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

August 2021

Abstract

With "2020 hindsight," the 2000s housing cycle is not a boom-bust but rather a boom-bust-rebound at both the national level and across cities. We argue this pattern reflects a larger role for fundamentally-rooted explanations than previously thought. We construct a city-level long-run fundamental using a spatial equilibrium regression framework in which house prices are determined by local income, amenities, and supply. The fundamental predicts not only 1997-2019 price and rent growth but also the amplitude of the boom-bust-rebound and foreclosures. This evidence motivates our neo-Kindlebergerian model, in which an improvement in fundamentals triggers a boom-bust-rebound. Agents learn about the fundamentals by observing "dividends" but become over-optimistic due to diagnostic expectations. A bust ensues when over-optimistic 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 fundamental shock and accounts quantitatively for cross-city patterns in the dynamics of prices and foreclosures.

*Contact information: chodorowreich@fas.harvard.edu, guren@bu.edu, tmcquade@berkeley.edu. We thank seminar participants at BU, Harvard, Emory/Atlanta Fed, BI Business School, the University of Copenhagen, San Francisco Fed, UCSD, Chicago, and the Boston Fed 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.

1 Introduction

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 the areas experiencing the most rapid price growth during the boom also having the largest price declines during the bust. The sharpness of this cycle and the role it played in triggering the Great Recession have made it among the most consequential and studied asset price cycles in recent history. The 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 and instead due to changes in credit or sun-spot beliefs (Shiller, 2008; Charles et al., 2018; Kaplan et al., 2020).

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 explain cross-city variation in long-run house price growth in reduced-form and structural supply regressions as well as the amplitude of the boom-bust-rebound and severity of the foreclosure crisis. In a second step, we introduce a "neo-Kindlbergererian" model of a fundamentally-rooted house price cycle in which belief over-reaction amplifies

[1] We consider developments up to 2020 due to the COVID-19 Pandemic dramatically altering the economy and housing market.

the boom and a foreclosure spiral exacerbates the bust, and discipline the model with our empirical moments. The estimated model generates the boom-bust-rebound from a single fundamental shock and quantitatively matches the cross-city patterns.

Our analysis starts in Section 2 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 from 1997-2006, bust from 2006-2012, and rebound from 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.

Section 3 introduces a simple spatial equilibrium empirical framework to shed light on whether price growth over the boom-bust-rebound reflects fundamental forces. Building on the work of Saiz (2010), we derive a structurally-interpretable long-run supply regression of house price growth as a function of population growth, the land share of prices, housing supply regulation, and their interactions. In this framework, omitted variables such as cost shocks that increase prices but deter population pose a threat to identification. We follow Saiz (2010), Diamond (2016), and others in selecting excluded instruments, including land unavailability, initial population density, shift-share predictors of local wages and employment, amenities, and taste-based determinants of regulation.

Section 4 contains the empirical analysis of our framework. The excluded instruments, which are all determined prior to the boom, strongly predict house price growth over the boom-bust-rebound as well as the house price determinants of population growth, land share, and regulation. In reduced form, they explain nearly 60% of the cross-city variation in house price growth over 1997-2019. The structural IV relationship also has strong explanatory power and the implied supply elasticities are in line with the existing literature. We show extensive robustness to the measurement of prices and population,

the excluded instrument set, included covariates, and specification.

We define a city's long-run fundamental as the linear combination of excluded instruments that is the second stage fitted value. Areas with higher fundamentals have larger booms, deeper busts, stronger rebounds, and more severe foreclosure crises during the bust. Importantly, our empirical approach in no way requires a correlation between fundamentals and the amplitude of the cycle or the magnitude of foreclosures.

We validate our measurement of fundamentals by examining whether increases in price-rent ratios in the boom forecast subsequent rent growth or price declines, motivated by the user cost approach in urban economics (Poterba, 1984) and the Campbell-Shiller decomposition in asset pricing (Campbell and Shiller, 1988a,b). Unconditionally, areas with higher price-rent growth in the boom have higher subsequent rent growth and relative price declines in comparison to other areas. Strikingly, the part of price-rent growth correlated with the long-run fundamental predicts only subsequent rent growth and no price decline, consistent with our labeling this component a fundamental.

Section 5 contains our neo-Kindlebergerian interpretation of these new empirical facts. Kindleberger (1978) presents a historical narrative of asset cycles that start with an improvement in an economic fundamental, boom as the result of over-optimism and credit expansion, and crash when the optimism corrects and levered investors are forced to sell. We formalize this narrative in a continuous time, dynamic model of a city in spatial equilibrium. House prices clear a Walrasian market, with demand emanating from potential entrants and supply coming from the the construction of new homes and foreclosures. An endogenous boom-bust-rebound cycle occurs in response to a single change in the city's fundamental, 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, both potential entrants

and mortgage lenders learn about the true growth rate from observing the path of dividends, but, in line with survey evidence (Case et al., 2012), are over-optimistic, which we obtain by assuming agents have *diagnostic expectations* (Bordalo et al., 2019). The over-shooting of beliefs generates a boom in construction and prices as beliefs update. Eventually, the over-optimism peaks and beliefs endogenously start to correct and approach the true long-run growth rate, triggering the bust phase of the cycle.

On their own, however, the 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 the empirical literature (Foote et al., 2008; Ganong and Noel, 2020), a foreclosure occurs when an under-water homeowner experiences a liquidity shock. Foreclosures add to housing supply, which further depresses prices, putting more owners under-water, and leading to more foreclosures in a price-foreclosure spiral that pushes prices well below their long-run level (Guren and McQuade, 2020). Finally, foreclosures recede, ongoing dividend growth causes new buyers to enter, and prices rebound toward the new balanced growth path.

The model quantitatively matches the cross-section of boom-bust-rebounds in the data. Exploiting the fact that it nests our empirical regression specification for long-run analysis, we form quartiles of CBSAs grouped by their empirical long-run fundamental. We externally calibrate several parameters, including the long-run supply elasticity estimated in Section 4 and the pre-boom house price growth rate in each quartile. We estimate the remaining parameters — holding deep parameters related to learning and preferences fixed across quartiles — to match the size of the boom, bust, and rebound, the peak foreclosure rate in and speed of the bust, and the mortgage equity distribution in each quartile. Despite being over-identified, the model fits these moments extremely well. We conduct counterfactual exercises that turn off the role of over-optimism or foreclosures and conclude

that both features are required to generate a fundamentally-rooted boom-bust-rebound.

Related literature. Our paper relates to literatures in macroeconomics, housing economics, urban economics, and finance. Most directly, it builds on a literature on the causes of the 2000s boom and bust which has predominantly focused on non-fundamental determinants such as credit and sun-spot expectations and debated how to disentangle the relative importance of these forces.[2]

The credit view of the boom focuses on a credit supply expansion driven by changes in lending technology, securitization, and secular decline in interest rates (Justiniano et al., 2017; Johnson, 2020; Foote et al., 2019). Favilukis et al. (2017), Justiniano et al. (2019), Landvoigt et al. (2015), Greenwald (2018), and Greenwald and Guren (2021) quantify the role of credit using applied macroeconomic models. A related empirical literature shows that credit expansion differentially impacted house prices in cities based on their supply elasticity (Mian and Sufi, 2009; Glaeser et al., 2008), while another strand studies the impact of credit on house prices (Favara and Imbs, 2015; Loutskina and Strahan, 2015; Di Maggio and Kermani, 2017).[3]

In the expectations view, the boom resulted from misplaced expectations about the growth of future fundamentals and the crash occurred when such growth did not materialize (Piazzesi and Schneider, 2009; Burnside et al., 2016; Kaplan et al., 2020; Gennaioli and

[2]Glaeser and Nathanson (2015) survey the literature on housing bubbles. 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." Similarly, Glaeser (2013) points out that the " economy was not growing particularly swiftly, nor was it obvious that there were any tectonic shifts in the geography of American enterprise" and that "there were few obvious changes in economic fundamentals that set off the bust." Sinai (2013) argues that while high-frequency fundamentals are correlated with the recent housing cycle, they cannot explain the magnitude or amplitude of fluctuations. Kaplan et al. (2020) and Greenwald and Guren (2021) calibrate models in which sun-spot expectations fully explain the portion of the boom-bust unexplained by credit.

[3]A closely related literature focuses on speculation and particularly its role in the boom (Chinco and Mayer, 2016; Nathanson and Zwick, 2018; Gao et al., 2020; Bayer et al., 2020; DeFusco et al., 2017; Mian and Sufi, 2017). We discuss how this literature relates to our work in Appendix B.1.

Shleifer, 2018). Case et al. (2012) and Soo (2018) provide measures of over-optimism in the boom, and Kuchler and Zafar (2019) and Armona et al. (2019) analyze the formation of expectations and their impact on house prices. Case et al. (2012) also document a gradual decline in long-run optimism during the bust, a feature our quantitative model will match. Glaeser and Nathanson (2017) microfound extrapolation in the housing market.

Our paper recognizes a role for both credit and expectations in the boom. However, we emphasize fundamentals as triggering the over-optimism of expectations and credit expansion as endogenous to lenders' expectations and important insofar as it allowed buyers to obtain mortgages despite rising house prices. Our work also connects to literatures focusing on the role of tightening financial constraints (Mian and Sufi, 2014; Garriga and Hedlund, 2020) and foreclosures (Guren and McQuade, 2020; Mian et al., 2015; Hedlund, 2016) in the bust, with foreclosures the most important quantitatively in our model. In contrast to the existing literature that requires a second exogenous change in expectations or credit conditions to trigger the bust (Kaplan et al., 2020; Favilukis et al., 2017; Greenwald and Guren, 2021), our model of learning about fundamentals instead generates this turning-point and the transition from bust to rebound endogenously.

A few papers ascribe a role to fundamental factors in the 2000s cycle as we do. Writing near the peak of the boom, Himmelberg et al. (2005) found "little evidence of a housing bubble" because of fundamental growth, undervaluation in the 1990s, and low interest rates. Ferreira and Gyourko (2018) estimate the timing of the boom across cities and show that the beginning of the boom was "fundamentally based to a significant extent" but that fundamentals revert in roughly three years. We similarly conclude that fundamentals played a significant role in the boom, but based on different methods that focus on long-term fundamentals rather than short-term income growth. More recently, Howard and Liebersohn (2021) propose an explanation for housing cycles based on divergence in

regional income growth, in which fluctuations in fundamentals fully explain the cycle, and Schubert (2021) identifies spillovers of fundamentals across cities via migration networks.

Our empirical analysis builds on an urban economics literature on the long-run determinants of housing costs and location choice (Rosen, 1979; Roback, 1982). Our approach most closely follows Saiz (2010), although we extend his framework to incorporate observed land shares, relax constraints, and apply to a different period.[4]

Our theoretical framework builds on work in macroeconomics and finance that focuses on the role of learning and behavioral expectations in financial markets and asset bubbles. Our emphasis on learning follows work in asset pricing pioneered in David (1997) and Veronesi (1999) and in particular builds on the role of diagnostic expectations formalized in Bordalo et al. (2019, 2020). Kindermann et al. (2021) also study learning in housing markets but emphasize heterogeneity by tenure status. Our model contains several features that may prove useful in the study of housing dynamics and are required to make it tractable to estimate, including a new characterization of the impulse response of diagnostic beliefs and a Monte Carlo method for determining mortgage pricing.

Finally, our interpretation of the 2000s housing cycle echoes research on asset bubbles and crashes beginning with the seminal work of Kindleberger (1978). For instance, Kindleberger writes that "virtually every economic mania is associated with a robust economic expansion" and that "during...economic expansions investors become increasingly optimistic...while lenders become less risk-averse. Rational exuberance morphs into irrational exuberance." As in Kindleberger, our focus on long-run fundamentals does not imply that there was no "housing bubble," nor does it refute any role for credit or expectations in the

[4]Related work focuses on estimating the role of housing supply constraints and supply elasticities in determining house prices (Glaeser and Gyourko, 2018; Mayer and Somerville, 2000; Glaeser et al., 2005a,b; Glaeser and Gyourko, 2005; Gyourko et al., 2013; Gyourko and Molloy, 2015; Baum-Snow and Han, 2020) and on the construction cost and land share of house prices (Davis and Heathcote, 2007; Davis and Palumbo, 2008; Gyourko and Saiz, 2006; Larson et al., 2021; Albouy et al., 2018).

boom or expectations, credit, and foreclosures in the bust. It does, however, suggest that these forces were rooted in long-run fundamentals. Our work differs from Kindleberger in emphasizing the rebound and its role in diagnosing the driving forces of the full cycle.

2 Boom, Bust, and Rebound

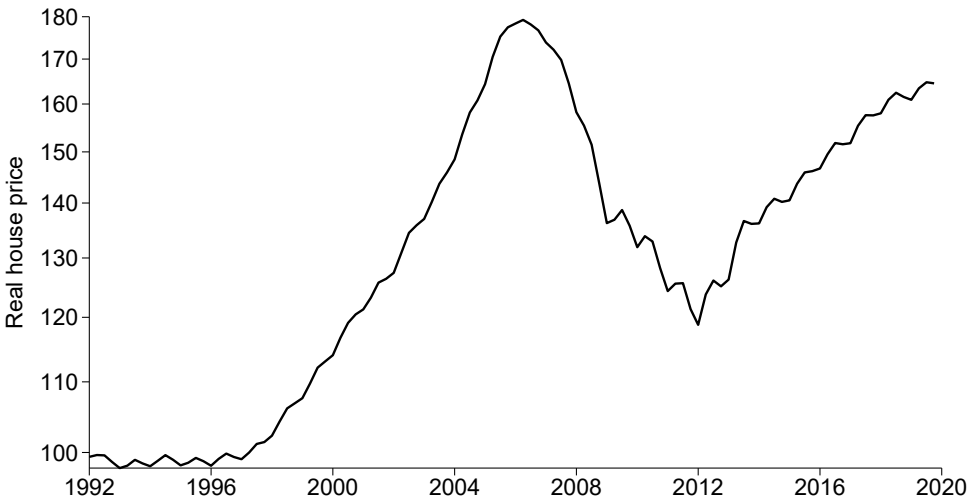
In this section we document the boom, bust, and rebound phases of the 2000s housing cycle in both national and local data.

Figure 1 shows the national Case-Shiller house price index, deflated using the GDP price index. After a period of zero real growth, 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. In what follows, we measure the boom as house price growth between 1997 and 2006, the bust as price growth between 2006 and 2012, and the rebound as price growth between 2012 and 2019. We start the boom in 1997 because very few cities have booms that start before that year.[5]

Figure 2 shows the correlation of the boom, bust, and rebound at the local level, using ZIP Code house prices from FHFA. Each blue circle represents one ZIP Code. The overlaid red circles show the mean value of the y-axis variable in each of 50 quantiles of the x-axis variable (the so-called binned scatter plot). 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

[5]Appendix Figure B.3 displays a histogram across cities in our sample of the first break in the house price growth rate series between 1992:Q1 and 2006:Q2, based on the structural break test of Andrews (1993) and Bai and Perron (1998, 2003) as implemented in Ditzen et al. (2021). While almost no areas have booms that start before 1997, about 11% of cities (23% weighted by population) have breaks sometime in 1997. Ferreira and Gyourko (2018) implement a similar procedure at the local level and also find fairly wide variation in the timing of the boom start, but with essentially no areas with booms that start before 1997. Applying the break test to the national time series identifies the national boom start as 1998Q2. We measure the city-level boom from 1997 to capture the full price growth in all cities. Appendix Figure B.3 shows much more uniformity in the timing of the price peak and trough across cities.

Figure 1: National Boom, Bust, and Rebound



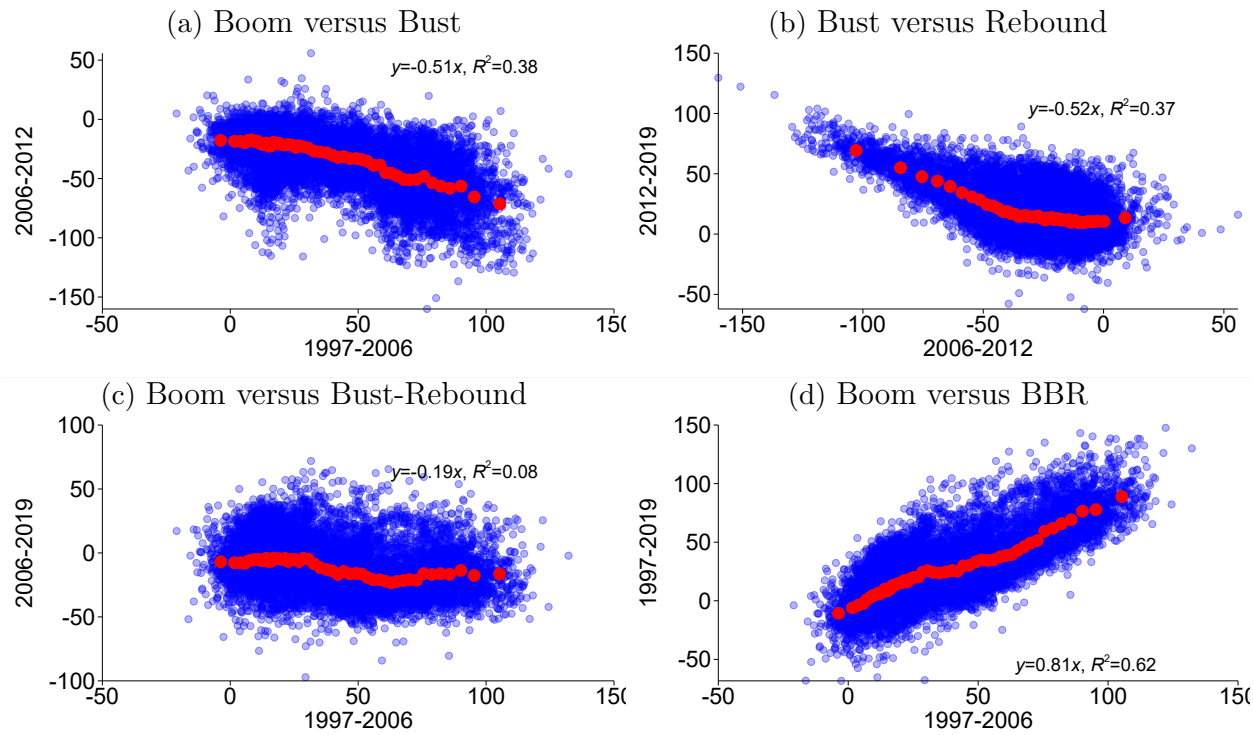
Notes: The figure shows the national Case-Shiller index deflated by the GDP price index, normalized so that 1997=100.

and the R^2 of this relationship is 0.38. Mayer (2011) refers to this boom-bust cycle at the local level as characteristic of a housing bubble.

