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

January 2022

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, rents, and foreclosures.

*Contact information: chodorowreich@fas.harvard.edu, guren@bu.edu, tmcquade@berkeley.edu. We thank numerous seminar and conference participants and discussants for useful feedback. Sage Belz, Stefano Pica, and Lei Ma provided excellent research assistance. Chodorow-Reich thanks the Alfred P. Sloan Foundation and Harvard Ferrante Fund for generous financial support. Guren thanks the National Science Foundation (grant SES-1623801) for financial support. The computational work for this paper was performed on Boston University's Shared Computing Cluster.

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 areas experiencing the most rapid price growth during the boom also having the largest price declines during the bust. The sharpness of the cycle and its role 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 (Shiller, 2008; Charles et al., 2018).

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

We argue that this pattern reflects a larger role for fundamentals than previously thought. In a first step, we use a standard spatial equilibrium framework to motivate fundamental determinants of location choice, land costs, housing supply, and house prices. We find that these 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-Kindlebergerian" model of a fundamentally-rooted house price cycle in which belief over-reaction amplifies the boom and a foreclosure spiral exacerbates the bust, and discipline the model with our

[1] We consider developments until 2019 due to the COVID-19 pandemic dramatically altering the economy and housing market.

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 (1997-2006), bust (2006-2012), and rebound (2012-2019). At the ZIP Code level, the boom correlates negatively with the bust but the bust correlates negatively with the rebound, each with an R^2 above 0.35. As a result, the boom correlates strongly with price growth between 1997 and 2019, with an R^2 of 0.62.

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 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 produces supply elasticities in line with the existing literature. We define a city's long-run fundamental as the second stage fitted value, which depends only on the excluded instruments. This measure combines both demand and supply-side determinants of house prices using the structure of the spatial equilibrium framework. 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 these correlations. We validate our measurement following the logic of Poterba (1984) and Campbell and Shiller (1988a,b) by showing that the part of price-rent growth correlated with the long-run fundamental predicts subsequent rent growth and no price decline while the residual predicts price declines and no rent growth.

Finally, we use our hindsight framework to investigate the nature of the shocks that triggered the onset of the boom. Housing demand in our framework reflects the discounted sum of the income/amenity prospects or "dividends" to living in a city. Accelerating demand growth could arise because discount rates decline or the growth rate of the dividend for urban living accelerates, and a dividend shock could be either common or stronger in cities with a stronger cycle. We present three pieces of evidence supporting a primary role for a common dividend shock, which has a larger impact on price growth in high long-run fundamental areas already growing rapidly, consistent with the cross-sectional results. First, rent growth accelerates around the start of the price boom. Second, price growth remains higher even after the interest rate decline has stabilized. Third, price growth acceleration at the end of the boom-bust-rebound is similar for cities with different long-run fundamentals. These findings motivate the dividend growth rate shock in our model.

Section 5 contains our neo-Kindlebergerian interpretation of the boom-bust-rebound. 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 cities 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 a city's fundamental, an increase in the growth rate of the dividend from living in the city.

Two model ingredients are instrumental to this result. First, agents 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 due to *diagnostic expectations* (Bordalo et al., 2019), which make them over-react to news about fundamentals during the boom. Eventually, over-optimism peaks and beliefs start their covergence toward the true long-run growth rate, triggering the bust phase of the cycle.

On their own, however, a turn in beliefs cannot generate a bust in which prices fall below their full-information value or as steeply as in the data. This motivates the model's second key ingredient, mortgage borrowing and foreclosures. Consistent with the empirical literature (Foote et al., 2008; Ganong and Noel, 2020; Gupta and Hansman, 2021; Gupta et al., 2019), a foreclosure occurs when an under-water homeowner experiences a liquidity shock. Foreclosures add to housing supply, further depressing prices, putting more owners under-water, and leading to more foreclosures in a price-foreclosure spiral that pushes prices below their long-run level (Guren and McQuade, 2020). Finally, foreclosures recede, ongoing dividend growth causes new buyers to enter, and prices rebound.

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 empirically-estimated long-run supply elasticity 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. It also does a good job explaining long-run rent growth, which we do not target.

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. Without over-optimism, an increase in fundamentals generates a smooth, monotonic rise in prices as they converge to a higher growth path. Diagnostic expectations instead produce a slow build-up and run-down of over-optimism, consistent with survey data. The correction in beliefs on its own generates a counterfactually-small price decline, as expectations do not overshoot on the downside so prices converge smoothly toward their long-run path. Foreclosures push prices below the level that would prevail with rational learning, creating a pronounced bust and rebound consistent with the data.

The structural model also sheds further light on the nature of the fundamental shock and the circumstances under which a diagnostically-driven boom-bust-rebound is particularly severe. The estimated parameters imply a similar-sized increase in the growth rate of dividends across cities, consistent with the empirical evidence discussed already. The shock nonetheless generates a sharper acceleration of growth along the transition and a more pronounced boom-bust-rebound in high fundamental areas already growing faster in the pre-boom period. These places have a larger share of the present value accounted for by dividends farther into the future, making them more sensitive to changes in the dividend growth rate. Thus, our interpretation can account both for the national character of the housing cycle as well as the cross-sectional patterns. An important corollary is that diagnostically-driven boom-bust-rebounds are more severe when interest rates are low or when initial price growth is high. Our analysis therefore implies that a low *level* of interest rates was critical to the severity of the cycle.

Related literature. Our paper differs from other work on asset cycles due to our focus on the rebound – both in using the rebound as an important data point to infer the source of the cycle and in presenting a model that generates a boom-bust-rebound rather than a boom-bust or persistent oscillation.

Theoretically, prior work has shown that belief disagreement, over-optimism, or extrapolation in asset markets can generate a boom-bust and that learning and certain types of extrapolative expectations can generate oscillations that dampen over time, but to our knowledge our paper is the first to endogenously generate a boom-bust-rebound.[2] The most closely related work on housing is Glaeser and Nathanson (2017), who present a behavioral microufoundation for extrapolative expectations that lead to price oscillations, and Burnside et al. (2016), who study boom-busts due to optimism that spreads through "social dynamics." Furthermore, our model of learning about fundamentals generates a boom-bust-rebound endogenously from a single shock, 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),

Empirically, prior work on housing has primarily considered the contribution of specific factors in a particular phase of the cycle such as changes in credit conditions, speculators, over-optimism, or foreclosures.[3] Structural work has also sought to ascertain the relative roles of credit, speculators, foreclosures, financial conditions, and house price expectations in the boom-bust (Kaplan et al., 2020; Favilukis et al., 2017; Greenwald and Guren, 2021;

[2]See for instance Barberis et al. (1998), Scheinkman and Xiong (2003), Adam et al. (2017), Barberis et al. (2018) for examples of a boom-bust and Hong and Stein (1999) and Glaeser and Nathanson (2017) for examples of oscilations. Barberis (2018) reviews this literature.

[3]Justiniano et al. (2017), Johnson (2020), Foote et al. (2019), Mian and Sufi (2009), Glaeser et al. (2008), Favara and Imbs (2015), and Di Maggio and Kermani (2017) study changes in credit conditions in the boom. Chinco and Mayer (2016), Nathanson and Zwick (2018), Gao et al. (2020), Bayer et al. (2020), DeFusco et al. (2017), and Mian and Sufi (2017) study the role of speculators. 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. Guren and McQuade (2020), Mian et al. (2015), and Hedlund (2016) study the role of foreclosures in the bust.

Justiniano et al., 2019; Landvoigt et al., 2015).

Most of this literature has not focused on the role of fundamentals.[4] There are a few notable exceptions. 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) show that the timing of the beginning of the boom in each city was "fundamentally based to a significant extent" but that fundamentals revert in roughly three years. More recently, Howard and Liebersohn (2021a) and Howard and Liebersohn (2021b) examine divergence in demand across areas and Schubert (2021) identifies spillovers of fundamentals across cities via migration networks. Finally, a literature following Kaplan et al. (2020) and Glaeser and Nathanson (2017) includes shocks to future fundamentals to microfound changing expectations but do not treat fundamentals as a measured and central feature of the cycle.

Methodologically, our empirical analysis builds on an urban economics literature on the long-run determinants of housing costs and location choice. Our approach most closely follows Saiz (2010), although we extend his framework to incorporate observed land shares, relax constraints, and apply it to a different time period. Our theoretical framework builds on work in macroeconomics and finance that focuses on the role of learning and behavioral expectations in financial markets (David, 1997; Veronesi, 1999) and in particular builds on the role of diagnostic expectations formalized in Bordalo et al. (2019, 2020).[5] Solving

[4]For instance, Charles et al. (2018) describe a "consensus that much of the variation in housing prices during the boom and bust derived from a speculative 'bubble' and not from changes in standard determinants of housing values such as income, population, or construction costs." Similarly, Glaeser (2013) points out that the " economy was not growing particularly swiftly [in the boom]" and "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.

[5]In related work, Jacobson (2020) generates overreaction in prices in the boom from learning together with a lack of historical data on the boom state, which leads agents to under-predict the mean reversion in house prices. She does not consider the bust or rebound and relies on exogenous changes in the aggregate state to generate turning points in the cycle. Kindermann et al. (2021) also study learning in housing markets but emphasize heterogeneity by tenure status.

our model necessitates methodological innovations that may prove useful in other contexts 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 the seminal work of Kindleberger (1978). For instance, Kindleberger writes that "virtually every economic mania is associated with a robust economic expansion."[6] As in Kindleberger, our focus on long-run fundamentals does not imply that there was no "housing bubble," nor does it refute important roles for other factors. It does, however, suggest that other factors such as optimism, credit, and foreclosures were rooted in long-run fundamentals. Our work expands on Kindleberger by emphasizing the rebound and its role in diagnosing the driving forces of the full cycle and applying these ideas to a structural model of the 2000s housing 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.

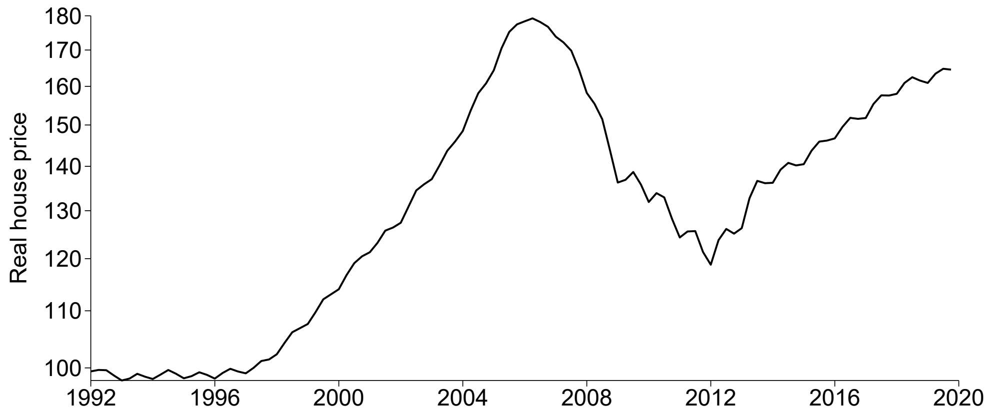
Figure 1 shows the national Case-Shiller house price index, deflated using the GDP price index. The series begins to rise in the late 1990s, peaks in 2006Q2, reaches a local trough in 2012Q1, and then grows again through the end of our sample in 2019. We use this timing to define the boom, bust, and rebound periods for the remainder of the paper.[7]

Figure 2 shows the correlation of the boom, bust, and rebound at the local level. Each blue circle represents one ZIP Code and the overlaid red circles show the mean value of

[6] Barberis et al. (2018) also emphasize this feature of Kindleberger, but they interpret the fundamental improvement as a sequence of positive news about fundamentals (which may or may not be wrong) rather than an actual fundamental improvement.

[7] Appendix Figure B.3 reports the timing of the boom start across the cities in our sample, using a procedure similar to Ferreira and Gyourko (2018). Because very few cities have booms that start before the year 1997, we use that year as the boom start for our cross-region analysis. Appendix Figure B.3 displays much more uniformity in the timing of the price peak and trough across cities.

Figure 1: National Boom, Bust, and Rebound



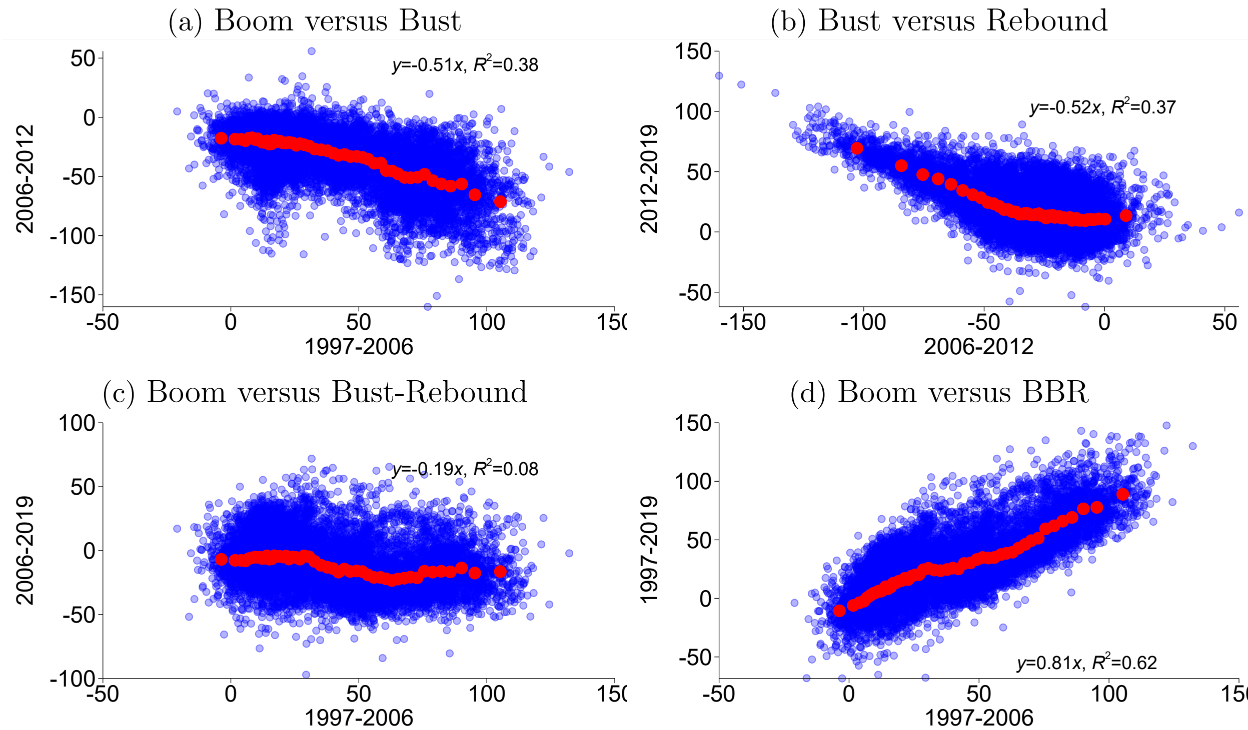
Notes: The figure shows the national Case-Shiller index deflated by the GDP price index, normalized so that 1997=100.

the y-axis variable in each of 50 quantiles of the x-axis variable. Panel (a) shows the correlation of price growth in the boom and the bust. Each additional percentage point of house price appreciation in the boom is associated with an additional decline of 0.51 percentage point in the bust and the R^2 of this relationship is 0.38.

