth will result from an increase in μ but not from a decline in the discount rate ρ. We refer to this result in Section 2 where we show that rent growth accelerated during the BBR.

Lemma 1 *Suppose up to time* 0, $\rho_t = \bar{\rho}$ *and* $\mu_i = \bar{\mu}_i$. *Let* x_{0-} *be the left-limit value of a variable just before time* 0 *and* x_T *the value after convergence to a new balanced growth path.*

1. *Following a change in the discount rate at date* 0 *from* $\bar{\rho}$ *to* ρ_0, *rents continue to growth at* $\bar{\mu}_i$: $\log\left(R_{i,T}/R_{i,0-}\right) = \bar{\mu}_i t$.
2. *Following a change in the growth rate at date* 0 *from* $\bar{\mu}_i$ *to* $\mu_{i,0}$, *rents grow at the rate* μ_0 *but shift down due to the decline in the price-rent ratio:* $\log\left(R_{i,T}/R_{i,0-}\right) = \log\left(\left(\mu_0/\bar{\mu}\right)^{-1/\gamma}(1-\zeta)+\zeta\right) + \mu_0 t$.

Proof: Recall that along a balanced growth path we have that $P_t^{\text{BGP}} = AH_t^{1/\eta}$ and $I_t H_t = g_H H_t x_m^\gamma \left(\frac{V_t}{P_t}\right)^\gamma$.[12] Using $\dot{V}_t/V_t = \dot{P}_t/P_t = \mu_t$ and $\dot{H}_t/H_t = \eta\mu_t$ and explicitly accounting for the costs of owning ζD_t, the Gordon Growth representation is:

$$P_t^{\text{BGP}} = \left(\frac{g_H}{\eta\mu_t}\right)^{1/\gamma} x_m \frac{D_t(1-\zeta)}{\rho_t - \mu_t}. \tag{C.16}$$

Substituting equation (C.16) into equation (C.15) and grouping terms, we have:

$$R_t^{\text{BGP}} = \left[\left(\frac{g_H}{\eta\mu}\right)^{1/\gamma} x_m (1-\zeta) + \zeta\right] D_t. \tag{C.17}$$

The first claim in the lemma follows immediately, since this expression does not depend on ρ_t. For the second claim, we can normalize $P_{0-} = V_{0-}$, which gives $x_m = \left(\eta\bar{\mu}/g_H\right)^{1/\gamma}$ and hence $\log\left(R_T^{\text{BGP}}/R_{0-}^{\text{BGP}}\right) = \log\left(\left(\mu_0/\bar{\mu}\right)^{-1/\gamma}(1-\zeta)+\zeta\right) + \mu_0 t$.

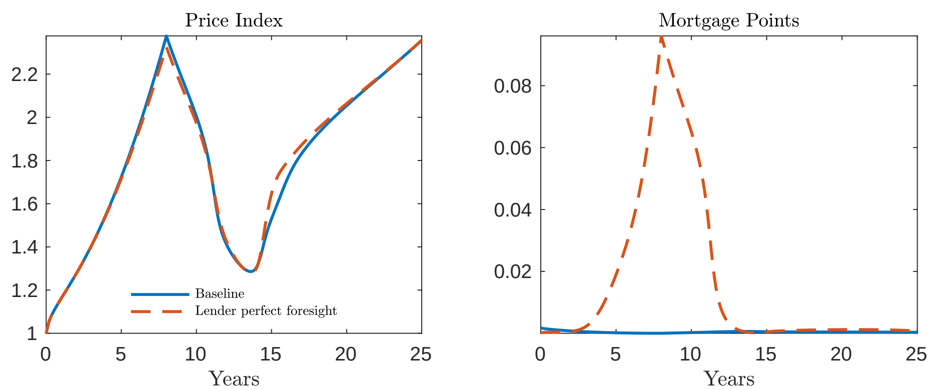
C.6 Lender Perfect Foresight and Role of Credit Markets

Figure C.1 shows the paths of price, P_t, and upfront mortgage cost as a share of the price, W_t/P_t, when lenders have perfect foresight over the path of dividends. Even when lenders perfectly anticipate the peak in buyers' beliefs and hence in prices, the rise in W_t/P_t of 9.6 p.p. has a small impact on prices. This insensitivity reflects the fact that mortgage costs W_t are small relative to the price P_t so that even large changes in W_t shift the demand

[12] Note that the demand equation for $I_t H_t$ coincides with equation (4) in Section 3 for $G = g_H x_m^\gamma$. That is, nothing in this proof requires any of the structure of Section 5 not already imposed in Sections 3 and 4. We ignore foreclosures for simplicity and all of what follows holds in the more general case.

curve by only a small amount relative to changes in V_t. Why do perfect foresight lenders not raise W_t by even more so as to choke off the boom-bust? With the double-trigger for default, the estimated liquidity shock frequency of roughly 5% per year, and the empirical recovery rate of roughly 65% on foreclosures, lenders receive substantial cash flows even on mortgages made just prior to a price peak. The 8p.p. rise in W_t/P_t is exactly sufficient to compensate for the anticipated wave of foreclosures.