Panel (b) reveals an equally strong correlation between the magnitude of 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. Putting the bust and rebound together in Panel (c), house price growth in the boom is nearly uncorrelated with total price growth after 2006. Panel (d) displays the corrollary 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.

Although suggestive, examining house prices in isolation cannot answer whether growth over the full BBR period reflects fundamentals. We thus turn next to a more rigorous econometric investigation of the role of fundamentals.

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.

3 Role of Fundamentals: Framework

In this section we introduce a long-run supply-and-demand framework for house prices and describe our data and measurement. The payoff to a structural regression model is that we can associate each city with a housing fundamental and thereby relate the boom-bust-rebound pattern to fundamental determinants of house prices.

3.1 Structural System

The supply framework 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 construction cost of the structure, and $L_{i,t}$ is the land cost. Equation (1) is a housing supply equation that equates the price of a home to the cost of the land and structure. Gyourko and Saiz (2006) argue that the construction sector is sufficiently competitive to justify this assumption, at least in the long run that we consider in this and the next section. The equation also assumes away other costs such as broker fees that are proportional to the cost of a new home and drop out when we take log changes below.

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

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

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

Here, $H_{i,t}$ is the total population in the city (H for houses), α_i and β_i are city-specific elasticities, and $A_{i,t}$ and $B_{i,t}$ denote city-specific cost shifters. These isolastic cost functions correspond to the long-run costs in the dynamic model in Section 5, as they omit any short-run adjustment costs stemming from accelerating or decelerating population growth that occur off the balanced growth path. 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 the Alonso (1964), Muth (1969), and Mills (1967) 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) must rise to induce new housing in less desirable locations.

Taking logs of equation (1), differencing over time, and letting $s_{i,t} = L_{i,t}/P_{i,t}$ denote the land share of prices and lower case p, c, ℓ, a, b denote the log differences of their respective

upper case letters, we obtain:

$$p_{i,t} = (1 - s_{i,t})\, c_{i,t} + s_{i,t}\ell_{i,t} \tag{4}$$

$$= a_{i,t} + s_{i,t}\,(b_{i,t} - a_{i,t}) + (\alpha_i + s_{i,t}\,(\beta_i - \alpha_i))\, h_{i,t}. \tag{5}$$

Equation (4) decomposes house price growth into a weighted average of the growth of construction costs and land costs and appears for example in Davis and Palumbo (2008). Equation (5) imposes the functional forms of construction and land costs.

Equation (5) has too many free parameters to estimate. We 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 costs and land costs to increase more steeply with the number of houses in places with stricter land-use regulations. We also demean $a_{i,t}$ and $b_{i,t}$ with respect to the cross-city average and let $\bar{x}_t = E_i[x_{i,t}]$ for any variable x, $\hat{x}_t \equiv x_t - \bar{x}_t$, and $\epsilon_{i,t} = \hat{a}_{i,t} + s_{i,t}\left(\hat{b}_{i,t} - \hat{a}_{i,t}\right)$ to obtain:

$$\begin{aligned} p_{i,t} = \bar{a}_t + s_{i,t}\left(\bar{b}_t - \bar{a}_t\right) + \alpha_0 h_{i,t} + (\beta_0 - \alpha_0)\, s_{i,t} h_{i,t} \\ + \alpha_1 m_i h_{i,t} + (\beta_1 - \alpha_1)\, m_i s_{i,t} h_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{6}$$

We can identify the parameters in equation (6) using the regression equation:

$$p_{i,t} = c_0 + c_1 s_i + c_2 h_{i,t} + c_3\,(s_i \times h_{i,t}) + c_4\,(m_i \times h_{i,t}) + c_5\,(m_i \times s_i \times h_{i,t}) + e_{i,t}, \tag{7}$$

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

Equation (7) is a long-run supply equation. It generalizes earlier work such as Saiz (2010) by treating the land share as observable and by allowing for construction costs to respond endogenously to population.[6] The coefficient c_1 identifies the average excess

[6]The Saiz (2010) model starts with equations (1) to (3) with $\alpha_i = 0\ \forall i$. In our notation, the final

secular (i.e. not driven by population growth) increase in land prices over construction costs. The coefficient c_2 identifies the average long-run elasticity of construction costs to population. The coefficient c_3 identifies the difference in the average long-run elasticities of land costs and construction costs to population growth. The coefficient c_4 identifies the marginal increase in the construction cost elasticity from higher regulatory strictness. The coefficient c_5 identifies the difference in the marginal increases in the land cost and construction cost elasticities. Finally, the supply inverse elasticity is $c_2 + c_3 \times s_i + c_4 \times m_i + c_5 \times s_i \times m_i$.

As in Saiz (2010), the endogeneity of location choice (which impacts population growth), land share, and regulatory strictness all pose a hurdle to consistent estimation of the parameters. More specifically, the structural residual $\epsilon_{i,t}$ contains city-specific, secular changes in land costs or construction costs. Such secular cost shocks increase house prices but also affect land share, possibly induce new zoning regulation, and (in spatial equilibrium) deter population growth. We therefore estimate equation (7) using instrumental variables, as described next.

3.2 Data, Measurement, and Excluded Instruments

We estimate equation (7) across 308 metropolitan core-based statistical areas (CBSAs) over the period 1997-2019.[7]

Outcome and endogenous variables. We measure CBSA house prices using the Freddie Mac house price indexes (HPIs) deflated by the national GDP price index and show

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 follows the motivation in Saiz (2010, p. 1264) in that we will treat Λ_i and $\ln H_{i,0}$ as excluded instruments that help to identify the land share terms in the supply elasticity.

[7] CBSAs consist of a set of adjacent counties with economic and commuting ties and a population of at least 50,000 and are the smallest geographic unit at which we can measure all of the variables in our analysis. Appendix Figure B.4 replicates at the CBSA level the boom-bust-rebound correlations in Figure 2.

robustness to using other HPIs below. We measure population growth by aggregating to the CBSA level the Census intercensal county population estimates. We obtain data on CBSA land share starting in 2012 from Larson et al. (2021). They use appraisal data that separate house prices into land costs and construction costs, exactly as in equation (1). We use the 2012 value in our regressions, making it important to treat land share as endogenous rather than pre-determined at the start of the boom. We equate regulatory strictness with the 2006 Wharton Residential Land Use Regulatory Index (WRLURI) developed in Gyourko et al. (2008) and again recognize the endogeneity of the 2006 values by treating the WRLURI as potentially endogenous to house price growth. The WRLURI is based on 15 questions on land use regulations at the town level and standardized to have zero mean and unit variance; we map towns to 2018 CBSA delineations, average across towns within a CBSA using population weights, and re-standardize to create a CBSA-level variable. Table B.2 presents summary statistics.

Excluded instruments. Equation (7) contains three endogenous variables — population growth, land share, and regulatory strictness — and their interactions. We heuristically group excluded instruments by which endogenous variable we expect them to predict and then interact these groups as well to produce our final excluded instrument set.

We motivate excluded instruments for population growth by appealing to the Rosen (1979) and Roback (1982) model of spatial equilibrium in population. In that framework, households choose location to maximize their earning potential and amenity value net of housing costs. Therefore, labor demand and amenities constitute valid demand shifters to identify the long-run supply elasticity. We follow Saiz (2010) and use shift-share instruments for labor demand and climate variables for amenities. Specifically, we construct shift-shares for the growth of employment and the average wage at the NAICS 4 digit

level by aggregating data from the County Business Patterns (CBP) up to the CBSA level as detailed in Appendix B.2. We obtain data on mean January temperature, mean January sunlight, and mean July humidity from the Department of Agriculture. These variables capture population movement toward areas with warmer January temperatures, more winter sunlight, and less July humidity. We also use two amenity variables proposed by Diamond (2016), the share of the population with a Bachelor's degree, measured as of 1990 using Census data, and the density of restaurants, measured as of 1997 using CBP restaurant employment relative to the population.

We also follow Saiz (2010) in choosing instruments relevant for land share and regulatory strictness. For land share, the Alonso (1964), Muth (1969), and Mills (1967) intra-city spatial equilibrium microfoundation of equation (3) predicts that the average price of land increases in the unbuildable share of an area (see equation (A.8) in Appendix A.2). We use as an instrument the share of land available for development (not water or a slope steeper than 15%) in a polygon containing the CBSA plus a 5% buffer (Lutz and Sand, 2019). Lutz and Sand argue that this measure improves on the land unavailability in Saiz (2010), although we confirm robustness to using Saiz's original measure. We also use as an instrument population density in 1997, measured as the ratio of population to land area. For the WRLURI index of regulatory strictness, we use as instruments the ratio of public expenditure on protective inspection to total tax revenue as measured in the 1992 Census of Governments and the share of Christians in nontraditional denominations as measured in the 1990 Census. Saiz (2010) motivates protective expenditure as revealing an area's overall taste for regulation and the non-traditional Christian share because the ethos of individualism in these denominations leads to reduced regulation.

It merits remarking that many of these instruments are either quite persistent or time invariant. Indeed, we purposely follow Saiz (2010) in choosing many of the excluded in-

struments, despite his using them to explain population and house price growth over 1970-2000, almost entirely previous to our period of analysis. For now, we note that nothing in the econometric setup precludes persistent instruments; we only require instruments that shift population, land share, and regulatory strictness and are orthogonal to unobserved, location-specific supply shifters during our period of study. Section 4.3 reports robustness specifications that explicitly control for lagged population or house price growth and yield very similar results. Finally, our calibrated model in Section 5 will appropriately account for differences in pre-boom house price growth rates across CBSAs.

4 Role of Fundamentals: Results

This section contains our main empirical results. Section 4.1 presents first-stage and reduced form regressions that show that the excluded instruments strongly predict both fundamental drivers of house prices and long-run house price growth over 1997-2019. Section 4.2 reports the IV results. Section 4.3 conducts robustness exercises. Section 4.4 defines the long-run fundamental as a linear combination of the excluded instruments and shows that a higher fundamental predicts a larger boom-bust-rebound in an area. Section 4.5 relates the long-run fundamental to growth in rents and shows that the fundamental component of rising price-to-rent ratios in the boom predicts only subsequent rent growth. Section 4.6 shows that areas with higher long-run fundamentals had more severe foreclosure crises during the bust.

4.1 First-Stage and Reduced-Form Results

This section shows that the excluded instruments have strong explanatory power for the endogenous variables and for long-run house price growth. Table 1 reports first-stage-type regressions for each of the endogenous variables separately. For each variable, we show

the explanatory power using only the excluded instruments motivated by that variable and also using the full set of uninteracted instruments. Let $\mathcal{H}, \mathcal{L}$, and $\mathcal{M}$ denote the sets of instruments heuristically assigned to population growth, land share, and WRLURI, respectively. The actual IV will also include $\mathcal{H} \times \mathcal{L}$, $\mathcal{H} \times \mathcal{M}$, and $\mathcal{H} \times \mathcal{L} \times \mathcal{M}$, where $\times$ denotes element-wise cross-set multiplication. In that sense, Table 1 contains regressions useful for establishing the explanatory power of the instruments without broaching many-instrument asymptotics, a subject we address in the robustness section.

Column (1) shows that more land unavailability and higher initial population density both predict higher land share, with an R^2 of 0.37 and joint effective F-statistic of 64.5.[8] Column (2) shows that their explanatory power persists after adding other excluded instruments. Columns (3) and (4) show predictors of population. Climate amenities — higher January temperature, higher January sunlight, and lower July humidity — all predict higher population growth, as do greater restaurant density and a higher college share of the population. Bartik-predicted employment and wage growth enter somewhat noisily, although in unreported results these variables have stronger predictive power in a specification without the amenity variables. Columns (5) and (6) show predictors of WRLURI. Consistent with the results in Saiz (2010), a higher share of Christians in nontraditional denominations negatively predicts regulation while a higher ratio of public expenditure on protective inspection to total tax revenue positively predicts regulation.

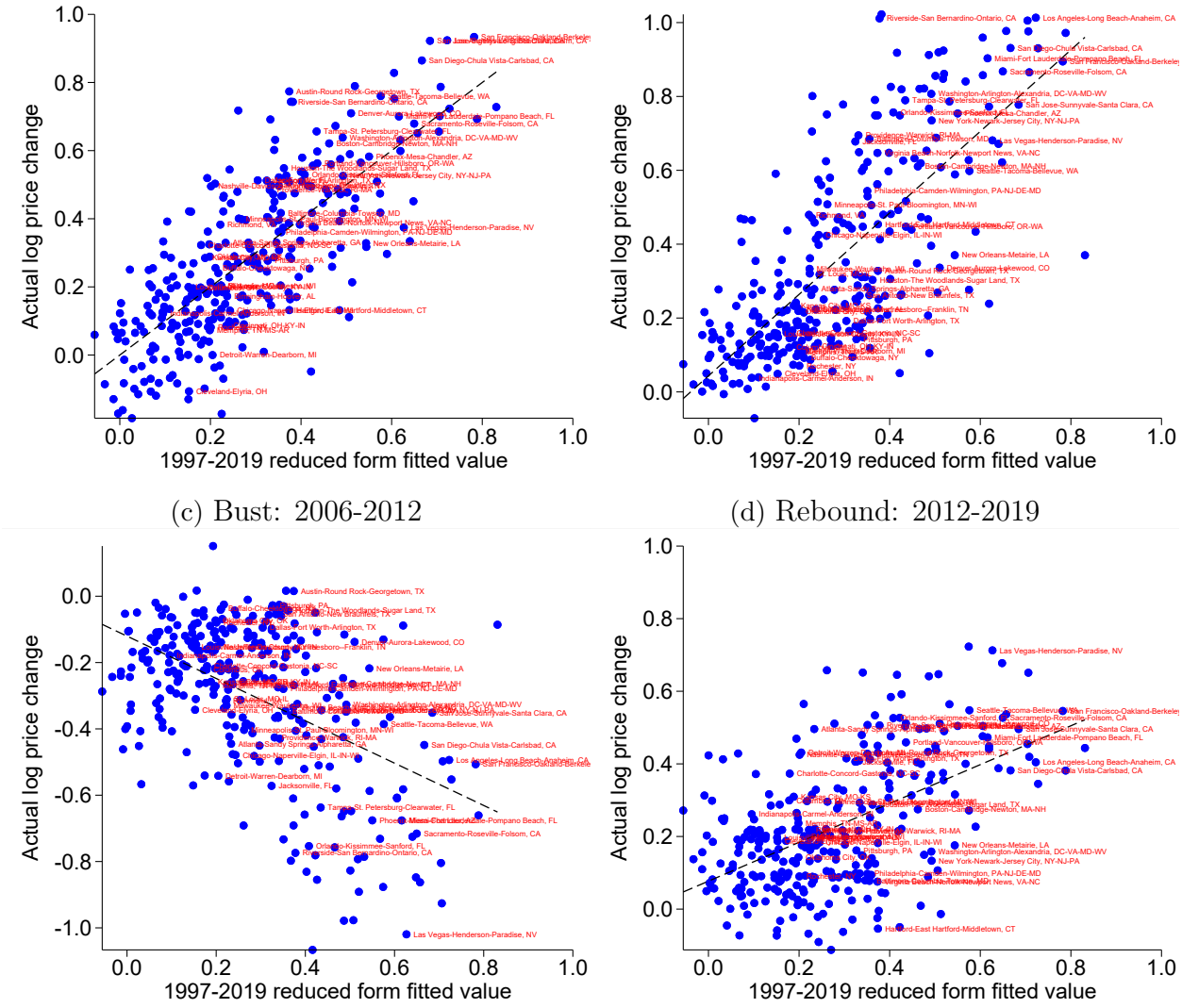
The final column of Table 1 reports the reduced form for long-run house price growth, using all of the uninteracted instruments. The instruments jointly explain 59% of the variation in house price growth. This column illustrates that fundamental drivers of location choice, land share, and regulation, all measured prior to the start of the boom, explain a substantial amount of the variation in house price growth over the entire BBR.

[8]Montiel Olea and Pflueger (2013) introduce the effective F-statistic as the proper metric of first stage

Figure 3: Correlation with Long-run Fitted Value

(a) BBR: 1997-2019

(b) Boom: 1997-2006



(c) Bust: 2006-2012

(d) 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 based on column (7) of Table 1. CBSAs with more than 1 million persons in 1997 are labeled in red.

Table 1: Pseudo First-stage and Reduced Form Regressions

Dep. var.:	Land share		Pop. growth		WRULRI		HPI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unavailability	3.03**	2.83**		−0.75		21.61**	10.07**
	(0.49)	(0.50)		(0.99)		(4.50)	(1.08)
Population density	4.81**	4.18**		−2.33**		15.91**	2.17+
	(0.51)	(0.47)		(0.61)		(5.52)	(1.13)
Bartik wage		−0.20	0.92	0.65		−0.54	1.38
		(0.46)	(0.93)	(0.92)		(4.52)	(1.00)
Bartik emp.		−0.63	−0.46	0.74		5.41	1.53
		(0.47)	(0.89)	(0.94)		(4.42)	(0.96)
January temp.		0.81	6.82**	5.53**		0.56	4.06**
		(0.59)	(0.99)	(1.54)		(6.17)	(1.55)
January sunlight		0.25	2.48*	2.64*		9.08+	2.93*
		(0.46)	(1.07)	(1.11)		(4.88)	(1.18)
July humidity		−0.23	−5.97**	−4.57**		−14.38**	−6.76**
		(0.44)	(0.95)	(1.06)		(4.84)	(1.22)
Restaurants		1.19**	3.47**	2.18+		−7.77	−0.23
		(0.42)	(1.10)	(1.24)		(4.94)	(0.98)
College share		2.51**	2.31*	3.21**		15.55**	6.25**
		(0.51)	(1.00)	(1.05)		(5.63)	(1.18)
Nontrad. Christian		−1.15*		2.25+	−26.70**	−23.11**	−2.75**
		(0.56)		(1.19)	(4.34)	(4.99)	(1.05)
Inspection/tax		0.74		−0.24	21.87**	7.12*	3.41**
		(0.50)		(0.94)	(3.43)	(3.54)	(1.02)
Effective F	64.5	24.5	19.9	14.9	39.3	14.3	
R^2	0.366	0.475	0.344	0.377	0.179	0.361	0.586
Observations	308	308	308	308	308	308	308

Notes: Columns (1), (3), and (5) report regressions of an endogenous variable on the group of excluded instruments associated with that variable. Columns (2), (4), and (6) report regressions of an endogenous variable on all excluded instrument main effects. Column (7) reports the reduced form regression of house price growth on all excluded instrument main effects. Heteroskedastic-robust standard errors in parentheses. **,* ,+ denote significance at the 1, 5, and 10 percent levels, respectively.