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

In the remainder of this paper, we link these patterns to fundamental determinants of house price growth. The next two sections describe and implement a long-run framework

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.

to measure an area's fundamental and show that a higher long-run fundamental correlates with a larger boom-bust-rebound cycle. We then provide a quantitative account of why fundamental growth may give rise to a boom-bust-rebound.

3 Long-run Fundamentals: Framework

This section introduces a long-run supply-and-demand empirical framework for house prices and describes our data. The framework disciplines the choice and weighting of the variables that form our long-run house price fundamental, allowing us to go beyond correlations and to associate areas that grew faster over the full 1997-2019 period with fundamental determinants of house prices. We in turn show that the long-run fundamental correlates positively with a stronger boom-bust-rebound pattern. As a by-product, we recover a long-run housing supply elasticity that we use to calibrate the model in Section 5.

3.1 Structural System

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

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

where $P_{i,t}$ is the price of a house in area i at date t, $C_{i,t}$ is the construction cost of the structure, and $L_{i,t}$ is the land cost.[8] Gyourko and Saiz (2006) argue that the construction sector is sufficiently competitive to justify this decomposition.

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

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

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

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

[8] Other proportional costs such as broker fees drop out when we take log changes below.

[9] 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) rises to induce new housing in less desirable locations.

reflect long-run fundamentals.

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

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

Here G_i is a constant, γ is the elasticity of demand, ρ_t is the discount rate, and $D_{i,t}$ is the "dividend" that captures the value of income/amenities in city i at date t. Equation (4) is a dynamic Rosen (1979) and Roback (1982) spatial equilibrium condition in which households choose location to maximize their earning potential and amenity value net of housing costs. Section 5 provides a microfoundation for this functional form (see equation (18)); intuitively, $V_{i,t}/P_t$ is the Tobin's Q associated with an area. If the dividend grows at a constant rate μ_i, then the present value has the Gordon growth representation $V_{i,t} = D_{i,t}/(\rho - \mu_i)$. The Gordon growth representation clarifies that $V_{i,t}$ increases if discount rates fall or if dividend growth rates rise. Furthermore, $V_{i,t}$ is *convex* in $\rho - \mu_i$, so that a given change in ρ or μ_i has a bigger effect when ρ is initially lower or μ_i is initially higher. This duration effect will prove important in our analysis of the nature of the fundamental shock in Section 4.5.

We next turn to a regression specification that allows us to measure the extent to which heterogeneous growth in $V_{i,t}$ and in the long-run supply elasticity η_i explain the cross-section of price growth over 1997-2019.

3.2 Regression Specification

We start by translating equations (1) to (4) into a regression specification for house prices. Taking logs of equation (1), differencing over time, and letting lower case p, c, ℓ, a, b denote the log differences of their respective upper case letters, we obtain:

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

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

Equation (5) decomposes house price growth into a weighted average of the growth of construction costs and land costs and appears for example in Davis and Palumbo (2008). Equation (6) imposes our functional forms for construction and land costs.

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 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, $\hat{x}_{i,t} \equiv x_{i,t} - \bar{x}_t$, and $\epsilon_{i,t} = \hat{a}_{i,t} + s_{i,t}\left(\hat{b}_{i,t} - \hat{a}_{i,t}\right)$. This yields:

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

Equation (7) corresponds to the regression equation:

$$p_{i,t} = c_0 + c_1 s_i + c_2 h_{i,t} + c_3\,(s_i \times h_{i,t}) + c_4\,(m_i \times h_{i,t}) + c_5\,(m_i \times s_i \times h_{i,t}) + e_{i,t}, \tag{8}$$

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

Equation (8) 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.[10] The coefficient c_1 identifies the average excess

[10]The Saiz (2010) model starts with equations (1) to (3) with $\alpha_i = 0\ \forall i$. In our notation, the final specification in Saiz (2010) is $p_{i,t} - c_{i,t} = k_1\,(1 - \Lambda_i) \times h_{i,t} + k_2 \ln H_{i,0} \times (1 - \Lambda_i) \times h_{i,t} + k_3 \times m_i \times h_{i,t} + e_{i,t}$, where Λ_i denotes the share of land available for development. Our approach instead follows the motivation

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

The endogeneity of population growth, land share, and regulatory strictness pose a hurdle to consistent estimation of the parameters. A simultaneity problem arises because $\epsilon_{i,t}$ in equation (8) contains city-specific, secular (i.e. unrelated to population) changes in land costs or construction costs that increase house prices and reduce population growth via the spatial equilibrium relationship (4). Moreover, if land share or zoning regulation respond to population or prices, these variables also are endogenous in equation (8).

3.3 Data, Measurement, and Excluded Instruments

We estimate equation (8) across 308 metropolitan core-based statistical areas (CBSAs) over the period 1997-2019.[11]

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 robustness to using other HPIs. 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), who use appraisal data that separate

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.

[11] 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.

house prices into land costs and construction costs. 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.[12] Table B.3 presents summary statistics.

Excluded instruments. Equation (8) contains three endogenous variables — population growth, land share, and regulatory strictness — and their interactions. We heuristically group excluded instruments by which endogenous variable they most closely affect and then interact these groups as well to produce our final excluded instrument set. We closely follow the existing literature and in particular Saiz (2010) and Diamond (2016) in choosing excluded instruments to make the point that a fundamentally-rooted interpretation of the 2000s cycle does not depend on our introducing novel instruments.

Equation (4) motivates excluded instruments for population growth, which form the basis for the demand-side components of the long-run fundamental. Specifically, labor demand and amenities that shift the dividend growth rate $\mu_{i,t}$ across areas constitute valid demand shifters to identify the long-run supply elasticity. We follow Saiz (2010) and use shift-share predictors of employment and wage growth as labor demand shifters (see Appendix B.2 for details). We use mean January temperature, January sunlight, and July humidity from the Department of Agriculture as climate-related amenity shifters that capture population movement toward the "sunbelt." We finally use two amenity variables proposed by Diamond (2016), the share of the 1990 population with a Bachelor's degree

[12] 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.

and the 1997 share of employment in restaurants.

We also follow Saiz (2010) in choosing instruments relevant for the land share of the price and regulatory strictness, which form the basis for the supply-side components of the long-run fundamental. For land share, we use the fraction of land available for development (not water or a slope steeper than 15%) from Lutz and Sand (2019) and 1997 population density.[13,14] 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.[15]

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

Instrument Persistence and Interpretation The excluded instruments are either quite persistent or time invariant.[16] Consistent with this fact, areas with higher predicted house price growth over 1997-2019 based on these measures also had on average faster house price and population growth over the period from 1975 to 1997.

Nothing in the econometric setup precludes using persistent instruments to identify the long-run supply elasticity. Instead, the framework simply requires demand instruments

[13]The Alonso (1964), Muth (1969), and Mills (1967) intra-city spatial equilibrium microfoundation of equation (3) provides a formal motivation for these instruments (see equation (A.8) in Appendix A.2).

[14]Lutz and Sand define the CBSA boundary as the polygon containing the CBSA plus a 5% buffer and argue that this improves on Saiz (2010). We also confirm robustness to using Saiz's original measure.

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

[16]Even the labor demand shift-share instruments exhibit high persistence, reflecting slow-moving industry trends. A regression of the predicted employment growth instrument on a shift-share for predicted employment growth over 1986-1996 has an R^2 of 0.35 (0.64 if population-weighted).

that shift population growth and are orthogonal to unobserved, location-specific supply shifters. Persistence in the instruments associated with land share and regulation occurs naturally, since supply elasticity is a persistent feature of areas. We show in Appendix Figures B.5 and B.6 that our results do not only reflect differences in supply elasticity and hence sensitivity to common housing demand shocks and defer to Section 4.5 a discussion of the more subtle issue of what made the income/amenity present value grow faster starting in the late 1990s.

4 Long-run Fundamentals: Results

This section contains our main empirical results. Section 4.1 shows that the excluded instruments strongly predict population growth, land share, regulatory strictness, and long-run house price growth over 1997-2019. Section 4.2 reports the IV results. Section 4.3 defines the long-run fundamental as as a linear combination of the excluded instruments and shows that a higher fundamental predicts: (i) faster long-run price growth; (ii) a larger amplitude of the boom-bust-rebound; and (iii) more severe foreclosure crisis. Section 4.4 validates that our measure reflects fundamentals. Finally, Section 4.5 presents evidence that the fundamental improvement was in large part due to an increase in the growth rate of the dividends that is common across cities.

4.1 Reduced-Form Results

Appendix Table B.4 reports first-stage-type regressions for each endogenous variable separately, using only the excluded instruments motivated by that variable and also using the full set of uninteracted instruments.[17] In brief, the instruments act as expected and strongly predict their respective endogenous variables, with effective F-statistics ranging

[17]This establishes the explanatory power of the instruments without broaching many-instrument asymptotics, a subject we address in robustness exercises in the appendix.

from 14.3 to 64.5.

Figure 3 plots the fitted value from regressing house price growth over 1997-2019 on all of the uninteracted instruments (column (7) of Table B.4) against actual house price growth in various sub-periods. Panel (a) shows a strong reduced form fit over the full BBR, illustrating that fundamental drivers of location choice, land share, and regulation measured prior to the start of the boom explain a substantial amount of the variation in house price growth over 1997-2019. Panels (b)-(d) show that higher predicted long-run growth correlates positively with 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.

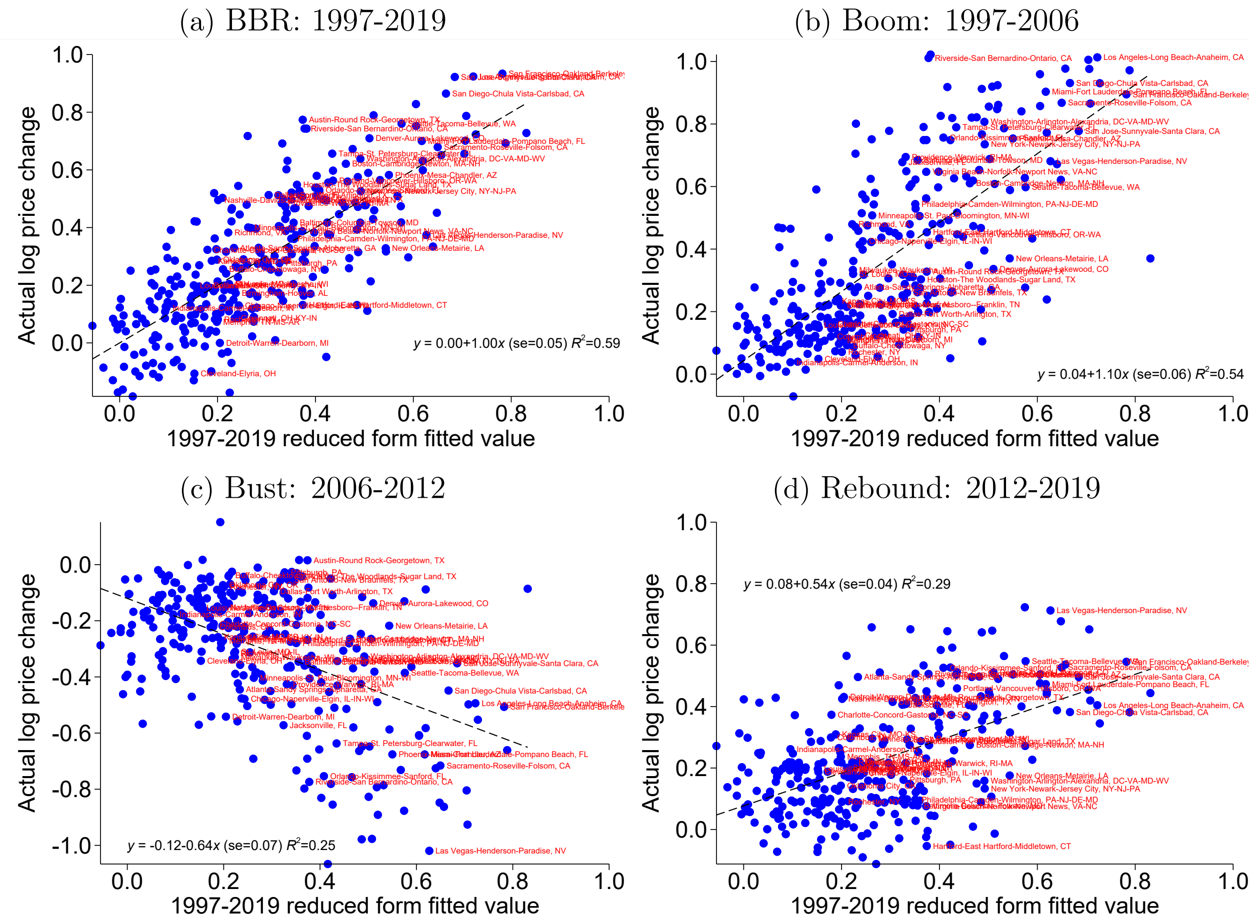
4.2 Structural IV Results

The full IV results impose additional structure 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 use to calibrate the model in Section 5.

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

Columns (2) and (3) report our main IV results. Column (2) reports the IV specification using all of the excluded instruments. Given the large number of interaction terms and instruments, the data are unable to tightly identify each interaction in column (2). Column (3) consequently constrains the coefficients on land share $\times$ population growth and WRLURI $\times$ population growth, which have insignificant coefficients close to zero in

Figure 3: Correlation with Long-run Fitted Value



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 B.4. CBSAs with more than 1 million persons in 1997 are labeled in red.

column (2), to be zero.[18] Imposing these restrictions results in tightly identified coefficients without sacrificing fit: The remaining coefficients remain relatively unchanged with much smaller standard errors, and the R^2 and inverse supply elasticity are unchanged. Column (3) is thus our preferred specification.

Several features merit comment. The impact of population growth and the average inverse elasticity are slightly larger than OLS, consistent with the expected bias of OLS

[18]The supply framework interprets zero coefficients on these variables as $\beta_0 = \alpha_0$ and $\alpha_1 = 0$ in equation (7), 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.

Table 1: 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

Notes: 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 as in equation (8). The standard error of the elasticity at the mean of land share is computed using the delta method. Heteroskedastic-robust standard errors in parentheses. **,* ,+ denote significance at the 1, 5, and 10 percent levels, respectively.

due to area-specific cost shocks. The R^2 value of 0.40 reveals strong explanatory power of land share, population growth, and WRLURI when imposing the IV coefficients. The IV results thus again illustrate the central message 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.

The values of the structural coefficients are also noteworthy. The coefficient on the 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 coefficient 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 106 log points suggests a significant role for nationwide forces affecting land prices, such as the secular decline in interest rates or an increase in the premia to living in the more expensive city center, the latter for example potentially due to the widespread fall in crime rates in the mid-1990s (Pope and Pope, 2012).[19] In Section 4.5 we use price and rent data to distinguish between these classes of explanations. The ability of our *cross-sectional* specification to reveal an *aggregate* trend in land prices stems from having started with a structural system.