Figure C.1: Prices and Mortgage Costs with Lender Perfect Foresight

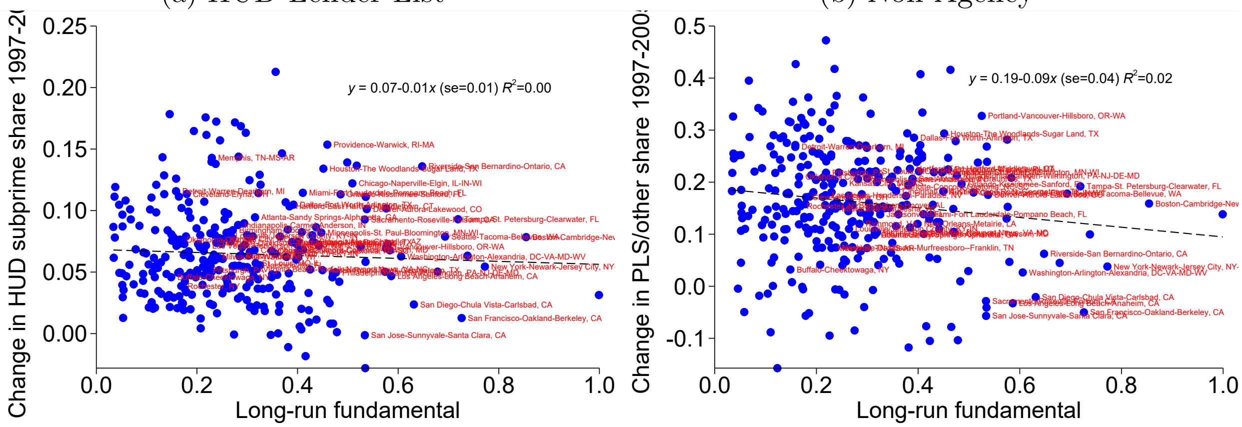


Notes: The figure shows the paths of price, P_t, and upfront mortgage cost as a share of the price, W_t/P_t, for the third quartile of CBSAs when lenders have perfect foresight over the path of dividends in the dashed orange line. The baseline model third quartile of CBSAs is shown in the solid blue line for comparison.

Changes in credit that affect approval rates on the extensive margin offer greater potential to impact prices in our model. We consider an extension in which each potential entrant first draws income y from a CDF $G(y)$ and gets approved for a mortgage only if $y > c_t P_t$. The cutoff parameter c_t encompasses a variety of mechanisms including down-payment constraints and payment-to-income constraints (Greenwald, 2018). With this modification, the parameter g_H becomes instead $(1 - G(c_t P_t))\, g_H$. With some abuse of notation, we can therefore accommodate such policies by replacing g_H in equation (17) with a time-varying potential buyer share $g_{H,t}$. In fact, in the presence of an approval constraint $y > c_t P_t$ that binds in at least part of the distribution of y prior to the boom, our

Figure C.2: Non-Prime Credit and Fundamentals

(a) HUD Lender List (b) Non-Agency

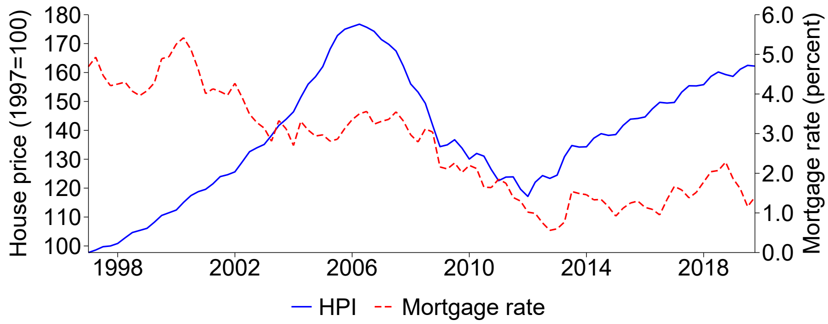


Notes: Panel (a) plots the share of purchase mortgages originated by lenders flagged by the the Department of Housing and Urban Development as subprime lenders against the long-run fundamental. Panel (b) plots the share of purchase mortgages below the jumbo threshold and purchased by non-Agency institutions (private securitization (HMDA code 5), commercial bank, savings bank or savings association (HMDA code 6), life insurance company, credit union, mortgage bank, or finance company (HMDA code 7), affiliate institution (HMDA code 8), and other purchasers (HMDA code 9)) against the long-run fundamental. The data include all first-lien purchase mortgages in HMDA not backed by manufactured housing or buildings with more than four units. CBSAs with more than 1 million persons in 1997 are labeled in red.

calibration with constant g_H *requires* an expansion of credit on the extensive margin (or a rightward shift in the distribution of y), as otherwise an increasing number of potential buyers would get denied mortgage approval as P_t rises (see also Foote et al., 2021). Even so, Figure C.2 shows that the long-run fundamental is essentially uncorrelated with the change in subprime share during the boom, suggesting the role of other types of credit relaxation (e.g. low interest rates that ease payment-to-income constraints) and rising incomes in keeping house prices affordable in high fundamental areas.