Figure 3 plots the fitted values from the reduced form regression in column 7 of Table 1 against actual house price growth in various sub-periods. Panel (a) shows a strong correlation with price growth over the full BBR, consistent with the high R^2 in column (7). The figure labels in red CBSAs with more than 1 million persons in 1997; these larger CBSAs

strength with non-iid standard errors. See Andrews et al. (2019) for further discussion.

have a reduced form fit similar to the full sample. Panels (b)-(d) show the correlation with each sub-period. Higher predicted long-run growth correlates positively with higher growth during the boom, negatively with growth during the bust, and positively with growth during the rebound. Thus, the reduced-form evidence is consistent with long-run fundamental growth producing a boom-bust-rebound cycle.

4.2 Structural Results

We now present the full IV results. These impose additional structure beyond the reduced form by forcing the instruments to act through the endogenous variables in the model. They also yield a structurally interpretable long-run housing supply elasticity that we will use to calibrate the model in Section 5.

Table 2 presents the results from estimating equation (7). Column (1) shows OLS. CBSAs with higher land share and faster population growth have 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 0.58 with a standard error of 0.06 using the delta method.

Column (2) reports the IV specification using all of the excluded instruments shown in Table 1 as well as the interactions of each of the instrument groups. Several features merit comment. The impact of population growth and the average inverse elasticity are slightly larger in the IV specification, consistent with the expected bias of OLS due to area-specific cost shifters. The coefficient on the main effect on land share maps to the average excess secular (i.e. not driven by population growth) increase in land prices over construction costs, which causes house prices to rise faster in areas where land is a larger share of the overall price. The value of 107 log points over 1997-2019 reflects nationwide forces such as a secular decline in interest rates and an increase in the premia to living in the more

Table 2: Main OLS and IV Results

Dep. var.:	Log House Price Growth 1997-2019		
	(1)	(2)	(3)
Land Share	0.87**	1.07**	1.06**
	(0.23)	(0.25)	(0.17)
Pop. Growth 1997-2019	0.30	0.71+	0.63**
	(0.29)	(0.41)	(0.08)
Land Share × Pop. Growth	1.02	−0.26	
	(1.08)	(1.41)	
WRLURI × Pop. Growth	−0.23	−0.22	
	(0.16)	(0.35)	
Land Share × WRLURI × Pop. Growth	1.61**	2.63*	1.88**
	(0.57)	(1.25)	(0.27)
Constant	−0.09	−0.16*	−0.15**
	(0.06)	(0.07)	(0.05)
Estimator	OLS	2SLS	2SLS
Elasticity at $\bar{s}_j$	0.58	0.63	0.63
Standard error of elasticity	0.06	0.08	0.08
R^2	0.48	0.40	0.40
Observations	308	308	308

The table reports OLS (column 1) and IV (columns 2 and 3) regressions of real CBSA house price growth over 1997-2019 on land share, population growth over 1997-2012, WRLURI, and interactions of these variables. 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.

expensive city center, the latter for example due to the widespread fall in crime rates in the mid-1990s (Pope and Pope, 2012).[9] Two coefficients, corresponding to the bivariate interactions of land share × population growth and WRLURI × population growth, have values close to zero and statistically insignificant.[10]

[9]Equation (A.9) illustrates how a reduction in interest rates, which increases the capitalization rate, or increase in the city center premium, which increases the population-weighted land price, raises average land prices in a city. Nichols (2019) uses prices of raw land transactions to also find a rising price of land over this period, albeit a smaller increase than implied by Table 2. If the 2012 land share in each CBSA is a fraction less than one of the 1997 land share due to over-shooting during the bust, then the land share coefficient in Table 2 will overstate the contribution of rising land prices commensurately. Note however that the second stage fitted value and elasticity would not change.

[10]The supply framework interprets zero coefficients on these variables as $\beta_0 = \alpha_0$ and $\alpha_1 = 0$ in equation (6), that is, the intercepts of the elasticities of land and construction are the same and the elasticity of construction costs does not vary with land-use regulation.

Column (3) is our preferred specification. Relative to column (2), it omits the land share × population growth and WRLURI × population growth variables. The remaining coefficients remain relatively unchanged, but with much smaller standard errors. Perhaps not surprisingly given the large number of interaction terms and instruments, the data appear unable to tightly identify each interaction in column (2). Imposing zero restrictions alleviates this difficulty. Importantly, the overall fit as measured by the IV R^2 and the inverse supply elasticity both remain unchanged between columns (2) and (3), indicating that both specifications fit the data equally well. The coefficient on the surviving interaction term land share × WRLURI × population growth indicates a larger inverse elasticity (price growth more sensitive to population) in areas with both high land share and high regulation. The R^2 value of 0.40 reveals strong explanatory power of land share, population growth, and WRLURI when imposing the IV coefficients. Thus, this column again illustrates the central result that fundamentally-driven population growth, land share, and heterogeneous long-run supply elasticities explain a substantial amount of the variation in house price growth over the entire BBR.

4.3 Robustness

Table 3 collects several alternative specifications that address various potential concerns with the baseline IV regression. Each row reports the coefficients and standard errors from a separate specification. For convenience, the first row reproduces the baseline coefficients from column (3) of Table 2.

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.[11] Row (5) replaces the land unavailability instrument with the measure

[11]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

Table 3: IV Robustness

Specification	s		h		$s \times WRLURI \times h$		Obs.
	Coef.	SE	Coef.	SE	Coef.	SE	
1. Baseline	**1.06**	0.17	**0.63**	0.08	**1.88**	0.27	308
2. FHFA HPI	**0.94**	0.17	**0.55**	0.08	**1.84**	0.26	306
3. CoreLogic HPI	**0.99**	0.18	**0.42**	0.09	**1.90**	0.32	308
4. Zillow HPI	**1.13**	0.18	**0.43**	0.10	**1.42**	0.26	225
5. Saiz unavail.	**1.13**	0.17	**0.71**	0.08	**1.57**	0.27	260
6. Housing units	**1.01**	0.19	**0.65**	0.10	**1.46**	0.31	308
7. Pop. weighted	**1.67**	0.22	**0.78**	0.11	**1.23**	0.35	308
8. Drop pop.$< 150K$	**1.01**	0.19	**0.64**	0.09	**1.60**	0.24	219
9. Drop shrinking	**1.13**	0.18	**0.54**	0.09	**1.75**	0.26	277
10. GMM	**1.10**	0.12	**0.66**	0.06	**1.74**	0.21	308
11. Bias-adjusted 2SLS	**0.88**	0.26	**0.62**	0.12	**3.05**	0.69	308
12. JIVE	**0.99**	0.33	**0.66**	0.15	**2.78**	0.82	308
13. No climate instr.	**0.96**	0.19	**0.67**	0.11	**2.34**	0.39	308
14. No lifestyle instr.	**1.15**	0.19	**0.67**	0.10	**1.98**	0.33	308
15. No Bartik instr.	**1.08**	0.19	**0.63**	0.08	**2.01**	0.33	308
16. No land unavail. instr.	**0.95**	0.19	**0.76**	0.09	**1.92**	0.31	308
17. No pop. density instr.	**1.12**	0.20	**0.58**	0.09	**2.15**	0.34	308
18. Control lag pop.	**1.02**	0.17	**0.47**	0.16	**1.88**	0.29	308
19. Control lag HPI	**1.01**	0.18	**0.62**	0.08	**1.92**	0.27	308

Notes: Each row reports coefficients and standard errors from a separate modification of the specification in column (3) of Table 2. Coefficients in bold font are statistically different from 0 at the 5% level.

from Saiz (2010). Because Saiz (2010) developed his measure for 1999 MSA definitions, we lose 16% of the sample, but the coefficients change little. Row (6) replaces population growth with the growth of housing units. In our model these variables will be equivalent; in the data the growth rates are highly correlated such that the results change little.

Rows (7) to (9) explore robustness to the sample, in row (7) by weighting the regression 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 coefficients remain stable,

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.

with the largest difference being that the weighted specification has a higher loading on the main effect on land share and a smaller loading on the triple interaction term.[12]

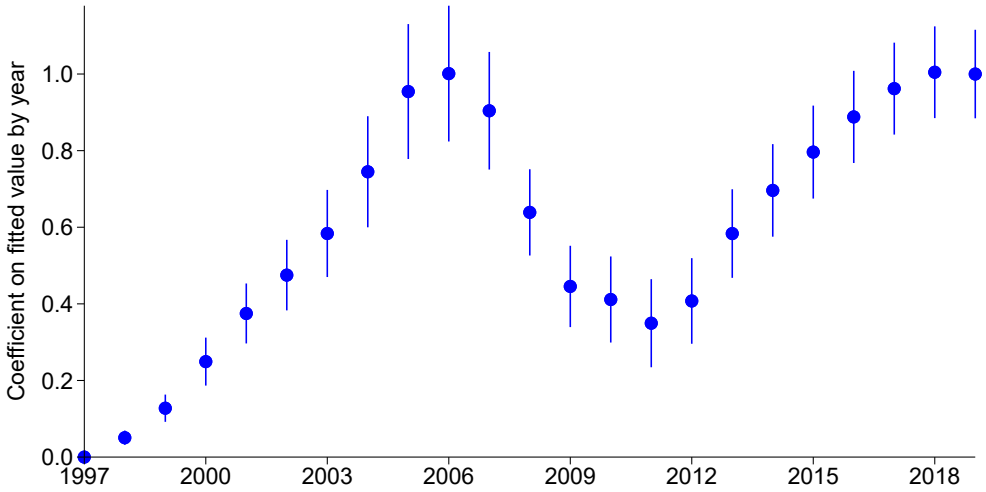
Rows (10) to (12) explore robustness to the estimator, in row (10) by replacing two-stage least squares with GMM, in row (11) with the JIVE estimator of Angrist et al. (1999), and in row (12) with the biased-adjusted estimator of Donald and Newey (2001). The JIVE and bias-adjusted estimators address a particular concern of bias due to the many instruments arising from the interactions of the population growth, land share, and WRLURI instrument sets.[13] While these estimators produce somewhat larger standard errors, the point estimates change little and remain highly statistically significant.

Rows (13)-(17) selectively exclude groups of excluded instruments. The estimation does not critically depend on any particular set of instruments for population or land, with similar results omitting the climate variables (row 13), college share and restaurants (row 14), shift-shares (row 15), land unavailability (row 16), and population density (row 17). Finally, rows (18) and (19) show that despite any persistence in the excluded instruments and endogenous variables, the results change little after controlling for the growth of population or house prices over 1987-1997.

[12]This change largely reflects the influence of the two largest CBSAs, New York and Los Angeles, both of which have high land shares and experienced high price growth despite below-median population growth. A population-weighted regression excluding the New York and Los Angeles CBSAs yields a land share coefficient of 1.23.

[13]As is well known (Bekker, 1994; Bound et al., 1995), with many weak instruments 2SLS will over-fit the first stage, biasing the second stage estimates toward OLS. JIVE avoids the overfitting problem 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. One reason these alternative estimators produce results similar to 2SLS is that unlike in the canonical many weak instrument case of Angrist and Krueger (1991), Table 1 shows that the instruments are generally strong predictors of the endogenous variables. Also different from the returns to education setting of Angrist and Krueger (1991), here economic theory suggests that OLS is biased toward zero. We do not report LIML results in the table because of the difficulty of interpretation when there is treatment heterogeneity (Kolesár, 2013), but find a somewhat smaller land share coefficient and larger interaction term in that case.

Figure 4: 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 2.

4.4 Long-Run Fundamental

We define the long-run fundamental as the second stage fitted value corresponding to column (3) of Table 2. Associating the fundamental with the second stage fitted value (i.e. using the fitted values of the endogenous variables from the first stage) has the attractive property of making it a linear combination of the excluded instruments, none of which depends on local characteristics that evolve after the start of the boom.

Figure 4 shows that areas with higher long-run fundamentals have more pronounced booms, busts, and rebounds. Each point corresponds to the coefficient $\{\beta_{1,h}\}$ from a cross-sectional regression of log real price growth between 1997 and 1997+h on the long-run fundamental. The boom-bust-rebound cycle in these coefficients accentuates our emphasis on a single, fundamental, driver of the whole cycle.[14]

[14]One would obtain very similar results defining the fundamental as the IV fitted value when estimating equation (7), but the fundamental would then be based on some variables realized over the 1997-2019 period rather than only pre-determined variables.

4.5 Fundamentals, Rents, and Price-to-Rent Ratio

The behavior of rents relative to prices can help to validate our measurement of fundamentals. Indeed, the rise in the ratio of house prices to rents was perhaps the feature of the boom most responsible for real-time diagnoses of bubble-like characteristics (see e.g. Shiller, 2008). Here, we revisit the increase in price-rent ratios with the benefit of 2020 hindsight and show that the component associated with long-run fundamentals predicts subsequent rent growth and not future price decline, consistent with our interpretation of this component as fundamentally-based.

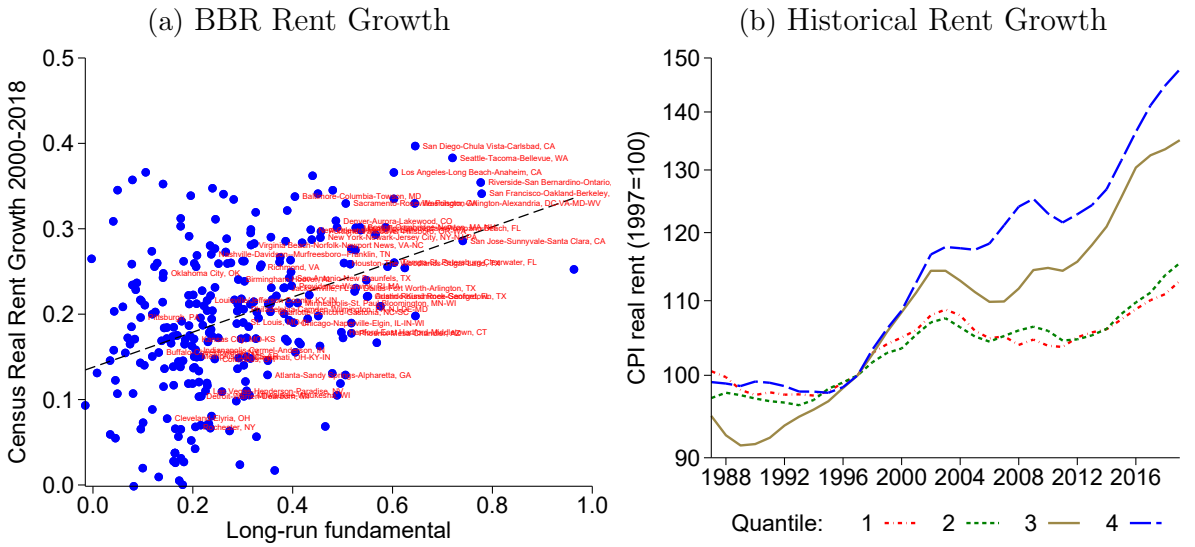
We start by characterizing the behavior of rents. Panel (a) of Figure 5 plots the growth rate of real rents in each CBSA between the 2000 Census and the 2018 American Community Survey (ACS) against the long-run fundamental, again measured as the second stage fitted value from column (3) of Table 2. Areas with higher fundamentals experienced larger rent growth over the BBR, with the relationship especially strong for larger CBSAs.

Panel (b) shows the timing of rent growth using BLS CPI rent data for the 22 CBSAs with annual data since 1987, grouped into population-weighted quartiles of the long-run fundamental. Rents rise fastest in areas with the highest fundamentals. Moreover, this growth appears to represent a break from the pre-boom period, consistent with a change in the income or amenity part of the fundamental at the start of the boom. This change will be the driving force in the neo-Kindlebergerian model.

Column (1) of Table 4 shows that CBSA-level price-rent increases in the boom correlate positively with the long-run fundamental. We measure the log growth in the price-rent ratio using 2000 Census and 2006 ACS mean rent and the same house price data as above. The bivariate relationship has an R^2 of 0.29.

Columns (2) and (5) report the correlation of price-rent growth in the boom with subsequent rent and price growth over the 2006-18 period. This user-cost decomposition

Figure 5: Rent Growth and Long-run Fundamental



Notes: The left panel plots the log change in rent in each CBSA as measured from the 2000 Census to the 2018 ACS, deflated by the national GDP price index, against the second stage fitted value from column (3) of Table 2, with CBSAs with more than 1 million persons in 1997 labeled in red. The right panel plots an index of annual average rent, deflated by the national GDP price index, for 22 CBSAs with CPI rent data available from 1987, grouped into quartiles based on the second stage fitted value from column (3) of Table 2, with the quartiles defined using the full sample.

(Poterba, 1984) closely resembles the Campbell and Shiller (1988b) exercise of decomposing variation in the price-dividend ratio of a stock into future dividend growth and returns, with rents replacing dividends as the cash-flow measure.[15] Larger price-rent growth during the boom forecasts both higher subsequent rent growth and future relative price declines.

Columns (3) and (6) restrict the variation in the price-rent ratio in the boom to the part associated with the long-run fundamental. Specifically, these columns report regressions of subsequent rent and price growth on the fitted value of the growth of the price-rent ratio from column (1). Strikingly, the rise in price-rent ratios associated with long-run fundamental growth predicts even faster subsequent rent growth than in column (2) and no subsequent price decline, validating our labeling of this component as a fundamental.

Columns (4) and (7) show that the part of price-rent growth not explained by long-

[15]This analogy is also recognized in Gallin (2008) and Campbell et al. (2009).

Table 4: What Does Price-Rent in the Boom Forecast?

Dep. var.:	Price-Rent 2000-06	Rent 2006-2018			HPI 2006-2018		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Long-run fundamental	0.68**						
	(0.08)						
Price-rent 2000-2006		0.07**			−0.40**		
		(0.02)			(0.05)		
Predicted price-rent			0.20**			0.03	
			(0.06)			(0.13)	
Residual price-rent				0.01			−0.58**
				(0.03)			(0.05)
R^2	0.29	0.03	0.06	0.00	0.25	0.00	0.36
Observations	272	272	272	272	272	272	272

Notes: Column (1) regresses the log change in the ratio of prices (Freddie Mac) to mean rents (Census, ACS) between 2000 and 2006 on the fitted value from column (3) of Table 2. Column (2) regresses rent growth between 2006 and 2018 (ACS) on the log change in the price-rent ratio between 2000 and 2006. Columns (3) and (4) regress rent growth between 2006 and 2018 on the fitted value and residual from column (1), respectively. Columns (5)-(7) mirror columns (2)-(4) but replacing the dependent variable with house price growth between 2006 and 2018. Heteroskedastic-robust standard errors in parentheses. In columns (3), (4), (5), (6), standard errors computed by re-estimating column (3) of Table 2 and column (1) in 250 bootstrap samples. ** denotes significance at the 1 percent levels.

run fundamentals predicts no subsequent rent growth and large subsequent price declines. These columns make clear that our empirical evidence admits the possibility of aspects of the housing boom not associated with long-run fundamentals; in fact, the part of price-rent increases not correlated with long-run fundamentals looks very bubble-like *ex post*. Candidate forces behind this component include speculation (Chinco and Mayer, 2016; Nathanson and Zwick, 2018; Gao et al., 2020; Bayer et al., 2020; DeFusco et al., 2017) and belief contagion (Bailey et al., 2018; DeFusco et al., 2018). Instead, our objective is to highlight the fundamental component and how it can give rise to a boom-bust-rebound.