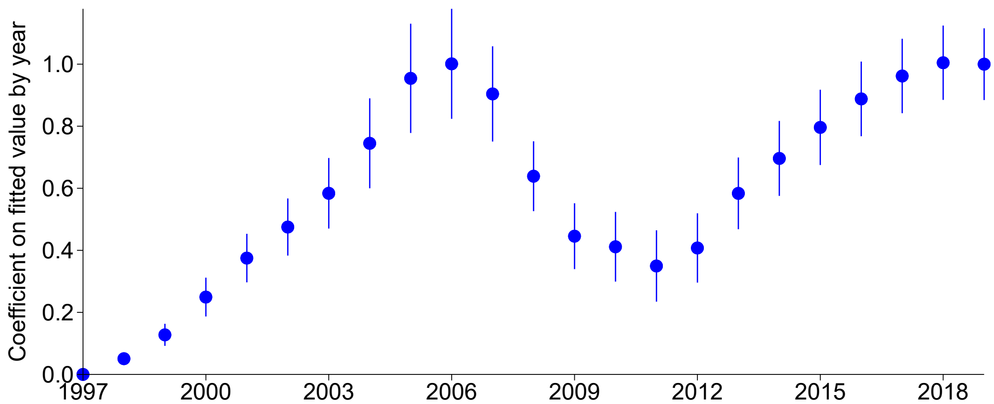
Appendix Table B.2 collects several alternative specifications that address various potential concerns with the baseline IV regression. In brief, the coefficient estimates remain similar (i) using alternative house price indexes; (ii) using the Saiz (2010) measure of land unavailability or housing units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

4.3 Long-Run Fundamental and Boom-Bust-Rebound

We define the long-run fundamental as the second stage fitted value corresponding to column (3) of Table 1. 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

[19] 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 1. 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 1 will overstate the contribution of rising land prices commensurately. Note however that the second stage fitted value and elasticity would not change.

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 1.

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.

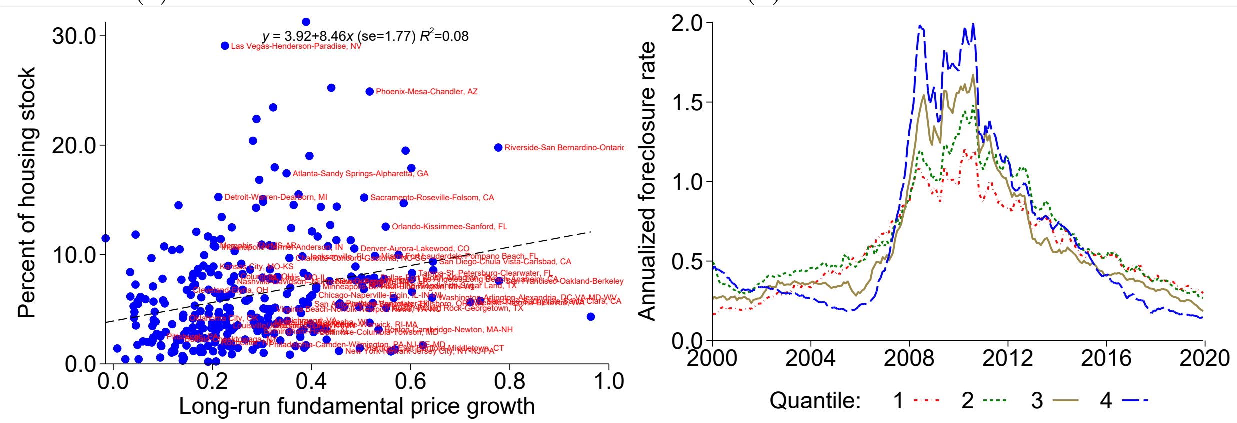
Figure 4 shows that areas with higher long-run fundamentals have more pronounced booms, busts, and rebounds. Each point corresponds to the coefficient $\{\beta_{1,h}\}$ from a cross-sectional regression of log real price growth between 1997 and 1997+h on the long-run fundamental. The boom-bust-rebound cycle in these coefficients accentuates our emphasis on a single, fundamental, driver of the whole cycle.[20]

We next show that a higher fundamental correlates with a more severe foreclosure crisis. We obtain proprietary data on completed foreclosures at the CBSA level over 2000-2019 from CoreLogic. Panel (a) of Figure 5 plots total foreclosures over 2007-2012, relative to the 2006 housing stock, against the long-run fundamental. 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

[20] In Appendix B.4 we show a version of Figure 4 that controls for a city's sensitivity to the regional house price index, calculated as in Guren et al. (2020) for the 1978-1996 period. The results are largely unchanged, indicating that our finding is not driven by cities that are historically more sensitive to aggregate dynamics.

Figure 5: Foreclosures and Long-run Fundamental

(a) Foreclosures 2007-2012 (b) Foreclosures Time Series



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 1, 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 1. The foreclosure data come from CoreLogic, whose coverage 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, extended using the growth rate of units from the Census.

relationship exceeds 4.5.[21] 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.

4.4 Validating Our Long-Run Fundamental Using Prices and Rents

The behavior of rents can validate our measurement of fundamentals. Indeed, rents are closely related to the "dividend" of living in the city and the rise in the ratio of house prices

[21] 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 Section 5.7.

to rents was perhaps the feature of the boom most responsible for real-time diagnoses of bubble-like characteristics (Shiller, 2008). We show that the part of the price-rent ratio associated with long-run fundamentals predicts subsequent rent growth and not future price decline, consistent with this component being fundamentally-based.

We draw on two sources of rent data: the decennial Census/ACS and the BLS CPI's measure of tenant rent. Each has advantages and flaws. The Census/ACS cover a large number of CBSAs but contain raw rents without quality adjustment and no data exist between 1990 and 2000. CPI rent indexes use repeat-sampling to quality adjust but are available for only 17 CBSAs over the period 1997-2019 and contain important methodological deficiencies prior to 1988 (Crone et al., 2010). Nonetheless, the two data sets give consistent implications.

We first establish that areas with higher long-run fundamentals also have faster rent growth over the full BBR. Columns (1) and (2) of Table 2 show this relationship using the Census and CPI data, respectively. Our model in Section 5 will match this pattern.

We next turn to the price-rent ratio. Column (3) shows that CBSA-level price-rent increases in the boom correlate positively and strongly with the long-run fundamental.[22] Columns (4) and (6) show that unconditionally, higher price-rent growth during the boom forecasts both higher subsequent rent growth and future relative price declines. Columns (5) and (7) 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 our empirical long-run fundamental predicts even faster subsequent rent growth than in column (4) and

[22] Columns (3)-(9) use Census/ACS rent data to maximize coverage. The BLS CPI data also show a strong relationship between the long-run fundamental and the rise in the price-rent ratio between 1997 and 2006, with an R^2 of 0.61 across the 23 cities with rent data in those years.

Table 2: Rent and Price-Rent Growth

Dep. var.:	Census Rent 2000-2019	CPI Rent 1997-2019	Price-Rent 2000-06	Rent 2006-2019			HPI 2006-2019		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Long-run fundamental	1.41^{**}	3.50^{**}	0.68^{**}						
	(0.15)	(0.45)	(0.08)						
Price-rent 2000-2006				0.11^{**}			-0.40^{**}		
				(0.03)			(0.05)		
Predicted price-rent					0.30^{**}			0.01	
					(0.06)			(0.12)	
Residual price-rent						0.03			-0.57^{**}
						(0.03)			(0.05)
R^2	0.25	0.78	0.29	0.07	0.15	0.00	0.24	0.00	0.35
Observations	272	17	272	272	272	272	272	272	272

Notes: Column (1) regresses the annualized log point change in rents between 2000 and 2019 using Census/ACS on the long-run fundamental. Column (2) regresses the annualized log point change in rents between 1997 and 2019 using BLS CPI on the long-run fundamental. Column (3) regresses the log change in the ratio of prices (Freddie Mac) to mean rents (Census, ACS) between 2000 and 2006 on the long-run fundamental. Columns (4) and (5) regress rent and price growth between 2006 and 2019 on the log change in the price-rent ratio between 2000 and 2006. Columns (6) and (8) regress rent growth between 2006 and 2019 on the fitted value and residual from column (1), respectively. Columns (7)-(9) mirror columns (2)-(4) but replacing the dependent variable with house price growth between 2006 and 2019. Heteroskedastic-robust standard errors in parentheses. In columns (3), (4), (5), (6), standard errors are computed by re-estimating column (3) of Table 1 and column (1) in 250 bootstrap samples. ** denotes significance at the 1 percent level.

no subsequent price decline, which is consistent with it reflecting fundamentals.

Columns (4) and (7) show that the part of price-rent growth not explained by long-run fundamentals predicts no subsequent rent growth and large subsequent price declines. Thus, the part of price-rent increases not correlated with long-run fundamentals looks very bubble-like *ex post*. These columns make clear that our empirical evidence admits the possibility of other aspects of the housing boom not associated with long-run fundamentals, such as speculation (Chinco and Mayer, 2016; Nathanson and Zwick, 2018; Gao et al., 2020; Bayer et al., 2020; DeFusco et al., 2017) and belief contagion (Burnside et al., 2016;

Bailey et al., 2018; DeFusco et al., 2018).

4.5 The Nature of the Shock Driving the Cycle

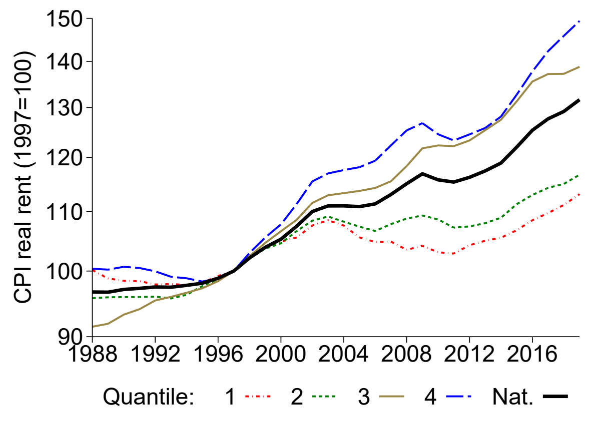
Our main result is that housing fundamentals — the present value of income/amenities $V_{i,t}$, the long-run supply elasticity η_i, and their interaction — explain the cross-section of price growth over 1997-2019 and which areas had larger boom-busts-rebounds. However, it is reasonable to ask if the data can elucidate more about the nature of the shock driving the cycle. Here we show that the behavior of prices and rents help distinguish among classes of explanations. We frame our discussion using the Gordon growth representation $V_{i,t} = D_{i,t}/(\rho_t - \mu_{i,t})$, which shows that declining interest rates or rising dividend growth could cause $V_{i,t}$ to increase. Across cities, heterogeneous changes in dividend growth can cause variation in the growth of $V_{i,t}$, but so can common declines in discount rates or increases in dividend growth due to heterogeneity in initial growth rates interacting with the convexity of $V_{i,t}$ in $\rho_t - \mu_{i,t}$. Intuitively, changes in ρ or μ have the largest impact on long-dated dividends and long-dated dividends are a larger share of $V_{i,t}$ when $\rho - \mu$ is small, similar to the duration effect of bond and equity pricing. After introducing rents into our model, we present evidence supportive of a common change in dividend growth.

Adding Rents to the Framework We first augment the supply and demand framework (1) to (4) to incorporate rents and give a useful result that distinguishes changes in dividend growth from declining discount rates. To do so, we add a Poterba (1984) style user cost condition that determines rents on the balanced growth path:

$$R_{i,t} = (\rho_t - \mu_{i,t})\, P_{i,t} + \zeta D_{i,t}, \tag{9}$$

where $R_{i,t}$ is rent, ρ_t is the discount rate, $\mu_{i,t}$ is the growth rate of the dividend on a balanced growth path and hence also expected house price growth (see Appendix C.3),

Figure 6: Rent Growth By Quartile of the Fundamental



Notes: The figure plots indexes of annual average rent deflated by the national GDP price index for 22 CBSAs with CPI rent data available from 1988, grouped into quartiles based on the long-run fundamental, with the quartiles defined using the full sample.

and ζ represents costs of owning relative to renting such as taxes or maintenance that are assumed to be proportional to the dividend. Equation (9) arises from indifference between renting and owning; see Appendix C.4 for a derivation.

Lemma 1 characterizes how a shock to either the discount rate ρ or the dividend growth rate μ_i affects the balanced growth path of rents (see Appendix C.4 for a proof):

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

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

Changes in the discount rate only do not impact rent growth on the balanced growth path as the increase in the price (due to the jump in $V_{i,t}$) exactly offsets the decline in the rent-price ratio. By contrast, changes in the income/amenity growth rate increase the growth rate of rents in the new balanced growth path. One can thus use the behavior of rents to distinguish between discount rates and dividends.

Elucidating the Nature of the Shock We now present three pieces of evidence that point to a common change in income/amenity dividends, which affected cities with higher initial dividend growth more due to convexity. First, we use Lemma 1, which states that an acceleration in rent growth signifies a change in the income/amenity part of the fundamental. Figure 6 shows time series for rents using the BLS CPI real rent data for all cities and for quartiles of CBSAs grouped by their long-run fundamental. There is a visible trend break in the late 1990s that persists through 2019 and is stronger in higher-fundamental quartiles. Table 3 formally investigates the acceleration in rent growth using a Bai and Perron (1998) structural break test. The upper panel shows that the national break in rent growth in 1997Q3 closely coincides with the break in price growth in 1998Q2. The lower panel shows that among the 17 CBSAs with a statistically significant price break between 1994Q1 and 2000Q4 and with CPI rent data, 10 have statistically significant breaks in rent growth within two years on either side of the price break. Moreover, these cities have larger and more statistically significant mean jumps in price at the timing of their breaks. Rents thus point towards fundamental improvement due to higher dividend growth.