Finally, we circumscribe the potential for mortgage rate changes to affect house prices. First, Figure C.3 shows that the timing of rate declines does not generally coincide with the periods of rising prices. We next quantify the impact of the decline in the real mortgage rate from 4.25% in 1998Q1 to 1.4% in 2019Q4. A standard fixed payment mortgage sets a constant dollar payment of $c\Delta$ per time interval Δ and amortizes to 0 after H

Figure C.3: House Prices and Mortgage Rates



Notes: The blue solid line shows the national Case-Shiller index deflated by the GDP price index. The red dashed line shows the Freddie Mac 30 year fixed mortgage rate less median inflation expectations from the Michigan Survey of Consumers.

years. The mortgage balance therefore evolves as $M_{t+h+\Delta} = M_{t+h} - (c\Delta - i_t \Delta M_{t+h}) = (1 + i_t\Delta)\, M_{t+h} - c\Delta$. Taking the limit as $\Delta \to 0$ and using the boundary conditions $M_t = \phi P_t$ and $M_{t+H} = 0$ gives $c = i_t \phi P_t / (1 - e^{-i_t H})$. As most buyers move before final payment, we also allow for early termination after B years. The present value of mortgage costs is therefore $\int_{h=0}^{B} e^{-\rho h} c dh + e^{-\rho B} M_{t+B} = \left(\left(\frac{1-e^{-\rho B}}{1-e^{-i_t H}} \right) \left(\frac{i_t}{\rho} \right) + e^{-\rho B} \left(\frac{1-e^{-i_t(H-B)}}{1-e^{-i_t H}} \right) \right) \phi P_t$. Combining this expression with the down payment of $(1-\phi)P_t$, for a 30 year mortgage with initial LTV of 0.8 and an actual horizon of 15 years (approximately modal numbers for the U.S.), the decline in i causes a decline in the present value of housing costs of 22 log points. This magnitude is not especially sensitive to parameter assumptions.

References Not in Main Text

Alonso, William (1964). *Location and Land Use*. Harvard University Press.

Angrist, Joshua, Guido Imbens, and Alan Krueger (1999). "Jackknife Instrumental Variables Estimation". *Journal of Applied Econometrics* 14 (1): 57–67.

Angrist, Joshua and Alan Krueger (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106 (4): 979–1014.

Bai, Jushan and Pierre Perron (1998). "Estimating and Testing Linear Models with Multiple Structural Changes". *Econometrica* 66 (1): 47–78.

Bai, Jushan and Pierre Perron (2003). “Computation and analysis of multiple structural change models”. *Journal of Applied Econometrics* 18 (1): 1–22.

Bekker, Paul A. (1994). “Alternative Approximations to the Distributions of Instrumental Variable Estimators”. *Econometrica* 62 (3): 657–681.

Bound, John, David A. Jaeger, and Regina M. Baker (1995). “Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak”. *Journal of the American Statistical Association* 90 (430): 443–450.

DeFusco, Anthony, Charles Nathanson, and Eric Zwick (2017). “Speculative Dynamics of Prices and Volume”. NBER Working Paper No. 23449.

Diamond, Rebecca, Tim McQuade, and Franklin Qian (2019). “The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco”. *American Economic Review* 109 (9): 3365–94.

Ditzen, Jan, Yiannis Karavias, and Joakin Westerlund (2021). “xtbreak: Estimating and testing for structural breaks in Stata”. Working Paper, University of Birmingham.

Donald, Stephen G. and Whitney K. Newey (2001). “Choosing the Number of Instruments”. *Econometrica* 69 (5): 1161–1191.

Eckert, Fabian, Teresa Fort, Peter Schott, and Natalie Yang (2020). “Imputing Missing Values in the US Census Bureau’s County Business Patterns”. NBER Working Paper No. 26632.

Edlund, Lena, Cecilia Machado, and Maria Micaela Sviatschi (2015). *Gentrification and the Rising Returns to Skill*. Working Paper 21729. NBER.

Foote, Christopher, Lara Loewenstein, and Paul Willen (2021). “Cross-Sectional Patterns of Mortgage Debt During the Housing Boom: Evidence and Implications”. *The Review of Economic Studies* 88 (1): 229–259.

Greenwald, Daniel (2018). “The Mortgage Credit Channel of Macroeconomic Transmission”. Working Paper, MIT.

Holian, Matthew and Matthew Kahn (2012). “The Impact of Center City Economic and Cultural Vibrancy on Greenhouse Gas Emissions from Transportation”.

Mills, Edwin S. (1967). “An Aggregative Model of Resource Allocation in a Metropolitan Area”. *The American Economic Review* 57 (2): 197–210.

Montiel Olea, José Luis and Carolin Pflueger (2013). “A Robust Test for Weak Instruments”. *Journal of Business & Economic Statistics* 31 (3): 358–369.

Muth, R.F. (1969). *Cities and Housing. The Spatial Pattern of Urban Residential Land Use.* The University of Chicago Press.