4.6 Foreclosures

Our final empirical result concerns foreclosures, which play an important role in our theory. We obtain proprietary data on completed foreclosures at the CBSA level over 2000-2019 from CoreLogic. Panel (a) of Figure 6 plots total foreclosures over 2007-2012, relative to the 2006 housing stock, against the long-run fundamental.[16] Areas with higher fundamentals had larger foreclosure crises. The slope of the best fit line of 8.5 indicates that a one standard deviation (0.16) higher fundamental implies additional foreclosures of 1.4% of the housing stock, and the t-statistic of this relationship exceeds 4.5.[17] Panel (b) shows the time series of the annualized foreclosure rate in four population-weighted quartiles of CBSAs grouped by their long-run fundamental. Foreclosures rose during the bust in all quartiles. Notably, the size of the increase and the peak foreclosure rate both increase monotonically in the long-run fundamental. The sharp rise in foreclosures and the cross-sectional pattern inform our quantitative model, in which a price-foreclosure spiral explains why house prices in the bust fell so drastically.

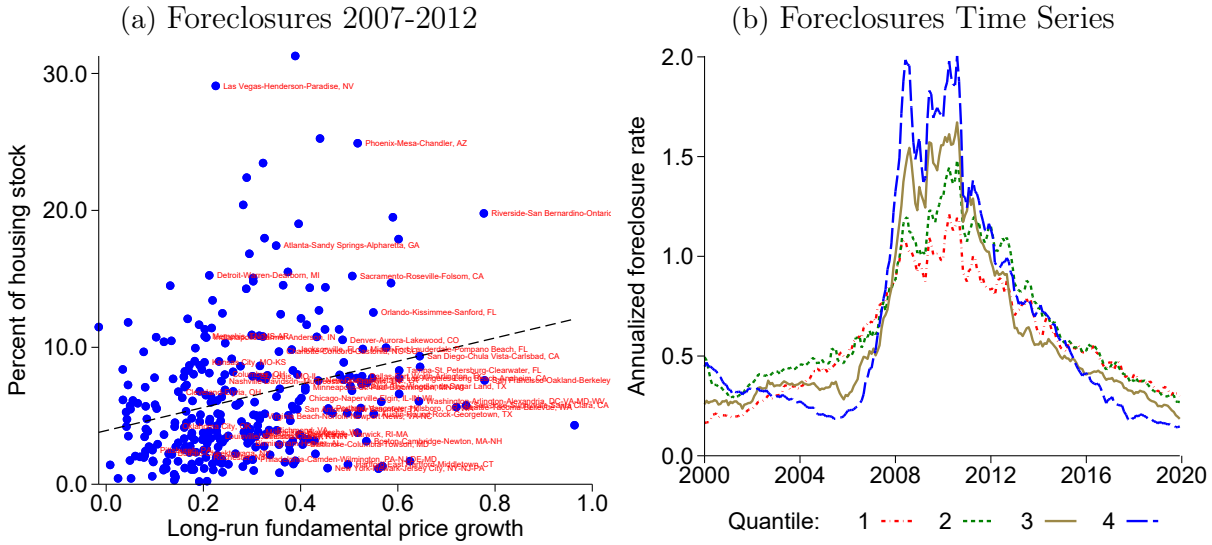
5 A Neo-Kindlebergerian View

Motivated by our empirical finding that areas with stronger long-run fundamental growth experienced a larger boom-bust-rebound and a larger foreclosure crisis, we offer a "neo-Kindlebergerian" interpretation of the boom-bust-rebound. As in Kindleberger's cele-

[16]The coverage of CoreLogic expands over time, with 90% of the population in our CBSA sample covered in 2000 and more than 99% covered by 2006. In Panel (a), we include only CBSAs with foreclosure data starting in 2007 or before. In Panel (b), within each quartile we extend backwards the rate in the sample of CBSAs with data starting in 2006 using the ratio of the foreclosure rate in knot months when new CBSAs enter the sample. The housing stock data also come from CoreLogic and we extend backward using the growth rate of units from the Census.

[17]There is also variation around the trend, with the largest outlier being Las Vegas, which had roughly 30% of its residential housing stock go through foreclosure. Our model in Section 5 accommodates this dispersion as the result of transitory growth in fundamentals; in the case of Las Vegas and some other areas, it might also reflect the role of speculation during the boom, which we discuss in greater detail in Appendix B.1.

Figure 6: Foreclosures and Long-run Fundamental



Notes: The left panel plots total foreclosures in each CBSA over the period 2007-2012 as a percent of the 2006 residential housing stock against the second stage fitted value from column (3) of Table 2, with CBSAs with more than 1 million persons in 1997 labeled in red. The right panel plots the three-month moving average annualized foreclosure rate in quartiles of CBSAs grouped by the second stage fitted value from column (3) of Table 2.

brated *Manias, Panics, and Crashes*, a single change in the economy's fundamentals sets off a boom and bust. In our urban setting, this change takes the form of an increase in the growth rate of the "dividend" from living in the city. Agents learn about the growth rate by observing the history of dividends. However, they become overly optimistic, which we formalize using diagnostic expectations as in Bordalo et al. (2019). Construction and prices boom as optimistic buyers enter the city. 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. Finally, ongoing growth of the dividend causes new buyers to enter and prices to rebound toward their new balanced growth path. 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

The model is set in continuous time. The economy consists of a single city with population H_t and residual "hinterland."

Dividend and beliefs. The driving force in the economy is the "dividend", D_t, from living in the city. 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:

$$dD_t = \mu_t D_t dt + \sigma_D D_t dW_{D,t}, \tag{8}$$

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

$$d\mu_t = \theta\left(\bar{\mu} - \mu_t\right) dt + \sigma_\mu dW_{\mu,t}, \tag{9}$$

where θ determines the rate of convergence to the unconditional mean $\bar{\mu}$ and $dW_{\mu,t}$ is a standard Wiener process uncorrelated with $dW_{D,t}$.[18] 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 rational 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 agent's 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:

$$dm_t = \theta\left(\bar{\mu} - m_t\right) dt + K dB_t, \tag{10}$$

[18] For readers unfamiliar with continuous time stochastic calculus, equation (8) is the analog of a random-walk with drift and equation (9) is the analog of an AR(1) process with mean $\bar{\mu}$.

and where the surprise innovation dB_t follows:

$$dB_t = \sigma_D^{-1} \left(dD_t / D_t - m_t dt \right). \tag{11}$$

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.[19]

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 (10) and (11). Bordalo et al. (2019) 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 peoples' estimate of the share of Irish with red hair: red hair is more prevalent among Irish than non-Irish, and as a result people vastly overestimate the share of Irish with red hair. In the context of asset price cycles, the reference population consists of the history of observed dividends and the class consists of recently-observed dividends, with the 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:

$$h_t^{\varphi} \left(\mu_t \right) = h_t \left(\mu_t | \mathcal{F}_t \right) \left[\frac{h_t \left(\mu_t | \mathcal{F}_t \right)}{h_t \left(\mu_t | \mathcal{F}_{t-k} \right)} \right]^{\varphi}. \tag{12}$$

The diagnostic distribution over-weights states that have become relatively more likely in light of recent dividend news, that is, where $h_t \left(\mu_t | \mathcal{F}_t \right) > h_t \left(\mu_t | \mathcal{F}_{t-k} \right)$. This "kernel-of-truth" property causes over-reaction to good news when the distribution of states satisfies

[19] We assume the asymptotic variance in the posterior drift rate, which satisfies $\sigma_m^2 = \sigma_\mu^2 / \left(2\theta + K/\sigma^D \right)$.

the monotone-likelihood property – as is the case with normally-distributed innovations – because higher growth states become more likely following positive news. The parameter φ controls the magnitude of departure from rational expectations and nests the rational case when $\varphi = 0$. One can further show that the diagnostic distribution takes a particularly simple form, with the mean simply shifted from the rational case by a term $\varphi\mathcal{I}$:

$$h_t^{\varphi}(\mu_t) \sim N\left(m_t^{\varphi}, \sigma_m^2\right),\ m_t^{\varphi} = m_t + \varphi\mathcal{I}_t, \tag{13}$$

where:

$$\mathcal{I}_t \equiv m_t - \mathbb{E}_{t-k} m_t = K \int_{t-k}^{t} e^{-\theta(t-s)} dB_s \tag{14}$$

is the information the diagnostic agent neglects in forming her background context.[20]

The expected present value of dividends will be the key driving force in the model. Letting ρ denote the discount rate, we have:

$$P^*\left(D_t, m_t^{\varphi}\right) = \int_{-\infty}^{\infty} \mathbb{E}_t \left[\int_t^{\infty} e^{-\rho(s-t)} D_s ds | \mu_t\right] h_t^{\varphi}(\mu_t)\, d\mu_t, \tag{15}$$

where we show in Appendix C.1 that this integral depends only on $D_t, m_t^{\varphi} = \mathbb{E}_t^{\varphi}[\mu_t] = m_t + \varphi\mathcal{I}$, and parameters.[21]

The present value $P^*\left(D_t, m_t^{\varphi}\right)$ encodes all of the relevant information coming from the evolution of the dividend and of beliefs. Any belief process over the path of dividends that produces the same path of P^* will produce the same results in our model. In this sense,

[20]Equation (14) extends the Bordalo et al. (2019) implementation of "slow-moving" information to our continuous time setting. Maxted (2021) proposes an alternative formulation that does not truncate at horizon k and instead generalizes the decay parameter in equation (14) not to necessarily equal mean reversion θ. We prefer the formulation in equation (14) because it has a direct interpretation in terms of the information received over a recent horizon and because the parameter k controls the timing of the peak of mean beliefs. While one can also choose the decay parameter to exactly time the length of the boom, this parameter also impacts the length of the bust. Azeredo da Silveira et al. (2020) provide an alternative model of costly information recall that also generates over-reaction to recent news.

[21]Convergence of this present value requires $\bar{\mu} + \frac{1}{2}\frac{\sigma_\mu^2}{\theta^2} < \rho$. 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.

our model does not depend on diagnostic expectations in particular, but only requires over-optimism that eventually recedes.[22] As an example, in Appendix C.3 we provide an alternative foundation for P^* based on a fully rational agent with an over-optimistic prior over the new drift rate distribution.

To summarize, dividends follow a geometric Brownian motion with a drift rate that follows an Ornstein-Ulhenbeck process. When dividends rise unexpectedly, diagnostic agents over-weight the likelihood of high trend growth. Eventually, the positive surprises fall out of the representativeness window and expected dividend 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. One may think of the potential entrant growth rate of g_H as reflecting births in the city or a "cumulative advantage" tendency for migrants to consider potential destinations in proportion to their size. We normalize to 0 the dividend and hence the house price from living in the hinterland. The total cost of purchasing in the city sums the price of the house, P_t, and the cost of a mortgage in up front origination fees or "points", W_t.[23]

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, $V_t = P^*\left(D_t, m_t^{\varphi}\right)$.[24] The idiosyncratic component ξ is drawn from a Pareto distribution

[22] We adopt diagnostic expectations both because it tractably models how over-optimism grows and subsides and because it matches a key feature of the data: The length of the boom, bust, and rebound are independent of their amplitude in the cross-section.

[23] We assume the lender charges borrowers up-front points rather than interest rate spreads in order to economize on a state variable, much like Kaplan et al. (2020).

[24] We make the assumption that agents expect to live in the city forever for simplicity. More generally, one can motivate $V_t = P^*\left(D_t, m_t^{\varphi}\right)$ as a 0^{th} order approximation to the value incorporating the possibility of selling and leaving the city, where the approximation is around the probability of moving, and potentially

with $P(\xi > x) = \left(\frac{x_m}{x}\right)^{\gamma}$.

In spatial equilibrium, a potential entrant purchases a house in the city if ξV_t exceeds the cost $P_t + W_t$. Substituting the distributional assumption on ξ, the total demand for houses from new entrants, Q_t, is given by:

$$Q_t = g_H H_t P\left(\xi > \left(P_t + W_t\right)/V_t\right) = g_H H_t x_m^{\gamma} \left[V_t/\left(P_t + W_t\right)\right]^{\gamma}. \tag{16}$$

Thus, the Pareto coefficient γ governs the slope of the housing demand curve.

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

$$C_t = \left[A H_t^{1/\eta}\right] \exp\left(\frac{I_t - \bar{I}_t}{\chi}\right), \tag{17}$$

where $I_t \equiv \dot{H}_t/H_t$ denotes the construction rate and $\bar{I}_t \equiv \frac{1}{\delta}\int_{t-\delta}^{t} I_s ds$ is the trailing moving average of the construction rate. Equation (17) parallels the empirical specification for housing costs given in equations (1) to (3), with two differences. First, for simplicity we do not separately model the cost of land and construction and instead let η govern the overall long-run elasticity of supply. Second, we introduce a short-run adjustment cost $\exp\left(\left(I_t - \bar{I}_t\right)/\chi\right)$. We write it as a trailing moving average so that it disappears and we recover our exact empirical specification at long horizons when $I_t \approx \bar{I}_t$, even though I_t converges to different values depending on the value of μ.[25] Economically, one may think of the cost as reflecting deviations from the existing capacity of the construction sector, which evolves over time with actual construction rates. The parameter χ is the short-run (instantaneous) elasticity of supply of new construction.

consider higher order terms. In the data, the probability of moving across cities is approximately 2%.

[25] We set δ to be 20 years.

Mortgages and foreclosures. A home-buyer at date t obtains a mortgage of $M_t = \phi P_t$, where the loan-to-value (LTV) ϕ is idiosyncratic for each buyer and drawn from a distribution. Mortgages are interest-only, although this assumption can easily be relaxed.[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 liquidity shock and has negative equity, $M_t > RP_\tau$. Thus, as in Guren and McQuade (2020), the model features a "double trigger", wherein both a liquidity shock and negative equity must occur for an owner to default, consistent with empirical evidence in Foote et al. (2008), Bhutta et al. (2017), Gerardi et al. (2018), Ganong and Noel (2020), and Gupta and Hansman (2021).[27] Foreclosed homes enter supply in the instant after the default occurs.[28]

Competitive, risk-neutral lenders provide mortgages. These lenders have the same beliefs as buyers; namely, they observe the path of dividends D_t but do not know the true drift rate μ_t.[29] They set points W_t, charged to buyers as an up-front fee at origination, to make zero expected profits on each loan given that they only recover a fraction ψ of a

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

[27] As in Greenwald et al. (2021), the idiosyncractic house price shock R smoothes the cliff function for default probability as a function of LTV to be consistent with the data.

[28] Formally, a foreclosure at time t enters at time step $t + \Delta$ where in the continuous time limit $\Delta \to 0$. This timing assumption eliminates multiple equilibria that can arise if foreclosures and prices are determined jointly, since the dependence of foreclosures on prices can create a backward-bending supply curve.

[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. We discuss below an alternative where lenders have perfect foresight over the path of dividends.

house's value in a foreclosure:

$$W_t = \mathbb{E}_t^{\varphi}\left[e^{-\rho(\tau-t)}\max\left\{M_t - \psi R P_\tau, 0\right\}\right], \tag{18}$$

where τ denotes the first date at which the ι liquidity shock hits.

Finally, the mortgage balance measure density, $g(M,t)$, evolves according to the Fokker-Planck equation:

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

5.2 Equilibrium

An equilibrium 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 reflects their beliefs:

$$V_t = P^*\left(D_t, m_t^{\varphi}\right), \tag{20}$$

where $P^*\left(D_t, m_t^{\varphi}\right)$ satisfies equation (15).

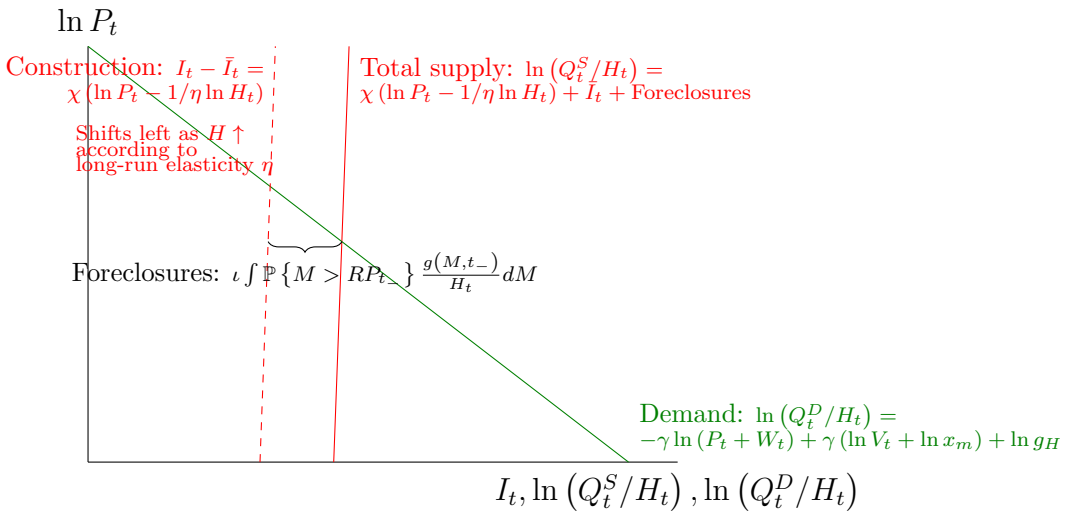
(ii) Price equals the marginal cost of construction given by equation (17):

$$P_t = \left[AH_t^{1/\eta}\right]\exp\left(\frac{I_t - \bar{I}_t}{\chi}\right), \tag{21}$$

where $I_t = \dot{H}_t/H_t$.

(iii) Lenders make zero expected profits such that W_t satisfies equation (18).

Figure 7: Supply and Demand Diagrammatic Treatment of Equilibrium



(iv) Housing demand (equation (16)) 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).

Equations (21) and (22) are the equilibrium conditions not previously stated.