Second, real price growth at the end of the rebound remained higher than before 1997 even though interest rates had stabilized. Panel (a) of figure 7 plots the long run fundamental on the x axis and annualized house price growth from 2016 to 2019 relative to 1975 to 1997 on the y axis. The positive intercept indicates that real price growth at the end of the rebound remained faster than in the pre-boom period in almost all CBSAs. Discount rates could rationalize higher growth at the end of the sample if they continued

Table 3: Rent Break Timing

National price growth break:			1998Q2	
National CPI rent growth break:			1997Q3	

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

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

to fall, but real interest rates *rose* slightly between 2016 and 2019. More generally, interest rate changes co-move positively with house price changes during the 1997-2019 period.[23]

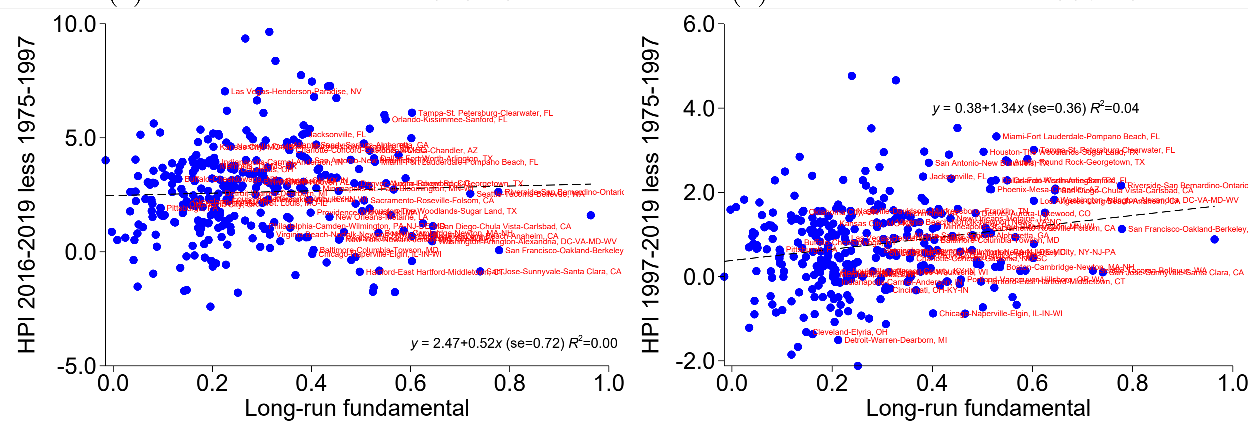
Third, Panel (a) also shows that areas with higher long-run fundamental growth were not on higher price growth trajectories relative to their pre-1997 trajectories at the end of the sample (when transition dynamics have mostly dissipated), while Panel (b) shows that price growth over the full 1997-2019 period accelerated by more relative to 1975-1997 in high fundamental areas. These two facts are consistent with convexity causing more price growth along the transition in these areas in response to a common change in dividend growth.

These facts all support a key role for a common shock that generated higher growth along the transition in already high-growth areas due to convexity. This conclusion echoes

[23]The real 10 year rate as calculated by the Federal Reserve Bank of Cleveland falls by 2.7p.p. between 1997 and 2019, with 2.3p.p. of the decline occurring between 2006 and 2012 while house prices also fell. TIPS data attribute an even larger decline to the 2006-2012 period. Similarly, Glaeser et al. (2012) and Adelino et al. (2020) argue that the semi-elasticity of prices to mortgage rates is too small to rationalize a dominant role for interest rates in the boom given the observed magnitude of the decline in mortgage rates. We consider extensive margin credit changes as a separate channel in Section 5.7.

Figure 7: Price Acceleration Relative to Pre-BBR

(a) Price Acceleration 2016-19 (b) Price Acceleration 1997-19



Notes: The figure plots annualized log price growth in 2016-2019 (left panel) or 1997-2019 (right panel) less 1975-1996 against the long-run fundamental, with CBSAs with more than 1 million persons in 1997 labeled in red.

the finding from Table 1 of a common change in land value across areas. Candidate forces that may have changed in the late 1990s include the decline in crime and revitalization of the center city mentioned above and the rising urban wage premium for educated workers (Autor, 2019). Our model will encompass these forces by focusing on a common change in the income/amenity growth rate μ as the driving force at the start of the boom.

5 A Quantitative Neo-Kindlebergerian Model

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

As in Kindleberger's celebrated *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 a city. Agents learn about the growth rate by observing the history of dividends but become overly optimistic, which we formalize using diagnostic expectations as in Bordalo et al. (2019). Construction and prices boom as optimistic buyers enter. 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. 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 set of cities each with population $H_{i,t}$ and a residual "hinterland." In what follows we describe the joint determination of population, house prices, mortgage distribution, and foreclosures in each city and suppress the i subscript for convenience.

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

$$dD_t = \mu_t D_t dt + \sigma_D D_t dW_{D,t}, \tag{10}$$

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

$$d\mu_t = \theta\left(\bar{\mu} - \mu_t\right)dt + \sigma_\mu dW_{\mu,t}, \tag{11}$$

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

A Bayesian agent would form beliefs over μ_t from the path of observed dividends. Let $\mathcal{F}_t = \sigma\left\{D_s : -\infty \leq s \leq t\right\}$ denote the information set at time t. Applying the Kalman-Bucy filter, the Bayesian posterior belief has a normal distribution, $h_t\left(\mu_t|\mathcal{F}_t\right) \sim \mathcal{N}\left(m_t, \sigma_m^2\right)$, where the current mean belief, m_t, follows the process:

$$dm_t = \theta\left(\bar{\mu} - m_t\right)dt + KdB_t, \tag{12}$$

and where the surprise innovation dB_t follows:

$$dB_t = \sigma_D^{-1}\left(dD_t/D_t - m_t dt\right). \tag{13}$$

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.[24]

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 (12) and (13). Bordalo et al. (2019) propose *diagnostic expectations* as one

[24]We assume the asymptotic variance in the posterior drift rate, which satisfies $\sigma_m^2 = \sigma_\mu^2/\left(2\theta + K/\sigma^D\right)$. Recalling that we have suppressed the i subscript indexing cities, the notation here reflects an assumption that agents learn only from the dividends of a single city, even if the Wiener shocks across cities were correlated. Because learning from correlated shocks has an effect similar to a rescaling of the posterior variance and hence the Kalman gain, this assumption does not materially impact the results that follow.

such departure that formalizes the representativeness heuristic of Tversky and Kahneman (1983). The representativeness heuristic causes agents to overweight the likelihood of a trait in a class when that trait has a higher likelihood in the class than in a reference population. Bordalo et al. (2018) give as an example the share of Irish with red hair: Red hair is more prevalent among Irish than non-Irish, and as a result people overestimate the share of Irish with red hair. In the context of asset price cycles, the reference population is the history of observed dividends and the class is 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:

$$h_t^{\varphi}\left(\mu_t\right) = h_t\left(\mu_t|\mathcal{F}_t\right)\left[\frac{h_t\left(\mu_t|\mathcal{F}_t\right)}{h_t\left(\mu_t|\mathcal{F}_{t-k}\right)}\right]^{\varphi}. \tag{14}$$

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 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 $\varphi\mathcal{I}$:

$$h_t^{\varphi}\left(\mu_t\right) \sim N\left(m_t^{\varphi}, \sigma_m^2\right), \; m_t^{\varphi} \equiv \mathbb{E}_t^{\varphi}[\mu_t] = m_t + \varphi\mathcal{I}_t, \tag{15}$$

$$\text{where: } \mathcal{I}_t \equiv m_t - \mathbb{E}_{t-k}m_t = K\int_{t-k}^{t} e^{-\theta(t-s)}dB_s. \tag{16}$$

$\mathcal{I}$ is the information the diagnostic agent neglects in forming her background context.[25]

The expected present value of dividends P^* will be the key driving force in the model. Appendix C.1 proves that P^* depends only on D_t, m_t^{φ}, and parameters:

$$P^*\left(D_t, m_t^{\varphi}\right) \equiv \int_{-\infty}^{\infty} \mathbb{E}_t\left[\int_t^{\infty} e^{-\rho(s-t)} D_s ds | \mu_t\right] h_t^{\varphi}\left(\mu_t\right) d\mu_t, \tag{17}$$

where ρ denotes the discount rate.[26]

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. As an example, in Appendix C.5 we provide an alternative foundation for over-optimism in P^* based on a fully rational agent with an over-optimistic prior over the new drift rate distribution. We prefer diagnostic expectations for three reasons. First, they concord with evidence from psychology and are highly tractable. Second, they deliver a key feature of the 2000s housing cycle where rational learning would struggle: The lengths of the boom, bust, and rebound are independent of their amplitude in the cross-section.[27] Third, they accord with empirical evidence that house price expectations did not overshoot on the downside during the bust.[28] This rules out models of learning from prices rather than dividends,

[25]Equations (15) and (16) extend 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 (16) not to necessarily equal mean reversion θ. We prefer the formulation in equation (16) 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.

[26]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.

[27]Rational learning does not have this property because learning occurs more quickly when growth states are further apart.

[28]Most "modern" surveys of house price expectations came too late to study the boom and bust. The exception is Case et al. (2012), who from 2003 to 2012 surveyed homebuyers in Alameda County, CA, Milwaukee County, WI, Middlesex County, MA, and Orange County, CA. They find that in the bust, price expectations generally fall, both at the one year and ten year horizons. However, the 10-year expectations

which create oscillations by having expectations overshoot.

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. We normalize to 0 the dividend and hence the house price from living in the hinterland.[29] 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 as in Kaplan et al. (2020).

The value to a potential entrant from purchasing in the city has a common component, V_t, and an idiosyncratic component, ξ, which creates a downward-sloping demand curve. The common component comprises the expected dividends received while living in the city, $V_t = P^* \left(D_t, m_t^{\varphi}\right)$.[30] The idiosyncratic component ξ is drawn from a Pareto distribution with $P\left(\xi > x\right) = \left(\frac{x_m}{x}\right)^{\gamma}$, such that γ governs the slope of the housing demand curve.

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

never go below 3% annual growth and the one-year expectations only rarely go negative. Furthermore, even when prices are falling rapidly, the vast majority of homebuyers say "it's a good time to buy because prices are likely to increase." Indeed, "in every single survey in every county, the share agreeing with [this] statement was never less than 67 percent, and in most cases it was over 80 percent."

[29] This implies that potential entrants consider only a single city (recall that the notation suppresses the i subscript indexing cities). This assumption arises naturally in a continuous time setting where "offers" from multiple cities never arrive in the same instant. More substantively, the normalization to 0 of the hinterland house price requires that potential entrants do not value the possibility of receiving a future offer to move into a city with a higher idiosyncratic component ξ (introduced below). We may justify this neglect either by assuming that a potential entrant always makes a once-and-for-all decision or by assuming a negative dividend to living in the hinterland that exactly offsets the option value to yield a house price of 0.

[30] We assume 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 consider higher order terms. In the data, the probability of moving across cities is approximately 2%.

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:

$$Q_t = g_H H_t P\left(\xi > \left(P_t + W_t\right)/V_t\right) = g_H H_t x_m^{\gamma} \left[V_t / \left(P_t + W_t\right)\right]^{\gamma}. \tag{18}$$

Without foreclosures and points, this demand equation reduces to equation (4).

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

$$C_t = \left[A H_t^{1/\eta}\right] \exp\left(\frac{I_t - \bar{I}_t}{\chi}\right), \tag{19}$$

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.[31] Equation (19) parallels the empirical specification for housing costs given in equations (1) to (3), with two differences. First, for simplicity we do not separately model land and construction and instead directly set η as the overall long-run elasticity of supply. Second, we introduce a short-run adjustment cost $\exp\left(\left(I_t - \bar{I}_t\right)/\chi\right)$. Economically, this cost may reflect 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. The adjustment cost term disappears at long horizons when $I_t \approx \bar{I}_t$ (even though I_t converges to different values depending on the value of μ), leaving only the long-run supply equation underlying the cross-city framework in Section 3.

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.[32]

[31] We set δ to be 20 years.

[32] 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

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), default requires a "double trigger," consistent with empirical evidence in Foote et al. (2008), Bhutta et al. (2017), Gerardi et al. (2018), Ganong and Noel (2020), Gupta and Hansman (2021), and Gupta et al. (2019).[33] Foreclosed homes enter supply in the instant after the default occurs.[34] The mortgage balance measure density, $g(M,t)$, evolves according to the Fokker-Planck equation:

$$\frac{\partial}{\partial t} g(M,t) = \underbrace{(I_t + \iota) H_t \phi (M/P_t)/P_t}_{\text{New originations}} - \underbrace{\iota g(M,t)}_{\substack{\text{Refis/}\\ \text{foreclosures}}}. \tag{20}$$

Competitive, risk-neutral lenders provide mortgages. These lenders have the same beliefs as buyers.[35] They set points W_t to make zero expected profits on each loan given that they only recover a fraction ψ of a house's value in a foreclosure:

$$W_t = \mathbb{E}_t^{\varphi}\left[e^{-\rho(\tau - t)} \max\left\{M_t - \psi RP_\tau, 0\right\}\right], \tag{21}$$

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

---

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.

[33]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.

[34]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.

[35]Consistent with this assumption, Gerardi et al. (2008) show that lenders during the boom understood the consequences of falling house prices but put little weight on this possibility and Cheng et al. (2014) show that mortgage lenders behaved similar to the rest of the population in their own housing choices. Figure C.1 shows an alternative where lenders have perfect foresight over the path of dividends.

5.2 Equilibrium

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

(i) Buyers' common valuation reflects their beliefs:

$$V_t = P^*\left(D_t, m_t^{\varphi}\right), \tag{22}$$

where $P^*\left(D_t, m_t^{\varphi}\right)$ satisfies equation (17).

(ii) Price equals the marginal cost of construction given by equation (19):

$$P_t = \left[AH_t^{1/\eta}\right]\exp\left(\frac{I_t - \bar{I}_t}{\chi}\right), \tag{23}$$

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

(iii) Lenders make zero expected profits such that W_t satisfies equation (21).

(iv) Housing demand (equation (18)) equals new construction plus foreclosures:

$$g_H H_t x_m^{\gamma}\left[V_t/\left(P_t + W_t\right)\right]^{\gamma} = \dot{H}_t + \iota\int \Phi_R\left(M/P_{t-}\right) g\left(M, t_-\right) dM, \tag{24}$$

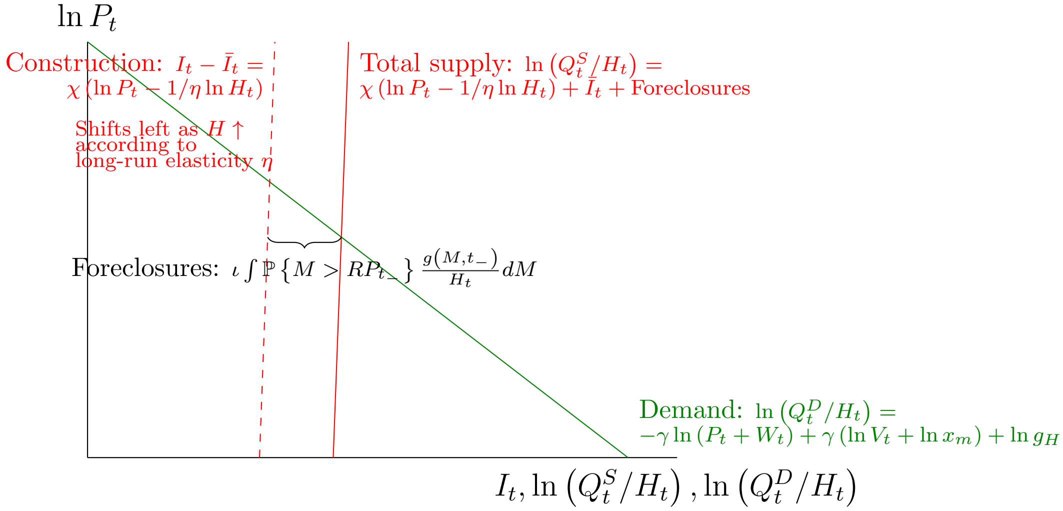
where $\Phi_R(.)$ denotes the cumulative density of a mean-one normal distribution with standard deviation σ_R.

(v) The mortgage density distribution $g(M,t)$ satisfies the Fokker-Planck equation (20).

Equations (23) and (24) 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 8. 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

Figure 8: Supply and Demand Diagrammatic Treatment of Equilibrium



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 (21) 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.

5.3 Calibration

We choose city-level targets for our calibration. Since we compute impulse responses of variables without the idiosyncratic Wiener shocks, we treat the model paths as representing groups of cities. Accordingly, we compute data moments after grouping the CBSAs in our sample into four (1997-population-weighted) quartiles based on their their long-run fundamental from Section 4.3.

We calibrate several parameters externally and set the remaining eight parameters by simulated method of moments (SMM). Table 4 lists the model parameters, their value 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. We first estimate all eight parameters using moments for the third quartile only, as this quartile is representative of the national path of prices and foreclosures. 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. In a balanced growth path, prices grow at the rate μ (see Appendix C.3). 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.[36] We set the diagnostic window, k, to 8 to achieve a boom length of 8 years.

We next turn to externally-calibrated parameters relating to housing supply, mortgages, and foreclosures. We set the long-run supply elasticity, η, to the inverse of the empirically-estimated value from Table 1 obtained by multiplying the column (3) population growth and triple interaction coefficients by the respective endogenous variables

[36]Mean reversion must be slow for a shock at the start of the boom to continue to drive house prices during the rebound. A value of 0.005 means that even 25 years after the boom start, μ_t has declined only $1 - e^{-.005\times 25} \approx 12\%$ of the distance from μ_0 to $\bar{\mu}$. Glaeser and Nathanson (2017) suggest using the first and second autocorrelations of the dividend process as moments to determine the variance and mean reversion. A value of 0.005 (in combination with the calibrated values of σ_D and σ_μ) gives that the first two autocorrelations of the annual change in D are 0.63 and 0.47, which is in the range of empirical values for the autocorrelation of demand changes reported in Glaeser and Nathanson.

Table 4: 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 5: 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	-38.0	-40.1	-41.1	-51.0	-49.0
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

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).[37] 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 in Table 5, with each moment computed separately for each quartile. We treat 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.

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

[37] 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.

the peak annualized foreclosure rate, as shown in Figure 5. 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.

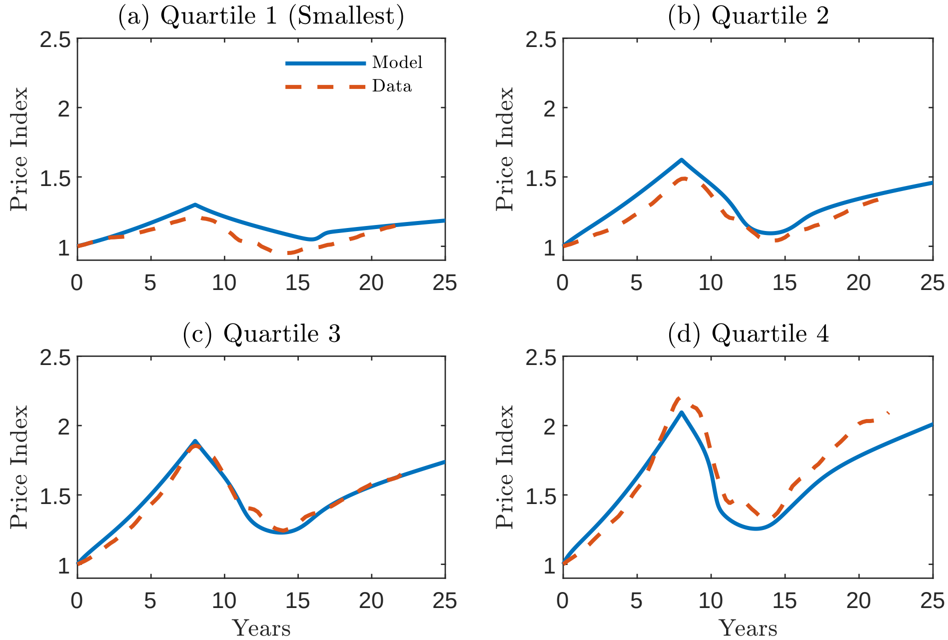
We choose the set of parameters that minimizes the weighted sum of squared residuals between the model and the data moments. We weight all moments equally except the three equity distribution moments, to which we assign 1/3 of the weight of the other moments, and the bust and rebound speed moments, to which we assign 1/2 of the weight of the other moments. We calculate the model moments for the impulse response to a one-time increase from $\bar{\mu}$ to μ_0, using the analytic mean path of beliefs. We define the boom as time zero to the price peak, the bust as the price peak to the price trough and the rebound as the seven years following the price trough given the lack of an end date for the rebound.

5.4 Model Fit

Table 5 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 also fits the remaining quartiles well despite the deep parameters being

Figure 9: 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 4.3.

held fixed in estimation, leading to many fewer degrees of freedom. Figure 9 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 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 increase in the drift rate $\mu_0 - \bar{\mu}$ is roughly 1.3p.p. per year higher than in the pre-boom for all four quartiles, not a radical change. The similarity across quartiles supports our finding in Section 4.5 that the change in dividend growth was common and affected higher fundamental areas more due to convexity. The 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.

The values of χ imply substantially steeper short-run than long-run supply curves. The value of γ of 1.1 aligns well with the estimate in Saiz (2003) of the response of housing rents in Miami to the Mariel boatlift. The liquidity shock ι being in the neighborhood of 5% accords with evidence on mortgage pre-payment rates during the boom, supporting our modeling of a single liquidity shock in both periods.[38]

5.5 Implications for Rent Growth

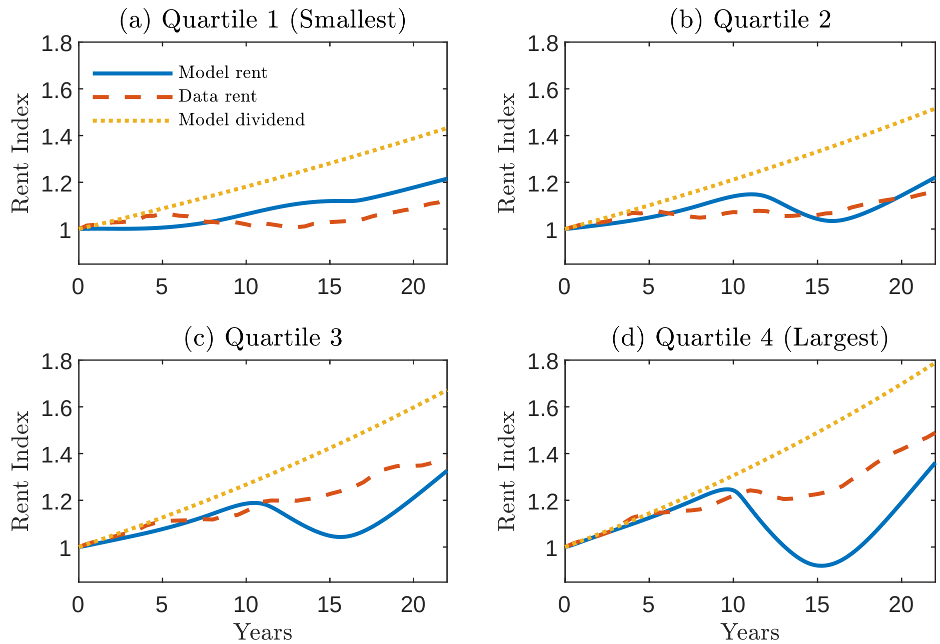
We introduce rents by specifying a user cost indifference relationship as in Section 4.5.[39] Doing so along a transition path and in a manner consistent with important empirical regularities requires a more involved expression than equation (9). We make the following additional assumptions: (i) households re-optimize their rent/own decision with Poisson intensity λ, (ii) during a rental contract signed at date t, rents grow at a fixed rate m_t^{φ}, the nowcast of the dividend drift rate, (iii) households use the path of expectations for prices and dividends at each point in time. These assumptions capture the long-term nature of rent/own decisions, the stickiness of contract rents, and the volatility of expected house price appreciation off the balanced growth path. In Appendix C.4, we show that the path of average rent R_t is characterized by a "reset" rent $R_{t|t}$ for contracts signed at date t and laws of motion for R_t and the average within-contract growth rate g_t (see equations (C.12) to (C.14)). We parameterize $\lambda = 1/6$ to match tenure rates and choose ζ, the proportional maintenance and property tax component of owning, to match an average price-rent ratio over the period of 13 from Begley et al. (2021).

Figure 10 shows rents in the model and data. The model explains rent growth in

[38]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.

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

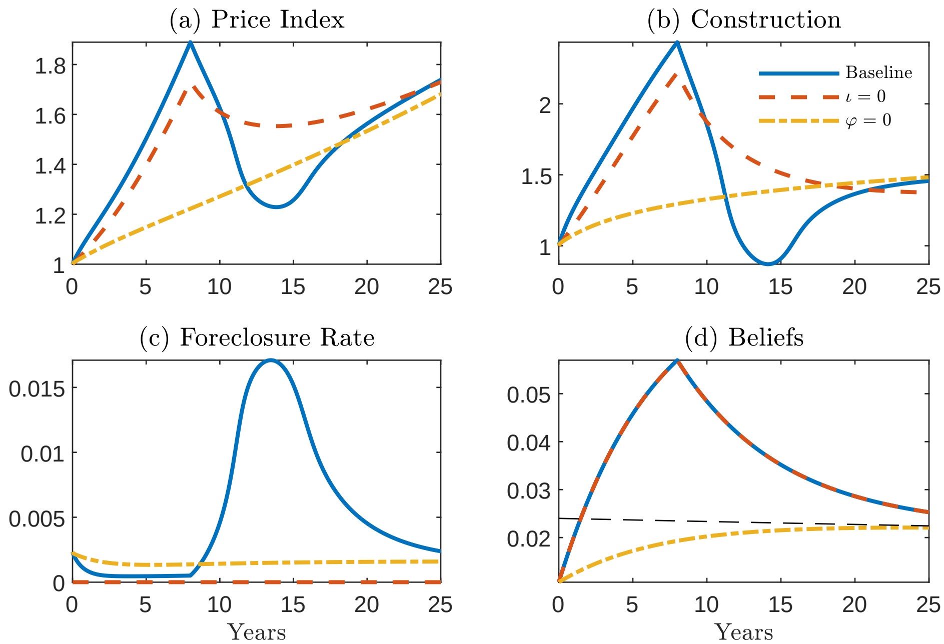
Figure 10: Quartile Rent Paths in Model and Data



Notes: The solid line plots model simulated rents for each quartile. The dashed line figure plots indexes of annual average rent deflated by the national GDP price index for 22 CBSAs with CPI rent data available from 1988, grouped into quartiles based on the long-run fundamental. The dotted line plots dividends.

each quartile over the entire 1997-2019 period reasonably well despite the fact that this moment is untargeted. The fit is within 0.2 p.p. per year for the middle two quartiles while slightly over-predicting rent growth in quartile 1 and under-predicting rent growth in quartile 4. Notably, rent growth lags dividend growth over the full period (dotted gold line), reflecting the decline in the rent-price ratio when expected price growth rises (see Lemma 1). The model features a counterfactually large rent drop in the bust, particularly in quartiles 3 and 4. This occurs because of the significant overshooting of prices due to foreclosures, as low prices and high expected price growth both drag down user-cost implied rent. Factors outside of the model such as downward stickiness of rents or a failure of rent-price arbitrage in the crisis could explain why rents in the data do not fall as much. Overall, the model's ability to explain rent growth across quartiles over the full BBR supports a fundamentals-based explanation of the path of prices.

Figure 11: 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.

5.6 Unpacking the Mechanism

We now discuss the model features that generate a boom-bust-rebound. For parsimony, we use the parameters estimated for the third quartile. Figure 11 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 the increase in the capitalization rate $m_t - \rho$ and the increase over time in the dividend.

Diagnostic learning generates an over-shooting of beliefs about the drift rate μ, as shown in Panel (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 or in models of price extrapolation that generate overshooting of expectations on the downside, with diagnostic expectations the convergence is gradual and monotonic. The over-shooting of diagnostic beliefs in the boom and gradual convergence in the bust have a natural counterpart in survey evidence on beliefs about house price growth from Case et al. (2012).[40]

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

Foreclosures generate a much larger bust in prices, to below the level that would prevail with rational learning.[41] The over-shooting occurs because of a price-foreclosure spiral

[40] In the Case et al. (2012) house price expectations surveys discussed previously, annualized price growth over 10-years exceed 10% in all four counties during the boom before falling gradually to between 3 and 5 percent by 2012. If we calculate 10 year price expectations, we obtain a similar qualitative pattern but quantitatively expectations do not peak as high, settle lower in line with the balanced growth level of price growth, and peak at the trough of the bust when the agents expect a rebound.

[41] 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

(Guren and McQuade, 2020): foreclosures add to housing supply, which further depresses prices, putting more owners under-water, leading to more foreclosures.[42] Reflecting this 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.

5.7 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 (18): $\ln\left(Q_t^D/H_t\right) = -\gamma\ln\left(P_t + W_t\right) + \gamma\left(\ln V_t + \ln x_m\right) + \ln g_H$. 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.[43]

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.

[42] 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.

[43] 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

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 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 (18) 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.[44] On the other hand, the rise in rents shown in Figure 6 militates against attributing the initial shift out of the demand curve primarily 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

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.

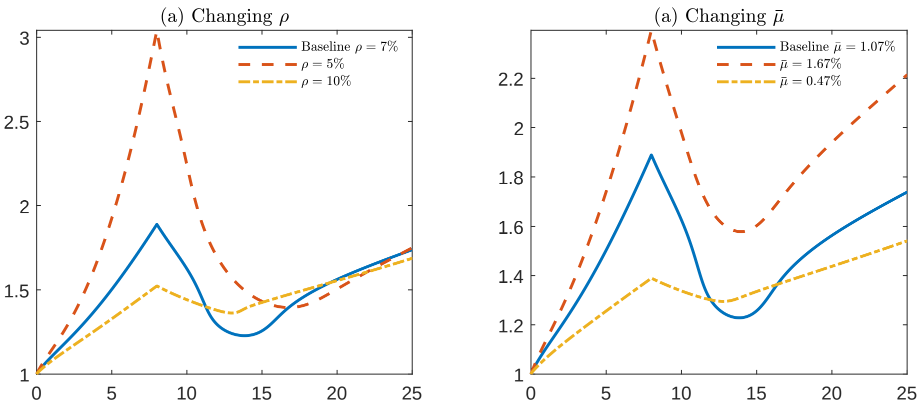
[44] In related work, Foote et al. (2021) argue that without extensive margin credit expansion, the distribution of mortgage credit would have counterfactually shifted toward high income groups. Appendix Figure B.8 shows that the long-run fundamental is essentially uncorrelated with the change in subprime share during the boom, suggesting the role of other types of credit relaxation (e.g. low interest rates that ease payment-to-income constraints) and rising incomes in keeping house prices affordable in high fundamental areas. Notably, 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.

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 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.