The five equilibrium conditions describe a supply-and-demand framework that determines the equilibrium house price at any instant, as illustrated in Figure 7. 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 with elasticity χ due to construction, 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 (18) by Monte Carlo. Appendix D contains details of our solution method. 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^{φ} to a one-time increase that we use to compute the path of $P^*\left(D_t, m_t^{\varphi}\right)$. This expression allows the efficient calculation of impulse responses for prices, construction, and other key model outcomes that we use in our estimation routine.

5.3 Calibration

Our model is of homebuyers choosing whether to live in a single city or the hinterland. Therefore, we choose city-level targets for our calibration. Since we compute impulse responses of variables after averaging out the idiosyncratic Wiener shocks, we group the CBSAs in our sample into four (population-weighted) quartiles based on their their long-run fundamental from Section 4.4. We average house price growth and other variables across the CBSAs within each quartile using 1997 population as weights.

We calibrate several parameters externally and set the remaining eight parameters for each quartile by simulated method of moments (SMM). Table 5 lists the model parameters, their calibrated values for each quartile of CBSAs, and the rationale for each parameter. We discipline the calibration by fixing across quartiles deep parameters related to learning or preferences as well as mortgage design, while allowing parameters related to fundamentals such as the drift rate of dividends and the supply elasticity to vary across quartiles. We first estimate all eight parameters using moments for the third quartile only, as this quartile is representative of the national boom-bust-rebound in prices and the magnitude of the foreclosure crisis. We then hold the deep parameters related to learning, preferences, and mortgage design fixed and match the same moments letting the remaining parameters

vary across quartiles.

We begin by describing the externally-calibrated parameters relating to beliefs. The market-clearing condition (22) and valuation function (C.1) together give that along a balanced growth path with a fixed μ, constant rate of construction and foreclosures, and constant ratio of points to price, prices grow at the rate μ. Accordingly, we set the long-run mean drift rate $\bar{\mu}$ in each quartile to the annualized average log change in real house prices over the 1975:Q1-1996:Q4 period. We set the mean reversion rate θ to 0.005 so that the increase in the drift rate at date 0 is highly persistent.[30] We set the diagnostic window, k, to 8 to achieve a boom length of 8 years.[31]

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 2 obtained by multiplying the column (3) population growth and triple interaction coefficients by the respective endogenous variables and taking quartile averages. We set the distribution of LTVs at mortgage origination, ϕ, to $N(0.87, 0.09)$ based on evidence from Adelino et al. (2018).[32] We set the foreclosure recovery rate, ψ, to 64.5% as in Guren et al. (2021). We set the growth rate of potential entrants, g_H, to 8.0% annually; this parameter has no impact on the impulse responses.

We estimate the remaining eight parameters – the drift rate at date 0, μ_0, the diagnostic over-shooting parameter, φ, the Kalman gain relative to the dividend noise, K/σ_D, the discount rate, ρ, the demand elasticity, γ, the short-run supply elasticity, χ, the liquidity shock, ι, and the idiosyncratic house price shock, σ_R^2 – to match 10 moments summarized

[30] Using equation (C.5), 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 $\bar{\mu}$.

[31] The Ditzen et al. (2021) break test procedure finds a boom of 8 years in quartiles 1 and 3 and of 8.25 in quartiles 2 and 4. Since the value of k determines the boom length almost exactly and has minimal impact on other parameters, we set $k = 8$ for all quartiles for parsimony.

[32] These values correspond to the national average over the period 1996-2012. Adelino et al. (2018) show that the distribution of LTVs at origination remains extremely stable throughout the boom and bust.

Table 5: Model Parameterization

Symbol	Description	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Rationale
Beliefs						
$\bar{\mu}$	Pre-boom drift rate	0.28%	0.67%	1.07%	1.50%	1975-96 growth rate
μ_0	New drift rate	1.7%	1.95%	2.4%	2.7%	SMM by quartile
φ	Diagnostic over-shooting	4.375	4.375	4.375	4.375	SMM in Q3
K/σ_D	Normalized Kalman gain	0.14	0.14	0.14	0.14	SMM in Q3
k	Diagnostic window	8	8	8	8	Boom length in Q3
θ	Mean reversion in drift	0.005	0.005	0.005	0.005	
Preferences						
ρ	Discount rate	7.0%	7.0%	7.0%	7.0%	SMM in Q3
γ	Demand elasticity	1.1	1.1	1.1	1.1	SMM in Q3
Construction and foreclosures						
η	Long-run supply elasticity	2.27	1.46	0.99	0.78	Empirical regressions
χ	Short-run supply elasticity	0.079	0.034	0.031	0.023	SMM by quartile
g_H	Potential entrants	0.08	0.08	0.08	0.08	Normalization
ι	Liquidity shock	5.78%	4.62%	5.0%	5.5%	SMM by quartile
R	House price shock	$N(1, 0.08)$	$N(1, 0.08)$	$N(1, 0.08)$	$N(1, 0.08)$	SMM in Q3
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)

Table 6: SMM Moments

		Quartile 1		Quartile 2		Quartile 3		Quartile 4	
Moment	Source	Data	Achieved	Data	Achieved	Data	Achieved	Data	Achieved
House price growth (log points unannualized)									
1. Boom	Freddie Mac	19.3	25.3	39.7	47.1	61.8	61.8	81.2	72
2. Bust	Freddie Mac	24.7	20.2	36.1	38	40.1	41.1	51	49
3. Rebound	Freddie Mac	17.9	10.4	24.7	22.4	25.7	26.1	42.8	34.9
Phase length (years)									
4. Bust	Freddie Mac	6	7.81	5.75	6.02	6	5.83	6	5.02
Role of foreclosures									
5. Max Foreclosure Rate (% of Housing Stock)	CoreLogic	1.21	1.5	1.48	1.65	1.67	1.7	2	1.97
6. % Jan-07 Equity< 20%	Beraja et al. (2019)	33.4	27.1	24.6	19.2	17.7	18.5	17.9	19.2
7. % Jan-07 Equity< 10%	Beraja et al. (2019)	17.5	9.9	11.8	8.3	8.2	8.6	8.2	9.4
8. % Jan-07 Equity< 0%	Beraja et al. (2019)	5.1	1.9	3	1.9	2.5	2.2	2.8	2.7
9. Bust Speed (log points)	Freddie Mac	-7.8	-3.4	-11.4	-11.2	-13	-12.5	-22.4	-22.9
10. Rebound Speed (log points)	Freddie Mac	3.5	3.9	5.4	5.9	6.3	6.2	13	7.4

in Table 6, with each moment computed separately for each quartile. We estimate the learning parameters φ and K/σ_D, the preference parameters ρ and γ, and the house price shock σ_R^2 as deep parameters and fix their values at the levels estimated using the moments for the third quartile. We then estimate quartile-specific values of the new drift rate μ_0, the liquidity shock ι, and the short-run elasticity χ, giving three parameters to fit ten quartile-specific moments in quartiles 1, 2, and 4.[33]

The first three moments measure the size of the boom, bust, and rebound in house prices. Heuristically, the size of the full boom-bust-rebound most directly informs μ_0, ρ, χ, and γ as it reflects the actual path of dividends and their impact on the price. The over-shooting of the boom relative to the long-run informs the learning parameters φ and K/σ_D. The size of the bust additionally depends on ι, χ, and γ, as these parameters govern the magnitude of the price-foreclosure spiral.

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 remaining moments characterize the role of foreclosures. The fifth moment is the peak annualized foreclosure rate, as shown in Figure 6. The next three moments describe the equity distribution near the start of the bust and come from the Beraja et al. (2019) data set: the share of properties with equity less than 20%, 10%, and underwater in January 2007. These moments all inform 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 as moments the maximum four-quarter price decline in the bust and price increase in the rebound. In addition to χ and γ, the idiosyncratic house price shock variance σ_R^2 impacts these speed moments.

[33]We constrain the short-run elasticity χ, for which we do not have an empirical counterpart, to have a similar ratio across quartiles to the empirically-estimated long-run elasticity η.

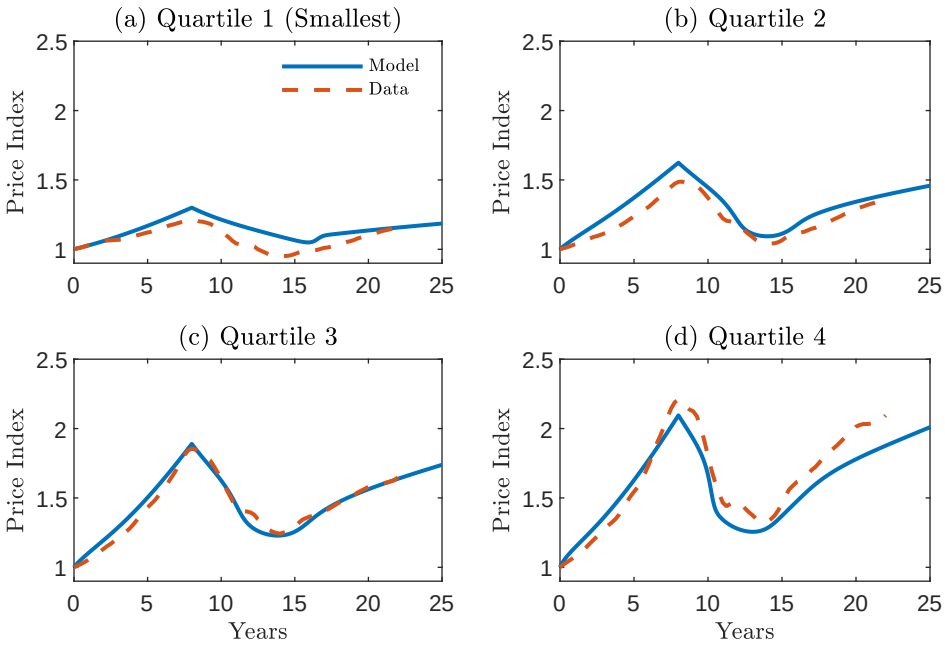
We choose the set of parameters that minimizes the weighted sum of squared residuals between the model and the data moments. We weight all moments equally except the three equity distribution moments, to which we assign 1/3 of the weight of the other moments, and the bust and rebound speed moments, to which we assign 1/2 of the weight of the other moments. To calculate the model moments, we first solve the model and compute the impulse response to a one-time increase from $\bar{\mu}$ to μ_0, using the analytic mean path of beliefs. We define the boom as time zero to the price peak and the bust as the price peak to the price trough. Because there is no natural end date for the rebound but our data end in 2019, we define the rebound as the seven years following the price trough. We calculate the log changes in house prices, the length of each phase, the maximum four-quarter price decline in the bust and price rise in the rebound using a moving average of the price (to mimic how HPIs time smooth), the moments of the home equity distribution 9 years after the boom start, and the maximum annualized foreclosure rate.

5.4 Model Fit

Table 6 shows that the estimated model fits the moments extremely well. For the third quartile for which we estimate all parameters, the model matches the sizes of the boom, bust, and rebound to within one log point, the bust length almost exactly, all three points of support of the January 2007 equity distribution to within one percentage point, the maximum foreclosure rate almost exactly, and the maximum speed of the bust and rebound to within one-half log point.

The model fits the remaining quartiles well despite the deep parameters being held fixed in estimation, leading to many fewer degrees of freedom. Figure 8 compares the price paths in model and data in each quartile. The quality of fit despite the strong over-identification serves as a validation of the model. Most important, the model reproduces

Figure 8: 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 fundamental from Section 4.4.

the empirical pattern of a higher long-run fundamental (larger μ_0, smaller η) producing a larger boom, bust, and rebound and a more severe foreclosure crisis.

Several of the estimated parameters that deliver this fit merit comment. The estimated drift rate μ_0 rises monotonically across quartiles of fundamental growth. The value of 2.4% in the third quartile implies new annual trend growth 1.3p.p. per year higher than in the pre-boom, not a radical departure. The estimated diagnostic parameter φ of 4.375 is somewhat larger than the value of 1.9 that Bordalo et al. (2019) estimate for stock returns, perhaps reflecting the difference between the households in our setting and the professional forecasters in theirs. We show below that this value delivers expectations in line with untargeted survey data. The estimated values of χ imply substantially steeper short-run than long-run supply curves. The values of the liquidity shock ι in the neighborhood of 5% accord remarkably well with evidence on mortgage pre-payment rates during the boom,

supporting our modeling of a single liquidity shock in both periods.[34]

The model also reproduces the empirical pattern shown in Table 4 that higher price-dividend growth during the boom forecasts higher future dividend growth and not large relative price declines.[35] For this exercise, we collapse the model time series by year and regress the log changes in dividends and price between model years corresponding to 2006 and 2019 on the log change in the price-dividend ratio during the boom. An additional log point of growth in the price-dividend ratio in the boom predicts additional 0.41 log point of dividend growth in the bust-rebound (compared to 0.20 in column (3) of Table 4) and additional 0.03 log point of price growth (compared to 0.03 in column (6) of Table 4).

5.5 Unpacking the Mechanism

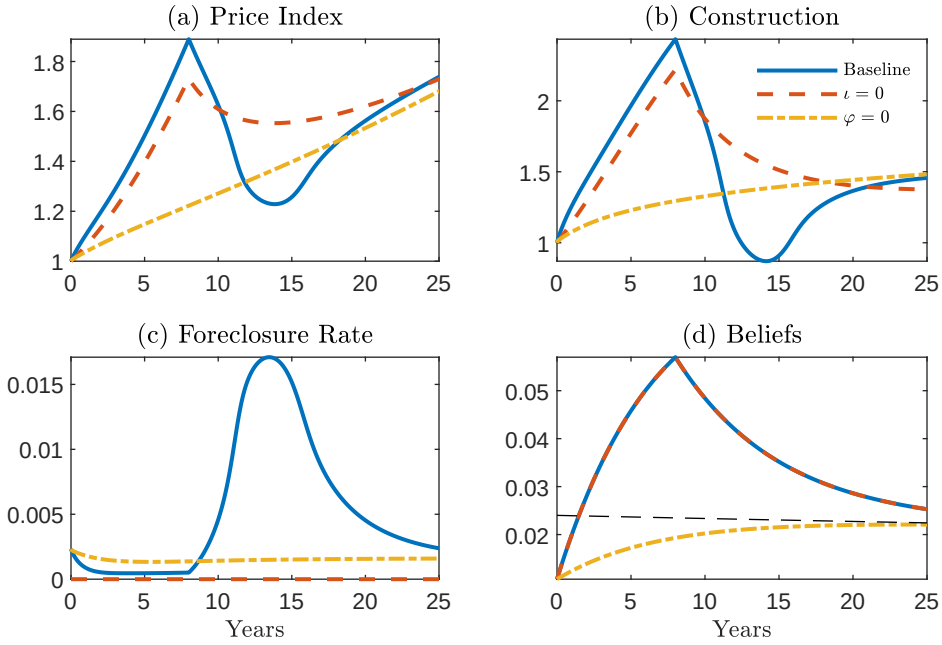
We now discuss the model features that generate a boom-bust-rebound. For parsimony, we use the parameters estimated for the third quartile. Figure 9 shows the evolution of prices, foreclosures, construction, and beliefs following a change from $\bar{\mu}$ to μ_0. The solid blue lines show the paths at the estimated parameter values, the dash-dot gold lines show the paths without diagnostic expectations ($\varphi = 0$), and the dashed red lines show the paths with no foreclosures ($\iota = 0$), holding other parameters fixed.

The comparisons across lines in each panel illustrate the importance of both belief overshooting and foreclosures to generating a realistic boom-bust-rebound. Without diagnostic learning, beliefs, prices, and construction all rise smoothly and monotonically. The steady price appreciation reflects three forces: (i) the increase in the capitalization rate $m_t - \rho$,

[34]Figure A-1 of Berger et al. (2021) shows the mortgage prepayment rate broken into interest rate refinancings, cash out refinancings, and purchase prepayment using CRISM data. The authors project this backwards using origination shares. They find that this series peaks at an annual rate of roughly 4.5% in the boom when essentially everyone was above water.

[35]To formally justify equating dividends in the model with rents in the data, we can augment the model to include a measure 0 of renters who live in units that fully depreciate each instant. Such individuals are willing to pay up to D_t in rent.

Figure 9: Model Boom-Bust-Rebound



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 third quartile of CBSAs. The dashed red line shows the model without foreclosures, that is with all of the same parameters except $\iota = 0$. The dash-dot gold line shows a parameterization with foreclosures but no diagnostic expectations, that is with all of the same parameters as the baseline calibration except $\varphi = 0$. Panel (a) shows the price index normalized to 1 in the instant before the drift rate changes to μ_0. Panels (b) and (c) show the foreclosure rate in annualized percent of the housing stock and the construction rate as an annualized growth rate. Panel (d) shows beliefs, with the dashed black line showing the true drift rate.

(ii) the increase over time in the dividend, and (iii) the decline in foreclosures relative to the pre-boom period.

Diagnostic learning generates an over-shooting of beliefs about the drift rate μ, as shown in Panel (d). These beliefs rise from their initial level of 1.1% up to above 5% before nearly converging to the true value (shown by the thin dashed line) by the end of the sample window. The turning point coincides with the peak of the boom. Unlike in models with "Minsky" moments where beliefs and asset prices crash suddenly, with diagnostic expectations the convergence is gradual, consistent with the six-year length of the bust. The over-shooting of diagnostic beliefs in the boom has a natural counterpart in survey

evidence on beliefs about house price growth. The best available evidence comes from Case et al. (2012), who show a similar degree of over-optimism in surveys of homebuyers.[36]

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: (i) credit expands in the baseline case as lenders perceive a decline in default risk, (ii) in the baseline path, the increase in prices causes foreclosures to decline relative to the pre-boom period, contracting supply and further pushing up prices. The boom length is nearly identical with and without foreclosures, reflecting the role of the parameter k in determining the peak in beliefs and 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.[37] 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.[38] Reflecting this

[36]Case et al. (2012) survey four counties (Alameda California, Milwaukee Wisconsin, Middlesex Massachusetts, and Orange California) from 2003 to 2012 and report 1-year and 10-year expectations of annual house price growth. These measures do not map straightforwardly into beliefs over μ, both due to the nature of the survey and because μ corresponds to price growth on the balanced growth path but not necessarily to price growth at any point in the transition. Nonetheless, expectations of annualized price growth over a 10-year horizon exceed 10% in all four counties during the boom before falling gradually to between 3 and 5 percent by 2012, very much in the ballpark implied by the beliefs of diagnostic agents in our model.

[37]The price-foreclosure spiral in our model is quantitatively large: The bust size without foreclosures is 10.9 log points relative to 41.1 in the baseline model. The speed and magnitude of the bust and rebound explain why our exercise requires a substantial impact of foreclosures to account for the overshooting during the bust. For comparison, in a structural exercise with a bursting bubble but no rebound, Guren and McQuade (2020) find that the supply-and-demand effects of foreclosures account for 32% of the bust. Their model, however, also includes another channel whereby foreclosures cause losses on bank balance sheets. Without that channel, they would infer a much larger role for supply-and-demand 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 house prices due to foreclosures compared to 30 log points in our model.

[38]This dynamic parallels more general fire sale episodes (Shleifer and Vishny, 2011), where the idiosyncratic location preference ξ plays the role of reallocating capital to a second-best use, thereby depressing prices.

dynamic, the increase in foreclosures in Panel (b) and decline in construction in Panel (c) reach 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.

5.6 Comparison to Other Mechanisms

To clarify how our focus on over-optimism about fundamental determinants of house prices compares to the existing literature, in this subsection we relate our analysis to narratives that emphasize changes in credit supply or the role of speculators.

Credit. In our model, changes in mortgage costs, including credit spreads and fixed costs of underwriting, have relatively small effects on prices. To see why, recall the demand equation (16): $\ln\left(Q_t^D/H_t\right) = -\gamma \ln\left(P_t + W_t\right) + \gamma\left(\ln V_t + \ln x_m\right) + \ln g_H$. In the baseline calibration, higher anticipated dividends shift the demand curve by $\gamma \Delta \ln V_t \approx 140$ log points from date 0 to the peak of the boom. If the present value of mortgage costs, W_t, is small relative to the price, P_t, then even large changes in W_t shift the demand curve by only a small amount relative to the change in V_t.[39]

Changes in credit that affect approval rates on the extensive margin offer greater potential for credit to impact prices in our model. 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

[39] 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 exercise, W_t/P_t rises by more than 8p.p., but the price path changes little. 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.

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 (16) 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 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.[40] On the other hand, the rise in rents in high fundamental areas shown in Figure 5 militates against attributing the initial shift out of the demand curve to easing credit.

Speculation. Appendix B.1 explores the relationship between the long-run fundamental and speculative activity in the data. We make five observations: (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) speculators did not contribute disproportionately to selling pressure during the bust; and (v) the degree of speculative activity in the late boom is uncorrelated with the long-run fundamental in an area. 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

[40] In related work, Foote et al. (2021) argue that without extensive margin credit expansion, low income borrowers would have been priced out of housing markets and the distribution of mortgage credit would have counterfactually shifted toward high income groups. Appendix Figure B.5 shows that the long-run fundamental is essentially uncorrelated with the change in subprime share during the boom. This pattern suggests the role of other types of credit relaxation (including low interest rates that ease payment-to-income constraints) and rising incomes (indeed, rising incomes are part of the fundamental) in keeping house prices affordable in high fundamental areas. Note also that credit relaxation need not raise prices in areas without growing fundamentals if credit constraints did not bind already in these areas, perhaps because the initial house price level was lower.

pressure during the bust. In sum, speculation complements our focus on fundamentals as a source of the late boom without impacting the quantitative analysis of our model.

6 Conclusion

We revisit the 2000s housing cycle with "2020 hindsight." At the city level, the areas with the largest price increases during the housing boom from 1997 to 2006 had the largest busts from 2006 to 2012 but also the fastest growth after the trough, and as a result have had the largest price appreciation over the full cycle. We present a standard spatial equilibrium framework of house price growth determined by local income, amenities, and supply determinants and show this framework 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-Kindlbergererian interpretation emphasizes the role of economic fundamentals in setting off this asset price cycle. In our model, the boom results from over-optimism about an increase in the "dividend" growth rate, the bust ensues when beliefs of home buyers and lenders correct, exacerbated by a price-foreclosure spiral that pushes prices below their full-information level, and eventually a rebound emerges as the economy converges to a price path commensurate with fundamental growth. We also acknowledge other features of the episode, including changes in credit supply and speculation, but conclude that these forces cannot substitute for the role of fundamentals as a driving force.

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 the 2000s cycle were only due to sunspot expectations or changes in credit supply, a macroprudential policy maker might want to aggressively stamp out any such boom. Our findings suggest that while policy may want

to temper over-optimism and aggressively mitigate foreclosures, it is also important not to suffocate fundamentally-driven growth. 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.

Albouy, David, Gabriel Ehrlich, and Minchul Shin (2018). "Metropolitan Land Values". *Review of Economics and Statistics* 100 (3): 454–466.

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

Andrews, Donald (1993). "Tests for Parameter Instability and Structural Change With Unknown Change Point". *Econometrica* 61 (4): 821–856.

Andrews, Isaiah, James Stock, and Liyang Sun (2019). "Weak Instruments in IV Regression: Theory and Practice". *Annual Review of Economics* 11: 727–753.

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.

Armona, Luis, Andreas Fuster, and Basit Zafar (2019). "Home Price Expectations and Behaviour: Evidence from a Randomized Information Experiment". *The Review of Economic Studies* 86 (4): 1371–1410.

Azeredo da Silveira, Rava, Yeji Sung, and Michael Woodford (2020). "Optimally Imprecise Memory and Biased Forecasts". National Bureau of Economic Research Working Paper No. 28075.

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

— (2003). "Computation and analysis of multiple structural change models". *Journal of Applied Econometrics* 18 (1): 1–22.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018). "The Economic Effects of Social Networks: Evidence from the Housing Market". *Journal of Political Economy* 126 (6): 2224–2276.

Baum-Snow, Nathaniel and Lu Han (2020). "The Microgegraphy of Housing Supply". Working Paper, University of Toronto.

Bayer, Patrick, Christopher Geissler, Kyle Mangum, and James W Roberts (2020). "Speculators and Middlemen: The Strategy and Performance of Investors in the Housing Market". *The Review of Financial Studies* 33 (11): 5212–5247.

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

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* Forthcoming.

Bhutta, Neil, Jane Dokko, and Hui Shan (2017). "Consumer Ruthlessness and Mortgage Default During the 2007 to 2009 Housing Bust". *The Journal of Finance* 72 (6): 2433–2466.

Bordalo, Pedro, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer (2020). "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, and Andrei Shleifer (2018). "Diagnostic Expectations and Credit Cycles". *The Journal of Finance* 73 (1): 199–227.

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.

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

Campbell, John and Robert Shiller (1988a). "Stock Prices, Earnings, and Expected Dividends". *The Journal of Finance* 43 (3): 661–676.

— (1988b). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". *The Review of Financial Studies* 1 (3): 195–228.

Campbell, Sean, Morris Davis, Joshua Gallin, and Robert Martin (2009). "What Moves Housing Markets: A Variance Decomposition of the Rent-Price Ratio". *Journal of Urban Economics* 66 (2): 90–102.

Case, Karl E., Robert J. Shiller, and Anne K. Thomson (2012). "What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets". *Brookings Papers on Economic Activity*: 265–315.

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.

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

David, Alexander (1997). "Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility". *Journal of Financial and Quantitative Analysis* 32 (4): 427–462.

Davis, Morris A and Jonathan Heathcote (2007). "The Price and Quantity of Residential Land in the United States". *Journal of Monetary Economics* 54 (8): 2595–2620.

Davis, Morris A. and Michael G. Palumbo (2008). "The Price of Residential Land in Large US Cities". *Journal of Urban Economics* 63 (1): 352–384.

DeFusco, Anthony, Wenjie Ding, Fernando Ferreira, and Joseph Gyourko (2018). "The Role of Price Spillovers in the American Housing Boom". *Journal of Urban Economics* 108: 72–84.

DeFusco, Anthony, Charles Nathanson, and Eric Zwick (2017). "Speculative Dynamics of Prices and Volume". National Bureau of Economic Research Working Paper No. 23449.

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

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.

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". National Bureau of Economic Research Working Paper No. 26632.

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". National Bureau of Economic Research 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.

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.

Foote, Christopher L., Lara Lowenstein, and Paul S. Willen (2019). "Technological Innovation in Mortgage Underwriting and the Growth in Credit, 1985-2015". Woorking Paper, Federal Reserve Bank of Boston.

Gallin, Joshua (2008). "The Long-Run Relationship Between House Prices and Rents". *Real Estate Economics* 36 (4): 635–658.

Ganong, Peter and Pascal J Noel (2020). "Why Do Borrowers Default on Mortgages? A New Method For Causal Attribution". National Bureau of Economic Research Working Paper No. 27585.

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

Garriga, Carlos and Aaron Hedlund (2020). "Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession". *American Economic Review* 110 (6): 1603–34.

Gennaioli, Nicola and Andrei Shleifer (2018). *A Crisis of Beliefs*. Princeton University Press.

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 and Joseph Gyourko (2018). "The Economic Implications of Housing Supply". *Journal of Economic Perspectives* 32 (1): 3–30.

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

Glaeser, Edward L and Joseph Gyourko (2005). "Urban Decline and Durable Housing". *Journal of Political Economy* 113 (2): 345–375.

Glaeser, Edward L, Joseph Gyourko, and Raven E Saks (2005a). "Why Have Housing Prices Gone Up?" *American Economic Review* 95 (2): 329–333.

Glaeser, Edward L and Charles G Nathanson (2015). "Housing Bubbles". In: *Handbook of Regional and Urban Economics*. Vol. 5. Elsevier: pp. 701–751.

— (2017). "An Extrapolative Model of House Price Dynamics". *Journal of Financial Economics* 126 (1): 147–170.

Glaeser, Edward L., Joseph Gyourko, and Albert Saiz (2008). "Housing Supply and Housing Bubbles". *Journal of Urban Economics* 64 (2): 198–217.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks (2005b). "Urban Growth and Housing Supply". *Journal of Economic Geography* 6 (1): 71–89.

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

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 and Christopher Hansman (2021). "Selection, Leverage, and Default in the Mortgage Market". *The Review of Financial Studies*. hhab052.

Guren, Adam M., Arvind Krishnamurthy, and Timothy J. Mcquade (2021). "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, Christopher Mayer, and Todd Sinai (2013). "Superstar Cities". *American Economic Journal: Economic Policy* 5 (4): 167–99.

Gyourko, Joseph and Raven Molloy (2015). "Regulation and Housing Supply". In: *Handbook of Regional and Urban Economics*. Vol. 5. Elsevier: pp. 1289–1337.

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.

Hedlund, Aaron (2016). "The Cyclical Dynamics of Illiquid Housing, Debt, and Foreclosures". *Quantitative Economics* 7 (1): 289–328.

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). "Regional Divergence and House Prices". Working Paper, University of Illinois.

Johnson, Stephanie (2020). "Mortgage Leverage and House Prices". Working Paper, Rice University.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti (2019). "Credit Supply and the Housing Boom". *Journal of Political Economy* 127 (3): 1317–1350.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti (2017). "The Mortgage Rate Conundrum". National Bureau of Economic Research Working Paper No. 23449.

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.

Kindermann, Fabian, Julia Le Blanc, Monika Piazzesi, and Martin Schneider (2021). "Learning About Housing Cost: Survey Evidence from the German House Price Boom". National Bureau of Economic Research Working Paper No. 28895.

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

Kolesár, Michal (2013). "Estimation In An Instrumental Variables Model With Treatment Effect Heterogeneity". Working Paper, Princeton University.

Kuchler, Theresa and Basit Zafar (2019). "Personal Experiences and Expectations About Aggregate Outcomes". *The Journal of Finance* 74 (5): 2491–2542.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2015). "The Housing Market(s) of San Diego". *American Economic Review* 105 (4): 1371–1407.

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.

Lipster, Robert and Albert Shiryaev (1977). *Statistics of Random Processes*. Springer.

Loutskina, Elena and Philip E. Strahan (2015). "Financial Integration, Housing, and Economic Volatility". *Journal of Financial Economics* 115 (1): 25–41.

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

Maxted, Peter (2021). "A Macro-Finance Model with Sentiment". Working Paper, Harvard University.

Mayer, Christopher (2011). "Housing Bubbles: A Survey". *Annual Review of Economics* 3 (1): 559–577.

Mayer, Christopher J. and C.Tsuriel Somerville (2000). "Residential Construction: Using the Urban Growth Model to Estimate Housing Supply". *Journal of Urban Economics* 48 (1): 85–109.

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.

— (2014). "What Explains the 2007–2009 Drop in Employment?" *Econometrica* 82 (6): 2197–2223.

— (2017). "Credit Supply and Housing Speculation". National Bureau of Economic Research Working Paper No. 24823.

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

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.

Nathanson, Charles and Eric Zwick (2018). "Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market". *The Journal of Finance* 73 (6): 2587–2633.

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

Piazzesi, Monika and Martin Schneider (2009). "Momentum Traders in the Housing Market: Survey Evidence and a Search Model". *American Economic Review* 99 (2): 406–11.

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.

Poterba, James M (1984). "Tax Subsidies to Owner-Occupied Housing: an Asset-Market Approach". *The Quarterly Journal of Economics* 99 (4): 729–752.

Roback, Jennifer (1982). "Wages, Rents, and the Quality of Life". *Journal of Political Economy* 90 (6): 1257–1278.

Rosen, Kenneth T. (1979). "A Regional Model of Multifamily Housing Starts". *Real Estate Economics* 7 (1): 63–76.

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.

Shleifer, Andrei and Robert Vishny (2011). "Fire Sales in Finance and Macroeconomics". *Journal of Economic Perspectives* 25 (1): 29–48.

Sinai, Todd (2013). "House Price Moments in Boom-Bust Cycles". In: *Housing and the Financial Crisis*. Ed. by Edward Glaeser and Todd Sinai. Chicago: University of Chicago Press: pp. 19–68.

Soo, Cindy K. (2018). "Quantifying Sentiment with News Media across Local Housing Markets". *The Review of Financial Studies* 31 (10): 3689–3719.

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

Veronesi, Pietro (1999). "Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model". *The Review of Financial Studies* 12 (5): 975–1007.

The 2000s Housing Cycle with 2020 Hindsight

Online Appendix

Gabriel Chodorow-Reich Adam Guren Tim McQuade

August 2021

A Microfounding Construction and Land Costs

A.1 Construction

Assume a construction function for producing new 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:

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

where:

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

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

As in Alonso (1964), Muth (1969), and Mills (1967) and Saiz (2010), consider a city with population $H_{i,t}$ laid out on a disk with radius $\Phi_{i,t}$. A fraction Λ_i of the disk is buildable land, giving:

$$\Phi_{i,t} = \sqrt{\frac{H_{i,t}}{\Lambda_i \pi}}, \tag{A.4}$$

where we have normalized lot size to 1. The rental cost of a plot of land $\nu_{i,t}$ depends on its distance τ from the city center:

$$\nu_{i,t}(\tau) = \kappa_{i,t}\left(\Phi_{i,t}^{\chi_i} - \tau^{\chi_i}\right). \tag{A.5}$$

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 $\chi_i > 0$ is the elasticity of the premium to living in the city center relative to living 1% of the city radius outside of the center; denoting $\hat{\nu}_{i,t}(\tau) \equiv \frac{\nu_{i,t}(0)-\nu_{i,t}(\tau)}{\nu_{i,t}(0)} = \left(\frac{\tau}{\Phi_{i,t}}\right)^{\chi_i}$, $\chi_i = \frac{\partial \ln \hat{\nu}_{i,t}(\tau)}{\partial(\tau/\Phi_{i,t})}$. The term $\left(\Phi_{i,t}^{\chi_i} - \tau^{\chi_i}\right)$ has the literal intepretation of offsetting the reduction in the cost of commuting to the city center; more generally it reflects any gradient in the desirability of different neighborhoods. As a city grows, the premium to living in the city-center rises to preserve intra-city spatial equilibrium.

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

$$L_{i,t}(\tau) = \int_t^{\infty} e^{-\rho(s-t)}\kappa_{i,s}\left(\Phi_{i,s}^{\chi_i} - \tau^{\chi_i}\right) ds. \tag{A.6}$$

We consider a balanced growth path with $\kappa_{i,s} = \kappa_i$ and population growth of I_i, giving $\Phi_{i,s} = \sqrt{\frac{H_{i,s}}{\Lambda_i \pi}} = \sqrt{\frac{e^{I_i(s-t)}H_{i,t}}{\Lambda_i \pi}} = e^{\frac{I_i}{2}(s-t)}\Phi_{i,t}$. Along this path, the price of a unit of land at distance τ from the city center is:

$$\begin{aligned} L_{i,t}(\tau) &= \int_t^{\infty} e^{-\rho(s-t)}\kappa_i\left(e^{\frac{I_i}{2}(s-t)}\Phi_{i,t} - \tau\right)^{\chi_i} ds \\ &= \kappa_i\left(\frac{\Phi_{i,t}^{\chi_i}}{\rho - \chi_i I_i/2} - \frac{\tau^{\chi_i}}{\rho}\right). \end{aligned} \tag{A.7}$$

Note that while the rental value of land at the city boundary $\nu_{i,t}(\Phi_{i,t})$ is zero, with positive population growth the price $L_{i,t}(\Phi_{i,t})$ is strictly positive, reflecting the capitalization of future non-zero rents.

The average price of a plot of land in the city at time t integrates over the available land at each distance τ:

$$\begin{aligned} L_{i,t} &= \frac{1}{H_{i,t}}\int_0^{\Phi_{i,t}} L_{i,t}(\tau)\,\Lambda_i 2\pi\tau d\tau \\ &= \kappa_i\left(\frac{1}{\rho - \chi_i I_i/2} - \frac{1}{\rho(\chi_i/2+1)}\right)\left(\frac{H_{i,t}}{\Lambda_i \pi}\right)^{\frac{\chi_i}{2}} \end{aligned}$$

$$= B_{i,t} H_{i,t}^{\beta}, \tag{A.8}$$

where:
$$B_{i,t} = \kappa_i \left(\frac{1}{\rho - \chi_i I_i/2} - \frac{1}{\rho\left(\chi_i/2 + 1\right)} \right) \left(\frac{1}{\Lambda_i \pi} \right)^{\frac{\chi_i}{2}}, \tag{A.9}$$

$$\beta = \chi_i/2. \tag{A.10}$$

Equation (A.8) extends equation (1) in Saiz (2010) to allow for non-zero population growth ($I_i > 0$) and arbitrary city center premium ($\chi_i \neq 1$), both of which feature in our empirical analysis. It nonetheless retains the crucial and intuitive prediction that the average price of land is higher in places with a smaller share Λ_i of land available for development.