5.8 Implications For When Strong Housing Cycles Are Likely

An important implication of our analysis is that housing cycles like the 2000s cycle are more likely when interest rates are low or initial dividend growth is high due to the convexity of $V_{i,t}$. In these conditions, the present value of dividends is dominated by terms further out in the future, so a shock to the dividend growth rate has a larger effect on $V_{i,t}$. To illustrate, Figure 12 shows price paths with alternate values of the discount rate ρ or the initial growth rate of dividends $\bar{\mu}$, holding fixed the increase in μ at time 0 and all other parameters as estimated for the third quartile of CBSAs. The red dashed line in panel (a) shows that a lower ρ results in a much larger boom-bust due to the higher sensitivity of $V_{i,t}$ to changes in μ, while the gold dash-dot line shows a dampened cycle with higher ρ. Thus, while Section 4.5 expressed some skepticism about the role of *changes* in discount rates during the 1997-2019 period, our analysis confirms the importance of the low *level* of discount rates throughout the 2000s. Panel (b) shows that raising $\bar{\mu}$ while keeping the jump in μ fixed causes a larger boom while lowering $\bar{\mu}$ causes a smaller boom. Interestingly, the bust size is not monotonic, as higher trend growth counteracts the correction in beliefs.

Figure 12: Changing ρ and $\bar{\mu}$



Notes: This figure shows the price index normalized to 1 at time zero for a 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. In panel (a), the dashed red line leaves all parameters unchanged but drops the discount rate ρ from 7% to 5%, while the dash-dot gold line raises ρ to 10%. In panel (b), the dashed red line leaves all parameters unchanged but raises the initial drift rate of dividends $\bar{\mu}$ from 1.07% to 1.67%, while the dash-dot gold line lowers $\bar{\mu}$ to 0.47%. In doing so, we adjust μ_0 so that the model has the same jump in μ just different initial growth rates.

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-Kindlebergerian 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 exogenously-changing 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. Conversely, the consequences of over-optimism appear most dire when initial price growth is high and interest rates are low. Of course hindsight is 20-20; distinguishing fundamental growth from over-optimism in real time rather than after observing a full boom-bust-rebound poses a formidable task. Nonetheless, our findings imply that policy makers should heed Kindleberger's dictum that essentially all manias are to some degree grounded in fundamentals.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017). "Stock Price Booms and Expected Capital Gains". *American Economic Review* 107 (8): 2352–2408.

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.

— (2020). *Credit Supply and House Prices: Evidence from Mortgage Market Segmentation.*

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

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.

Autor, David H. (2019). "Work of the Past, Work of the Future". *AEA Papers and Proceedings* 109: 1–32.

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.

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.

Barberis, Nicholas (2018). "Psychology-Based Models of Asset Prices and Trading Volume". In: *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 1. Elsevier: pp. 79–175.

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

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

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.

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

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.

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.

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.

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

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 Michael G. Palumbo (2008). "The Price of Residential Land in Large U.S. 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.

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

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". Working Paper, Federal Reserve Bank of Boston.

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.

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

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

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

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

Glaeser, Edward L., Joshua D. Gottlieb, and Joseph Gyourko (2012). “Can Cheap Credit Explain the Housing Boom?” In: Glaeser, Edward L. and Todd Sinai. *Housing and the Financial Crisis*. University of Chicago Press: pp. 301–359.

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

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

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

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jon Steinsson (2020). “Housing Wealth Effects: The Long View”. *The Review of Economic Studies* 88 (2): 669–707.

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 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.

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

Howard, Greg and Jack Liebersohn (2021a). "Regional Divergence and House Prices". Working Paper, University of Illinois.

— (2021b). "Why is the rent so darn high? The role of growing demand to live in housing-supply-inelastic cities". *Journal of Urban Economics* 124: 103369.

Jacobson, Margaret M. (2020). "Beliefs, Aggregate Risk, and The U.S. Housing Boom". Working Paper, Federal Reserve Board of Governors.

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.

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.

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.

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.

— (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.

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.

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 (2003). "Room in the Kitchen for the Melting Pot: Immigration and Rental Prices". *The Review of Economics and Statistics* 85 (3): 502–521.

— (2010). "The Geographic Determinants of Housing Supply". *The Quarterly Journal of Economics* 125 (3): 1253–1296.

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

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

January 2022

This online appendix is split into four sections. In Section A, we microfound the construction and land costs in the empirical framework in Section 3. In Section B, we present additional empirical results on the role of investors and speculators, on the Bartik instrument, on the robustness of our IV results, and a number of additional empirical results. In Section C, we present a number of model derivations and proofs. Finally, in Section D, we detail the computational methods used to solve the model.

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:

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

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

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

A.2 Land

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:

$$L_{i,t}(\tau) = \int_t^{\infty} e^{-\rho(s-t)} \kappa_i \left(e^{\frac{I_i}{2}(s-t)} \Phi_{i,t} - \tau \right)^{\chi_i} ds$$

$$= \kappa_i \left(\frac{\Phi_{i,t}^{\chi_i}}{\rho - \chi_i I_i/2} - \frac{\tau^{\chi_i}}{\rho} \right). \tag{A.7}$$

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:

$$L_{i,t} = \frac{1}{H_{i,t}} \int_0^{\Phi_{i,t}} L_{i,t}(\tau) \Lambda_i 2\pi\tau d\tau$$

$$= \kappa_i \left(\frac{1}{\rho - \chi_i I_i/2} - \frac{1}{\rho(\chi_i/2+1)} \right) \left(\frac{H_{i,t}}{\Lambda_i \pi} \right)^{\frac{\chi_i}{2}}$$

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

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

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

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. Furthermore, 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, both raise average land prices in a city.

B Empirical Appendix

This section of the appendix presents additional empirical results. Appendix B.1 provides an analysis of investors and speculators and their role in the housing cycle, fleshing out the argument in Section 5.7. Appendix B.2 provides details on the construction of the Bartik instruments. Appendix B.3 provides a robustness analysis of our main IV results. Finally, Appendix B.4 provides a number of additional empirical results.

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 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]

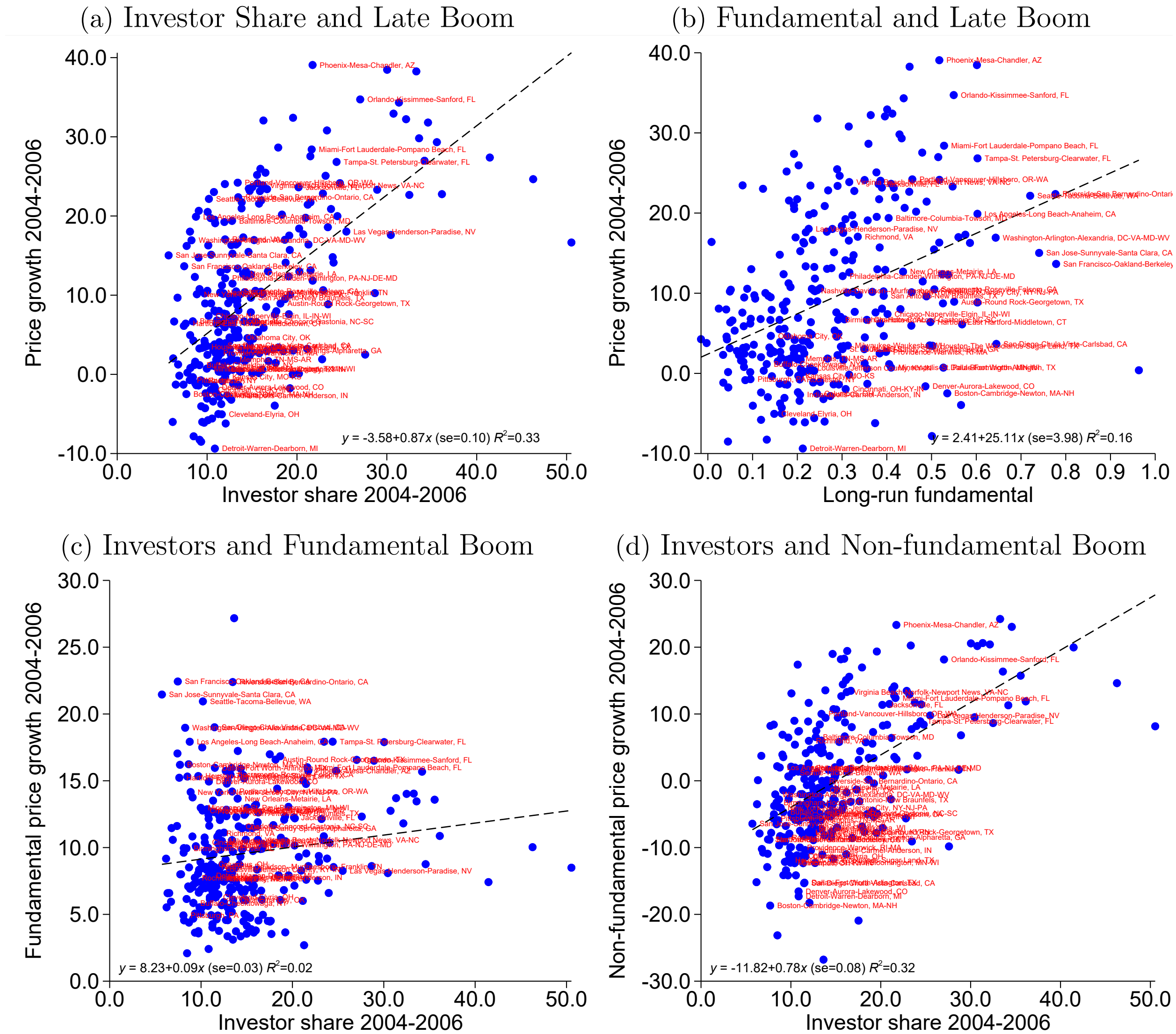
Panel (b) shows that house price growth during 2004-2006 also correlates strongly with the long-run fundamental, measured as usual as the fitted value from the second stage regression in column (3) of Table 1. Thus, even at the end of the boom when speculative activity was plausibly most rampant, long-run fundamentals continue to explain house price growth.

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

[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.

[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' Role in the Late Boom and Relation to Fundamentals



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 ($\times$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

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

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 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.

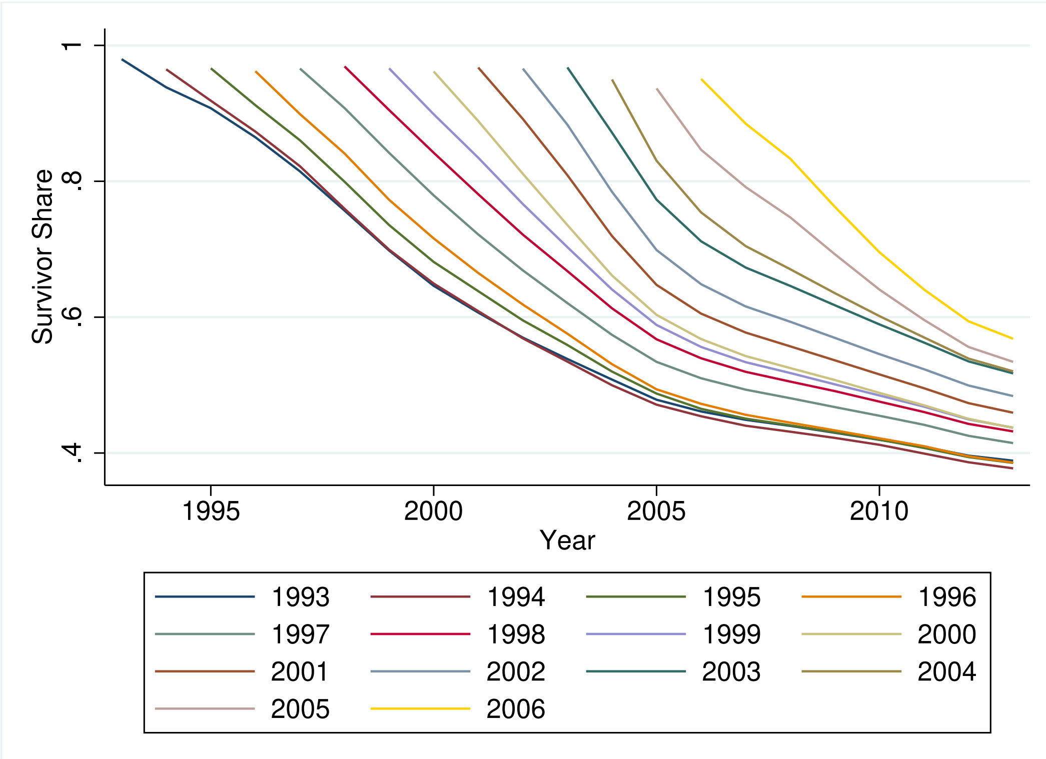
Panel (d) plots the investor share against the non-fundamental component of house price growth in 2004-2006. Areas with late price booms not explained by their long-run fundamental had higher investor shares of purchases, and this relationship explains essentially all of the overall correlation shown 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 below their long-run level in the bust. Thus, while adding speculators to our model could partially substitute for diagnostic behavior in amplifying the boom, it 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

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.

magnitude of the boom, indicating a larger role for speculation in driving the boom-bust cycle than in the typical area. This pattern is representative; a regression of 2006-2019 house price growth on the 2004-2006 investor share indicates 5 percentage points lower house price growth in the bust-rebound for each 10 percentage points higher investor share (Table B.1 column (7) less sum of columns (1) and (4)).

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 2 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.

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 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. For area-industries with suppressed wage data, we replace $w_{-i,j,t}$ with $w_{j,t}$, where $w_{j,t}$ is the national wage in industry j at date t.

B.3 IV Robustness

Table B.2 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 1.

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

Table B.2: 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 1. Coefficients in bold font are statistically different from 0 at the 5% level.

similar results.[3] Row (5) replaces the land unavailability instrument with the measure 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

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

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, 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.[4]

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).[5] The JIVE and bias-adjusted estimators address a particular concern that many instruments will over-fit the first stage, biasing the second stage toward OLS (Bekker, 1994; Bound et al., 1995). 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 B.4 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. 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).

[4]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.

[5]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. 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.

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.

B.4 Additional Empirical Results

This section presents 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. This analysis is discussed in Footnote 7 and motivates how we choose the timing of the start of the boom, bust, and rebound.

Table B.3 reports summary statistics for our CBSA-level data set.

Table B.4 reports first-stage-type results. These results, which are discussed in Section 4.1, show the instruments are strong and enter with the expected sign. The final column shows the reduced form of the instruments (uninteracted) on house price growth over the BBR, which corresponds to Figure 3.

Figures B.5 and B.6 repeat Figure 3 using only the supply or demand instruments separately and show that both contribute to higher long-run growth and to the boom-bust-rebound pattern. As discussed in Section 3.3, this is important because it shows our results do not only reflect differences in the supply elasticity.

Figure B.7 reproduces the analysis in Figure 4, in which 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, adding a control for the pre-1997 cyclical sensitivity of each CBSA based on Guren et al., 2020. One can see that adding this control does not affect our conclusion that a high fundamental predicts a boom-bust-rebound pattern, suggesting that this is not driven solely by more cyclically-sensitive cities having higher

Table B.3: 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

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

exposure to the 2000s cycle.