B Empirical Appendix

B.1 Investors/Speculators

In this appendix we explain how our focus on long-run fundamental determinants of house prices relates to work emphasizing the role of speculation in the boom. We make five observations: (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 use Home Mortgage Disclosure Act (HMDA) data to measure the share of purchase mortgages to non-owner occupiers in each CBSA and year. Chinco and Mayer (2016), Gao et al. (2020) and DeFusco et al. (2017) associate the non-owner occupier share with speculative activity by investors.[1] Gao et al. (2020) show that the investor share in 2004-06 predicts house price growth in the same period, including

[1] 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 simply rescale the investor measure and because the comparisons across periods in Table B.1 hold fixed the investor share.

when instrumented with state tax treatment of capital gains, while Chinco and Mayer (2016) focus on out-of-town buyers. 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.[2]

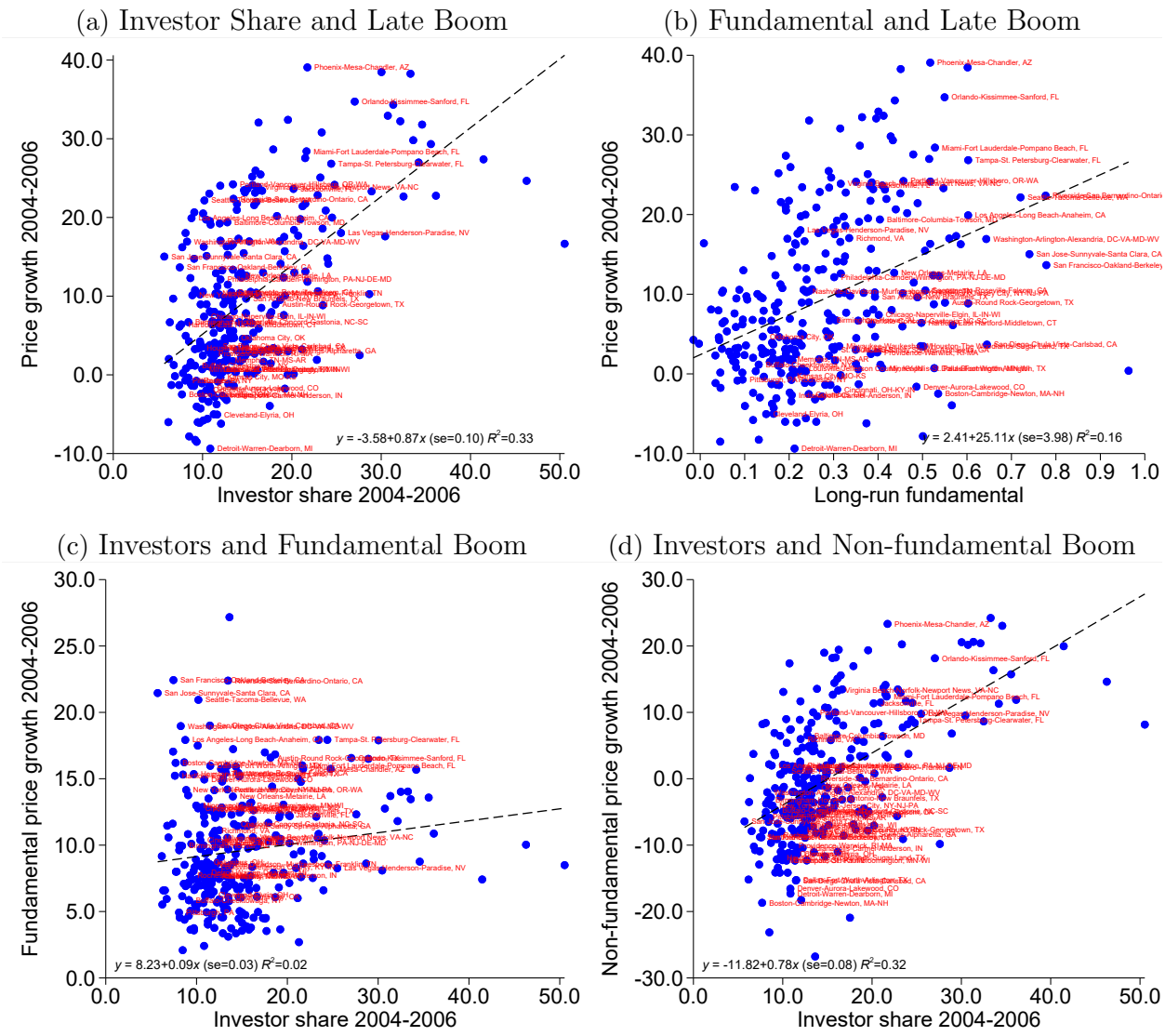
Panel (b) shows that house price growth during 2004-2006 also correlates strongly with the long-run fundamental, measured as usual as the fitted value from the second stage regression in column (3) of Table 2. Thus, even at the end of the boom when speculative activity was plausibly most rampant, long-run 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, 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 ($R^2 = 0.32$, column (5)). The R^2 of the long-run fundamental for house price growth over the full 1997-2019 period of 0.48 (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, as can be seen by the downward slope of the CBSAs labeled in red which contain more than 1 million persons. 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, which has received substantial attention for the role of speculation in

[2]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



Notes: Panel (a) plots the investor share of purchases averaged over 2004-2006 against house price growth in 2004-2006. The investor share of purchase mortgages comes from HMDA. Panel (b) plots the long-run fundamental against house price growth in 2004-2006. Panel (c) plots the investor share averaged over 2004-2006 against the fitted value from a regression of 2004-2006 house price growth on the long-run fundamental. Panel (d) plots the investor share averaged over 2004-2006 against the residual from a regression of 2004-2006 house price growth on the long-run fundamental. 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.

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**		0.80**	0.57**		0.32+	0.92**		0.57**
	(0.10)		(0.08)	(0.16)		(0.17)	(0.16)		(0.13)
Long-run fundamental		0.25**	0.20**		0.73**	0.71**		1.01**	0.97**
		(0.04)	(0.03)		(0.08)	(0.08)		(0.06)	(0.06)
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 (×100)	16.0	16.0	16.0	16.0	16.0	16.0	16.0	16.0	16.0
R^2	0.330	0.159	0.432	0.034	0.318	0.328	0.068	0.477	0.503
Observations	301	301	301	301	301	301	301	301	301

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 2. Robust standard errors in parentheses. **,+ denote significance at the 1, and 10 percent levels, respectively.

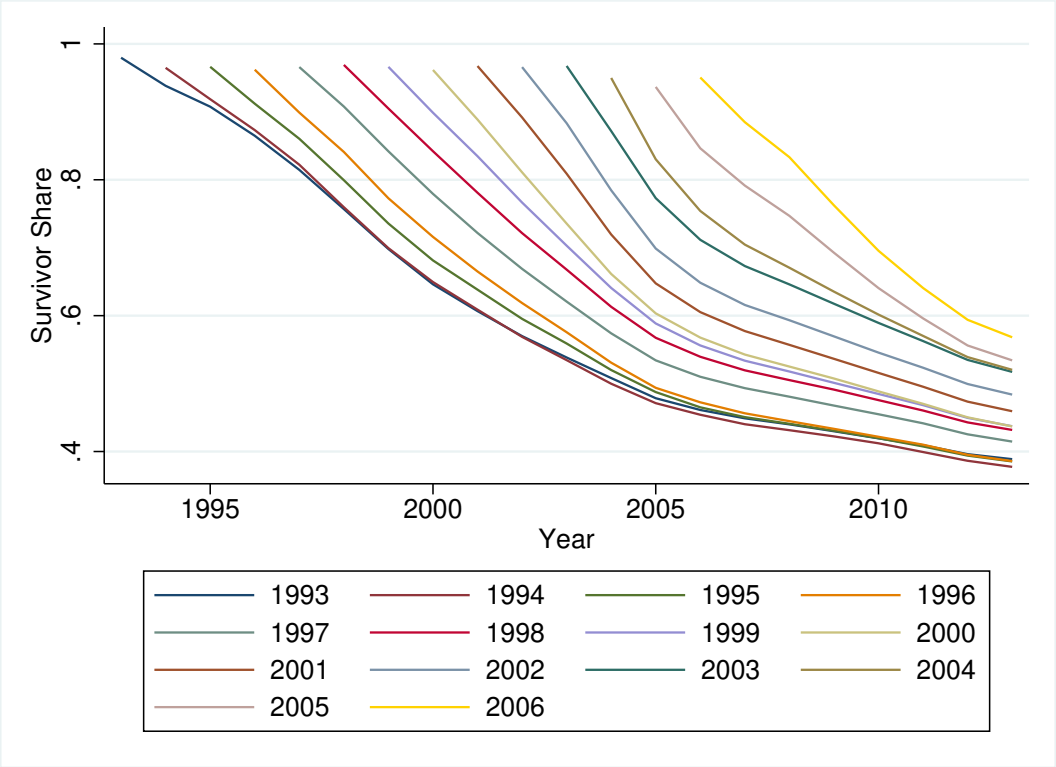
its housing boom (Chinco and Mayer, 2016; Nathanson and Zwick, 2018), provides an example of a CBSA with a high investor share but a relatively low long-run fundamental and hence a small predicted value for 2004-2006 house price growth.

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, and this relationship explains 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. To do so, we start with a merge of HMDA and the DataQuick deeds data from Diamond et al. (2019). We use the same indicator for an investor from the HMDA data as above. Using the deeds data, we identify the next arms-length transaction on a property previously purchased by an investor. We match 87% of residential transactions in the DataQuick deeds data from 1993 to 2009 to a HMDA record.

Figure B.2 shows the survival functions for continuing to own the property for investors who purchased prior to the bust. The survival function flattens around 2006 for all cohorts. This indicates that investors did not dump their properties *en masse* in the bust and in fact became *less* likely to sell. Consequently, while the emergence and receding of investor demand can potentially explain a rise and fall in prices, it cannot explain why prices fell far

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.

below their long-run level in the bust. This key feature of the data, which we infer from the subsequent rebound in prices, explains why adding speculators to our model could partially substitute for diagnostic behavior in amplifying the boom but cannot substitute for the role of foreclosures in the bust in driving prices below their long-run level. 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 growth for each 10 percentage points higher investor share (Table B.1 column (7) less sum of columns (1) and (4)).

Overall, these results are consistent with speculation playing a role late in the house price boom in areas such as Las Vegas and echo the result from Table 4 that price-rent growth in the boom not associated with long-run fundamentals strongly predicts

subsequent price decline. 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, all of which are the focus of our paper.

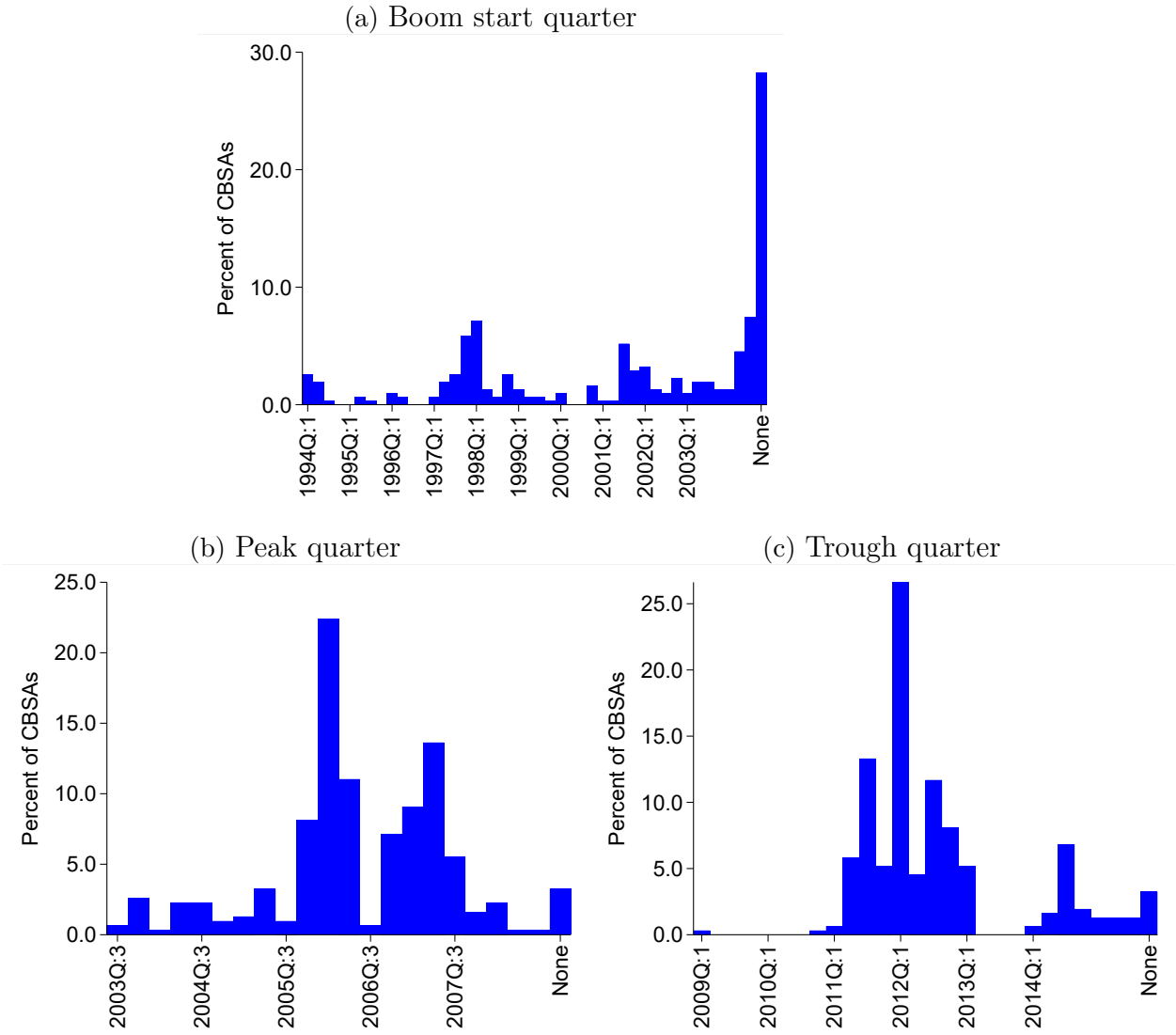
B.2 Bartik Instrument Details

We combine the CBP files provided by the Census with the files from Eckert et al. (2020) that optimally impute suppressed employment and wage 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 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.

B.3 Additional Empirical Results

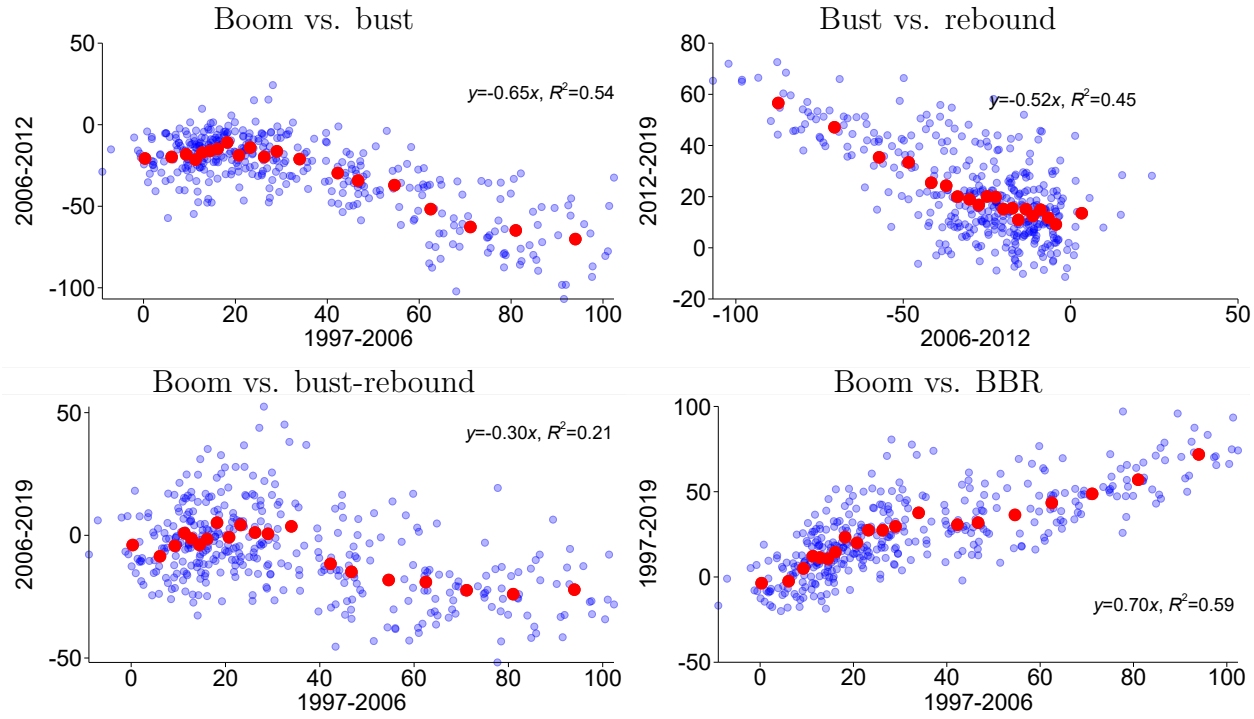
Figure B.3 plots the timing of boom starts, peaks, and bust troughs across CBSAs. Figure B.4 plots correlations of house price growth during the boom, bust and rebound across CBSAs. Table B.2 reports summary statistics. Figure B.5 displays the correlation of the long-run fundamental with measures of non-prime lending.

Figure B.3: Local Peak and Trough Timing



Notes: Panel (a) reports a histogram across CBSAs in our final sample 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 procedure in Ditzen et al. (2021). Areas without a break identified at the 95% confidence level are shown in the bar labeled "None." Panel (b) reports a histogram of the quarter with the peak in real house prices between 2003:Q2 and 2009:Q2. Areas without an interior peak in this period are shown in the bar labeled "None." Panel (c) reports a histogram of the quarter with the trough in real house prices between 2007:Q1 and 2015:Q4. Areas without an interior trough in this period are shown in the bar labeled "None."