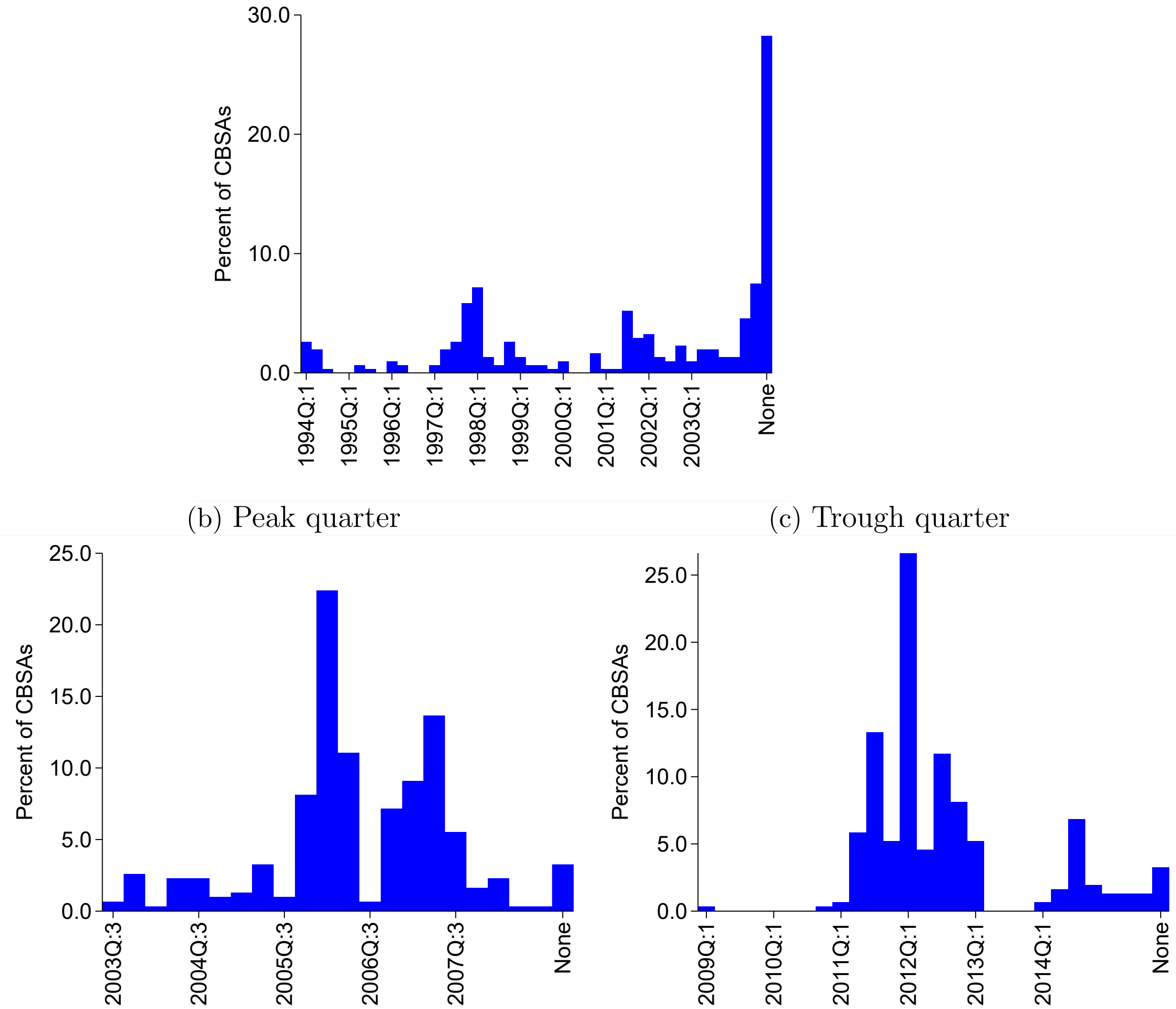
Finally, Figure B.8 displays the (non-)correlation of the long-run fundamental with measures of non-prime lending, as discussed in Section 5.7.

Figure B.3: Local Peak and Trough Timing

(a) Boom start quarter

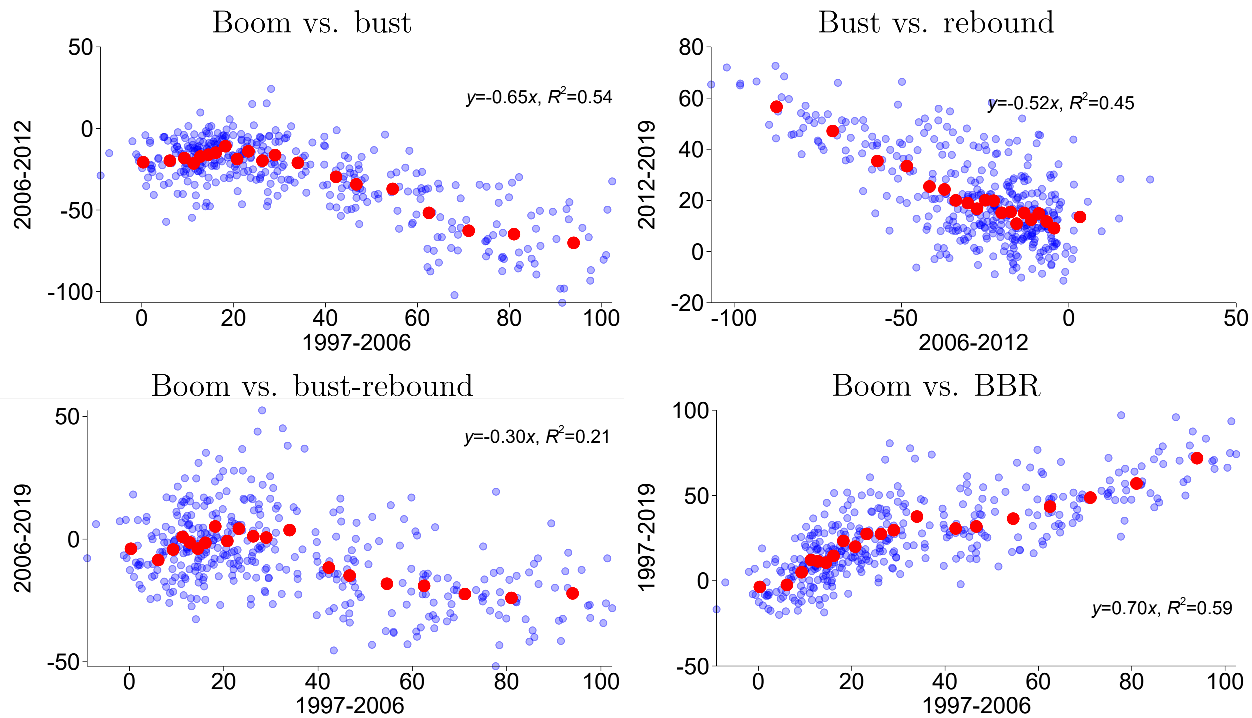
(b) Peak quarter

(c) Trough quarter



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 structural break test of Bai and Perron (1998, 2003) as implemented in Ditzen et al. (2021). Areas without a break identified at the 95% confidence level are shown in the bar labeled "None." 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



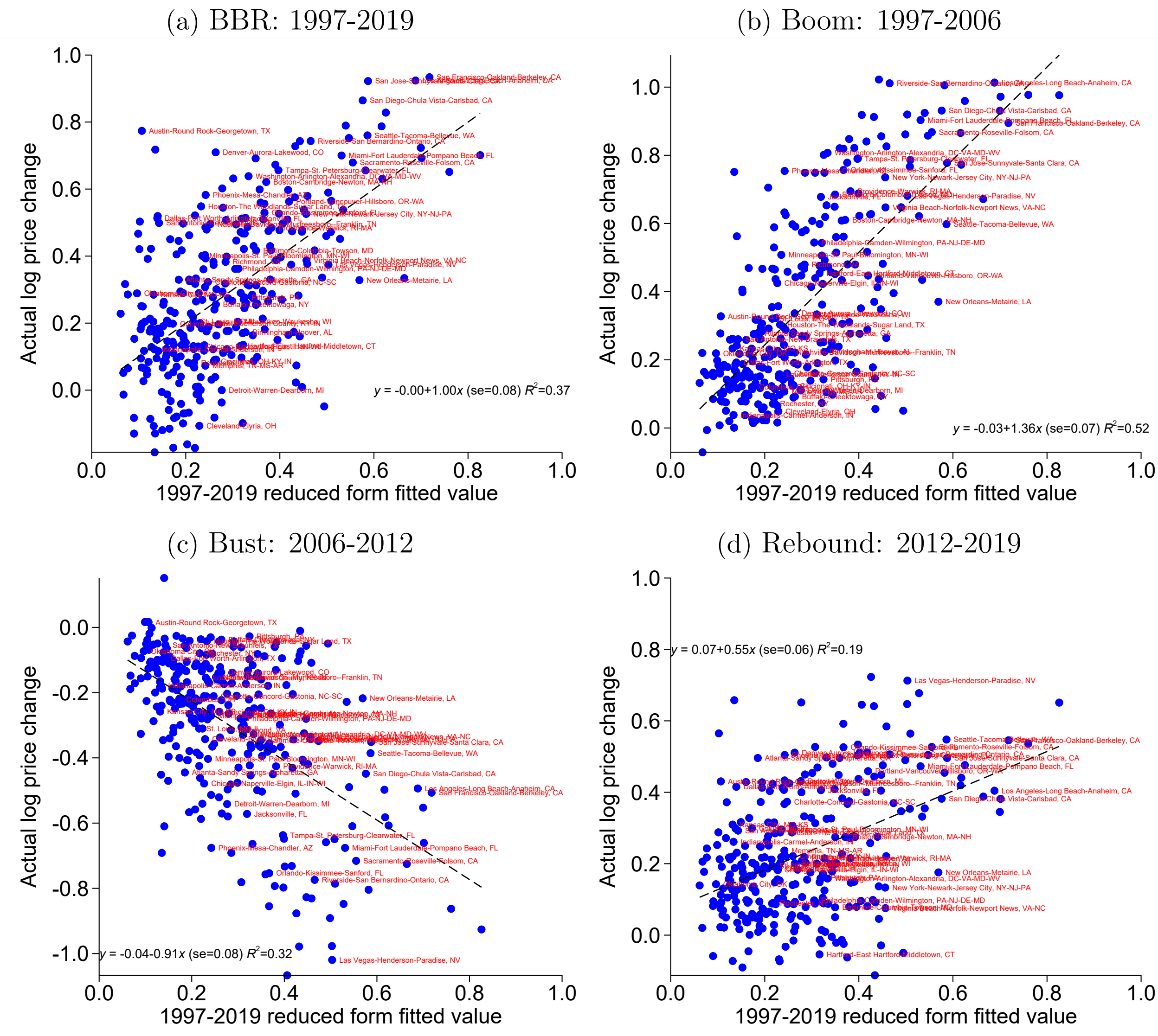
Notes: Each blue circle represents one CBSA. The red circles show the mean value of the y-axis variable for 20 bins of the x-axis variable. CBSA house price data from Freddie Mac deflated using the national GDP price index.

Table B.4: 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

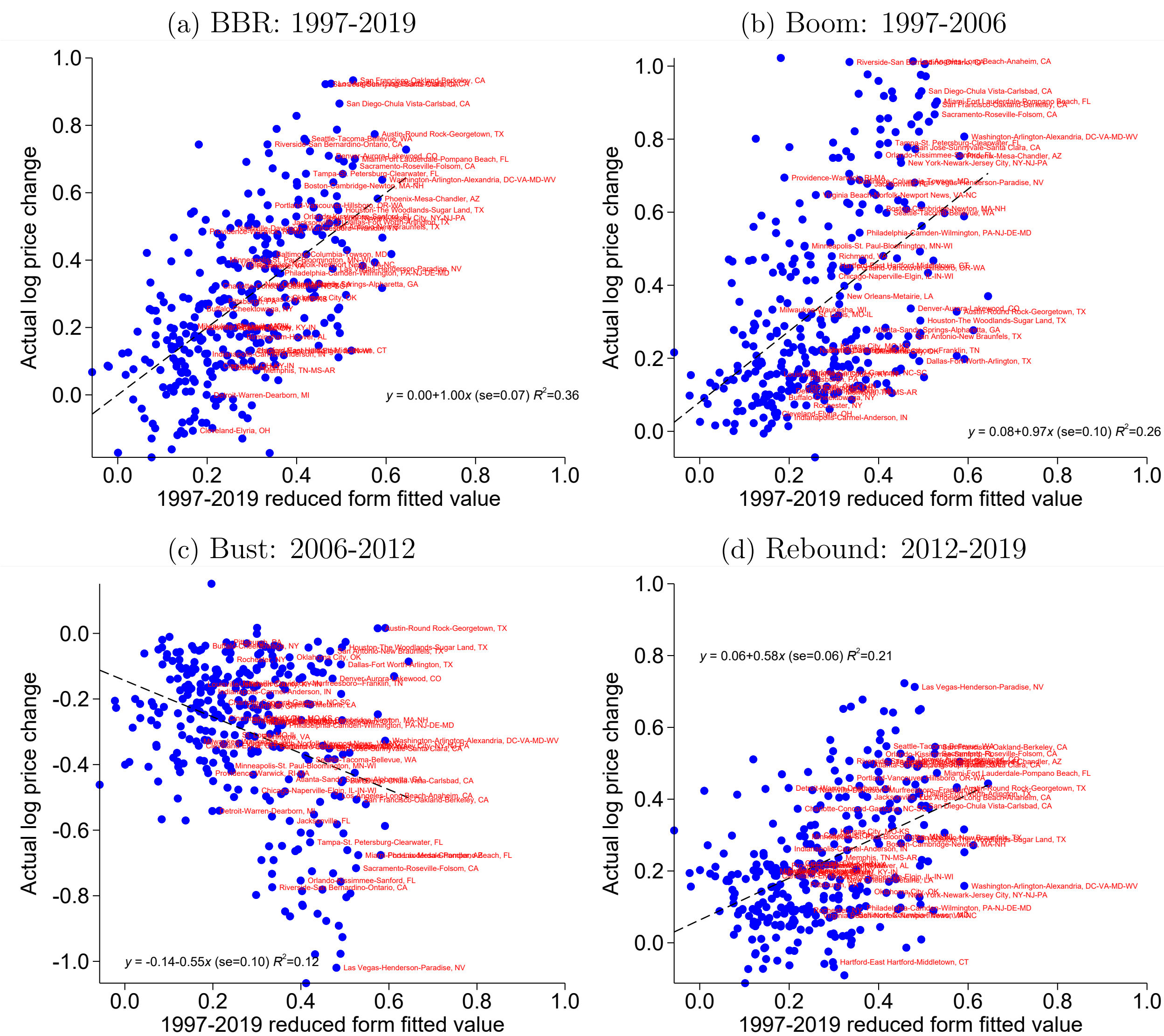
Notes: Columns (1), (3), and (5) report regressions of an endogenous variable on the group of excluded instruments associated with that variable. Columns (2), (4), and (6) report regressions of an endogenous variable on all excluded instrument main effects. Column (7) reports the reduced form regression of house price growth on all excluded instrument main effects. Heteroskedastic-robust standard errors in parentheses. The effective F-statistic is computed as in Montiel Olea and Pflueger (2013). **,*,$^{+}$ denote significance at the 1, 5, and 10 percent levels, respectively.