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

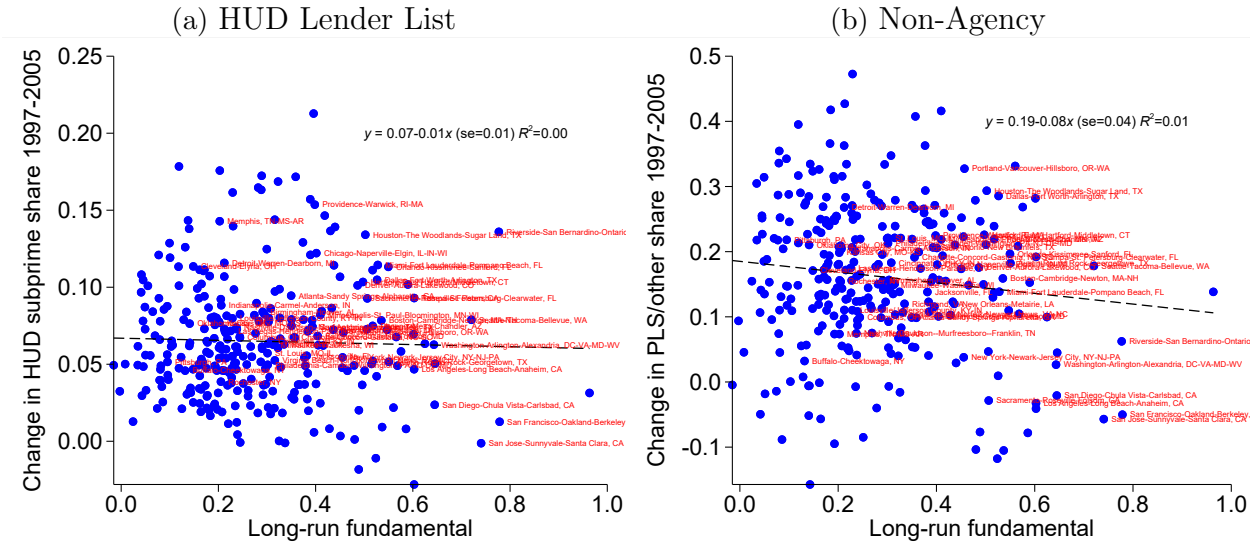


Notes: Each blue circle represents one CBSA. The red circles show the mean value of the y-axis variable for 20 bins of the x-axis variable. CBSA house price data from Freddie Mac deflated using the national GDP price index.

Table B.2: 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
House price growth 1997-2006	35.2	26.8	8.1	26.0	79.0	308
House price growth 2006-2012	−29.9	22.8	−65.9	−24.2	−5.8	308
House price growth 2012-2019	22.7	17.8	2.4	19.3	49.6	308
Census rent growth 2000-2018	54.3	8.5	41.7	54.4	64.8	272
Population growth 1997-2019	19.9	17.2	−0.1	17.7	42.3	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
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

Figure B.5: Non-Prime Credit and Fundamentals



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.

C Model Appendix

C.1 Present Value of Dividends

We restate equation (15) for convenience:

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

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

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

One can show:

$$\begin{aligned} \mathbb{E}_t \left[\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_t \right] &= \bar{\mu}(s-t) + \frac{1}{\theta}\left(1 - e^{-\theta(s-t)}\right)(\mu_t - \bar{\mu}), \\ \mathrm{Var}\left(\int_{\mathrm{t}}^{\mathrm{s}} \mu_\tau \mathrm{d}\tau | \mu_t \right) &= \frac{\sigma_\mu^2}{\theta^2}(s-t) - \frac{3\sigma_\mu^2}{2\theta^3} + \frac{\sigma_\mu^2}{2\theta^3}\left[4e^{-\theta(s-t)} - e^{-2\theta(s-t)} \right], \end{aligned}$$

giving:

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

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

Then:

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

giving the desired result.

C.2 Analytic Path of Beliefs

We solve for the mean path of beliefs m_t^φ starting from the initial condition $m_0 = \bar{\mu}$ and the initial drift rate μ_0. That is, we solve for m_t^φ if all subsequent Wiener shocks are equal to 0. From equation (13), we have:

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

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

We first solve the SDE for m_t. Substituting equations (8) and (11) into equation (10), we have:

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

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

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

Note that equation (9) implies:

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

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

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

We next solve for the mean path of $\varphi \mathcal{I}_t$. Using equations (8), (11) and (14), we have:

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

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

$(\mu_0 - m_0)\, e^{-(\kappa+\theta)s}$. Therefore:

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

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

C.3 Alternative Belief Formation

Diagnostic expectations have several attractive features for our analysis: they are a simple departure from rational learning, they are easily parameterized, different sized changes in fundamentals lead to a boom-bust-rebound of a different size but not a different length as in the data, and they are motivated by evidence in the behavioral economics literature. However, the boom-bust-rebound in response to a change in fundamentals really only requires over-optimism and learning, and any belief microfoundation that generates the same path of expected discounted dividends would generate the same results.

As an example, in this Appendix we provide an alternative rationalization of V_t based on rational Bayesian updating but an over-optimistic prior. To simplify the exposition, we consider a drift rate that can take $n = 3$ possible values μ_1, μ_2, μ_3. As in the main text, agents observe D_t but do not know the instantaneous drift rate. Each instant a shock arrives with Poisson intensity φ that changes the drift rate, with f_i the probability the drift rate changes to μ_i if a regime-shift shock occurs.

We consider a one-time change from μ_1 to μ_2. To generate over-optimism, we set $\{f_i\}$ such that agents perceive a higher probability of changing to μ_3 than to μ_2 if a regime shock occurs. Since we consider a one-time change, one may interpret each f_i as agents' subjective belief about the probability of jumping to regime i conditional on a shift occurring, giving rise to an over-optimistic posterior relative to the the true objective probability of switching to each regime.

Following Lipster and Shiryaev (1977), agents' belief $\pi_i(t)$ that $\mu(t) = \mu_i$ evolves as:

$$d\pi_i(t) = \varphi\left(f_i - \pi_i(t)\right) dt + \pi_i(t)\left(\mu_i - \bar{\mu}(t)\right) \sigma_D^{-1} d\tilde{W}(t), \tag{C.9}$$

where $\bar{\mu}(t) \equiv \sum_i \pi_i(t)\mu_i$ is the expected drift and $d\tilde{W}(t) \equiv \frac{dD(t)}{D(t)} - \bar{\mu}(t)$ is the surprise innovation given agents' current beliefs. The first term of equation (C.9) reflects the

possibility of a regime change. The second term shows that $\pi_i(t)$ will increase if the drift μ_i exceeds the current expected drift $\bar{\mu}(t)$ and the dividend increases by more than the expected drift, with the amount of updating determined by the precision of the Wiener process σ_D^{-1}.

The expected present value of dividends $P^*(D, \Pi)$, where Π denotes the current set of beliefs, satisfies:

$$P^*(D, \Pi) = E\left[\int_t^\infty e^{-\rho s} D(s) | \mathcal{F}(t)\right] = \sum_{i=1}^{n} \frac{\pi_i D}{(\rho + \varphi - \mu_i)(1 - \varphi\Delta)}, \tag{C.10}$$

where $\Delta \equiv \sum_{i=1}^{n} \frac{f_i}{\rho+\varphi-\mu_i}$. To interpret equation (15), suppose for the moment $\varphi = 0$. Without the possibility of a regime shift, the expected present value is the expectation of the Gordon growth formula, $E_i[D/(\rho - \mu_i)]$, across possible drift rate regimes. The additional terms in equation (15) account for the possibility of regime shifts.

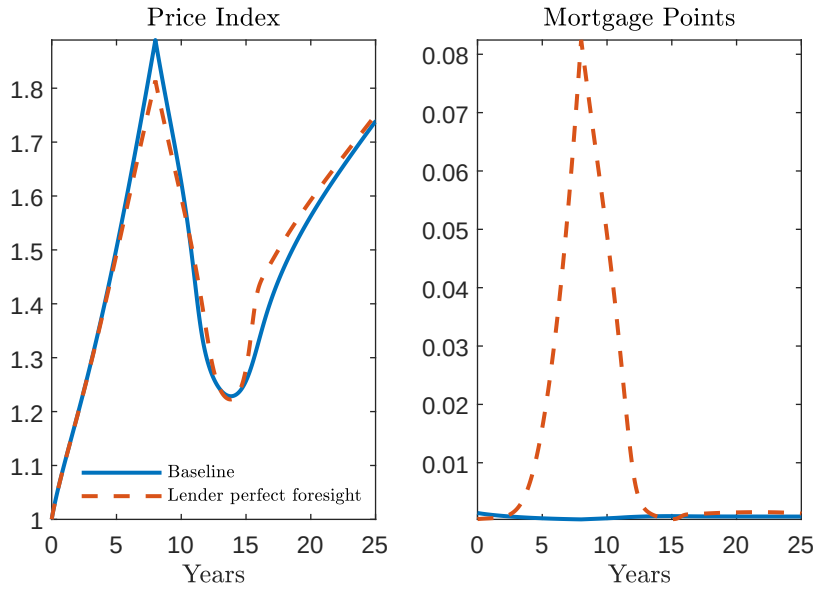
To connect equation (C.10) to equation (15), consider starting from a near-certain belief that $\mu(t) = \mu_1$ and that if the regime changes it is very likely to change to regime 3, i.e. $\pi_1(0)$ close to 1 and $f_3 >> f_2$. An agent with such a prior who observes dividend growth higher than μ_1 will update her mean belief of the growth rate but overshoot due to the optimistic prior, giving a path for P^* similar to the diagnostic case.

C.4 Lender Perfect Foresight

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.

As discussed in footnote 39, this exercise explains why changes in the cost of credit have a small effect on prices. Even when lenders perfectly anticipate the peak in buyers' beliefs and hence in prices, the rise in W_t/P_t of more than 8p.p. has a small impact on the path of prices. 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, when lenders have perfect foresight over the path of dividends.

D Computational Appendix

In this appendix, we detail the computational approach we employ for solving the model. We first discuss approximating the Fokker-Plank partial differential equation via polynomial projection methods. We then discuss how we use a sparse grids approach for our global solution method. Finally, we discuss how Monte-Carlo simulation methods are employed in our iterative global solution method to compute equilbirum mortgage pricing.

D.1 Approximating the Fokker-Plank Equation

One of the key state variables in our model is the infinite-dimensional distribution of mortgage balances. The measure density $g(M,t)$ of this distribution follows the Fokker-Plank equation:

$$\frac{\partial}{\partial t}g(M,t) = (I_t + \iota) H_t \phi(M/P_t)/P_t - \iota g(M_t)$$

where i is the arrival rate of liquidity shocks and $\phi(\cdot)$ is the LTV origination distribtuion. There is no closed form solution for this initial value problem, so we must rely on numerical techniques.

The key idea is to approximate the loan balance distribution with a Chebyshev series:

$$g\left(M,t\right)=\sum_{i=0}^{N}\alpha_{i}\left(t\right)T_{i}\left(M\right)$$

with:

$$T_{i}\left(M\right)=\cos\left(i*\arccos\left(\frac{M-\left(M^{u}+M^{l}\right)/2}{\left(M^{u}-M^{l}\right)/2}\right)\right)$$

the (scaled) i^{th} Chebyshev polynomail of the 1st kind, and M^{u},M^{l} the upper and lower bounds of the mortgage balance distribution respectively. The coefficients $\alpha_{i}\left(t\right)$ are determined by requiring the polynomial series to interpolate the true measure density at $\left(M_{1},...,M_{N}\right)$ collocation points. We set the collocation nodes to be the (scaled) Chebyshev-Gauss-Lobato (CGL) points:

$$M_{i}=\frac{1}{2}\left(M^{u}+M^{l}\right)-\frac{1}{2}\left(M^{u}-M^{l}\right)\cos\left(\frac{i\pi}{N}\right),$$

equal to the $\left(N-1\right)$ extrema of the N^{th} Chebyshev polynomial plus the endpoints. We thus require:

$$\sum_{i=1}^{N}\alpha_{i}\left(t\right)T_{i}\left(M_{j}\right)=g\left(M_{j},t\right)$$

for $j=1,...,N$.

Using the Fokker-Plank equation we find the system of differential equations governing the coefficients:

$$\sum_{i=0}^{N}\alpha_{i}^{\prime}\left(t\right)T_{i}\left(M_{j}\right)=\left(I_{t}+\iota\right)H_{t}\phi\left(M_{j}/P_{t}\right)/P_{t}-\sum_{i=0}^{N}\iota\alpha_{i}\left(t\right)T_{i}\left(M_{j}\right)$$

for $j=1,...,N$. Letting:

$$\mathcal{A}^{*}T_{i}\left(M\right)=\iota\alpha_{i}\left(t\right)T_{i}\left(M\right)$$

we have:

$$\boldsymbol{\alpha}^{\prime}\left(t\right)=T\left(\boldsymbol{M}\right)^{-1}\left[\mathcal{A}^{*}T\left(\boldsymbol{M}\right)\boldsymbol{\alpha}\left(t\right)+\left(I_{t}+\iota\right)H_{t}\phi\left(\boldsymbol{M}/P_{t}\right)/P_{t}\right],$$

where:

$$\boldsymbol{\alpha}(t) = \begin{bmatrix} \alpha_0(t) \\ \vdots \\ \alpha_N(t) \end{bmatrix}$$

$$T(\boldsymbol{M}) = \begin{bmatrix} T_0(M_0) & \cdots & T_N(M_0) \\ \vdots & \ddots & \vdots \\ T_0(M_N) & \cdots & T_N(M_N) \end{bmatrix}$$

which provides a finite-dimensional system of differential equations governing the evolution of the coefficients of the Chebyshev expansion.

D.2 Sparse Grids

Using a Chebyshev series to approximate the loan balance distribution, the state variables in the model are the current dividend D_t, current beliefs about the dividend growth rate m_t^{φ}, the current housing stock H_t, the moving average of past investment rates $\bar{I}_t$, and the vector of Chebyshev coefficients $\boldsymbol{\alpha}(t)$. Full grid methods for the global solution quickly run into the curse of dimensionality. We thus employ sparse grid techniques to get the global solution of the model. Our approach follows Judd et al. (2014).

We use the Smolyak construction for the sparse grids, once again utilizing Chebyshev-Gauss-Lobatto (CGL) points, that is extrema of Chebyshev polynomials of the 1st kind. In particlar, let d denote the number of state variables. The Smolyak construction proceeds as follows. We first extract a subsequence of unidimensional grid points $S_1, S_2, ...$ from the extrema of the Chebyshev polynomials satisfying $|S_1| = 1$ $|S_i| = 2^{i-1} + 1$ for $i > 1$ and $S_i \subset S_{i+1}$. The first such four nested sets are:

$$S_1 = \{0\}$$

$$S_2 = \{0, -1, 1\}$$

$$S_3 = \left\{0, -1, 1, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}$$

$$S_4 = \left\{0, -1, 1, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{-\sqrt{2+\sqrt{2}}}{2}, \frac{-\sqrt{2-\sqrt{2}}}{2}, \frac{\sqrt{2-\sqrt{2}}}{2}, \frac{\sqrt{2+\sqrt{2}}}{2}\right\},$$

equal to the extrema of the 1st, 3rd, 5th, and 7th Chebyshev polynomials of the 1st kind.

To form multidimensional grid points, we can take d-fold products of the unidimensional sets above. In particular, let:

$$\mathcal{K}^{\boldsymbol{i}} = \prod_{j=1}^{d} S_{i_j}$$

for $\boldsymbol{i} = (i_1, ..., i_d)$. Finally, let $\mu \geq 1$ be the order of the approximation. Then, the Smolyak sparse grid is formed as:

$$\mathcal{H}^{d,\mu} = \bigcup_{d \leq |\boldsymbol{i}| \leq d+\mu} \mathcal{K}^{\boldsymbol{i}}$$

where $|\boldsymbol{i}| = i_1 + \cdots + i_d$.

We then construct an approximation to the true global solution $f(x) : \mathbb{R}^d \to \mathbb{R}$ as:

$$\hat{f}(x) = \sum_{n=1}^{M} c_n \Upsilon_n(x)$$

where $\Upsilon_n : \mathbb{R}^d \to \mathbb{R}$ for $n = 1, ..., M$ are a set of d-dimensional basis functions. We then set the coefficients $(c_n)_{n=1}^{M}$ by minimizing the L^2-norm:

$$\boldsymbol{c} = \arg\min_{\boldsymbol{c}} \left\| f\left(\mathcal{H}^{d,\mu}\right) - \sum_{n=1}^{M} c_n \Upsilon_n\left(\mathcal{H}^{d,\mu}\right) \right\|_2,$$

where $f\left(\mathcal{H}^{d,\mu}\right), \Upsilon_n\left(\mathcal{H}^{d,\mu}\right) \in \mathbb{R}^{|\mathcal{H}^{d,\mu}|}$ are vectors which evaluate the respective functions at each element of the sparse grid $\mathcal{H}^{d,\mu}$.

D.3 Monte-Carlo Simulation and Global Solution Algorithm

The key challenge for the global solution is determining equilibrium mortgage pricing, which takes the form of mortgage points:

$$W_t = \mathbb{E}_t\left[e^{-r(\tau - t)}\left(M_t - \psi R P_\tau\right)\right].$$

The difficulty is that the mortgage points depend on the equilibrium house price function in a nonlinear way. But of course equilibrium house prices depend on equilibrium mortgage points.

We therefore follow an iterative procedure, in conjunction with Monte-Carlo simulation, to solve for the global solution. Let $W^0\left(\mathcal{H}^{d,\mu}\right)$ be an initial guess for equilibrium mortgage points on the sparse grid $\mathcal{H}^{d,\mu}$. Then at iteration j:

1. Given the current solution $W^j\left(\mathcal{H}^{d,\mu}\right)$ on the sparse grid, solve for the equilibrium price function $P^j\left(\mathcal{H}^{d,\mu}\right)$ and equilibrium investment function $I^j\left(\mathcal{H}^{d,\mu}\right)$ on the sparse grid.

2. Construct appproximants to the house price and investment functions $\hat{P}^j$ and $\hat{I}^j$ with coefficients $\boldsymbol{c}_P^j, \boldsymbol{c}_I^j$ using the procedure described in the previous subsection.

3. At each point of the sparse grid $\mathcal{H}^{d,\mu}$, use Monte-Carlo simulation with N trials to simulate house prices forward.

   (a) At each point of the sparse grid $\mathcal{H}^{d,\mu}$, simulate dividends and beliefs forward using Euler-Maruyama method for the SDE system.

   (b) Use the investment function approximant $\hat{I}^j$ to simulate forward the housing stock and the house price approximant $\hat{P}^j$ to construct house prices at each step of the simulation.

4. At each point of the sparse grid $\boldsymbol{x} \in \mathcal{H}^{d,\mu}$, compute:

$$W(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_P\left[e^{-r\tau}\left(M(\boldsymbol{x}) - \psi R P_\tau\right) 1\left[\tau_L < \tau_C\right]\right],$$

   where the expectation $\mathbb{E}_P[\cdot]$ is conditional on the simulated future house prices for that Monte-Carlo trial. This gives $W^{j+1}\left(\mathcal{H}^{d,\mu}\right)$.

If $\left\|W^{j+1}\left(\mathcal{H}^{d,\mu}\right) - W^j\left(\mathcal{H}^{d,\mu}\right)\right\| < \varepsilon$ for some specified tolerance level $\varepsilon > 0$, then terminate and construct the approximant $\hat{W}^*$ with coefficient $\boldsymbol{c}_W^*$. If not, move to iteration $j+1$ and return to Step 1. This procedure has the advantage of being highly parallelizable.

To estimate the parameters, we search over a grid which can be parallelized on a high-performance computing cluster.