Figure B.5: Correlation with Long-run Fitted Value, Land Share and Regulation Instruments Only



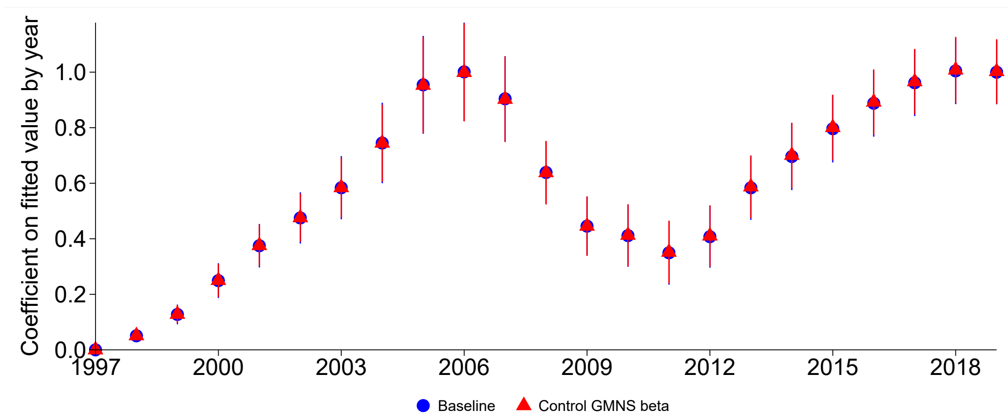
Notes: In each panel, each blue dot is the real house price growth in a CBSA over the period indicated on the vertical axis plotted against the predicted real house price growth over the period 1997-2019 using only the excluded instruments associated with land share and WRLURI (land unavailability, 1997 population density, non-traditional Christian share, and public expenditure on protective inspection). CBSAs with more than 1 million persons in 1997 are labeled in red.

Figure B.6: Correlation with Long-run Fitted Value, Population Growth Instruments Only



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 using only the excluded instruments associated with population growth (employment and wage shift-shares, climate variables, restaurant density, and college share). CBSAs with more than 1 million persons in 1997 are labeled in red.

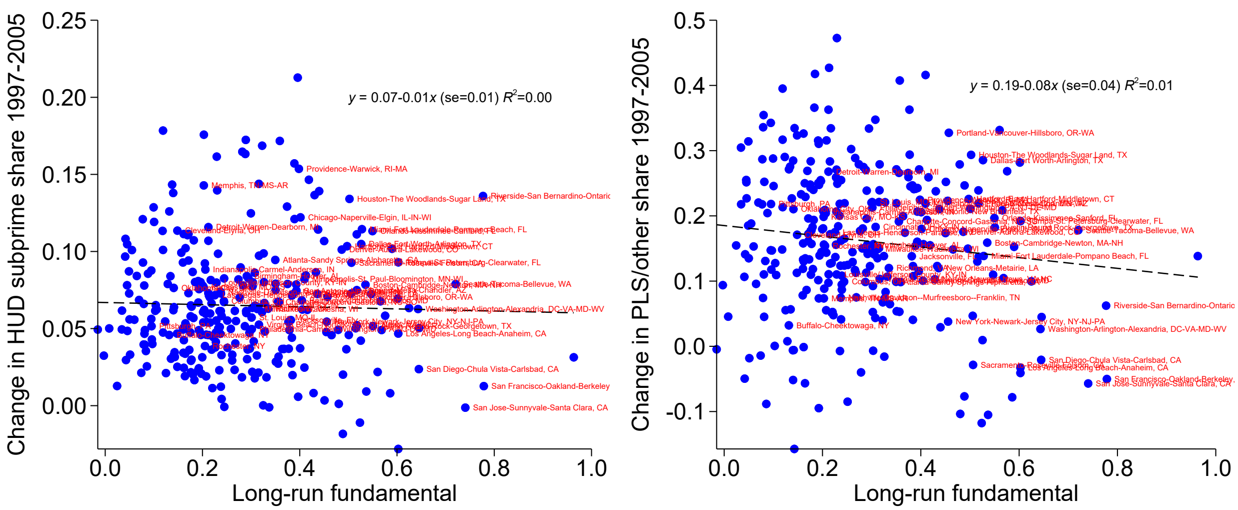
Figure B.7: Fundamentals and House Prices Over Time, Controlling for Cylical Sensitivity



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 $p_{i,t,t+h} = \beta_{0,h} + \beta_{1,h}\hat{p}_{i,t} + \gamma_h x_i + \nu_{i,h}$, where $\hat{p}_{i,t}$ denotes the second stage fitted value from column (3) of Table 1. The blue dots show the baseline coefficients when x_i is empty. The red dots show the coefficients when x_i is the local sensitivity estimated using the procedure in Guren et al. (2020). In particular, we use GMNS' code to create city-level sensitivities of city house prices to regional house prices controlling for city-level income and economic activity and city-level exposures to income and economic activity, based on equation (4) in their paper. This regression is run on data from 1978 to 1996 so that the 2000s housing cycle does not affect it.

Figure B.8: Non-Prime Credit and Fundamentals

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



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

C Model Appendix

This section of the appendix provides derivations and proofs for the model in Section 5. Appendix C.1 derives the present value of dividends. Appendix C.2 provides an analytic path of beliefs that we use in calculating impulse responses. Appendix C.3 defines the balanced growth path of our model. Appendix C.4 derives the user cost for rents and proves Lemma 1. Appendix C.5 discusses an alternative to diagnostic expectations for belief formation. Finally, Appendix C.6 derives and shows impulse responses for a version of the model where lenders have perfect foresight rather than the same diagnostic beliefs as the other agents in the economy.

C.1 Present Value of Dividends

We restate equation (17) for convenience:

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

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 $e^{-\rho(s-t)} D_s = D_t \exp\left(-\rho\left(s-t\right) - \frac{1}{2}\sigma_D^2\left(s-t\right) + \int_t^s \mu_\tau d\tau + \sigma_D \int_t^s dW_{D,\tau}\right)$. Taking an expectation:

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

One can show:

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

giving:

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

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:

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

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 (15), 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 (10) and (13) into equation (12), we have:

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

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 (11) 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 (10), (13) and (16), we have:

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

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

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

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 Balanced Growth Path

A balanced growth path consists of a fixed μ, constant rate of construction and foreclosures, and constant ratio of points to price. Let δ^f denote the balanced growth path ratio of foreclosures to population and ω denote the balanced growth ratio of points to price. Substituting this notation and the balanced growth definition into the market-clearing condition (24) gives:

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

Dividing through by H_t and recognizing that a constant rate of construction means $\dot{H}_t / H_t = I_t$ is constant on the balanced growth path, it is apparent that the ratio V_t/P_t is constant on the balanced growth path. Since V_t grows at rate μ (see e.g. the valuation function (C.1)), along a balanced growth path prices also grow at the rate μ.

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

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

confirming that the LTV distribution is stable on the balanced growth path.

C.4 User Cost Derivation and Proof of Lemma 1

C.4.1 Derivation of User Cost

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

With Poisson intensity λ a household gets a Calvo-like opportunity to re-optimize the rent-own decision. Let $R_{t+\tau|t}$ denote the rent paid in period $t + \tau$ for a contract signed in period t. An agent must be indifferent between renting and owning, where owning involves an outlay of the down payment $(1 - \phi) P_t$ at date t plus the discounted sum of interest payments $i_t \phi P_t$, other costs of owning such as maintenance or property taxes that are assumed to be proportional to the dividend, ζD_t, and expected cash-flow at sale $\mathbb{E}_t^{\varphi} [P_{t+j}] - \phi P_t$, where the notation $\mathbb{E}_t^{\varphi}$ indicates that expectations are taken using the diagnostic measure.[6] We have:

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

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

$$
\begin{aligned}
\int_0^{\infty} e^{-(\rho+\lambda)\tau} \mathbb{E}_t^{\varphi} \left[R_{t+\tau|t} \right] d\tau = &\left(1 + \phi \left(\frac{i_t + \lambda}{\rho + \lambda} - 1 \right) \right) P_t \\
&- \lambda \int_0^{\infty} e^{-(\rho+\lambda)\tau} \mathbb{E}_t^{\varphi} [P_{t+\tau}] \, d\tau + c_t \zeta D_t. \qquad \text{(C.11)}
\end{aligned}
$$

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

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

Equation (C.11) involves an indeterminacy since only the expected discounted present value of rents is pinned down by the expected cost of owning. We resolve this indeterminacy by assuming a contract where rents grow at a rate $g_{t|t}$ until the λ reoptimization shock hits: $R_{t+\tau|t} = e^{g_{t|t}\tau}R_{t|t}$ and $\int_0^\infty e^{-(\rho+\lambda)\tau}R_{t+\tau|t}d\tau = R_{t|t}/\left(\rho+\lambda-g_{t|t}\right)$. This assumption captures the empirical regularity that contract rents are sticky but grow at some trend rate. We assume in particular that $g_{t|t} = m_t^\phi$, the nowcast of the drift rate, but provide the derivation for a general $g_{t|t}$. Imposing this assumption on equation (C.11), the "reset" rent can then be written as:

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

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

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

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

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

Equation (C.13) says that the change in rent consists of the average growth rate of non-resetting rents plus the product of the Poisson intensity λ and the difference between the reset rent and average rent. Equation (C.14) says that the change in the average growth rate of non-resetting rents is equal to the product of the Poisson intensity λ and the difference between the reset growth rate and the average growth rate.

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

Along a balanced growth path with dividends growing at rate μ, $g_t = \mu$, and $i = \rho$, equations (C.12) to (C.14) imply:

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

Equation (C.15) coincides with equation (9) in the main text.

C.4.2 Proof of Lemma 1

Recall that along a balanced growth path we have that $P_t^{\mathrm{BGP}} = AH_t^{1/\eta}$ and $I_t H_t = g_H H_t x_m^{\gamma}\left(\frac{V_t}{P_t}\right)^{\gamma}$.[7] Using $\dot{V}_t/V_t = \dot{P}_t/P_t = \mu_t$ and $\dot{H}_t/H_t = \eta\mu_t$ and explicitly accounting

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

for the costs of owning ζD_t, the Gordon Growth representation is:

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

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

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

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

C.5 Alternative Belief Formation

Diagnostic expectations have several attractive features for our analysis: They are a simple departure from rational learning, they are easily parameterized and are motivated by evidence in the behavioral economics literature, 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 after overshooting beliefs gradually decline towards the rational belief without overshooting again or oscillating as in the data. 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:

$$d\pi_i(t) = \varphi\left(f_i - \pi_i(t)\right) dt + \pi_i(t)\left(\mu_i - \bar{\mu}(t)\right) \sigma_D^{-1} d\tilde{W}(t), \tag{C.18}$$

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.18) 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 $P^*\left(D, \Pi\right)$, where Π denotes the current set of beliefs, satisfies:

$$P^*\left(D, \Pi\right) = E\left[\int_t^{\infty} e^{-\rho s} D(s) | \mathcal{F}\left(t\right)\right] = \sum_{i=1}^{n} \frac{\pi_i D}{\left(\rho + \varphi - \mu_i\right)\left(1 - \varphi\Delta\right)}, \tag{C.19}$$

where $\Delta \equiv \sum_{i=1}^{n} \frac{f_i}{\rho + \varphi - \mu_i}$. To interpret equation (17), 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, $E_i\left[D/\left(\rho - \mu_i\right)\right]$, across possible drift rate regimes. The additional terms in equation (17) account for the possibility of regime shifts.

To connect equation (C.19) to equation (17), 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.

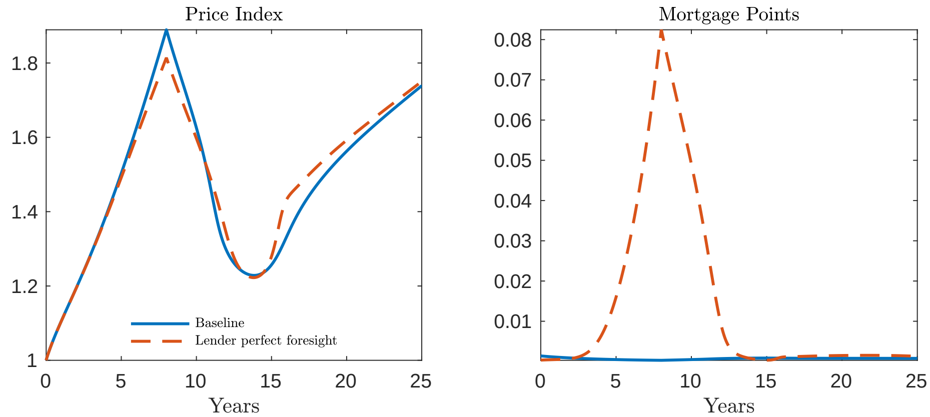
While qualitatively similar, this framework has three important shortcomings relative to diagnostic expectations. First, achieving a realistic length of the bust requires relatively fast learning, which in turn requires an extreme calibration of f_3 (over-optimism) to sustain an eight year boom. Second, the interaction between the size of the fundamental change (the relative values of μ_1, μ_2, and μ_3) and the speed of learning means that achieving a similar length of boom and bust across cities with different changes in fundamentals requires varying the belief and learning parameters. Diagnostic expectations do not require this additional degree of freedom in arguably deep learning parameters in a knife-edge way. Third, unlike diagnostic expectations there is no analytic result for the mean path of beliefs, which necessitates an additional simulation step and slows down the estimation routine.

C.6 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 43, 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

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



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

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.

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:

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

where i is the arrival rate of liquidity shocks and $\phi\left(\cdot\right)$ 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 $(N-1)$ 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:

$$\sum_{i=0}^{N} \alpha_i' (t) T_i (M_j) = (I_t + \iota) H_t \phi (M_j/P_t) / P_t - \sum_{i=0}^{N} \iota \alpha_i (t) T_i (M_j)$$

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

$$\mathcal{A}^* T_i (M) = \iota \alpha_i (t) T_i (M)$$

we have:

$$\boldsymbol{\alpha}' (t) = T (\boldsymbol{M})^{-1} [\mathcal{A}^* T (\boldsymbol{M}) \boldsymbol{\alpha} (t) + (I_t + \iota) H_t \phi (\boldsymbol{M}/P_t) / P_t],$$

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:

$$
\begin{aligned}
&S_1 = \{0\} \\
&S_2 = \{0, -1, 1\} \\
&S_3 = \left\{0, -1, 1, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\} \\
&S_4 = \left\{0, -1, 1, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{-\sqrt{2+\sqrt{2}}}{2}, \frac{-\sqrt{2-\sqrt{2}}}{2}, \frac{\sqrt{2-\sqrt{2}}}{2}, \frac{\sqrt{2+\sqrt{2}}}{2}\right\},
\end{aligned}
$$

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 $f\left(\mathcal{H}^{d,\mu}\right), \Upsilon_n\left(\mathcal{H}^{d,\mu}\right) \in \mathbb{R}^{\left|\mathcal{H}^{d,\mu}\right|}$ 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.

References

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

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

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

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

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

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

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

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

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

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

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

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

Eckert, Fabian, Teresa Fort, Peter Schott, and Natalie Yang (2020). "Imputing Missing Values in the US Census Bureau's County Business Patterns". National Bureau of Economic Research Working Paper No. 26632.

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

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jon Steinsson (2020). "Housing Wealth Effects: The Long View". *The Review of Economic Studies* 88 (2): 669–707.

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

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

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.

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